Quantum algorithms for quantum chemistry



Daniel Marti Dafcik

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2025

Abstract

Computer simulation is central to understanding the behaviour of matter. Yet, more than a century after the development of quantum theory, first-principles simulation of chemistry remains a daunting challenge, owing to the exponential growth in memory required to represent interacting quantum particles such as electrons in molecules.

Quantum computing offers, in principle, a fundamentally more efficient paradigm for simulating quantum systems, with quantum chemistry often cited as a flagship application. However, many quantum algorithms face intrinsic limitations, and any genuine quantum advantage must surpass the strong baseline set by decades of progress in classical algorithms and hardware. While classical methods often suffice for simulating weakly correlated systems, systems with many strongly correlated electrons remain a challenge for both classical and quantum approaches.

In this thesis, we develop quantum algorithms for solving the time-independent Schrödinger equation, with a focus on molecular systems with strongly correlated electrons. We introduce a framework for identifying a compact set of states that parametrise the many-body wavefunction by capturing the correlation due to spin coupling between localised electrons. This representation requires very few variational parameters but incurs a very high memory cost on classical computers. To overcome this, we design quantum circuits that efficiently prepare such states by exploiting their symmetry, encoding the dominant entanglement structure of electronic wavefunctions into the initial qubit states. Through theoretical analysis and classical simulations of quantum computers, we show that this substantially reduces the quantum resources required to accurately compute ground- and excited-state properties.

Finally, we develop an approach for approximating the optimal ansatz within the variational quantum eigensolver, an algorithm designed for near-term quantum devices. Our method constructs wavefunctions from products of symmetry-preserving unitary operators and couples discrete and continuous optimisation to yield highly compact representations of strongly correlated electrons using shallow quantum circuits.

Together, these results establish complementary directions for embedding problem-specific structure into quantum simulation algorithms, advancing the practical simulation of strongly correlated electrons.

Acknowledgements

Now that I work outside academia, I have come to appreciate more fully what makes it special: the freedom to pursue questions for their own sake, to travel, to meet curious people from around the world, and to exchange ideas without immediate practical demands. I am grateful to everyone who gave their time, attention, and support along the way.

David, thank you for being a kind, patient, and supportive advisor. I have greatly valued your ability to cut through hype and approach complex problems from first principles, your insistence on developing physical intuition for abstract mathematical methods, and your determination to pursue a deeper understanding of our results long after I thought they were settled.

Simon, thank you for welcoming me into your group and treating me as one of your own students. I was repeatedly struck by your ability to listen carefully to an unfamiliar piece of work and then ask a seemingly simple question that exposed a gap in my understanding. Our jargon-free discussions of electronic structure became a valuable test of whether I truly understood what I was doing.

To my friends in the Tew and Benjamin groups: Alex, thank you for being an excellent office partner during the unusual first year of the pandemic. You were my first port of call for questions about coupled-cluster theory, configuration interaction, group theory, and almost anything outside quantum computing. Your help was instrumental in getting several of the projects in this thesis started. Chris, thank you for the many welcome distractions over lunch, dinner, and coffee. Your passion for theoretical physics and intense focus are inspiring, and I look forward to seeing what you achieve working on "die Elektronen". Thanks also to Hans, Kesha, Nick, Annina, Andrew, Jack, Gurkan, Ollie, Arman, and Xiaoning for the conversations and meals that made the PTCL an enjoyable place to work. To Sam M., Arthur, Hamza, Cai, Tyson, Richard, Balint, Matt, Goh, Cica, Greg, Minjun, Armands, Tim, Adam, and Sam J., thank you for the many thoughtful discussions about quantum algorithms. I learned a great deal from working alongside you, both in person and remotely.

Hugh and Nick, thank you for your patience with my endless chemistry questions. I learned enormously from your command of electronic structure, and our discussions and collaborations shaped much of this thesis.

Yuan and Guang Hao, thank you for hosting me as an intern at Microsoft and for making my summer in Seattle so rewarding. Yuan, thank you for the many hours you spent at the whiteboard teaching me how to reason about algorithms, and for your persistence in working through research roadblocks. Guang Hao, thank you for making time to meet with me regularly and for sharing your remarkable breadth of knowledge across physics, mathematics, and computer science. I continue to follow your work with great interest.

To my Bay Area friends—Katie, Norm, and Nico—thank you for opening doors to international collaborations and for giving me opportunities to speak at SIAM conferences. I look forward to seeing you all next time I'm in San Francisco.

Priyank Kumar and David Norris, from my time at ETH—you inspired me to become a scientist and encouraged me to pursue a PhD. Thank you for fostering my interest in quantum physics, research, and teaching, and for creating opportunities that took me from Zurich to Los Angeles, Beijing, Boston, and eventually Oxford and London.

Everyone at Creator Fund, especially Jamie, Alexandra, and Adam, thank you for welcoming me to the team and sharing your fun London office with me during my final year. I joined CF for the professional opportunity and left with many new friends.

Sarah, thank you for your support and friendship across Boston, Oxford, and London, and for providing a much-needed sense of stability. I am grateful for everything we learned and experienced together.

Carolina, Adriana, Marta, and Monika, thank you for the Mediterranean summers and for reminding me that there is life beyond science. Marta, I especially appreciate your interest in my research and your many questions about quantum physics.

Finally, to my parents, Mark and Monica, to whom much of the credit for this thesis is ultimately due: thank you for supporting me throughout what must at times have seemed an endless academic path. More than a decade after I left home, you have always made it easy to remain close. I am grateful for your encouragement to pursue a PhD, for trusting my judgement, and for your unconditional support.

I thank the EPSRC Hub in Quantum Computing and Simulation for supporting this work and enabling much of the research presented in this thesis.

Featured publications and preprints

Chapters 3, 4, and 5 correspond to the following manuscripts, which have been published or are under review.

Chapter 3:

Spin-coupled molecular orbital theory: chemical intuition meets quantum chemistry,

D. Marti-Dafcik, N. Lee, H. G. A. Burton, D. P. Tew,

arXiv:2402.08858

(Currently under review. Subsequent revisions may appear on arXiv; this chapter reflects the version first posted in February 2024.)

Chapter 4:

Spin coupling is all you need: Encoding strong electron correlation in molecules on quantum computers,

D. Marti-Dafcik, H. G. A. Burton, D. P. Tew,

Physical Review Research **7**, 013191 (2025)

Chapter 5:

Exact electronic states with shallow quantum circuits from global optimisation,

H. G. A. Burton, D. Marti-Dafcik, D. P. Tew, D. J. Wales,

npj Quantum Information **9**, 75 (2023)

I was the primary contributor to the first two papers (Chapters 3 and 4) and a secondary contributor to the third paper (Chapter 5). A detailed statement of contributions is provided at the end of each chapter.

Contents

1 Introduction **1**

1.1 Motivation and thesis outline 1
1.2 Quantum algorithms: promise and challenges 4
1.3 Simulation of strongly correlated electrons 10
1.4 Summary of results 15

2 Classical and quantum algorithms for quantum chemistry **23**

2.1 The electronic structure problem 24
2.2 Classical algorithms for electronic structure 30
2.3 Quantum algorithms for quantum chemistry 48
2.4 Spin coupling and configuration state functions 62

3 Spin-coupled molecular orbital theory **73**

3.1 Introduction 76
3.2 Molecular hydrogen 77
3.3 Configuration State Functions 80
3.4 Homolytic bond cleavage 83
3.5 Hydrogen chains: H_4 91
3.6 Quantum computation using spin-coupled states 92
3.7 Conclusions and outlook 96

4 Spin coupling is all you need: Encoding strong electron correlation in molecules on quantum computers **101**

4.1 Introduction 104
4.2 Background and conceptual framework 109
4.3 This work: overview of results 118
4.4 Quantum circuits for preparation of spin eigenfunctions 120
4.5 Heuristic quantum algorithms 135
4.6 Avoiding the orthogonality catastrophe in quantum phase estimation 156
4.7 Conclusions and discussion 165

5 Exact electronic states with shallow quantum circuits from global optimisation 175
5.1 Introduction . . . 178
5.2 Results . . . 181
5.3 Discussion . . . 195

6 Conclusions 199

Appendices

A Appendix for Chapter 3: Spin-coupled molecular orbital theory 205
A.1 Spin eigenfunctions . . . 205

B Appendix for Chapter 4: Spin coupling is all you need: Encoding strong electron correlation in molecules on quantum computers 207
B.1 Spin eigenfunctions . . . 208
B.2 Quantum circuits for basis rotations . . . 208
B.3 Quantum circuits for spin eigenfunctions: explicit decompositions and cost . . . 211
B.4 Energy and gradients in Nonorthogonal VQE . . . 222
B.5 Matrix elements for QSD based on real-time evolution . . . 224
B.6 Quantum circuits for preparation of matrix product states . . . 225

C Appendix for Chapter 5: Exact electronic states with shallow quantum circuits from global optimisation 227
C.1 Universality of one-body and paired two-body operator pool . . . 228
C.2 Qubit operators and quantum circuit implementation . . . 231
C.3 Summary of the DISCO-VQE algorithm . . . 233
C.4 Linear and tetrahedral H_4 (STO-3G) . . . 244
C.5 Triplet ground-state of O_2 (STO-3G) . . . 249
C.6 Linear H_6 binding curve (STO-3G) . . . 249
C.7 Symmetric H_2O binding curve (STO-3G) . . . 250
C.8 N_2 binding curve (STO-3G) . . . 251
C.9 Two-dimensional Hubbard lattice . . . 252

References 257

1 Introduction

Contents

1.1 Motivation and thesis outline 1
1.2 Quantum algorithms: promise and challenges 4
1.3 Simulation of strongly correlated electrons 10
1.3.1 Classical algorithms 10
1.3.2 Quantum algorithms 13
1.4 Summary of results 15
1.4.1 Developing heuristics for strong electron correlation . . 15
1.4.2 Addressing the limitations of quantum hardware 18

1.1 Motivation and thesis outline

This thesis presents new methods for the simulation of quantum chemistry on quantum computers, a field driven by the promise of efficiently solving problems that remain intractable on classical machines. However, at the time of writing, no quantum computer has yet been able to outperform a classical (i.e., non-quantum) computer on a useful practical problem related to chemistry.[1] Moreover, classical algorithms for quantum chemistry continue to advance and new computing

[1] In contrast, there is strong evidence [1–3] that quantum computers have greatly outperformed classical computers on a problem known as random circuit sampling.[4, 5] It is unclear whether this has any useful application, other than the certified generation of random bits,[6] provided that the challenge of efficient verification of the circuits' output distribution can be solved.[7, 8]

technologies such as deep learning [9–14] and tensor networks [15–18] are further pushing the boundaries of quantum simulation methods, often accelerated by parallel and distributed implementations on supercomputers [19–21] and novel compute hardware such as graphical [21–25] and tensor [26] processing units. This raises a natural question: why spend years working on quantum algorithms for quantum chemistry, designing algorithms to be run on computers that do not yet exist?

The scientific component of this question appears easy to answer on the surface. Quantum mechanics, and therefore quantum chemistry, is challenging to simulate on classical processors due to the curse of dimensionality: the quantum state of N interacting particles (such as electrons in molecules) lives in a vector space that grows exponentially with N. No matter their architecture, classical computers quickly run out of memory even just attempting to represent such states. Exact simulations of molecular electronic systems remain limited to around 20-25 electrons even using the world's most powerful supercomputers.[20, 27][2] This is unlikely to improve significantly due to the fundamental limitations of classical processors. Quantum computers, by contrast, natively exploit quantum mechanical principles and offer a new model of computation believed to violate the Extended Church-Turing Thesis—a true paradigm shift with the potential of transforming quantum simulation and other fields.[1, 28, 29] The idea that a quantum-mechanical device might naturally be well suited for quantum simulation was first articulated in the 1980s by Y. Manin [30] and R. Feynman [31], and it continues to motivate modern research in quantum computation.

To date, the simulation of molecules and materials is widely believed to become the killer application of quantum computers.[32–34] The intuition is simple: a quantum device is itself a controllable quantum many-body system, so an approach that "fights fire with fire" intuitively appears advantageous. Despite this appealing symmetry, the extent to which quantum computers might deliver a practical advantage in quantum chemistry remains unclear. This is partly because—unlike what the simplistic dimensionality argument would suggest—existing classical algorithms

[2] For reference, paracetamol, a relatively small molecule, contains 80 electrons.

already deliver satisfactory answers for a great range of quantum chemical problems,[35–38] and quantum algorithms do not in general offer exponential speedups for the central task of computing ground state energies.[39–41] On top of these fundamental algorithmic challenges, errors in quantum hardware lead to significant runtime overheads that further complicate the prospects of quantum speedups, both in the short-term on noisy quantum devices [42–45] and in the long-term on error-corrected quantum devices.[46, 47]

This thesis is guided by the following questions: (1) Which features make certain chemical systems hard to simulate at the quantum-mechanical level? (2) What properties of a quantum computer make it uniquely well-suited to address these challenges? (3) How can these properties be exploited to reduce the cost and increase the accuracy of quantum algorithms for chemical simulation? We focus on the electronic structure problem, the central task in quantum chemistry. This corresponds to computing the low-lying spectrum of the electronic molecular Hamiltonian $\hat{H}_{\text{el}}$ by solving the time-independent Schrödinger equation $\hat{H}_{\text{el}} |\Psi\rangle = E |\Psi\rangle$ (see Chapter 2, Section 2.1 for details).

These questions naturally steer the research towards the simulation of systems of strongly correlated electrons. For such systems, classical algorithms generally remain inaccurate or inefficient, making them good candidates for a potential speedup on a quantum computer. Our approach centres on developing an intuitive understanding of the physical mechanisms causing strong electron correlation in molecules and the structure of the corresponding wavefunctions to inform the design of algorithms for their parametrisation. The main outcomes of the thesis are a conceptual framework for reasoning about strong correlation in molecules, and a set of quantum algorithms or algorithmic primitives for many-electron simulations.

This thesis is structured as follows. In the remainder of this chapter, we discuss the main developments in, and limitations of, quantum algorithms for quantum simulation (Section 1.2), as well as the unique challenges and opportunities in simulation of strongly correlated electrons (Section 1.3). Some of this material assumes an

elementary understanding of quantum information and quantum chemistry/many-body quantum physics. We refer readers unfamiliar with such topics to the standard text books,[37, 38, 48, 49] the more recent review papers,[32–34] or Chapter 2 below for a condensed summary. We summarise the main contributions of this thesis in Section 1.4. Chapter 2 provides the technical background on classical and quantum algorithms for quantum chemistry, as well as techniques for representing symmetry-adapted, spin-coupled states. Chapters 3-5 contain the main research contributions in this thesis and previously appeared as journal publications or preprints (Refs. [50–52]). Chapter 6 summarises and reflects on the main findings presented in the thesis, related ongoing and future work, as well as the broader state of quantum algorithms for chemistry.

1.2 Quantum algorithms: promise and challenges

By encoding information in two-level quantum systems (qubits)[3] and evolving them in time according to the Schrödinger equation, quantum computers can efficiently generate, store and transform vectors of dimension 2^N, which grows exponentially with the number of qubits N.[48, 54]

This offers a great advantage in the scaling of the number of steps required to solve certain problems as a function of the problem size, including for problems that have no direct connection to physics. Notable examples include Shor's exponential speedup for factoring integers into primes [29] and the quadratic speedup for unstructured search due to Grover.[55] Said results are highly celebrated, sparked great interest in the field, and provided algorithmic primitives still used to date,[54, 56–58] including the Quantum Fourier Transform,[29] or Grover's reflection operator,[55] later generalised to amplitude amplification.[59] However, it is unclear if Shor's algorithm will provide any direct positive benefit to society,[60] and it is highly unlikely that Grover's asymptotic speed-up will be sufficient for any practical runtime advantage in often-cited applications such as optimisation.[46, 61–63]

[3] Generalisations to d-level systems (qudits) are possible [53] but less common.

More hope for useful applications is commonly placed in quantum simulation.[63–65] In this context, the first mathematically-rigorous argument towards a quantum speedup was given by Lloyd in 1996.[66] He showed that quantum computers can reproduce the dynamics generated by evolution under any time-independent, k-local Hamiltonian[4] (of which the electronic structure Hamiltonian is a special case) to arbitrary accuracy in polynomial time. In other words, given a k-local Hamiltonian $\hat{H}$, a quantum computer can efficiently implement a unitary operator $\hat{U}$ such that

$$\|\hat{U} - \exp(-i\hat{H}t)\| \leq \epsilon, \tag{1.1}$$

where t is the evolution time, a real-valued scalar, $\|\cdot\|$ is the spectral norm and ϵ is the worst-case error over all input states. Provided an initial quantum state $|\psi_0\rangle$, this approximates the time-evolved state

$$|\psi(t)\rangle = \exp(-i\hat{H}t)\,|\psi_0\rangle \approx \hat{U}\,|\psi_0\rangle\,, \tag{1.2}$$

to within error ϵ, a task now known as Hamiltonian simulation.

The year before, Kitaev introduced the quantum phase estimation algorithm,[67] which enables finding the eigenvalues and eigenvectors of a unitary operator. This assumes that said operator can be implemented efficiently on a quantum computer and that the quantum computer can be initialised in an initial state with significant overlap with the target eigenstate. Combined with a polynomial-time quantum algorithm for antisymmetrising a quantum state [68, 69] or a method for applying antisymmetric fermionic operators,[70, 71] and Lloyd's polynomial-time algorithm for Hamiltonian simulation (Eqs. 1.1 and 1.2),[66] this enables simulating the dynamics of fermionic systems,[68] reading out their energies (Hamiltonian eigenvalues) and preparing the corresponding eigenvectors in polynomial time.[71, 72] In 2000, Fahri *et al.* introduced adiabatic quantum computation, a paradigm

[4]This notion of locality is different from the geometric notion of locality. A k-local Hamiltonian is a Hamiltonian that can be expressed as a sum of terms, $\hat{H} = \sum_i \hat{h}_i$, such that each term $\hat{h}_i$ acts non-trivially on at most k qubits, where k is a constant integer (independent of system size). For the electronic structure Hamiltonian, $\hat{H}_{\mathrm{el}}$, expressed in second quantisation (using fermionic field operators), $k \leq 4$ for every term.

of quantum computation in which a system is initialised in the ground state of a simple Hamiltonian.[73–75] The Hamiltonian is slowly morphed towards a Hamiltonian whose ground state is of interest, because it contains the solution to some computational problem. In 2005, Aspuru-Guzik *et al.* proposed the use of a digitised version of adiabatic quantum computation as a method for initialising a quantum register in a state close to the ground state of the simulated electronic system, known as adiabatic state preparation.[76] They combined this with Kitaev's quantum phase estimation routine and Lloyd's k-local Hamiltonian simulation method applied to implementing time-evolution under the electronic Hamiltonian, $\exp(-i\hat{H}_{\text{el}}t)$, resulting in an end-to-end algorithm to solve the electronic structure problem in polynomial time.[76]

Since these early developments, research has significantly brought down the cost of Hamiltonian simulation and quantum phase estimation,[56–58, 61, 64, 65, 77–90] and alternatives to quantum phase estimation have been developed to reduce the requirements on the size and quality of quantum hardware.[91–103] To date, quantum phase estimation and its variants [104, 105] are widely believed to become the leading candidates for electronic structure simulations on quantum computers.[65, 88, 90, 106] It is therefore instructive to examine the implications of these results for practical quantum simulation.

Unsurprisingly, these early results generated much excitement, as they appeared to promise an exponential quantum speedup for electronic structure. Classical algorithms that yield exact solutions, such as full configuration interaction, scale exponentially with systems size (see Chapter 2, Section 2.2.2), whereas the corresponding quantum algorithm scales only polynomially. Two decades later, it has become clear that this expectation was overly optimistic: exponential speedups are unlikely for most *ab initio* electronic structure problems.[41] However, this does not preclude meaningful quantum advantages. A closer examination reveals a more intricate balance of factors that determine whether a practical speedup can be achieved in realistic settings.[46] In practice, performance depends less on asymptotic scaling and more on heuristics and constant factors that govern concrete

runtimes for finite systems. It is nevertheless useful to summarise several rigorous results that delineate what is and is not provably achievable.

Statement 1. Time-evolution under any instance of the electronic structure Hamiltonian is efficient on a quantum computer.[5] This result follows from the proof by Lloyd [66] and the fact that the electronic structure Hamiltonian is 4-local.

Statement 2. For general instances of the electronic structure Hamiltonian, all existing classical algorithms for time-evolution are inefficient. Intuitively, this follows from the exponential growth of the Hilbert space dimension.[6]

Conclusion 1. From statements **1** and **2**, it follows that there exist electronic structure Hamiltonians for which there is an exponential advantage in applying the time-evolution operator $\exp(-i\hat{H}_{\mathrm{el}}t)$, compared to all existing classical algorithms. In other words, based on existing quantum and classical algorithms, there exist exponential quantum speedups for simulating quantum dynamics.

Statement 3. In its most common formulation (defined within a fixed basis and with a fixed number of electrons), the electronic structure problem is QMA-complete, as was proven by O'Gorman *et al.*,[40] who extended the more general result for the k-local Hamiltonian from Kempe *et al.*[39]

Conclusion 2. Statement **3** implies that there exist instances of the electronic structure Hamiltonian for which neither a quantum computer nor a classical computer can provide efficient solutions. In other words, for certain Hamiltonians, the electronic structure problem is intractable even on a perfect, large-scale quantum computer.

[5] Here, we use "efficient" in the standard computational-complexity sense: the space (number of qubits) and time (circuit depth) resources scale polynomially in the system size (number of electrons or basis functions).

[6] However, note that there is no proof for the impossibility of an efficient classical algorithm for generic quantum dynamics, although it is believed to be highly unlikely that such an algorithm would exist. The same can be said about factoring,[29] whereas Grover's algorithm is provably more efficient than any possible classical algorithm.[55]

Statement 4. Provided an initial state of N electrons with $1/\text{poly}(N)$ overlap with the exact eigenstate, the quantum phase estimation algorithm yields the exact eigenvalue and eigenvector in polynomial time.[67]

Conclusion 3. To obtain an exponential speedup using quantum phase estimation requires two conditions: (1) the cost of the best classical algorithm is exponential in N, (2) there exists a polynomial-time quantum algorithm that produces a state with $1/\text{poly}(N)$ overlap with the exact eigenstate.

The value of these asymptotic scaling statements is mixed. They give insight into the limitations of quantum (and classical) algorithms, which can guide the design and analysis of algorithms. For example, from the QMA-hardness result (Statement **3**) it is immediately clear that quantum phase estimation cannot provide an exponential speedup in the general case i.e. for every possible version of the electronic structure Hamiltonian. Since the time-evolution component of quantum phase estimation is always efficient on a quantum computer (Statement **1**), there exist Hamiltonians with eigenstates for which it requires an exponential amount of resources to prepare an approximate state with high overlap. Thus, it is not always correct to assume access to a state with sufficient overlap with the target eigenstate.

Nevertheless, one must be cautious when evaluating these statements. Understanding the asymptotic scaling of algorithms is helpful when analysing an algorithm on paper, or when numerical data is only available for small system sizes, as it helps predicting the algorithm's cost for larger systems. However, determining the exact cost of an algorithm requires going beyond the asymptotic analysis and including the constant factors.[46] To assess the practical advantage offered by a quantum algorithm, one must compare the total cost for a specific Hamiltonian.[7] Empirical evidence from quantum chemical calculations as well as quantum algorithm resource estimates reveals that the following counterintuitive conditions are at times fulfilled in practice:

[7] In this thesis, we only consider memory and runtime as the cost variables, although one could also consider the financial cost of building and operating the computer, or the energy consumption per calculation.

1. Polynomially scaling algorithms are often sufficiently accurate (and fast) despite not having rigorous success guarantees.[35]

2. Despite their formally prohibitive scaling, exponentially scaling algorithms can be tractable for systems with a small number of strongly correlated electrons, achieving accurate energy estimates using hours of runtime on modern supercomputers.[17, 18, 107, 108]

3. Conversely, polynomially scaling algorithms can be prohibitively expensive despite their favourable asymptotic scaling.[42, 109]

Indeed, the practical success of computational chemistry and materials science is often due to algorithms that satisfy Condition **1**. Single-reference, truncated coupled cluster theory—in particular, coupled cluster singles doubles with perturbative triples, CCSD(T), which scales like $\mathcal{O}(N^7)$—is often considered the gold standard algorithm for quantum chemistry because it can achieve high accuracy (< 1 kcal/mol for energy differences) for molecules with weakly correlated electrons, i.e., systems for which a mean-field approximation is qualitatively accurate.[35] Algorithms such as full configuration interaction quantum Monte Carlo (FCI-QMC) or the density matrix renormalization group (DMRG)[15, 110, 111] can solve some of the most challenging electronic structure problems despite scaling exponentially (Condition **2**).[17, 18, 107, 108] However, they require significant computational resources and remain insufficient for systems with many strongly correlated electrons such as transition metal clusters.[65, 106, 112] Finally, quantum phase estimation, despite its favourable polynomial scaling, requires roughly 10^{10} non-Clifford gates[8] for estimating a single energy of classically challenging molecules, leading to runtimes in the order of days under optimistic hardware assumptions,[65, 88, 106] or orders of magnitude more for the simulation of solids [109] (Condition **3**). This is particularly problematic in applications that require the evaluation of many (up

[8]These are the class of gates that dominate the cost of a fault-tolerant quantum computation under typical error-correcting codes, see Chapter 2, Section 2.3.4.

to millions or billions of) different energies such as drug discovery, and often-cited industrial use-case for quantum computers.[113]

The constant-factor analysis becomes particularly important when taking into account the hardware challenges inherent in quantum algorithms. Quantum processors are generally subject to much higher levels of noise than classical processors due to the delicate nature of quantum systems, which tend to quickly decohere. In implementations without active error correction, many runs of the same circuit are needed to statistically mitigate the effect of hardware noise, e.g. through quantum error mitigation techniques,[114–116] or to measure expectation values as part of variational optimisation routines.[42, 92] In the case of quantum error correction, the overhead required to make the quantum algorithm fault-tolerant is so large that it is expected that the clock cycles for fault-tolerant quantum gates will be several orders of magnitudes higher than for equivalent classical gates,[117, 118] greatly slowing down the speed of quantum algorithm executions.[46, 62, 63] Thus, it is possible that a linear-scaling quantum algorithm may have runtimes similar to those of an exponentially scaling classical algorithm for problem sizes of practical interest.

It is therefore pertinent to focus the development of quantum algorithms on problems where classical algorithms clearly fall short.[63] In electronic structure theory, this points naturally to systems containing strongly correlated electrons.[65, 112] This is precisely the focus of this thesis.

1.3 Simulation of strongly correlated electrons

In this section, we review the current state of methods for electronic structure and the challenges of simulating strongly correlated systems. We motivate our approach from an intuitive and historical perspective. We provide a more technical discussion of mathematical techniques for strong correlation in Chapter 2.

1.3.1 Classical algorithms

By the end of the 1920s, non-relativistic quantum mechanics had been fully developed, with exact analytical solutions to the Schrödinger equation derived for

simple systems like the hydrogen atom.[119] To extend quantum theory to larger systems, approximate methods became essential.

Some of these approximations, such as perturbation theory [120–123] and the Rayleigh–Ritz variational principle [124, 125] arose from reformulating quantum mechanics in linear-algebraic terms.[120, 126–128] Others, including the Born–Oppenheimer approximation,[129] relied on physical intuition. The Hartree–Fock self-consistent field method combined both: it applied the variational principle [130, 131] to a mean-field description while ensuring antisymmetry through Slater determinants.[132, 133] Yet even then it was already recognised that the discipline of quantum chemistry faced a seemingly intractable problem, as famously summarised by Dirac in 1929:[119] "The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble."

Nearly a century later, Dirac would likely be surprised by the state of quantum chemistry. The fundamental dimensionality problem remains, and most modern quantum chemistry methods rely on refinement and extension of the concepts and approximations introduced in the 1920s, listed above. Despite this seeming lack of any fundamental shifts, progress in computational chemistry has been remarkably successful, enabling the prediction of energies for small-to-medium-sized molecules at equilibrium with near-experimental accuracy [35] and routine computation of thermodynamic properties and equilibrium structures.[134–136] A major part of the progress is due to improvements in numerical linear algebra techniques, including eigenvalue solvers (such as the QR,[137] Lanczos,[138] and Davidson [139–141] algorithms), and matrix or tensor decompositions (including LU, Cholesky, and singular value decompositions, and tensor hypercontraction).[137, 142–151] Optimisation algorithms have also played a central role.[152, 153] Together with increasingly powerful classical processors, the systematic design and optimisation of basis sets [154–156] and the development of highly efficient integral evaluation protocols

[157–159] have extended the applicability of the concepts and approximations from many-body quantum theory to a wide range of chemical systems.

In essence, the success of quantum chemistry rests on three pillars:

- *A simplified physical model*; Molecular orbital (MO) theory, which underlies the Hartree–Fock method, transforms an N-body problem into a set of N coupled one-body problems.

- *Systematic many-body corrections;* Starting from a simplified model (such as a mean-field operator), methods like the variational principle and perturbation theory enable progressively increasing the complexity and accuracy of solutions through controlled expansions.

- *Efficient numerical techniques;* Developments in linear algebra and optimisation exploit ever-improving classical hardware to scale these methods to larger systems.

In the limit of weak entanglement or correlation between electrons, the combination of these techniques can give accurate answers for a broad range of chemical systems. The recipe is straightforward: Hartree–Fock provides a reference state towards which to bias the search within the vast Hilbert space, while truncated many-body corrections—such as low-order perturbation or coupled cluster theory—refine the wavefunction by locally exploring the surrounding Hilbert space, increasing accuracy while maintaining polynomial cost. Since this often works, an empirical lesson from a century of quantum chemistry is that a large number of molecular electronic states only occupy a very small region of the vast Hilbert space, namely the weakly-correlated region.

Unfortunately, both the underlying MO theory and the truncated many-body methods break down in the presence of strong electron correlation. Examples of strongly correlated molecular electronic systems include stretched chemical bonds and transition metal clusters.[17, 18, 160] Some polynomially scaling classical methods are accurate for small numbers of strongly correlated electrons,[160, 161]

but the only methods known to be accurate for more general strongly correlated systems scale exponentially. For example, the complete active space self-consistent field method (CASSCF)[162] is exact within the chosen active space (a subset of the single-particle states believed to contain the strong correlation effects), but it cannot be applied to large numbers of electrons. Notable success has been achieved by methods that scale exponentially but whose computational cost can be controlled, e.g. by stochastically selecting the number of determinants at different steps of the calculation in FCI-QMC,[107, 163] or by limiting the bond dimension in DMRG.[15, 17, 18, 108] However, the high cost of these calculations means that their accuracy is limited for severely correlated systems [65, 112] or systems that combine a small number of strongly correlated electrons with a larger number of less correlated electrons distributed across a large number of orbitals.[106]

1.3.2 Quantum algorithms

Quantum computers do not suffer from this memory bottleneck: by using entangled quantum systems as the processor hardware, quantum computers can efficiently perform certain linear-algebraic operations, such as matrix-vector multiplications, in an exponentially large Hilbert space.[164] Recently, a unified framework for quantum linear algebra has been established, enabling the implementation of increasingly general exponential matrix transformations across application domains.[54, 57, 58, 165–167] It is expected that machines leveraging these techniques will be capable of accessing highly entangled quantum states found in nature, or exploring the "entanglement frontier", as coined by John Preskill.[168]

However, the ability to create and transform highly entangled states does not directly translate into solving exponentially large eigenvalue problems.[39, 169] Quantum algorithms for eigenvalue problems need good heuristics to guide the solution within the exponential Hilbert space, e.g., by limiting the search to a small set of entangled states. In the absence of appropriate heuristics, quantum algorithms also suffer in the strong correlation regime, much like classical algorithms.[41]

For example, there has been remarkable progress in quantum phase estimation for electronic structure, or Hamiltonian simulation, more generally. This has largely been achieved through quantum algebraic techniques for efficiently implementing the time-evolution operator $\exp(-i\hat{H}t)$—including efficient methods for implementing quantum walks,[170, 171] simulation under sparse Hamiltonians,[172, 173] oblivious amplitude amplification,[173] linear combinations of unitary operators,[56] and Taylor series expansions,[81, 82, 174]—or by instead implementing a qubitized quantum walk [83, 85] through a unifying paradigm known as qubitization, which has optimal scaling in all the relevant parameters.[57, 58] Further improvements were achieved using matrix decompositions ported from the classical computational quantum chemistry literature,[96, 149, 151, 175, 176] including different low-rank factorizations [84, 85, 88] and tensor hypercontraction.[65] These quantum linear algebra breakthroughs have resulted in an impressive reduction in the scaling of quantum phase estimation for electronic structure from $\mathcal{O}(M^{11})$[177] to roughly $\mathcal{O}(M^2)$ in the number of orbitals M using second-quantised encodings,[9] or even sublinear scaling in the number of orbitals in first quantisation.[178–180]

As encouraging as these results are, quantum phase estimation requires an initial state with high overlap with the target eigenstate. The number of determinants of said state typically increases very rapidly with the number of strongly correlated electrons. One could approximate such states through exponentially scaling classical methods and then used them in quantum algorithms. However, beyond the usual limitations of classical methods, an additional challenge is that loading such states onto a quantum device through a state preparation circuit requires a number of gates scaling linearly with the number of determinants.[181, 182]

In the case of heuristic quantum algorithms, their performance as a function of the initial state is less rigorously understood, but intuition and empirical evidence suggest that they also suffer in strongly correlated regimes.[98, 100, 102, 183–185] For example, using a standard wavefunction *ansatz*, the variational quantum eigensolver typically requires a larger number of parameters to describe stretched

[9] See Chapter 2, Section 2.3, for information on different encodings.

chemical bonds compared to molecules at equilibrium.[183–185] Similarly, the cost of adiabatic state preparation depends inversely on the gap between the eigenstates, and it is known that the gap shrinks rapidly with the number of strongly correlated electrons.[41, 186–188]

1.4 Summary of results

This thesis addresses the limitations of both quantum and classical algorithms for simulating strongly correlated electrons. We tackle the problem from two complementary directions. First, we develop a physically motivated framework for parametrising strongly correlated electronic states in molecules (Section 1.4.1). Although this can in principle be used as part of classical computational methods, we show that our heuristic is particularly well-suited for simulating strongly correlated systems using quantum computers. Second, we design a variational quantum algorithm and associated wavefunction ansatz designed to mitigate the limitations imposed by imperfect quantum hardware (Section 1.4.2).

1.4.1 Developing heuristics for strong electron correlation

We take the view that, just like physically motivated approximations enabled unlocking the power of classical algorithms for electronic structure, we must develop heuristics for quantum computation rooted in physico-chemical insight. Classical heuristics exploited MO theory and mean-field states. To enable exploring the entanglement frontier in chemistry, quantum heuristics must depart from mean-field theories and directly encode strongly correlated phenomena, guided by entangled states. This might appear daunting because most such states are complex and hard to represent compactly.

Here, we exploit the insight that highly entangled states that appear in prototypical strongly correlated molecules are in fact highly symmetric. This makes it possible to efficiently select such states among the vast number of states in the Hilbert space, and to efficiently prepare them on quantum processors.

Chapter 3 results

Chapter 3 introduces *spin-coupled molecular orbital theory*, a conceptual and computational framework for rationalising strongly correlated electronic states in molecules and a wavefunction ansatz for their parametrisation. This generalised MO theory describes both mean-field effects and strong correlation effects using a small number of variational parameters. States are represented in terms of configuration state functions (CSFs). CSFs are symmetry-adapted linear combinations of Slater determinants which—unlike individual Slater determinants—are guaranteed to have a well-defined spin quantum number. These objects have long been studied in the quantum chemistry literature, and their construction based on group-theoretical techniques is well established.[16, 189, 190] However, their use as reference states in electronic structure methods on classical computers is limited due to their high-dimensional structure.[160]

CSFs offer a natural and informationally compact representation of strongly correlated electrons: provided an appropriate orbital basis, one or a few CSFs capture effects such as short-range and long-range entanglement within and between localised clusters of valence electrons. These correlations are the main many-body effect in stretched bonds or transition metal clusters, prototypical examples of strong electron correlation in molecules. In principle, the number of possible CSFs scales combinatorially with the number of valence electrons, as they span a full basis for the Hilbert space of fixed spin. However, the symmetric structure of molecules found in nature often makes it possible to predict a small subset of relevant CSFs, provided one understands the physical mechanisms underlying strong correlation: the coupling of localised electrons through their spin angular momentum, with particular spin couplings preferred according to a competition between delocalisation and exchange effects. For a range of small molecules with stretched bonds, we show that the dominant couplings can often be determined from simple chemical heuristics such as Hund's rules.

Our ansatz approximates wavefunctions by expanding them as a linear combination of CSFs, each represented in a bespoke orbital basis. The CSFs include

delocalised, mean-field states, such as the Hartree–Fock state, and spin-coupled states. This captures the dominant entanglement structure in strongly correlated electrons in molecules, yielding wavefunctions with high overlap with the exact ground state at the cost of a mean-field calculation. The material in this chapter first appeared in Ref. 50.

Chapter 4 results

Much like traditional MO theory, spin-coupled MO theory exploits the intuition that physical states in nature live in a small corner of the Hilbert space. The main difference is that, for strongly correlated electrons in molecules, it is a highly entangled corner of the Hilbert space, or a small number of different corners, each corresponding to a distinct electronic configuration but with a clear physical interpretation. Regardless, there is no need to generate an expansion over a large region of the Hilbert space, and a surgical, targetted approach suffices for obtaining accurate wavefunction parametrisations. Such findings are independent of quantum computing.[10] However, although our wavefunction parametrisation requires few degrees of freedom, the expansion of such wavefunctions in a Slater determinant basis scales exponentially with the number of strongly correlated electrons. This makes it hard to exploit spin-coupled MO theory on classical computers, since methods such as many-body perturbation theory cannot efficiently correlate highly entangled reference states.

On quantum computers, applying unitary operators, such as the time-evolution operator, is efficient regardless of the initial state. However, without further insight, incorporating spin-coupled states into quantum algorithms is inefficient due to the challenge of state preparation. To load such states onto quantum computers, one must generate them starting from a single computational basis state—corresponding to a single Slater determinant in second-quantised encodings (see Section 2.3.1)—by applying a sequence of quantum gates. In the general case,

[10]However, they are in agreement with a theoretical result from the quantum information literature: Poulin *et al.* proved that the majority of states within the Hilbert space are unphysical in the sense that they can only be generated by real- or imaginary-time-evolving a state under a k-local Hamiltonian for an exponentially long time.[191]

the number of gates scales linearly with the number of Slater determinants, and exponentially with the number of valence electrons.[181, 192]

Chapter 4 addresses the challenge of initial state preparation in quantum algorithms for strongly correlated electrons. We introduce quantum circuits of depth $\mathcal{O}(N)$ that prepare spin-coupled states with $\binom{N}{N/2}$ determinants, enabling the efficient use of these CSFs on quantum computers.

We propose methods for integrating spin-coupled initial states within existing quantum algorithms, e.g., by using one of the maximally localised spin-coupled states within adiabatic state preparation. We analyse the performance of quantum algorithms including the variational quantum eigensolver, quantum subspace diagonalisation, and adiabatic state preparation when boosted with such initial states. We show that the cost of all quantum algorithms is greatly reduced by using such reference states, compared to standard approaches that use mean-field initial states.

We examine the use of such CSFs in the context of quantum phase estimation, including the opportunities presented by extending our approach to more complex fermionic systems such as transition metal clusters. We conclude that our CSF-based approach for initial state preparation is ideally suited for quantum phase estimation due to its very low quantum circuit cost. The material in this chapter first appeared in Ref. 51.

1.4.2 Addressing the limitations of quantum hardware

In this section, we summarise a novel wavefunction ansatz and optimisation-based quantum algorithm specifically designed to circumvent or minimise the impact of imperfect quantum hardware.

Before, we provide a high-level description of the challenge of noise in quantum hardware, the main idea behind quantum error correction, and its implications. We refer the reader to Chapter 2, Section 2.3.4 for more details on fault-tolerant quantum computing.

Noise and quantum error correction

Digital quantum computing requires maintaining quantum many-body systems in coherent quantum states while evolving them by applying quantum gates. This is a formidable challenge currently preventing large-scale application of quantum algorithms. In any quantum computation, errors can occur during gate executions, measurement, or due to general decoherence even when the qubits are meant to remain idle. The collection of such errors is referred to as *noise*. The more quantum gates are executed in a circuit, the higher the probability that errors will occur, corrupting the outcome of a quantum algorithm.

There exists a powerful solution to this: quantum error correction, first proposed by Peter Shor.[193] In principle, quantum error correction protocols can exponentially suppress errors such that deep quantum circuits can be run with arbitrarily high accuracy.[48] Quantum algorithms implemented in this manner are referred to as being *fault-tolerant*. Quantum error correction encodes the state of a single *logical* qubit—such as a qubit that represents the occupation of an electronic spin-orbital—in the many-body state of multiple *physical* qubits. To check for errors, one measures some of the physical qubits in such a way that they reveal information about whether an error has occurred or not and, e.g., on which qubit this has occurred. Crucially, this is achieved without revealing the full quantum state of the logical qubit, as the act of measurement of the full logical qubit would collapse its quantum state to a classical state and destroy its quantum information before the end of the algorithm. Through this redundant encoding, errors can be corrected during the quantum computation by applying additional gates, effectively reducing the logical error rate.

For these protocols to be useful requires, among other things: (1) a much larger number of physical qubits than the number of logical qubits, i.e., than the number of qubits it would require to encode the problem on a quantum computer with no physical errors, (2) sufficiently low physical error rates, because if the physical error rate is above a certain threshold, the process of checking for errors introduces too many additional errors. The threshold theorem [193–196] establishes

that exponential error suppression is achieved if the physical error rate is below a certain threshold, enabling a *netto* benefit from the use of error-correcting protocols.

In principle, the full potential of quantum algorithms will be realised once quantum computers with sufficiently many physical qubits and sufficiently low physical error rates become available. In the context of quantum chemistry, it is expected that quantum phase estimation will become the default fault-tolerant quantum algorithm for computing ground and excited state properties on error-corrected quantum hardware.[64] Unfortunately, even if or once quantum hardware progresses to that state,[11] noise will severely impact the performance of fault-tolerant quantum algorithms: Due to the high overhead of implementing quantum error-correcting protocols, fault-tolerant quantum computers will be very slow, since a large number of additional gates and circuit repetitions will be required to check and correct for errors.[46]

For example, it is estimated that using the quantum phase estimation algorithm to compute the electronic ground state energy of molecules that are hard to simulate classically will require a runtime in the order of a few days and four million physical qubits.[65, 106] Considering that some of the often-proposed practical applications of quantum computers, such as drug discovery,[197, 198] require between millions and billions of energy evaluations,[113] it is desirable to develop alternatives to QPE that require fewer gates and circuits of lower depth to reduce the overall runtime.

Chapter 5 results

Existing quantum devices have generated much excitement because they have reached physical qubit counts in the order of 100 to 1000 and error rates below, or close to, the thresholds estimated for popular error correcting codes such as the surface code.[199, 200] Since executing fault-tolerant quantum algorithms will require scaling up devices to hundreds of thousands or millions of physical qubits, in the mean-time, it is intriguing to consider the possibility of whether a quantum speed-up is possible over classical algorithms using quantum devices without any

[11] The author's belief is that this is a question of *when* rather than *if*, but predicting the timeline is difficult, as it is most likely on the order of decades.

error correction. This has been answered in the affirmative for random sampling experiments,[1–3] with high confidence but no absolute certainty,[7, 8] a feat referred to as *quantum supremacy*. However, it is unclear whether quantum speed-ups will be achieved without quantum error correction for problems of practical use, a prospect often referred to as *quantum advantage*. Quantum processors of limited size and no error correction capabilities are referred to as *noisy intermediate quantum computers* (NISQ) computers.[201]

To avoid catastrophic errors during a NISQ computation, quantum circuits must have low circuit depth.[12] Furthermore, the use of two-qubit entangling gates must be minimised since these are usually the largest source of error.[199] These constraints render some of the most powerful quantum algorithms, such as quantum phase estimation, impractical or unfeasible on NISQ devices. Thus, there has been great interest in designing novel quantum algorithms with lower circuit depth requirements, often referred to as NISQ algorithms.[45]

A popular proposal is the variational quantum eigensolver (see Chapter 2, Section 2.3.3).[91–93] Here, a parametrised quantum state is prepared on the quantum computer by applying a product of N_U parametrised unitary operators,

$$\hat{U}(\boldsymbol{\theta}) = \hat{U}_1(\theta_1) \cdots \hat{U}_{N_U}(\theta_{N_U}), \tag{1.3}$$

onto some initial state $|\Phi_0\rangle$.[91, 92] The energy of the state $|\Psi(\theta)\rangle = \hat{U}(\boldsymbol{\theta})\,|\Phi_0\rangle$ is estimated through measurements of the qubits, and a classical computer is used to update the parameters $\boldsymbol{\theta}$ to minimise the energy (typically, but not necessarily, through gradient descent). The new parameters are fed back to the quantum computer and the loop is repeated until convergence is reached. The operators $\hat{U}_i(\theta_i)$ are chosen from a pool of operators, of which a wide range of options has been proposed.[45]

The choice of the parametrised ansatz is very flexible and it is unclear which ansatz, if any, is capable of efficiently approximating the eigenstates of strongly correlated electronic systems. Ansätze that preserve physical symmetries, such

[12]The quantum circuit depth is the number of sequential layers of quantum gates that must be applied to execute a quantum algorithm, assuming parallel execution of independent gates.[48]

as the spin or the number of fermions, typically require deeper circuits despite being more accurate than so-called *hardware-efficient* ansätze, which are more suited for NISQ hardware but might break symmetries. This results in a tradeoff between circuit depth and accuracy.[32, 183] Moreover, the popular unitary coupled cluster singles and doubles (UCCSD) ansatz preserves particle number and spin projection symmetry but does not preserve the spin quantum number, which is a good quantum number for electronic wavefunctions.[202] Finally, the operators in general do not commute, making the resulting state highly dependent on the order of the operators, and there is no established method for determining the ordering *a priori*.

In Chapter 5, we provide an ansatz and a discrete optimisation algorithm to address the strong electron correlation problem using low-depth quantum circuits, aiming to enable challenging electronic structure calculations on NISQ devices. We show that commonly-used one- and two-body operators can be combined to form a symmetry-adapted operator pool for the variational quantum eigensolver that preserves the number of electrons, spin projection, and total spin. These operators have a relatively efficient circuit implementation, and we prove that, despite the loss of flexibility due to the symmetry-adaptation, the ansatz is universal in the sense that, in principle, it is capable of parametrising any state within the relevant Hilbert space. However, obtaining exact eigenstates generally requires a very large number of operators $\hat{U}_i(\theta_i)$, leading to deep circuits. To avoid this, we frame the ansatz choice as a discrete optimisation problem, restricting the number of operators in the ansatz, N_U, while allowing different orderings and selections of these operators from the symmetry-preserving operator pool. Through extensive numerical simulations, we show that this yields extremely compact ansätze that accurately parametrise the ground states of strongly correlated electrons in molecules as well as the two-dimensional Fermi-Hubbard model.

The material in this paper first appeared in Ref. 52.

2

Classical and quantum algorithms for quantum chemistry

Contents

2.1 The electronic structure problem 24
2.1.1 Continuum derivation 24
2.1.2 Projection to a finite basis and second quantisation . . . 26
2.2 Classical algorithms for electronic structure 30
2.2.1 Basis sets and mean-field theory 30
2.2.2 Exact solutions through full configuration interaction . 36
2.2.3 Coupled cluster: high accuracy with polynomial scaling 37
2.2.4 Weak and strong electron correlation 40
2.2.5 Approximate ansätze with exponential scaling 45
2.3 Quantum algorithms for quantum chemistry 48
2.3.1 Mapping fermions to qubits 48
2.3.2 Quantum Phase Estimation 52
2.3.3 Variational Quantum Eigensolver 58
2.3.4 Fault-tolerant vs. NISQ implementation of quantum algorithms 60
2.4 Spin coupling and configuration state functions 62
2.4.1 The role of spin in quantum chemistry 63
2.4.2 Angular momentum in quantum mechanics 64
2.4.3 Genealogical construction of spin eigenfunctions 66
2.4.4 Configuration state functions 68
2.4.5 Role of the basis 69

This chapter begins with a derivation of the electronic structure Hamiltonian

from a series of standard assumptions and approximations (Section 2.1). We discuss methods for solving the electronic structure problem on classical computers in Section 2.2, followed by a summary of quantum computational methods in Section 2.3. We conclude with a discussion of the role of spin in electronic structure and the representation of spin-coupled states in terms of configuration state functions (Section 2.4).

2.1 The electronic structure problem

The goal of quantum chemistry is to explain, quantify and predict chemical properties and processes directly from quantum mechanics. This can involve obtaining both static and dynamic information about interacting quantum particles. The focus of this thesis, and of most of the discipline of quantum chemistry, is on static properties, which can be obtained by computing stationary states and values of observables for those states. A particular quantity of interest is the energy difference between two molecular structures, as this has a very direct impact on common chemical processes of interest. For example, chemical reaction rates depend exponentially on such energy differences. This exponential dependence is the reason why very high accuracy is often required from quantum chemical calculations.[64, 88]

Below, we summarise how one arrives at the electronic structure Hamiltonian by taking a series of well-established approximations. This Hamiltonian is the central object in quantum chemistry, as its eigenspace provides most of the desired chemical information. We begin with the derivation in the continuum formulation (Section 2.1.1) and then proceed to discuss the approximations introduced to project the Hamiltonian to a finite basis (Section 2.1.2).

2.1.1 Continuum derivation

For conceptual and computational ease, it is useful to introduce a set of approximations. First, we model atomic nuclei as point particles with a fixed mass and charge given by the periodic table. At the length, time, and energy scales relevant

to everyday chemistry, it is safe to assume that any additional effects due to the structure of nuclei (protons and neutrons, as well as their constituent particles) can be ignored.[203]

Second, we assume that the underlying physical theory is non-relativistic quantum mechanics. Due to the small speed of electrons relative to the speed of light, ignoring relativistic effects is a good approximation for many atoms in the periodic table. Relativistic effects can become significant for heavier atoms. A well-known example is the prediction of the colour of gold (atomic number $Z = 79$): non-relativistic calculations would predict a silver-like colour similar to that of most metals; accounting for special relativity correctly predicts a contraction of the $6s$ orbital due to an increase in the effective mass of its electrons, causing them to be closer to the nucleus and lowering the orbital's energy. This lowers the energy gap between the $5s$ and $6s$ orbitals to about 2.3 eV compared to a non-relativistic model, correctly predicting absorption of blue light, which corresponds to an emission spectrum of yellowish colour. In such situations, small relativistic effects can often be accounted for through perturbative corrections, whereas full relativistic treatment requires moving to the Dirac equation. Nevertheless, the majority of everyday matter involves lighter atoms, for which it is safe to ignore special relativity.[204]

Thus, the relevant equation of motion is the Schrödinger equation. Since we are interested in static properties, the task is to solve the time-independent Schrödinger equation, which can be expressed as an eigenvalue problem for the Hamiltonian operator $\hat{H}$:

$$\hat{H} \left|\Psi\right\rangle = E \left|\Psi\right\rangle . \tag{2.1}$$

The Hamiltonian contains potential energy terms due to the Coulomb force between charged particles (the positively charged nuclei and the negatively charged electrons), as well as kinetic energy terms for each particle. To simply things further, we apply the Born-Oppenheimer approximation. We decouple the electronic and nuclear motion, assume that the nuclei are fixed in space, and only consider the

electronic degrees of freedom as a function of the nuclear positions. This is justified because the nuclear mass is three orders of magnitude higher than the electronic mass, and therefore electron motion is much faster than nuclear motion. We model electrons as moving under a potential defined by the clamped nuclei, and parametrise the electronic wavefunction and Hamiltonian as a function of said nuclear positions. With this, we arrive at the electronic Hamiltonian:

$$\hat{H}_{\text{el}}(\mathbf{R}) = \frac{1}{2}\sum_i \nabla_i^2 - \sum_{i,I}^{N_n,N_e} \frac{Z_I}{\mathbf{r}_i - \mathbf{R}_I} + \sum_{i\neq j}^{N_e} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + E_{\text{nuc}}, \tag{2.2}$$

where $\mathbf{R}_I$ and Z_I are the nuclear coordinates and masses, $\mathbf{r_i}$ are the electron coordinates, all expressed in atomic units, N_n and N_e are the total number of nuclei and electrons, and the nuclear energy is a constant defined by

$$E_{\text{nuc}} = \sum_{I,J<I}^{N_n} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \tag{2.3}$$

The first term in Eq. 2.2 corresponds to the kinetic energy of each electron; the second term is the attraction between each electron and each nucleus; the third term corresponds to the repulsive interaction between the electrons.

2.1.2 Projection to a finite basis and second quantisation

It is in principle possible to solve the Schrödinger equation under $\hat{H}_{\text{el}}(\mathbf{R})$ without further assumptions, i.e., with a continuous representation of the space coordinates of the electrons, and representing the Hamiltonian and wavefunctions in an infinite-dimensional Hilbert space. This *continuum approach* is often taken in stochastic methods such as quantum Monte Carlo,[205, 206] which has recently seen a renaissance due to novel variational methods based on deep neural networks.[9–14]

Throughout this thesis, we take the more common *basis set approach*. Here, the spatial coordinates are discretised by introducing a finite basis of single-particle states, $\{\phi_i(\mathbf{x}_i)\}$, typically referred to as the orbitals. The orbitals are functions of the three spatial coordinates $\mathbf{r}_i = (x_i, y_i, z_i) \in \mathbb{R}^3$ as well as the spin coordinate σ_i, a binary variable that specifies the spin angular momentum state of the electron. We refer to this state as being either *up* or *down*, or α and β: $\sigma_i \in \{\alpha, \beta\}$. Thus,

$\mathbf{x}_i = (\mathbf{r}_i, \sigma_i)$ is the combined spatial-spin coordinate of the i-th electron. At times, it is more helpful to separate the spatial and spin components of the orbitals and refer to the spatial component as a *spatial orbital* and the combined function as a *spin-orbital*: $\phi_i(\mathbf{x}_i) = \chi_i(\mathbf{r}_i)\zeta_i(\sigma_i)$, where χ_i is the spatial wavefunction and ζ_i is one of two orthonormal functions specifying the spin state.

The spatial orbital can in principle be any function in three-dimensional space, including a delta function peaked at a grid point $\mathbf{r_i}$.[180, 207, 208] This discretisation is not often used in classical computational methods. However, it can be advantageous in quantum algorithms as it reduces the number of terms in the electronic structure Hamiltonian from $\mathcal{O}(M^4)$ to $\mathcal{O}(M^2)$ for a system with $2M$ orbitals, and consequently, the cost of Hamiltonian simulation.[180, 209] This is particularly helpful when working in first quantised mappings of fermionic states onto qubits (see Section 2.3.1). More commonly, and throughout this thesis, the orbitals are a linear combination of Gaussian functions centred around each nuclear position.

The introduction of this finite single-particle basis leads to a finite many-particle basis consisting of every possible way the electrons can occupy the set of orbitals. This allows projecting the infinite Hilbert space to a finite (albeit very large) Hilbert space.

To construct the many-particle basis, we must consider antisymmetry. According to the spin-statistics theorem, electrons are fermions, i.e., indistinguishable particles whose wavefunction is antisymmetric with respect to exchange of any pairs of particles. In the so-called first quantised formalism, antisymmetry is imposed by explicitly writing the wavefunction as an antisymmetric linear combination of all possible permutations of products of particles. For a system of N electrons, each many-body basis state in first quantisation takes the form of an antisymmetrised product of single-particle functions,

$$\hat{\mathcal{A}}\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)\cdots\phi_N(\mathbf{r}_N), \tag{2.4}$$

where $\hat{\mathcal{A}}$ is the antisymmetrisation operator, a normalised linear combination of all $N!$ permutation operators. For example, for a system with two particles in

two orbitals we have

$$\begin{aligned}\psi(\mathbf{r}_1,\mathbf{r}_2) &= \hat{\mathcal{A}}\,\phi_1(\mathbf{r}_1)\,\phi_2(\mathbf{r}_2)\\ &= \frac{1}{\sqrt{2}}\Big(\phi_1(\mathbf{r}_1)\,\phi_2(\mathbf{r}_2) - \phi_1(\mathbf{r}_2)\,\phi_2(\mathbf{r}_1)\Big)\\ &= \frac{1}{\sqrt{2}}\begin{vmatrix}\phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1)\\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2)\end{vmatrix}.\end{aligned} \tag{2.5}$$

and it is easy to check that $\psi(\mathbf{r}_1,\mathbf{r}_2) = -\psi(\mathbf{r}_2,\mathbf{r}_1) = \hat{\mathcal{A}}\phi_1(\mathbf{r}_2)\phi_2(\mathbf{r}_1)$. As shown in Eq. 2.5, one can conveniently express this state in terms of the determinant of a matrix, known as a Slater determinant. The generalisation of this object to N electrons in $2M$ orbitals is:

$$\psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix}\phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \cdots & \phi_{2M}(\mathbf{x}_1)\\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \cdots & \phi_{2M}(\mathbf{x}_2)\\ \vdots & \vdots & \ddots & \vdots\\ \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & \cdots & \phi_{2M}(\mathbf{x}_N)\end{vmatrix}. \tag{2.6}$$

The Slater determinant object ensures antisymmetry because exchanging the positions of any two electrons is equivalent to swapping two rows of the determinant, which induces the desired sign change in the wavefunction.

In second quantisation, this structure is exploited to provide a convenient shorthand notation for Slater determinants in terms of bitstrings known as occupation number vectors,

$$\psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) = |f_1,...,f_p,...,f_{2M}\rangle = |f\rangle \tag{2.7}$$

where $f_p = 0$ if ϕ_p is unoccupied and $f_p = 1$ if ϕ_p is occupied by an electron. The set of all possible occupation number vectors forms a complete basis for the antisymmetric many-electron Hilbert space, also known as the Fock space. While each occupation number vector corresponds to a Slater determinant, and we will refer to them as Slater determinants, the key benefit of second quantisation is that the determinant expansion is never explicitly written out. This is convenient because every determinant expansion is very large (due to the $N!$ permutations) but contains a lot of redundant information.

Instead, the antisymmetric exchange statistics are respected by ensuring that the operators acting on a fermionic wavefunction are antisymmetric with respect to

exchange of any two fermions. This can be achieved by introducing the canonical fermionic creation and annihilation operators: a creation operator $a_p^\dagger$ excites an electron in spin-orbital p, whereas an annihilation operator a_p removes an electron.[1] In the case that p is already occupied, acting on it with $a_p^\dagger$ deletes the state; if p is unoccupied, acting on it with a_p deletes the state. These operators obey the fermionic anticommutation relations:

$$\{a_p, a_q^\dagger\} = a_p a_q^\dagger + a_q^\dagger a_p = \delta_{p,q}, \quad \{a_p, a_q\} = \{a_p^\dagger, a_q^\dagger\} = 0. \tag{2.8}$$

Equipped with these operators, we finally arrive at the Hamiltonian in second quantisation:

$$\hat{H}_{\text{el}} = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_s a_r, \tag{2.9}$$

where the scalars h_{pq} and h_{pqrs} are integrals over the orbitals:

$$h_{pq} = \int d\mathbf{r} \phi_p^*(\mathbf{x}) \left(-\frac{\nabla^2}{2} - \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \right) \phi_q(\mathbf{x}), \tag{2.10}$$

$$h_{pqrs} = \int d\mathbf{x}_1 \int d\mathbf{x}_2 \frac{\phi_p^*(\mathbf{x}_1)\phi_q^*(\mathbf{x}_2)\phi_r(\mathbf{x}_1)\phi_s(\mathbf{x}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \tag{2.11}$$

The first term in Eq. 2.9 contains all one-body operators, corresponding to the kinetic energy of the electrons and their Coulomb attraction to the nuclei. The second term contains the repulsive electron-electron Coulomb interaction, consisting of two-body operators. Thus, in this general form, the total number of terms scales as $\mathcal{O}(M^4)$ for $2M$ spin-orbitals.

The electronic structure problem can be stated as follows: Given a set of atoms (each with fixed charge Z_I) fixed at a set of nuclear coordinates $\{\mathbf{R_I}\}$, a single-particle basis $\{\phi_i\}$, and a number of electrons N, solve the time-independent Schrödinger equation (Eq. 2.1) with $\hat{H}_{\text{el}}$ given by Eq. 2.9. As discussed above, we are rarely interested in the full set of eigenvalues and eigenvectors. Moreover, we are typically not interested in any wavefunction *per se*. For moderately large

[1] Although we generally use the hat notation $\hat{O}$ to denote operators, we avoid using the hat for the fermionic operators, for convenience.

systems, reading out any full wavefunction expansion would be unfeasible, even if this wavefunction might be efficiently encoded on a quantum computer. The precise outcome of the calculation thus depends on the application context: common properties of interest are the energy of the electrons, or the forces generated by the electrons on the nuclei, as a function of the nuclear coordinates, for a small set of the low-energy eigenstates, often only the ground state.

In this thesis, we always compute the energy of the ground state and often also of a small number of the low-lying excited states. When memory allows, we also explicitly compute and store the wavefunction for benchmarking it against the exact wavefunction.

Below, we review some of the classical algorithms for electronic structure relevant to this thesis.

2.2 Classical algorithms for electronic structure

Quantum chemistry is a mature yet active field with a century of history. Unlike in condensed matter theory, or some other fields involving simulation of physical systems, where establishing the underlying equation of motion or Hamiltonian is often part of the challenge, the electronic structure Hamiltonian in quantum chemistry is well-defined (Eq. 2.2 above). However, exact solutions are intractable beyond a small number of electrons. As such, the majority of research activity focuses on the development and application of approximate methods for solving the electronic structure problem. In the context of quantum computing, some of these techniques are used to design quantum algorithms whereas others are used as a classical benchmark for quantum algorithms. Below, we provide an overview of the methods and concepts most relevant to this thesis.

2.2.1 Basis sets and mean-field theory

Given N electrons and a single-particle basis of $2M$ orbitals, a complete many-body basis for the Hilbert space is formed by $\binom{2M}{N}$ Slater determinants. Thus, computing the exact wavefunction requires solving the Schrödinger equation with

$\binom{2M}{N}$ degrees of freedom. This combinatorial scaling with system size prohibits exact simulations of most molecules of interest.

To enable scalable computations, a simple yet surprisingly useful approximation is often made to start: the electronic coordinates are treated as if they were independent from each other. This turns an N-body problem into N coupled one-body problems. To solve said system of equations, the most commonly-used wavefunction-based approach is the Hartree–Fock (HF) method.[2] The goal of the HF method is to obtain the set of orbitals $\{\phi_i\}, i = 1, 2, ..., 2M$ that minimise the energy of a single Slater determinant.

Atomic orbitals and basis set

As we will see, the molecular orbitals going into the HF equations are linear combinations of atomic orbitals. Since an analytical solution to the Schrödinger equation is known for the hydrogen atom (as well as ions with a single electron but different nuclear masses), a common choice of atomic orbitals are so-called Slater-type orbitals (STOs). These have the same qualitative form as the hydrogenic orbitals but are parametrised such that they can be fitted to different atoms. In spherical coordinates, these are given by:

$$\psi_{\ell,m}(r, \theta, \phi)^{\mathrm{STO}} = N_\ell r^\ell e^{-\zeta r} Y_{\ell m}(\theta, \phi), \tag{2.12}$$

where $\ell \in 0, 1, ..., n-1$ is the orbital angular momentum quantum number for the energy level n, ζ is the fitting parameter, $Y_{\ell m}$ are spherical harmonic functions, and N_ℓ is a normalisation constant. Due to the structure of the spherical harmonics, each value of ℓ comes with $2\ell + 1$ orbitals which share the radial part of the wavefunction in Eq. 2.12 but differ in their orientations. In the absence of magnetic fields, all $2\ell + 1$ orbitals are degenerate. These different values of ℓ give different characteristic shapes to the orbitals: s-type orbitals have spherical symmetry ($\ell = 0$), p-type orbitals have a dumbbell shape ($\ell = 1$), d-type orbitals

[2]Density Functional Theory (DFT) is also a one-body theory which replaces the many-body wavefunction with the density, and it can in principle yield an exact solution to the many-body problem. However, it is not directly relevant to this thesis.

have a cloverleaf shape ($\ell = 2$), etc. The origin $r = 0$ is set such that the orbitals are centred around each atom.

In practice, the STOs are replaced by a linear combination of Gaussian-type orbitals (GTOs), each of the form

$$\psi_{\ell,m}(r,\theta,\phi)^{\mathrm{GTO}} = M_\ell e^{-\zeta r^2} Y_{\ell m}(\theta,\phi), \tag{2.13}$$

where M_ℓ and ζ again a constant and fitting parameter. GTOs are chosen for computational reasons: their Gaussian form allows accurate and fast evaluation of integrals between different orbitals, which is necessary to construct the Hamiltonian. Since STOs more accurately represent the electron density, the linear combination of GTOs is fitted to closely match the form of the STO. The atomic orbitals have the form

$$\tilde{\phi} = \sum_i \alpha_i \psi_i^{\mathrm{GTO}}. \tag{2.14}$$

In parts of this thesis, we make use of the STO-3G basis set, which is part of the STO-nG series, a series of very simple basis sets for which each atomic orbital is one STO and n Gaussians are used to represent each STO.[210] STO-nG basis sets are an example of minimal basis sets. In a minimal basis set, each atomic orbital corresponds to one basis function, with some exceptions. For example, in the STO-nG basis, all atoms on the first row of the periodic table (H and He) have a $1s$ orbital, and all atoms on the second row (Li through Ne) have a $1s$, a $2s$, and three $2p$ orbitals (p_x, p_y, p_z). This provides a basic description of the molecule and can be used as a first approximation. Apart from reducing the complexity of the Schrödinger equation, it also enables reasoning about the electronic structure through intuitive chemical heuristics, which is why we use it as a starting point in this thesis. However, the resulting discretisation error is too large to obtain quantitatively accurate computations of chemical properties.[211] To reduce the discretisation error, more basis functions are included per atomic orbital, progressively increasing the flexibility of the parametrisation. Nevertheless,

for simplicity, we will from here on refer to each GTO-based basis function as an *atomic orbital*, or AO.

The advantage of using these atomic orbitals over, say, delta-like functions on a grid, is that these already encode a large amount of information about the physics of the problem through the radial and spherical-harmonic shape of the wavefunction, including the symmetries of the spherical harmonic functions and the exponential decay in wavefunction amplitude as a function of distance from the nucleus. Thus, compared to more generic basis functions, significantly fewer basis functions are required to achieve a given accuracy for calculations on molecules.

Molecular orbital theory and Hartree–Fock theory

Beyond the hydrogen-like solutions for atoms with a single electron, there exist no exact analytic solutions to the Schrödinger equation for molecules. As such, the theory of molecules and the chemical bond is necessarily heuristic.

The two prevalent theories are molecular orbital (MO) theory and valence bond theory; both can be used as the basis of computational many-body methods, with MO theory clearly surpassing valence bond theory in popularity. In short, valence bond theory describes bonds as shared electron pairs formed when the AOs from different atoms locally overlap.[212] It provides a localised picture of chemical bonding, where electrons belong to individual bonds.

In contrast, molecular orbital theory is a delocalised picture of chemical bonds in which AOs of different atoms are combined to form MOs, linear combinations of atomic orbitals:

$$\phi_i = \sum_i C_{ij} \tilde{\phi}_j, \tag{2.15}$$

where ϕ_i is an MO and $\tilde{\phi}_j$ are AOs are centred around each atom, e.g., as given per Eq. 2.14. Typically, the MOs are fixed to be orthonormal to each other, and thus the matrix formed by the coefficients C_{ij} is unitary. From this definition, it is clear that MO theory predicts that electrons are generally not confined to a spatially local region around their atom of origin, but instead, they can delocalise

across the molecule when the atoms come together to form a bond. Bonding MOs are formed by constructive interference between the AOs, leading to an increased electron density between the nuclei. Antibonding MOs are formed from destructive interference between the MOs, decreasing the electron density between the nuclei. Thus, when bonds are weakened, the probability of occupying an antibonding orbital increases.

Hartree–Fock theory can be viewed as a computational implementation of MO theory. In HF theory, the many-electron wavefunction is approximated by a single Slater determinant (Eq. 2.6),[130, 131] namely the one that doubly occupies all the lowest-energy MOs: provided N electrons in M spatial orbitals ordered by increasing energy, the HF ansatz reads:

$$|\psi_{\mathrm{HF}}\rangle = \underbrace{|1,1,\ldots,1\rangle}_{N}\underbrace{|0,0,\ldots,0\rangle}_{2M-N}, \tag{2.16}$$

where in our notation, the product of two fermionic states is assumed to be an antisymmetrised product. The HF wavefunction can be obtained by rotating the orbitals through a unitary transformation, i.e., varying the coefficients C_{ij} in Eqn. 2.15 for each orbital ϕ, such that the energy of the determinant is minimised. This is equivalent to diagonalising the Fock operator, an effective one-body operator:

$$\hat{F} = \hat{h} + \hat{v} = \sum_{pq} F_{pq} a_p^\dagger a_q, \tag{2.17}$$

where the matrix elements are:

$$F_{pq} = h_{pq} + \sum_r \sum_s B_{rs}(h_{prqs} - h_{prsq}). \tag{2.18}$$

Here, h_{pq} is the one-body term from the original electronic Hamiltonian in second quantisation (Eq. 2.10). The second term in the Fock matrix (Eq. 2.18) is defined from the two-body integrals from the original Hamiltonian (Eq. 2.11). The first term, $J_{pq} = \sum_r B_{rs} h_{prqs}$, corresponds to the direct classical Coulomb repulsion between charges (electrons), and is referred to as the Coulomb term. The second term, $K_{pq} = \sum_r B_{rs} h_{prsq}$, is entirely of quantum-mechanical origin: it accounts for the fact that electrons are indistinguishable particles (fermions) and

therefore two electrons of the same spin cannot occupy the same spatial orbital, as per the Pauli principle, generating an effective repulsion. Both the Coulomb and exchange repulsion are summed over all electrons, as evident from the indices in Eqs. 2.17 and 2.18. Thus, each electron experiences the average or mean field created by all other electrons.

The Fock operator is approximately diagonalised in a self-consistent procedure: since the operator itself depends on the orbitals (as these define the integrals h_{pq}, h_{prqs} and h_{pqrs}), we cannot immediately solve for the orbitals. Instead, we apply an iterative procedure: we provide an initial guess of the orbital coefficients C_{ij}, construct the Fock matrix using these orbitals and solve the Roothaan-Hall equations:

$$\boldsymbol{FC} = \boldsymbol{SC}\boldsymbol{\epsilon}, \tag{2.19}$$

where $\boldsymbol{F}$ is the Fock matrix, $\boldsymbol{C}$ is the matrix with the MO coefficients, $\boldsymbol{S}$ is the overlap matrix between the AOs (necessary since these are generally nonorthogonal, even if the resulting MOs are). The orbitals are determined from the eigenvectors of the Fock operator (columns of $\boldsymbol{C}$); the eigenvalues of the Fock operator, $\boldsymbol{\epsilon}$, are the orbital energies.

From the above discussion, it is clear that MO and HF theory are incapable of describing localised phenomena: electrons are always delocalised across the molecules, as the HF orbitals are formed from the average potential caused by the electrons and inherit the point group symmetry of the molecule. This is what leads to the formation of, e.g., the σ and π MOs, which are delocalised functions that inherit the angular momentum structure of the AOs leading to bonding and antibonding orbitals depending on the symmetry of the coefficients in the linear combination of AOs. As we discuss in Section 2.4 of this Chapter and in Chapter 3 (Ref. 51), strong correlation effects usually occur when electrons localise around fragments of the molecule, which makes MO and HF theory unsuitable for the description of such phenomena. In such situations, many near-degenerate HF solutions can exist,[213] complicating the self-consistent field procedure. As a

mathematical minimisation problem, the HF problem is intractable in the worst case: it is **NP**-complete.[40]

Nevertheless, HF theory is often satisfactory in practice, both in the sense that the energy minimum can be found reliably and in that the HF state gives a qualitatively accurate description of electronic wavefunctions in molecules with weak electron correlation. It is cornerstone in quantum chemistry as it serves to define a set of orbitals as well as providing an initial guess for the many-electron wavefunction, the lowest-energy Slater determinant, which can serve as a reference for more sophisticated and accurate methods.

2.2.2 Exact solutions through full configuration interaction

From the HF method, we have obtained the HF wavefunction, $|\psi_{\mathrm{HF}}\rangle$, for a set of N electrons in M (spatial) molecular orbitals. Here, the $N/2$ spatial orbitals (N spin-orbitals) with lowest energy are occupied by two electrons each, for a closed-shell system. The remaining $M - N/2$ orbitals remain unoccupied. The full CI wavefunction can be constructed starting from the HF states as a reference, by applying a series of *excitation operators* that *excite* electrons from the occupied to the unoccupied (*virtual*) orbitals:

$$|\psi_{\mathrm{FCI}}\rangle = (1 + \hat{T})\,|\psi_{\mathrm{HF}}\rangle\,, \tag{2.20}$$

where

$$\hat{T} = \sum_{i=1}^{N} \hat{T}_i, \tag{2.21a}$$

$$\hat{T}_1 = \sum_{\substack{i\in\mathrm{occ}\\ a\in\mathrm{virt}}} t_a^i a_a^\dagger a_i \tag{2.21b}$$

$$\hat{T}_2 = \sum_{\substack{i>j\in\mathrm{occ}\\ a>b\in\mathrm{virt}}} t_{ab}^{ij} a_a^\dagger a_b^\dagger a_i a_j, \tag{2.21c}$$

...

Here, t_a^i, t_{ab}^{ij} are linear expansion coefficients, and occ and virt denote the occupied and unoccupied (virtual) orbitals in the Hartree-Fock reference state. The ground state wavefunction is obtained by variational minimisation of the energy functional:

$$|\psi_{\mathrm{FCI}}\rangle = \min_{\boldsymbol{t}} \frac{\langle\psi_{\mathrm{FCI}}| \hat{H}_{\mathrm{el}} |\psi_{\mathrm{FCI}}\rangle}{\langle\psi_{\mathrm{FCI}}|\psi_{\mathrm{FCI}}\rangle}, \tag{2.22}$$

where $\boldsymbol{t}$ is a vector with all the expansion coefficients.

The Full CI method yields the exact ground state of the electronic Hamiltonian, within a fixed basis. However, it requires including all $\binom{2M}{N}$ determinants that span the basis, which is intractable beyond small systems. In practice, the FCI wavefunction is found through subspace iteration, rather than building and diagonalising the full $\binom{2M}{N} \times \binom{2M}{N}$ matrix (this has some resemblance to the quantum subspace diagonalisation algorithms presented in this thesis).[214] However, due to the combinatoric scaling of the dimension, routine computations can only be achieved for roughly 18 electrons in 18 orbitals.[27] To our knowledge, the current world record is set at 26 electrons and 23 orbitals through sophisticated, massively parallel, distributed and multithreaded implementations on state-of-the-art supercomputers.[20]

A more commonly used configuration interaction method is CI singles and doubles (CISD), where the excitation operators are limited to be of order two:

$$|\psi_{\mathrm{CISD}}\rangle = (1 + \hat{T}_1 + \hat{T}_2) |\psi_{\mathrm{HF}}\rangle . \tag{2.23}$$

The resulting algorithm scales as $\mathcal{O}(N^6)$, but the linear expansion makes it unsuited for describing large amounts of (both weak or strong) electron correlation.

2.2.3 Coupled cluster: high accuracy with polynomial scaling

The coupled cluster method utilises the same excitation operators as the FCI method. However, a much more powerful wavefunction is produced by expanding it in terms of an exponential excitation operator

$$|\psi\rangle = e^{\hat{T}} |\psi_{\mathrm{HF}}\rangle , \tag{2.24}$$

where the operator $\hat{T}$ is the same one defined for FCI in Eq. 2.21a. The coupled cluster wavefunction is exact when it includes excitations up to the number of electrons, $\hat{T}_N$, as it becomes equivalent to the FCI wavefunction (Eq. 2.20).

In practice, the CC ansatz is typically truncated to a small number of operators. It can be shown that if only single excitations $\hat{T}_1$ are included, the resulting CCS (coupled cluster singles) wavefunction gives no improvement in the energy from the Hartree–Fock method, since the CCS ansatz is equivalent to performing an orbital rotation.[215]

A more correlated wavefunction can be parametrised by including both singles and doubles excitations (CCSD):

$$|\psi\rangle = e^{\hat{T}_1+\hat{T}_2}\,|\psi_{\mathrm{HF}}\rangle\,. \tag{2.25}$$

The CCSD(T) method uses the CCSD equations but adds a perturbative correction to the CCSD energy. CCSD(T) has been shown to be very accurate for a range of small-to-medium-sized molecules at equilibrium and is often referred to as the gold standard method for quantum chemical calculations.[35]

The CC method has deep connections to many-body perturbation theory. A diagrammatic formulation can be used to show that the ansatz has favourable properties such as size consistency and size extensivity. A method is size consistent if the energy of two non-interacting subsystems equals the sum of the energy of the individual subsystems. This ensures correct computation of dissociation energies, and is guaranteed by the exponential structure of the wavefunction, which is multiplicatively separable: Given two non-interacting subsystems A and B with a factorisable reference

$$|\Phi_{AB}\rangle = |\Phi_A\rangle \otimes |\Phi_B\rangle\,, \tag{2.26}$$

and disjoint cluster operators satisfying $[\hat{T}_A, \hat{T}_B] = 0$, we have

$$e^{\hat{T}_A+\hat{T}_B} = e^{\hat{T}_A}e^{\hat{T}_B} = e^{\hat{T}_A} \otimes e^{\hat{T}_B}. \tag{2.27}$$

Therefore,

$$\begin{aligned} |\Psi_{\mathrm{CC}}(A{+}B)\rangle &= e^{\hat{T}_A+\hat{T}_B}\Big(\,|\Phi_A\rangle \otimes |\Phi_B\rangle\,\Big) \\ &= \Big(e^{\hat{T}_A}\,|\Phi_A\rangle\,\Big) \otimes \Big(e^{\hat{T}_B}\,|\Phi_B\rangle\,\Big) \\ &= |\Psi_{\mathrm{CC}}(A)\rangle \otimes |\Psi_{\mathrm{CC}}(B)\rangle\,. \end{aligned} \tag{2.28}$$

With $\hat{H} = \hat{H}_A + \hat{H}_B$ it follows that

$$E_{\mathrm{CC}}(A{+}B) = E_{\mathrm{CC}}(A) + E_{\mathrm{CC}}(B). \tag{2.29}$$

Size extensivity is the requirement that the energy scale linearly with system size; for a homogeneous N-particle system,

$$\lim_{N\to\infty} \frac{E(N)}{N} = \varepsilon, \tag{2.30}$$

where ε is a finite constant (the energy per particle). This property ensures that the error grows systematically with system size. It is particularly helpful when computing energy differences, as some of the error made when approximating the energies of different systems can cancel out. This reduces the relative error, which is the most important quantity for determining chemical reaction rates. In contrast, CISD is not size extensive.

The power of the CC ansatz comes from the fact that higher-order excitations can be obtained from exponentiating lower-order excitations thanks to the exponential structure. This can be seen from the expansion

$$e^{\hat{T}_1+\hat{T}_2} = 1 + (\hat{T}_1 + \hat{T}_2) + \frac{1}{2!}(\hat{T}_1 + \hat{T}_2)^2 + \frac{1}{3!}(\hat{T}_1 + \hat{T}_2)^3 + \ldots. \tag{2.31}$$

In CCSD, besides the singles ($\hat{T}_1$) and doubles ($\hat{T}_2$) excitations that are included explicitly, triples excitations are obtained through products of singles and doubles ($\frac{1}{2}\hat{T}_1\hat{T}_2 + \frac{1}{2}\hat{T}_2\hat{T}_1$) and quadruples are included through products of two doubles ($\frac{1}{4}\hat{T}_2\hat{T}_2$). Apart from guaranteeing size extensivity and consistency, this effectively generates a wavefunction that contains a very large number of determinants but with a smaller number of parameters. For example, CCSD generates all the determinants up to quadruple excitations (the ones that also appear in CISDTQ) but with a smaller number of parameters (the same number as CISD). However,

the CC wavefunction is never explicitly stored. Instead one can solve for the coupled cluster amplitudes (coefficients in the excitation operator, t_i^a and t_{ij}^{ab} in (Eq. 2.21a)), and compute the energy from those coefficients by solving a projected Schrödinger equation.[35]

The CCSD(T) method, when based on the Hartree–Fock reference, gives accurate energies for systems with weak electron correlation while scaling like $\mathcal{O}(N^7)$. However, it fails catastrophically (diverges) in the presence of strong electron correlation, much like other methods based on many-body perturbation theories that use the Fock operator as the reference. An intuitive reason is that the orbitals become near-degenerate, leading to small energy differences in the denominator of the perturbative expansion of the Hamiltonian.

By inspecting the Slater determinant expansion of strongly correlated wavefunctions, as we do in this thesis (Chapter 3), it becomes clear that many higher-order excitations are needed from the HF determinant to describe such systems. These high-order excitations are not contained in the CCSD(T) wavefunction. Multireference coupled cluster theory aims to circumvent this by using a more general reference state composed of a linear combination of Slater determinants, or by expanding the wavefunction around multiple reference states. Unfortunately, implementations for systems with many strongly correlated electrons remain inaccurate or impractical,[160] although some progress is being made in this direction.[216] We briefly discuss these different notions of electron correlation below.

2.2.4 Weak and strong electron correlation

A rigorous breakdown of correlation into classical correlation, quantum correlation, and entanglement between orbitals can be obtained from quantum-information-theoretic measures.[217] However, defining a single quantity for the correlation contained in a many-electron state is challenging, since it depends both on the choice of the basis and the partitioning of the system into subsystems.

Here we review the more commonly used but less rigorously defined concepts of weak, strong, static and dynamic correlation in many-electron systems. These are

helpful for characterising approximate many-body wavefunctions and methods. We define the correlation energy as the difference in energy between the exact ground state energy (e.g., as obtained from FCI) and the HF energy

$$E_{\text{corr}} = E_{\text{FCI}} - E_{\text{HF}}. \tag{2.32}$$

Under this commonly-used definition, the HF state is completely uncorrelated. This is incorrect in the purely statistical sense of correlation between two variables, since the positions and spin variables of different electrons are certainly correlated: from the Pauli exclusion principle, their wavefunctions must be antisymmetric with respect to exchange. The correlation induced due to antisymmetry is also known as Fermi correlation.[218] However, since this is trivially incorporated by working in the Slater determinant basis, we ignore this type of correlation.

Weak and strong correlation are less well-defined, but can be understood qualitatively. Weakly correlated systems are those for which a single Slater determinant (the HF determinant) dominates the exact wavefunction expansion (Eq. 2.20), i.e., $|\langle\psi_{\text{FCI}}|\psi_{\text{HF}}\rangle|$ is large compared to the coefficients of all the other determinants. The contributions to the correlation energy coming from all the other determinants with small coefficients are referred to as weak correlation. For such systems, the HF state provides a qualitatively accurate description of the electronic structure; e.g., many small organic molecules at equilibrium can be described qualitatively by HF or MO theory. Consequently, the HF state forms a good reference state for correlated calculations, and methods such as CCSD(T) can provide highly accurate energy estimates at a reasonable $\mathcal{O}(N^7)$ cost.

In contrast, strongly correlated systems are those for which multiple Slater determinants contribute significantly to the wavefunction. In this situation, the HF state becomes a poor approximation. For example, for stretched bonds, the HF state has an overlap with the FCI ground state that decays exponentially with the number of valence electrons because all the valence orbitals become near-degenerate.[50] A similar phenomenon occurs in transition metal clusters even at

equilibrium geometries.[41] In Chapters 3 and 4, we analyse this wavefunction structure in detail and its consequences for quantum algorithms.

Apart from analysing the FCI wavefunction, which is unfeasible for large systems, it is possible to observe signatures of strong correlation from the results of approximate methods: (1) Multiple orbitals are close in energy. (2) The occupation of natural orbitals—the orbitals that diagonalise the one-particle reduced density matrix—shows a large number of singly occupied orbitals, also known as open-shell orbitals (as opposed to doubly occupied, i.e., closed-shell orbitals or empty orbitals). (3) Unrestricted HF (UHF) calculations yield broken-symmetry states with a high spin expectation value $\langle \hat{S}^2 \rangle$.[219] [3] (4) The entanglement entropy or mutual information between orbitals is high.[220] (5) Coupled cluster expansions fail.[35, 221]

The terms weak and strong corrrelation are often used in equivalence to dynamic and static correlation. Most systems contain some dynamic correlation, perhaps with the exception of bonds at dissociation in a minimal basis, which can be described exactly by a single configuration state function.[50] Dynamic correlation is particularly prevalent when employing large basis sets for the Hilbert space discretisation since, by greatly enlarging the Hilbert space beyond the valence orbitals, this generates a lot of states that have small (but nonzero) probability of being occupied.

Due to the need for exponential expansions of the wavefunction for strongly correlated systems, to date, there is no classical method that can describe such systems in a scalable manner (in the sense that the computational cost scales moderately with the number of strongly correlated electrons). However, in most molecules with strong correlation, only a subset of electrons near the HOMO-LUMO gap[4] are strongly correlated, whereas the remaining electrons remain weakly correlated. Thus, a common strategy is to partition the orbital space into three sets of orbitals, ordered by their energy: the core orbitals are those which are

[3] In a UHF calculation, each α and β spin-orbital can have a different spatial wavefunction, as opposed to Restricted HF calculations for which they share the same spatial function.

[4] HOMO stands for Highest Occupied Molecular Orbital, LUMO standard for Lowest Occupied Molecular Orbital.

occupied by two electrons with high probability; the active orbitals are occupied by the valence electrons, which can be strongly correlated, with varying occupations including many configurations that occupy each orbital with only one electron; the virtual orbitals are the highest energy orbitals which remain empty with high probability. Given this partitioning, a common strategy is to use an exponentially scaling method within the active space. The complete active space configuration interaction (CASCI) is precisely this: a FCI calculation is performed within the active space, while the core orbitals remain doubly occupied. The complete active space self-consistent field method (CASSCF) performs a CASCI expansion while self-consistently optimising the orbitals (including the core orbitals);[162] by including *orbital relaxation* effects, this can lower the energy as the orbitals are variationally optimised for a many-body state (the CASCI state in the active space) rather than the HF state.

To obtain quantitative accuracy, large basis sets (much larger than minimal basis sets) are required, which leads to many small contributions of electrons in higher-energy orbitals. Thus, it is necessary to combine an accurate active space calculation of the static correlation with an efficient method for computing dynamic correlation effects. Since CASCI and CASSCF only correlate the valence electrons, they cannot capture dynamic correlation which also requires excitations to orbitals outside of the active space. This can be achieved by combining CASSCF with second order pertubation theory in a method known as CASPT2.[222, 223]

The below nomenclature is used inconsistently in the literature,[160] but we define a series of concepts as follows. Computational methods or theories that use a single Slater determinant as the reference (typically, the Hartre–Fock state) and expand around it—such as the standard coupled cluster method presented in Section 2.2.3—are referred to as single reference methods. In contrast, methods that use several reference states are known as multireference methods.

Wavefunctions with a large contribution from multiple Slater determinants are referred to as being multideterminantal. At times, multideterminantal is also used synonymously with multiconfigurational, since each Slater determinant corresponds

to a different configuration of electrons in orbitals. However, there exist situations for which multiple Slater determinants have a large weight in the wavefunction expansion, but they can be expressed in terms of a single configuration state function (CSF), a symmetry-adapted linear combination of determinants, which we define more rigorously in Section 2.4. Since a single CSF can be thought of as a single electron configuration, in our view, wavefunctions for which a single CSF dominates should not be referred to as multiconfigurational, but only multideterminantal instead. Thus, truly multiconfigurational states are those for which multiple CSFs are required for a qualitative description. We reiterate that, apart from always being qualitative, these definitions are not always followed in the literature. However, we believe that this classification provides more clarity, and it has also been argued by others.[224]

In summary, the standard consensus in the field is that very large wavefunction expansions are required to describe the strong correlation within the active space and that this requires expensive, poorly scaling methods, whereas dynamic correlation can be obtained through less costly methods such as low-order perturbation theory or coupled cluster theory. Thus, many types of molecular electronic states remain challenging to resolve in a scalable fashion. For example, besides the cost limitations posed by expensive calculations within the active space, it can be challenging to combine a strongly correlated wavefunction in the active space with large amounts of dynamic correlation in the remaining orbitals.[160] Although some approaches achieve this for open-shell systems, they are typically not scalable beyond a few strongly correlated electrons.[160, 161, 225] Spin-adapted coupled cluster approaches for multideterminantal wavefunctions might provide a good solution for a subset of strongly correlated systems but remain inaccurate for truly multiconfigurational states that require multiple highly entangled reference CSFs.[216]

Due to the scaling limitations of FCI-like approaches for the active space and the need to systematically study strongly correlated systems in increasingly large bases, significant progress has been made in alternative methods for strong electron

correlation. These might still scale exponentially, but their implementation is more practical: their cost can be controlled by the user and they can exploit some structure in the wavefunction to avoid computing all variational coefficients present in an FCI expansion, effectively reducing the constant prefactors in the scaling. We summarise a few of these methods below.

2.2.5 Approximate ansätze with exponential scaling

Selected Configuration Interaction and FCI-QMC

Selected CI (SCI) is a popular approach for strongly correlated electrons. It comes in many variants, including adaptive sampling [226] and heat bath [227] configuration interaction as well as methods based on cluster expansions.[228–230] More recently, it has been adapted to quantum algorithms in a method known as quantum selected configuration interaction (QSCI).[231] SCI algorithms partition the Hilbert space into a *reference space* (e.g., at first, the Hartree–Fock determinant) and an *environment* that includes possible determinants that can be added to the wavefunction. At each iteration, the algorithm screens a set of determinants from the environment and decides whether to include them in the wavefunction ansatz or discard them. This screening is done by evaluating Hamiltonian matrix elements, which might include perturbative terms, and applying a deterministic or stochastic filtering criterion.[226, 227] At each iteration, the Hamiltonian is diagonalised within the space of all the reference and newly selected determinants, until convergence in the energy or coefficients is reached.

A popular alternative is the FCI-QMC method, which defines the search within the Hilbert space in terms of so-called *walkers*.[163] Each walker is assigned to a Slater determinant and the wavefunction coefficients are determined based on the population of walkers for a given determinant. The acceptance of new walkers is decided based on a Monte-Carlo sampling criterion (such as Metropolis–Hastings under Markovian dynamics). The algorithm effectively simulates the imaginary-time Schrödinger equation until the population of walkers reaches a steady state (or a maximum number of walkers specified by the user), and the ground state

energy is estimated through a stochastic sampling protocol.[163] Unlike other SCI methods, it does not use perturbative techniques to define the determinant space and is thus more successful for strongly correlated systems where the Hartree–Fock determinant is a poor approximation.[107]

The motivation behind SCI and FCI-QMC methods is to exploit sparsity in the wavefunction expansion while circumventing the limitations of truncated CI expansions. By including Slater determinants that are highly excited with respect to the reference determinants, the algorithms are in principle able to select determinants with very high excitation rank, unlike, e.g., the CISD method which only includes up to double excitations from the reference determinant. Moreover, stochastic approaches such as FCI-QMC never explicitly store the full determinantal expansion to estimate the energy, which enables exploring much larger Hilbert spaces than those of deterministic methods. Thus, these methods have been applied to study increasingly large strongly correlated systems.[19, 107, 163, 232] Nevertheless, their scaling is still exponential due to the exponential growth of the Hilbert space, and their accuracy remains limited for systems where a very large number of determinants is required. Besides the memory bottleneck, searching within this exponential space can lead to numerical complications.[233]

Density Matrix Renormalization Group

DMRG employs an alternative wavefunction parametrisation. Rather than working with an expansion in the Slater determinant basis, the wavefunction is expressed in terms of a matrix product state.[110, 111] This is also discussed in Appendix B.6 and Ref. 51, but we briefly define it here for completeness. The matrix product state takes the form

$$|\Psi\rangle = \sum_{\boldsymbol{m}} A_1^{[m_1]} A_2^{[m_2]} \cdots A_M^{[m_M]} \prod_{i=1}^{M} |m_i\rangle \,. \tag{2.33}$$

Here, the one-electron space is defined by a set of M sites (corresponding to molecular orbitals), and $m_i \in \{0, 1, ..., d-1\}$ is the so-called *physical* index that runs over the possible occupations of the Hilbert space of a single site (of *local*

dimension d, where $d = 4$ since orbitals can have the four occupations $\{0, \alpha, \beta, 2\}$). The vector $\boldsymbol{m} = (m_1, m_2, ..., m_M)$ defines the occupation number for each site.[15]. Each $A_i^{[m_i]}$ is a tensor of order χ_i, where χ_i is the bond dimension, and each bond dimension is bounded by the maximal bond dimension, $\chi_i \leq \chi$. The bond dimension directly determines the amount of entanglement that can be generated between subsystems through an MPS. Specifically, the maximum entanglement of an MPS of bond dimension χ scales as $\mathcal{O}(\log(\chi))$.[234]

The MPS representation is exact using a polynomially growing bond dimension for systems which fulfill so-called area laws.[235] This refers to situations in which the entanglement does not grow with the size (volume) of the system, but rather, with the boundary between subsystems, as is the case for one-dimensional Hamiltonians with local interactions, such as certain spin models [235] or one-dimensional models of chain-like molecules. In contrast, for systems whose entanglement grows with the system size, the bond dimension required to achieve high accuracy generally scales very steeply, at worst exponentially with N.[236]

The DMRG algorithm is used to variationally optimise the tensor coefficients to minimise the ground state energy, in a procedure that involves truncating the coefficient matrix after singular value decomposition. Like for all formally exact electronic structure methods, the computational cost (here, determined by the bond dimension χ) required to obtain the exact (FCI) wavefunction through this ansatz scales combinatorially with the number of orbitals/electrons. However, it has been shown that a MPS with a relatively modest bond dimension ($\chi \approx 1000 - 4000$) can be used to obtain highly accurate energy estimates for challenging strongly correlated systems including the chromium dimer and transition metal clusters.[17, 18, 108] To our knowledge, together with FCI-QMC, DMRG is the leading algorithm for systems with many strongly correlated electrons, and thus the most likely competitor for any quantum algorithm.[41]

2.3 Quantum algorithms for quantum chemistry

Below, we give an overview of the main techniques for simulating quantum chemistry on digital quantum computers as relevant to this thesis. We assume an elementary knowledge of quantum computation, as given by the first four chapters of the book by Nielsen and Chuang.[48] This section is less comprehensive than the previous section on classical algorithms for quantum chemistry because a lot of the background material on quantum algorithms is already provided in the publications (Chapters 3–5). We discuss how to map fermions to qubits (Section 2.3.1), the quantum phase estimation algorithm (Section 2.3.2)—including the subroutines of Hamiltonian simulation and initial state preparation—the variational quantum eigensolver (Section 2.3.3), and a brief description of the considerations required to account for the overheads due to hardware noise in both the NISQ as well as fault-tolerant quantum computing paradigms (Section 2.3.4).

2.3.1 Mapping fermions to qubits

The Pauli operators and the identity matrix

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.34}$$

form a complete basis for arbitrary Hermitian operators on a single qubit. This extends to systems of n qubits through tensor products of these operators. Thus, any observable, such as the electronic structure Hamiltonian, can be expressed as a linear combination of tensor products of Pauli operators.

Qubits are distinguishable two-level systems. Consequently, unlike for bosons or fermions, their states are not symmetric nor antisymmetric with respect to exchange. To encode the states of systems of multiple fermions in the states of qubits and apply operators onto them, we must enforce the antisymmetric structure. In second-quantised encodings, this is achieved by antisymmetrising the qubit (Pauli) operators; in first quantised encodings, the qubit wavefunction is explicitly antisymmetrised, and no antisymmetrisation needs to be enforced on the operators.[66, 68]

In second quantisation, the electronic wavefunction of $2M$ spin-orbitals is mapped to a quantum register of $2M$ qubits. The most commonly used encoding is the Jordan–Wigner encoding, in which the state of each qubit corresponds to the occupation of a spin-orbital: the qubit is in state $|1\rangle$ if the orbital is occupied and $|0\rangle$ if it is unoccupied:

$$|f_1, f_2, ..., f_{2M}\rangle \rightarrow |j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_{2M}\rangle = |j_1, j_2, ...j_{2M}\rangle = |j\rangle\,, \tag{2.35}$$

with $f_i, j_i \in \{0, 1\}$. Here the right hand side is the qubit vector, a standard tensor product of each qubit state, and the left hand side is the fermionic occupation number vector for a Slater determinant (Eq. 2.7), an antisymmetrised tensor product of one-fermion states.

To ensure that antisymmetry is satisfied, we need an exact isomorphism between the canonical fermionic operators and the qubit operators. The Jordan–Wigner transformation provides one such mapping between spin algebras and fermionic algebras, defined as:[70]

$$a_p = \frac{1}{2}(X_p + iY_p) \bigotimes_{k=1}^{p-1} Z_k, \tag{2.36a}$$

$$a_p^\dagger = \frac{1}{2}(X_p - iY_p) \bigotimes_{k=1}^{p-1} Z_k, \tag{2.36b}$$

where $\bigotimes_{k=1}^{p-1} Z_k = Z_{p-1} \otimes Z_{p-2} \otimes ... \otimes Z_1$.

The X and Y terms ensure that creation operators create a particle if they act on the $|0\rangle$ state and annihilation operators delete a particle if they act on the $|1\rangle$ state. The Z-string computes the parity of all spin-orbitals/qubits of index smaller than p, inducing a phase factor of -1 on the state if the parity is odd and leaving the state unchanged if the parity is even.

Despite the conceptual simplicity of the Jordan–Wigner mapping, it has the disadvantage of being highly non-local due to the string of Pauli operators, which act on $\mathcal{O}(M)$ qubits. This introduces an overhead in many of the standard quantum circuits required in quantum algorithms for fermionic simulation. For example,

ansatz-based quantum algorithms often require applying exponentials of excitation and de-excitation operators

$$a_p^\dagger a_q - a_q^\dagger a_p = \frac{i}{2}(Y_q X_p - X_q Y_p) \bigotimes_{k=q+1}^{p-1} Z_k. \tag{2.37}$$

This can be implemented directly in terms of known quantum circuits,[177, 237] but they require $\mathcal{O}(M)$ additional two-qubit entangling gates (e.g., CNOT gates) due to the Z-string. This is particularly problematic in NISQ algorithms where two-qubit gates are the largest source of error. This overhead can be avoided for certain two-body operators, as we discuss in Appendix C (Supplementary Material of Ref. 52). In the context of Hamiltonian simulation, where exponentials of the Hamiltonian terms such as $a_p^\dagger a_q$ must be implemented, an alternative circuit implementation in terms of fermionic swap networks can also reduce this overhead.[169, 238]

A popular alternative encoding is the Bravyi–Kitaev transformation, which was designed with the purpose of minimising the locality of fermionic operators on qubits.[239] Under this mapping, fermionic operators act non-trivially only on $\log_2(M)$ qubits.[240] There exist a wide range of other fermion-to-qubit encodings [32] as well as algorithms designed to create custom encodings, e.g., to optimise the circuits to a particular hardware architecture.[241] In this thesis, we only consider the Jordan–Wigner transformation, although most of the algorithms that we provide could be recompiled under a different mapping.

Although second-quantised encodings such as the Jordan–Wigner mapping are natural and easy to interpret, it is immediately clear that they are very inefficient in terms of space complexity for representing systems with a fixed particle number. Although the Hilbert space of $2M$ qubits has dimension 2^{2M}, only $\binom{2M}{N}$ of the qubit states contain the correct number of particles, N. Thus, a large portion of the qubit Hilbert space remains unused or is undesired as it would break physical symmetries. This effect is even more pronounced if we consider the spin projection $\langle \hat{S}_z \rangle$ or total spin $\langle \hat{S}^2 \rangle$ expectation values. These are also good quantum numbers for electronic wavefunctions and are associated with an even smaller symmetry subspace of the Hilbert space. A resulting problem is that, during the course of a

quantum algorithm, states with incorrect quantum numbers can be created due to hardware noise or due to the application of non-symmetry-adapted operators. The latter can be avoided by a careful choice of operators, as we discuss in Chapter 5; symmetry-breaking due to noise is a serious problem for NISQ algorithms,[231] although it might be mitigated through post-selection [231] or error mitigation techniques.[115]

In first-quantised encodings, the qubit Hilbert space is utilised more exhaustively. The wavefunction must be antisymmetrised explicitly on the quantum computer; this can be achieved by starting from a product state and applying a quantum circuit controlled by an ancilla register.[68, 69] First-quantised mappings can encode spin-orbitals (traditional Gaussian-type orbitals from quantum chemistry basis sets, or plane waves) or simpler grid points in real space. They require one quantum register per electron, where each electron is represented by $2M$ basis functions/grid points using only $\lceil \log_2(2M) \rceil$ qubits, resulting in $N\lceil \log_2(2M) \rceil$ qubits overall.[68, 207, 242–244] Although this is asymptotically much more efficient than second-quantised encodings, the number of qubits required to represent small systems is larger, and the quantum circuits required for implementations of fermionic operators can be more efficient or more expensive than for second-quantised representations, depending on the choice of the basis. Thus, while the first- and second-quantised representations are formally equivalent, encodings that use the two representations on quantum computers with a limited number of qubits (basis functions) offer alternative simulation methods with different space-time tradeoffs.

First-quantised encodings have been shown to have very good scaling in a basis of plane waves due to the reduced complexity of the Hamiltonian operator: the kinetic energy term is diagonal in the plane wave basis, whereas the Coloumb term is diagonal in real space, and one can efficiently map between the two bases using the quantum Fourier transform.[178, 180, 244] This results in quantum algorithms for Hamiltonian simulation scaling sublinearly in the basis set size.[178, 180] However, for molecular systems, the number of plane waves required to achieve

a certain discretisation error is much higher than the number of Gaussian-type orbitals. The intuition behind this is that a lot of information about the molecule is encoded in Gaussian-type atomic orbitals (see Section 2.1.2), whereas plane waves in momentum space or delta-like functions in real-space are more generic objects.[5] Therefore, the state-of-the-art quantum algorithms for computing molecular electronic eigenstates based on quantum phase estimation employ Gaussian orbitals in a second quantised encoding with relatively small active spaces.[65, 88]

2.3.2 Quantum Phase Estimation

Given a unitary operator $\hat{U}$ and an eigenstate $|\Psi\rangle$ such that $\hat{U}|\Psi\rangle = e^{2\pi i\theta}|\Psi\rangle$. the quantum phase estimation algorithm enables computing the corresponding eigenvalue $e^{2\pi i\theta}$ by reading out the phase $\theta \in \mathbb{R}$.[67] In the context of quantum simulation, given an initial state formed by a superposition of the energy eigenstates $|\psi\rangle = \sum_i c_i |E_i\rangle$ of some many-body Hamiltonian $\hat{H} = \sum_i E_i |E_i\rangle\langle E_i|$, we can use quantum phase estimation to prepare an eigenstate $|E_i\rangle$ and read out its energy E_i by performing measurements in the Hamiltonian eigenbasis. Although many variants of the algorithm have been developed to optimise its implementation,[32, 86, 87, 248, 249] here we describe the canonical implementation.[48]

Our setup requires a main quantum register, initialised in some state $|\psi\rangle$ with expansion in the Hamiltonian eigenbasis $|\psi\rangle = \sum_i c_i |E_i\rangle$, and an n-qubit ancilla register initialised in the zero state, $|0\rangle^{\otimes n}$. We assume that we can implement some operator $\hat{U}$ that is isospectral to the Hamiltonian $\hat{H}$. We implement the circuit in Figure 2.1, which can be broken down as follows:

1. We apply a Hadamard gate to each ancilla qubit, producing the state

$$|\psi_1\rangle = \frac{1}{2^{n/2}}\Big(|0\rangle + |1\rangle\Big)^{\otimes n} \otimes \sum_i c_i |E_i\rangle\,. \tag{2.38}$$

[5] In contrast, plane waves, such as Bloch wavefunctions,[245, 246] Wannier functions,[247] or Bloch orbitals,[109] are good objects for representing electrons in lattices such as solid-state materials or model Hamiltonians with periodic boundaries.

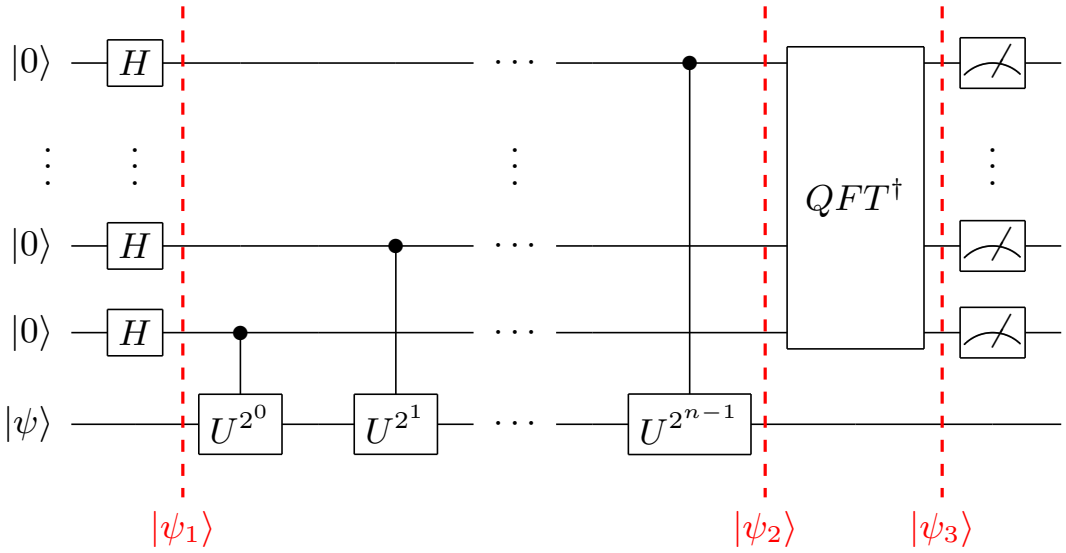


Figure 2.1: Quantum circuit for quantum phase estimation. The symbol H corresponds to the Hadamard gate, which is different from the Hamiltonian $\hat{H}$. In the setting where $\hat{U} = e^{-i\hat{H}t}$ and $|\psi\rangle = \sum_i c_i |E_i\rangle$, where the Hamiltonian has spectral decomposition $\hat{H} = \sum_i E_i |E_i\rangle \langle E_i|$, the output after measurement in the computational basis yields a binary encoding of a phase corresponding to the eigenvalue E_i of $|E_i\rangle$ with probability propertional to $|c_i|^2$.

The ancilla register is now in a superposition of all n-qubit computational basis states, which can also be written as

$$|\psi_1\rangle = \frac{1}{2^{n/2}} \sum_{j=0}^{2^n-1} |j\rangle \otimes \sum_i c_i |E_i\rangle , \tag{2.39}$$

where $|j\rangle = \bigotimes_{l=0}^{n-1} |j_l\rangle$. We use the ancilla register to encode the phase θ in a binary representation with n bits. Thus, the single-qubit states $|j_l\rangle$ correspond to the bits of the binary representation of j, $j = \sum_{l=0}^{n-1} j_l 2^l$, for each of the 2^n possible bitstrings.

2. We now apply a controlled version $C\hat{U}$ of the unitary $\hat{U}$, where the control qubits are the ancilla register and the target qubits are the main register. The controlled unitary operator can be written as $C\hat{U} = \sum_{j=0}^{2^n-1} |k\rangle \langle k| \otimes \hat{U}^k$. Here, $\hat{U}^k = \prod_{i=1}^{k} \hat{U}$, i.e., $\hat{U}^k$ can be applied by applying $\hat{U}$ k times, and the controlled unitary can be decomposed as $C\hat{U} = \prod_{l=0}^{n-1} C_l(\hat{U}^{2l})$, where the notation C_l indicates that qubit l is the control qubit. If the unitary

is the time-evolution operator $\hat{U} = e^{-i\hat{H}t}$, then the state after this series of controlled time-evolutions reads

$$|\psi_2\rangle = \frac{1}{2^{n/2}} \sum_i c_i \sum_{k=0}^{2^n-1} e^{-2\pi i k E_i} |k\rangle \otimes |E_i\rangle . \tag{2.40}$$

This technique is also known as phase kickback as it "transfers" the phase of the eigenstates in the main register to the control qubits.[250] The eigenvalues E_i are now encoded in the phase $\theta = -2\pi k E_i$ of a complex exponential, as bitstrings corresponding to the computational basis of the ancilla register.

3. To read out the phase, we apply the inverse quantum Fourier transform. The quantum Fourier transform has the following action on a computational basis state:[48] $|j\rangle \to \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} e^{2\pi i j k/2^n} |k\rangle$. By applying its inverse on the ancilla register, we are left with a binary representation `bin` of the energy eigenvalues:

$$|\psi_3\rangle = \sum_i c_i |\texttt{bin}(E_i)\rangle |E_i\rangle . \tag{2.41}$$

4. Finally, measuring the ancilla qubits in the computational basis collapses the wavefunction of the main register to one of the eigenstates $|E_i\rangle$, with probability $|c_i|^2$, and yields a bitstring corresponding to the binary representation of the corresponding eigenvalue E_i. Thus, we have prepared the Hamiltonian eigenstate $|E_i\rangle$ on the quantum computer, and we can classically reconstruct its energy eigenvalue E_i.

Note that we have assumed that the phase $|\theta| = |\frac{E_i t}{2\pi}|$ is represented exactly with n bits. In practice, if this is not the case, this procedure yields an approximation to θ with error bounded by $1/2^n$. Thus, this error can be suppressed exponentially with the number of ancilla qubits.

Since the quantum Fourier transform and its inverse have an efficient implementation that requires $\mathcal{O}(n \log n)$ gates, it is rarely a bottleneck in practical calculations. The main cost of implementing quantum phase estimation comes from the need to apply time-evolution under the Hamiltonian on the main register for different times $t = \{2\pi, 4\pi, ..., 2^n\pi\}$. Moreover, since the probability of measuring

the desired eigenstate is $|c_i|^2$, on average, we must repeat the procedure $1/|c_i|^2$ times. Thus, the two main factors determining the cost of quantum phase estimation are the cost of Hamiltonian simulation and initial state preparation. Below, we give a brief overview of the different approaches used for the these subroutines.

Initial state preparation

The goal is to prepare an initial state $|\psi\rangle = \sum_i c_i |E_i\rangle$ with significant overlap $|\langle\psi|E_i\rangle|^2 = |c_i|^2$ with the target eigenstate $|E_i\rangle$. In the worst case, the cost of preparing such a state for a k-local Hamiltonian is exponential in the size of the system,[39] including the electronic structure Hamiltonian.[40]

In practice, heuristic quantum or classical algorithms can be used to find good approximate initial states. One possible route it to use the adiabatic state preparation method.[73, 74, 76, 251] Here, we must find a simplified Hamiltonian, $\hat{H}_0$, and initialise the quantum computer in its ground state. In the context of electronic structure, a natural choice is the Fock operator and its corresponding ground state, the Hartree–Fock state (Section 2.2.1), for which the state preparation circuit is trivial both in second- and first-quantised encodings.[238] The Hamiltonian is gradually morphed from the initial Hamiltonian $\hat{H}_0$ to the full electronic structure Hamiltonian $\hat{H}_{\text{el}}$. By time-evolving the state under this varying Hamiltonian, this prepares the ground state of the desired final Hamiltonian $\hat{H}_{\text{el}}$, provided that the evolution is sufficiently slow relative to the spectral gap $E_1 - E_0$, to avoid jumping from the ground state to an excited state.

The Hartree–Fock state is not a good initial state in systems with strong electron correlation.[51, 186] In such systems, the spectral gap is smaller and the overlap of a mean-field state with the exact ground state is small. In Chapter 4 (Ref. 51), we provide an alternative initial state and Hamiltonian that directly encode strong correlation effects, making it well-suited for strongly correlated electrons in molecules.

Alternatively, one could use other quantum heuristics such as the variational quantum eigensolver,[91–93] quantum imaginary time evolution,[95] or quantum

subspace diagonalisation methods.[252] As we argue in Chapter 4, this is inefficient if these quantum algorithms are themselves initialised in a mean-field Hartree–Fock state.

Finally, one could run a classical algorithm to obtain a classical description of an approximate entangled state, and then prepare this on a quantum device, either in the form of a Slater determinant expansion or a matrix product state.[181, 253] This approach is likely practical for systems with a small number of strongly correlated electrons, but in general, both the classical algorithm and the quantum state preparation circuit scale exponentially with the number of strongly correlated electrons.

In Chapters 3 and 4, we address the initial state overlap problem for systems with strong electron correlation to reduce the cost of quantum phase estimation as well as heuristic quantum algorithms.

Hamiltonian simulation

Hamiltonian simulation is a key subroutine in many quantum algorithms, three of which are discussed in this thesis: quantum phase estimation, adiabatic state preparation, and the quantum Krylov method. It can be viewed as a compilation problem: Given a qubit operator $\hat{H}$, devise a quantum circuit $\hat{U}$ (usually compiled as a series of one- and two-qubit gates) such that $\|\hat{U} - e^{-i\hat{H}t}\| \leq \epsilon$ for a small error ϵ.

The simplest method is known as Trotterization. Given a time-independent Hamiltonian expressed as a linear combination of k-local operators, $\hat{H} = \sum_j \hat{h}_j$, where $\hat{h}_j$ act non-trivially on at most k qubits, there exist efficient and exact quantum circuits for exponentiating each term $e^{-i\hat{h}_j t}$ individually.[177] Thus, a first approximation is to simply implement the product of each of the exponentials, $\prod_j e^{-i\hat{h}_j t}$. However, since the Hamiltonian terms in general do not commute, $[\hat{h}_i, \hat{h}_j] \neq 0$, this introduces a large error. The first-order Trotter-Suzuki method aims to minimise this error by applying the evolution for many short times,

$$\hat{U} = e^{-i\hat{H}t} = \left(\prod_j e^{-i\hat{h}_j t/r} \right)^r := S_1(t, r), \tag{2.42}$$

where $r \in \mathbb{N}_{>0}$ is the number of Trotter steps. The resulting error can be bounded approximately by $\|\hat{U} - e^{-i\hat{H}t}\| = \mathcal{O}(t^2/r)$. Higher-order Trotter formulas increase

the accuracy of the approximation. A more precise characterisation of the error can be achieved by computing the norms of the commutators between all the different Hamiltonian terms $\{\hat{h}_j\}$, referred to as the commutator norm.[254] Specifically, for a p-th order Trotter decomposition with r Trotter steps, and Hamiltonian with commutator norm α, the error after time t is bounded by

$$\|e^{-i\hat{H}t} - S_p(t,r)\| = \mathcal{O}\Big(\frac{(\alpha t)^{p+1}}{r^p}\Big). \tag{2.43}$$

Although the accuracy can be increased by increasing p and r, the number of gates required to implement a p-th order Trotter-Suzuki formula is $\mathcal{O}(5^{p/2}r)$, where $r \sim \left(\frac{t^{p+1}}{\epsilon}\right)^{1/p}$ steps are required to achieve accuracy ϵ. Thus, the optimal formula depends on the application according to this accuracy vs. cost trade-off.

More modern approaches for Hamiltonian simulation include Taylor-series methods, which approximate the time-evolution operator by a truncated series expansion,[81, 82, 174] quantum-walk-based techniques, which encode Hamiltonian dynamics into a discrete walk on an enlarged Hilbert space,[170, 171] linear-combination-of-unitaries (LCU) methods, which express the Hamiltonian or propagator as a weighted sum of efficiently implementable unitaries,[56] and sparsity-exploiting algorithms, which leverage oracle access to the non-zero matrix elements of sparse Hamiltonians to simulate their dynamics efficiently.[172, 173]

An alternative approach, known as qubitization and quantum signal processing,[57, 58, 83, 85] implements a quantum walk whose eigenphases encode the spectrum of the Hamiltonian via $e^{\pm i \arccos(\hat{H}/\lambda)}$, where λ is a normalising constant satisfying $\|\hat{H}\| \leq \lambda$. Qubitization has optimal scaling in all the relevant parameters,[57, 58] including linear time complexity and logarithmic scaling in the required precision, $\mathcal{O}(\lambda t + \log(1/\epsilon))$. Even after a constant-factor analysis, it yields the best performance for current resource estimates of fault-tolerant quantum phase estimation of electronic structure.[65, 88]

However, for other algorithms that require less precision, Trotter formulas are often more useful due to their simpler implementation on quantum hardware and the fact that the empirical errors are typically much lower than what the theoretical

bounds predict.[255] In this thesis, we use the first-order Trotter approximation as part of the quantum Krylov algorithm in Chapter 4.

Quantum phase estimation is a very appealing eigenvalue solver due to its rigorous success guarantee (assuming access to an initial state with significant overlap). However, it requires a very precise implementation of the Hamiltonian simulation operator, leading to deep circuits that almost certainly require a fault-tolerant quantum computer for simulating systems of interest. Below, we discuss a quantum algorithm that requires shallower quantum circuits and thus can be implemented more feasibly on NISQ computers.

2.3.3 Variational Quantum Eigensolver

The variational quantum eigensolver (VQE) was first proposed and experimentally implemented in 2014.[91] Since then, a vast amount of activity has been dedicated to study and improve its performance.[45, 256] The main design choice is to remove as much complexity as possible from the quantum circuit implementation and use the quantum computer only for the task for which it is believed that it has an advantage: the preparation of a highly entangled quantum state.[91–93]

Specifically, VQE involves the following loop:

1. Prepare some initial state on the quantum computer, $|\psi_0\rangle$.

2. Apply a parametrised unitary transformation on the state to prepare a parametrised ansatz
$$\hat{U}(\boldsymbol{\theta})\,|\psi_0\rangle = |\psi(\boldsymbol{\theta})\rangle\,. \tag{2.44}$$

3. Compute its energy and/or energy gradients as an expectation value, through measurements of the qubit states:
$$E(\boldsymbol{\theta}) = \langle\psi(\boldsymbol{\theta})|\,\hat{H}\,|\psi(\boldsymbol{\theta})\rangle \tag{2.45}$$

4. On a classical computer, optimise the variational parameters to minimise the energy, i.e., solve
$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} E(\boldsymbol{\theta}). \tag{2.46}$$

5. Repeat until convergence is reached.

The hope is that, by offloading a large amount of the computational cost to the classical optimisation routine, a shallow quantum circuit can prepare a quantum state $|\psi(\boldsymbol{\theta})\rangle$ that is hard to compute on a purely-classical device.

This algorithmic framework is extremely flexible; there are a vast range of choices for any component of it: the choice of the parametrised unitary $\hat{U}(\boldsymbol{\theta})$ for step 2, the measurement scheme (step 3), the classical optimiser (step 4), and the initial state (step 1). Unfortunately, each of these components causes significant challenges.

For example, the computation of the energy or gradient is typically achieved via a technique called Hamiltonian averaging:[92] the Hamiltonian, given as a sum of terms $\hat{H} = \sum_j \hat{h}_j$, is split into sets of commuting operators and all commuting operators are measured simultaneously.[257] This must be repeated a large number of times for each step to statistically converge to the correct expectation value, scaling as $\mathcal{O}(1/\epsilon^2)$ for accuracy ϵ. Each energy/gradient evaluation needs to be run for many different parameters, since the optimisation loop requires a large number of parameter updates. Since the number of terms in the electronic Hamiltonian scales like $\mathcal{O}(N^4)$, the number of measurements required per evaluation of the objective function (Eq. 2.45) can be very large. For example, a study by Gonthier *et al.* estimates that, even for relatively small molecules, the expected runtime for evaluating a single energy is on the order of days to months.[42] This is further aggravated by the effects of hardware noise, which can corrupt both the measurement results and the prepared quantum state.[44]

The classical optimisation is also an extremely challenging task: the energy landscape is known to be highly non-convex and plagued with local minima far from the global minimum, often requiring exponentially many measurements for successful optimisation.[258] Furthermore, depending on the ansatz, the energy gradient can vanish exponentially with the size of the system, a phenomenon known as barren plateaus.[256, 259]

Finally, even ignoring roadblocks due to measurement or optimisation, it is not known whether any of the ansätze is capable of parametrising quantum states beyond what is possible on a classical computer in the presence of device noise. These challenges are often connected: for example, circuits $\hat{U}(\boldsymbol{\theta})$ that are capable of parametrising a large amount of Hilbert space are often harder to optimise.[258, 259]

This has led to the design of a myriad of ansätze with different properties.[202, 260–263] Two categories are so-called *physically motivated* vs. *hardware-efficient* ansätze. The latter are designed to minimise the quantum circuit depth by applying many parametrised gates in parallel as well as choosing gates that can be run natively on existing quantum hardware such as superconducting qubit processors.[260] However, these can often break physical symmetries such as particle number or spin, and more generally tend to explore a very large subspace of the Hilbert space, complicating the optimisation.[256] In contrast, physically-motivated ansätze are often designed to preserve physical symmetries [202, 264] but can lead to deeper quantum circuits with less expressive power.[265]

In this thesis, we aim to improve the practicality of the VQE algorithm through two angles. In Chapter 4, we show that the performance of VQE is greatly enhanced when the initial state is a linear combination of spin-coupled initial states such as those derived in Chapter 3. In Chapter 5, we present an ansatz that is provably universal (it can in principle parametrise any state in the Hilbert subspace) and symmetry-adapted (preserves spin, spin projection, and electron number symmetry), and an optimisation algorithm that can be used to design shallow quantum circuits with few parameters and that is less prone to converge to local minima.

2.3.4 Fault-tolerant vs. NISQ implementation of quantum algorithms

In principle, provided a sufficient number of qubits and circuit depth, any of the algorithms discussed in this thesis can be executed with or without quantum error

correction. In practice, deeper circuits such as the ones required for accurate Hamiltonian simulation and phase estimation, are expected to only be viable on fault-tolerant quantum devices; in contrast, shallow circuits such as those required by VQE or certain quantum subspace diagonalisation algorithms might already be executable with sufficient accuracy on non-error-corrected physical qubits. Equally, algorithms originally designed for NISQ devices might still be useful in the fault-tolerant setting. Despite this generality, when analysing and predicting the performance of a quantum algorithm, it is important to consider the different limitations posed by these two paradigms of quantum computation. We have already mentioned some of these throughout the previous chapters, but we synthesise them here.

Due to the high error rates of physical qubits, algorithms designed to run on NISQ devices benefit greatly from reducing the number of any gates. However, the number of two-qubit gates is a particularly good proxy for the rate/magnitude of error generation since these are typically much noisier than single-qubit gates.[199] Furthermore, the limited coherence time of physical qubits means that reducing the circuit depth, and therefore the runtime, is essential. Although error mitigation techniques can significantly alleviate the impact of hardware noise,[116] any NISQ computation is greatly constrained by noise.

In contrast, fault-tolerant quantum processors operating below the noise threshold can in principle run arbitrarily long computations without corrupting the algorithm output due to noise (or corrupting it only to a small, finite and constant amount).[193, 196] In terms of efficiency, the fault-tolerant implementation of gates within the Clifford set can be achieved without a large overhead. This includes the single-qubit Hadamard, Pauli and Phase/S gates, as well as the two-qubit CNOT gate. Thus, a fault-tolerant quantum algorithm with many CNOT gates is not necessarily inefficient.

In contrast, implementing arbitrary rotations on logical qubits—such as single-qubit rotations $R_i(\theta)$, where $i \in \{x, y, z\}$—is not directly possible on a fault-tolerant architecture. Instead, continuous rotations must be approximated through a discrete set of rotations with fixed angles. This rotation gate synthesis requires

an overhead in terms of circuit depth and gate counts. As a consequence, this can slow down the fault-tolerant implementation of the single-qubit rotations typical of many variational quantum algorithms such as VQE.[46, 256]

More generally, implementing any error-correction techniques—including syndrome measurement, i.e., checking for errors, applying error-correcting gates, and implementing magic state distillation for non-Clifford gates—introduces a large overhead to any fault-tolerant algorithm, greatly increasing the overall runtime.[46, 62, 266] As such, even when operating below the noise threshold, deep quantum circuits can be impractical. The same could be said about classical algorithms, which also require finite runtime and implement classical versions of error correction. However, the overhead of quantum error correction is many orders of magnitude larger due to (1) the higher error rates of quantum vs. classical hardware and (2) the fundamentally higher complexity of processing quantum information which does not allow, e.g., copying the state of a qubit according to the No-cloning theorem.[267]

2.4 Spin coupling and configuration state functions

This thesis focuses on simulating molecular systems with strong electron correlation, a many-body phenomenon whose description requires an entangled wavefunction. In molecules, the strong electron correlation often occurs due to spin coupling. The mathematical representation of such systems is most natural in terms of configuration state functions (CSF). Below, we briefly review the role of spin in quantum chemistry (Section 2.4.1), the treatment of angular momentum in quantum mechanics (Section 2.4.2), methods for the construction of spin eigenfunctions (Sections 2.4.3 and 2.4.4), and the role of the basis in representations of CSFs (Section 2.4.5). For details, we refer the interested reader to standard textbooks on quantum mechanics,[268] quantum chemistry,[38, 269] and CSFs.[189, 270]

2.4.1 The role of spin in quantum chemistry

Spin is a purely quantum phenomenon with no classical analogue. While it is central to the study of systems exposed to a magnetic field, it might appear counter-intuitive that spin plays a central role in quantum chemistry. The electronic structure Hamiltonian (Eqs. 2.2 or 2.9) contains no spin term and commutes with both the total spin and spin projection operators,

$$[\hat{H}, \hat{S}^2] = [\hat{H}, \hat{S}_z] = 0, \tag{2.47}$$

where $\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$. Therefore, any Hamiltonian eigenstates must simultaneously be an eigenfunction of both spin operators with corresponding quantum numbers,

$$\hat{S}^2 \left|S, M_S\right\rangle = \hbar^2 S(S+1) \left|S, M_S\right\rangle, \tag{2.48a}$$

$$\hat{S}_z \left|S, M_S\right\rangle = \hbar M_S \left|S, M_S\right\rangle, \tag{2.48b}$$

where S determines the total spin and M_S determines the projection of the spin angular momentum vector onto the z-axis.

Even in the absence of any spin term in the Hamiltonian, spin has an impact on the electronic structure of molecules due to its interplay with another purely quantum-mechanical property: the antisymmetry of fermions. Since the overall wavefunction of a fermionic system must be antisymmetric with respect to exchange, the symmetry of the spin wavefunction imposes a constraint on the symmetry of the spatial wavefunction. For example, if the spin part of the wavefunction is maximally antisymmetric, the spatial part must be maximally symmetric to satisfy the overall antisymmetry requirement. Through this symmetry-induced spin-spatial coupling, the spin state of the electrons can influence the eigenvalues of spin-independent observables such as the energy, which only depends directly on the spatial wavefunction through the Coulomb potential.

The objects of choice for expanding many-electron wavefunctions are typically Slater determinants (Eq. 2.6), since they have a simple product form and satisfy fermionic antisymmetry by construction. However, while all Slater determinants

are eigenfunctions of the spin projection operator $\hat{S}_z$, they are not necessarily eigenfunctions of $\hat{S}^2$. In particular, closed-shell determinants are spin eigenfunctions with $\langle \hat{S}^2 \rangle = 0$, whereas open-shell determinants correspond to a linear combination of spin eigenfunctions of different $\langle \hat{S}^2 \rangle$ values.

Strong electron correlation in molecules often arises due to the spin coupling between electrons localised around different regions of the molecule, such as atomic fragments. Such electron configurations correspond to Slater determinants with singly occupied orbitals. To form a spin eigenfunction requires a properly spin-adapted linear combination of multiple Slater determinants, also known as a Configuration State Function (CSF). Before discussing CSFs, we briefly review the quantum-mechanical treatment of angular momentum.

2.4.2 Angular momentum in quantum mechanics

Spin is a form of angular momentum. In quantum mechanics, the angular momentum corresponds to the expectation value of the angular momentum operator, a three-component vector operator, usually expressed as

$$\hat{\boldsymbol{J}} = (\hat{J}_x, \hat{J}_y, \hat{J}_z). \tag{2.49}$$

The components obey the following commutation relations:

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z, \quad [\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x, \quad [\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y, \tag{2.50}$$

which can be derived from the canonical commutation relations $[\hat{x}_i, \hat{p}_j] = \delta_{i,j}$, where $\hat{x}$ and $\hat{p}$ are the position and momentum operators.

The squared angular momentum operator can be expressed as

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2, \tag{2.51}$$

and it commutes with each of the components of $\hat{J}$:

$$[\hat{J}^2, \hat{J}_x] = [\hat{J}^2, \hat{J}_y] = [\hat{J}^2, \hat{J}_z] = 0. \tag{2.52}$$

Thus, there exists a common eigenbasis for $\hat{J}^2$ and one of the $\hat{J}_i$, where $i \in \{1, 2, 3\}$, forming the basis $\{|J, M_J\rangle\}$. The usual convention is to chose the component $\hat{J}_z$.

Thus, for the spin angular momentum, the eigenbasis of $\hat{S}^2$ and $\hat{S}_z$ is formed by the spin eigenfunctions $\{|S, M_S\rangle\}$ (Eq. 2.48).

For a system of N identical particles, the maximum possible angular momentum J is given by the number of particles times the angular momentum of each particle. Given J, the angular momentum projection can attain the values $M_J = -J, -J+1, ..., J-1, J$. Thus, given J, there exist $2J+1$ possible values for M_J.

Electrons are spin-$\frac{1}{2}$ fermions. Thus, for a system of N electrons, the possible spin values are

$$S = 0, \frac{1}{2}, 1, \frac{3}{2}, ..., \frac{N}{2} \tag{2.53}$$

and

$$M_S = -S, -S+1, ..., S-1, S. \tag{2.54}$$

Given two systems with angular momenta J_A and J_B and orthonormal eigenbases $\{|J_A, M_{J_A}\rangle\}$, $\{|J_B, M_{J_B}\rangle\}$ spanning the Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$, we can form an orthonormal basis for the combined Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ through a tensor product

$$\{|J_A, M_{J_A}\rangle \otimes |J_B, M_{J_B}\rangle\}. \tag{2.55}$$

In general, the states formed by such products (Eq. 2.55) are not eigenstates of the total angular momentum operator, defined as $\hat{\boldsymbol{J}} = \hat{\boldsymbol{J}}_A + \hat{\boldsymbol{J}}_B$ with $\hat{\boldsymbol{J}}^2 = \hat{\boldsymbol{J}}_A^2 + \hat{\boldsymbol{J}}_B^2 + \hat{\boldsymbol{J}}_A \hat{\boldsymbol{J}}_B$.

To construct a simultaneous eigenbasis for $\{\hat{\boldsymbol{J}}^2, \hat{J}_z, \hat{\boldsymbol{J}}_1^2, \hat{\boldsymbol{J}}_2^2\}$, where $\hat{J}_z = \hat{J}_{Az} + \hat{J}_{Bz}$, we must couple the individual subsystems through a linear combination of the tensor product states. Each basis state has the form

$$|J, M, J_A, J_B\rangle = \sum_{M_{J_A} + M_{J_B} = M} C_{J_A, M_A, J_B, M_B} \, |J_A, M_{J_A}\rangle \otimes |J_B, M_{J_B}\rangle \,, \tag{2.56}$$

where the expansion coefficients are known as the Clebsch–Gordan coefficients:

$$C_{J_A, M_A, J_A, J_B} = \langle J, M, J_A, J_B | J_A, M_{J_A}, J_B, M_{J_B}\rangle \tag{2.57}$$

The Clebsch–Gordan coefficients for the special unitary group SU(n) can be computed using recursive algorithms such as the Gelfand–Tsetlin calculus,[271] and analytical formulas are known for SU(2). In practice, relevant coefficients of SU(2) and SU(3) have been computed and are available in tables.[272]

We have so far discussed the general properties of angular momentum and the construction of angular momentum eigenfunctions for increasingly large systems. Since electrons are spin-$\frac{1}{2}$ fermions, this thesis focuses on the construction of spin eigenfunctions for systems of multiple spin-$\frac{1}{2}$ particles.

2.4.3 Genealogical construction of spin eigenfunctions

A straightforward approach to constructing eigenfunctions of the spin angular momentum operator is through genealogical coupling, which begins with the spin state of a single electron and couples one electron at a time. We specify the spin eigenfunction of a single electron $|S, M_S\rangle = |\frac{1}{2}, \pm\frac{1}{2}\rangle$ solely based on the spin projection quantum number, $|\sigma_i\rangle$, where $\sigma_i \in \{\alpha, \beta\}$. Starting from the state of a single electron, $|\sigma_1\rangle = |\alpha\rangle$, we can couple the second electron to form two possible states: the singlet state

$$|0,0\rangle = \frac{1}{\sqrt{2}}(|\alpha\beta\rangle - |\beta\alpha\rangle) \tag{2.58}$$

or the triplet state

$$\begin{aligned} |1,1\rangle &= |\alpha\alpha\rangle\,, \\ |1,0\rangle &= \frac{1}{\sqrt{2}}(|\alpha\beta\rangle + |\beta\alpha\rangle), \\ |1,-1\rangle &= |\beta\beta\rangle\,, \end{aligned} \tag{2.59}$$

where the tensor product is implied. Starting from the two-electron singlet ($S = 0$), coupling the third electron necessarily creates a state with spin $S = \frac{1}{2}$. If the two-electron state is the triplet ($S = 1$, Eq. 2.59), the third electron can be coupled such that it forms either a quartet ($S = \frac{3}{2}$) or doublet ($S = \frac{1}{2}$). Given a spin eigenfunction for $N-1$ electrons, $|N-1, S, M_S\rangle$, two possible N-electron states

can formed by coupling one electron and adding its spin angular momentum,

$$\begin{aligned} |N, S+\frac{1}{2}, M\rangle = C^{S+\frac{1}{2},M}_{S,M-\frac{1}{2},\frac{1}{2},\frac{1}{2}} |N-1, S, M-\frac{1}{2}\rangle |\alpha\rangle \\ + C^{S+\frac{1}{2},M}_{S,M+\frac{1}{2},\frac{1}{2},-\frac{1}{2}} |N-1, S, M-\frac{1}{2}\rangle |\beta\rangle \end{aligned} \tag{2.60}$$

or by subtracting it:

$$\begin{aligned} |N, S-\frac{1}{2}, M\rangle = C^{S-\frac{1}{2},M}_{S,M-\frac{1}{2},\frac{1}{2},\frac{1}{2}} |N-1, S, M-\frac{1}{2}\rangle |\alpha\rangle \\ + C^{S-\frac{1}{2},M}_{S,M+\frac{1}{2},\frac{1}{2},-\frac{1}{2}} |N-1, S, M-\frac{1}{2}\rangle |\beta\rangle \,. \end{aligned} \tag{2.61}$$

The specific Clebsch–Gordan coefficients in Eqs. 2.60 and 2.61 can be obtained analytically.[189]

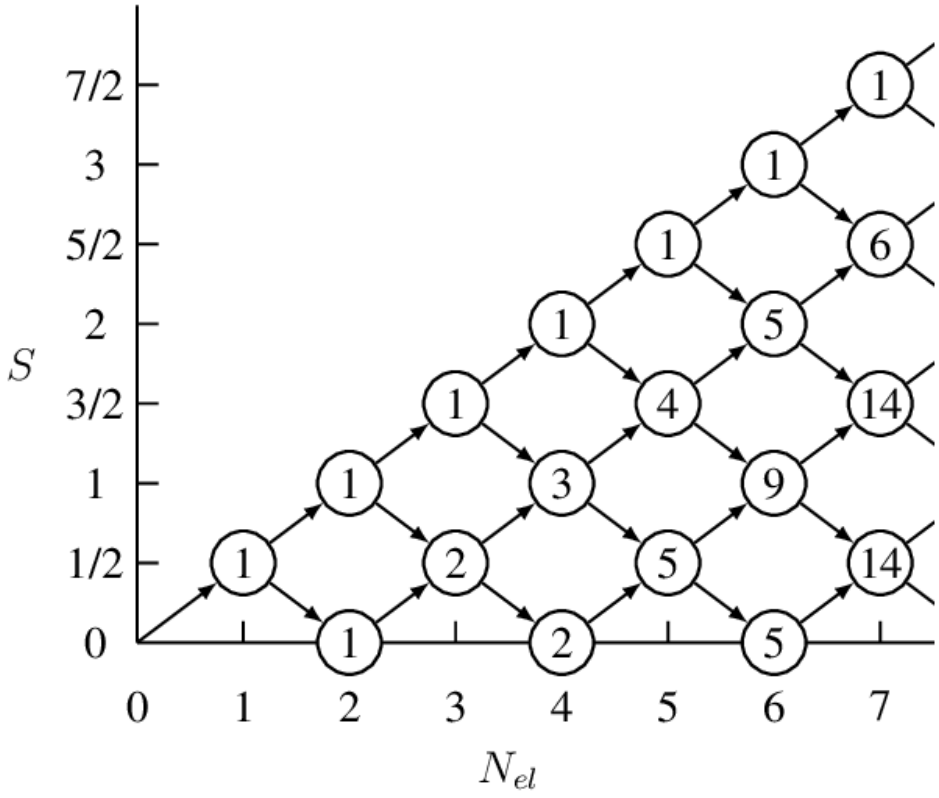


Figure 2.2: Genealogical branching diagram:[189] The number in each circle is the number of states with N_{el} and S, which corresponds to the number of paths $f(N_{el}, S)$ (Eq. 2.62).

In practice, rather than computing the states for one electron at a time and considering all projections at every step, one can simply specify the path required to reach a particular spin state $|S, M, N\rangle$ and apply Eqs. 2.60 and 2.61 iteratively to

compute the Clebsch–Gordan coefficients of the N-electron state in the uncoupled (tensor product) basis. The different paths can be illustrated by a genealogical branching diagram (Fig. 2.2).

The genealogical spin eigenfunctions constructed in this way are orthonormal.[189] Thus, the number of possible spin eigenfunctions for a system of N spin-$\frac{1}{2}$ particles is equal to the number of unique genealogical paths, given by

$$f(N_{el}, S) = \binom{N_{el}}{\frac{N_{el}}{2} - S} - \binom{N_{el}}{\frac{N_{el}}{2} - S - 1}. \tag{2.62}$$

Including all possible spin projections, the number of genealogical spin functions is $f(N_{el}, S)(2S + 1)$. By summing over all spins S, it can be shown that the total number of spin functions is

$$\sum_S f(N_{el})(2S + 1) = 2^{N_{el}} \tag{2.63}$$

where S ranges from $S = \frac{1}{2}$ for odd N_{el} and $S = 0$ for even N_{el} to $S = \frac{N_{el}}{2}$.[189] From their orthogonality and Eqn. 2.63, it is clear that the genealogical spin functions form a complete basis of the Hilbert space of N_{el} spins.

2.4.4 Configuration state functions

When expressed in a basis of Slater determinants, as common in quantum chemistry, spin eigenfunctions are known as configuration state functions (CSFs). The name comes from the fact that they correspond to a specific configuration of electrons in orbitals. Unlike for artificial spin-$\frac{1}{2}$ systems, which are formed from simple tensor products, the tensor products of one-electron states must be antisymmetrised. Furthermore, CSFs differ from the simple spin eigenfunctions in 2.4.3 in that the number of orbitals might be larger than the number of electrons.

Given N_{el} electrons and M spatial orbitals, the number of Slater determinants is $\binom{2M}{N_{el}}$, and each orbital can be occupied by at most two electrons, i.e., the possible occupations include $\{0, \alpha, \beta, 2\}$. In contrast, in the spin-$\frac{1}{2}$ case, there are only N *sites* for N spins, and each site is occupied by a single spin-$\frac{1}{2}$ particle. Thus, the spin eigenfunctions discussed in 2.4.3 can be used to describe the states of open-shell

(singly occupied) orbitals in a CSF. States with only closed-shell (doubly occupied) and virtual (unoccupied) orbitals correspond to (antisymmetrised) products of single-electron states, i.e, a single Slater determinant, with spin $S = 0$. In contrast, Slater determinants with open-shell occupation are not spin eigenfunctions. Thus, CSFs must be formed by products of closed-shell determinants and open-shell spin eigenfunctions, and they are defined by the nature of the orbitals and the couplings between the open-shell orbitals (e.g., as represented by a genealogical branching diagram). We formally define these states in Chapter 3. The total number of CSFs of spin S for system of N_{el} electrons in M orbitals is given by the Weyl-Paldus formula [189]

$$g(N_{el}, M, S) = \frac{2S+1}{M+1}\binom{M+1}{\frac{N_{el}}{2}-S}\binom{M+1}{M-\frac{N_{el}}{2}-S}. \tag{2.64}$$

From this perspective, one can view strong electron correlation in molecules as being necessary to simply achieve good spin quantum numbers for systems with localised electrons. In Chapter 3, we show that this picture can be used to greatly simplify the parametrisation of molecular eigenstates for strongly correlated systems.

2.4.5 Role of the basis

The states constructed through the genealogical coupling approach in Section 2.4.3 form the so-called Kotani–Yamanouchi basis.[273] Since a linear combination of spin eigenfunctions of the same spin S and particle number N are also spin eigenfunctions with the same quantum numbers, it is clear that there is a lot of freedom for constructing alternative bases.

Popular examples include the Serber and Rumer bases. In the Serber basis, electrons are coupled in pairs, following an iterative approach similar to that for the Kotani–Yamanouchi basis, with a corresponding branching diagram.[274] In the method by Rumer, spin eigenfunctions are constructed by first building pairs of singlets and additively coupling α electrons if $S > 0$.[275] This leads to a highly over-complete basis, where the number of Rumer functions is much larger than

the size of the basis (Eq. 2.62), as many of them are not linearly independent. Orthogonalising the Rumer functions using the Gram–Schmidt procedure yields the Kotani–Yamanouchi functions. The Serber functions can be recovered by applying the Gram–Schmidt orthogonalisation method while ensuring that (anti)-symmetry is enforced in pairs.[189]

These bases have different advantages for different systems. In particular, the choice of basis greatly affects the sparsity of the many-electron wavefunction. For example, the Serber basis is naturally suited for describing systems where the correlation between pairs of electrons is large, whereas the Kotani–Yamanouchi basis yields highly compact wavefunction expansions for stretched bonds.[50]

There also exist different methods for efficiently computing Hamiltonian matrix elements $H_{i,j} = \langle i| \hat{H} |j\rangle$, where $\{|i\rangle\}$ are the spin eigenfunctions forming the basis. For example, efficient computation of matrix elements can be achieved by a method based on representations of the unitary group which only tracks the genealogical path rather than explicitly building the expansion of the CSFs in the Slater determinant basis.[276, 277] In the graphical unitary group approach (GUGA), the CSFs are represented through a graph and the path corresponds to a walk on the graph.[278, 279] These methods for generating CSFs and computing the relevant matrix elements between them have enabled efficient spin-adapted implementations of classical algorithms for electronic structure, including FCI-QMC [190] and DMRG.[16]

In this thesis, rather than constructing a full (orthonormal) basis of CSFs for the Hamiltonian, we focus on finding wavefunctions formed from a linear combination of a small number of CSFs that approximate molecular eigenstates, where each CSF corresponds to a clear physical motif. The CSFs are constructed directly from the Clebsch–Gordan coefficients. Each such CSF is most efficiently represented in a bespoke orbital basis. Thus, we might include states that each correspond to a single spin eigenfunction in a particular CSF and orbital basis but that, when expressed within the same fixed basis, might correspond to a large number of CSFs. This *custom* approach has two benefits. First, each CSF maps clearly to a physical

picture. For example, localised electrons are best described by localised orbitals, and CSFs that couple different molecular subsystems are most naturally described by first coupling the electrons within each subsystem, and then coupling across the subsystems. Second, when working in a fixed basis, the molecular eigenstates might require a large number of CSFs for their description. In contrast, a bespoke basis chosen for each CSF according to physical principles can enable a compact description of the eigenstates in terms of a few (nonorthogonal) CSFs.

In the quantum algorithms discussed in Chapter 4, we use bespoke bases to discover and construct the CSFs but then rotate the CSFs to a common orbital basis, since the cost of basis rotations on a quantum processor is very small.[51, 238] The downside of this approach is that it does not allow directly applying clever techniques for evaluation of matrix elements such as the GUGA method.[278, 279] However, this is irrelevant in the context of the quantum algorithms discussed in Chapter 4, since we construct the Hamiltonian and states in a Slater determinant basis, for which implementation of Hamiltonian terms is efficient on a quantum computer. This inevitably leads to exponentially large wavefunction expansions even for a single CSF with open-shell orbitals, but this poses no problem for quantum processors due to the exponential growth of the Hilbert space with the number of qubits.

3

Spin-coupled molecular orbital theory

Contents

3.1 Introduction 76
3.2 Molecular hydrogen 77
3.3 Configuration State Functions 80
3.4 Homolytic bond cleavage 83
3.4.1 Breaking a triple bond: N_2 83
3.4.2 First row diatomics 87
3.4.3 Breaking two single bonds: H_2O 89
3.5 Hydrogen chains: H_4 91
3.6 Quantum computation using spin-coupled states 92
3.7 Conclusions and outlook 96

Abstract

Molecular orbital theory is powerful both as a conceptual tool for understanding chemical bonding, and as a theoretical framework for ab initio quantum chemistry. Despite its undoubted success, MO theory has well documented shortcomings, most notably that it fails to correctly describe diradical states and homolytic bond fission. In this contribution, we introduce a generalised MO theory that includes spin-coupled radical states. We show through archetypical examples that when bonds break, the electronic state transitions between a small number of valence configurations, characterised by occupation of both delocalised molecular orbitals and spin-coupled localised orbitals. Using a linear combination of such configurations, we recover a large amount of electron correlation from symmetry considerations and chemical intuition. Our theory provides a model for chemical bonding that is both intuitive and qualitatively accurate when combined with ab initio theory. Although exploitation of our theory presents significant challenges for classical computing, the symmetric structure of spin-coupled states is ideally suited to algorithms that exploit quantum computers. Our approach provides a systematic route to overcoming the initial state overlap problem in quantum algorithms for strongly correlated electrons in molecules and unlocking the potential of quantum computational chemistry.

3.1 Introduction

Solving Schrödinger's equation to characterise the electronic structure of molecules and compute the Born–Oppenheimer potential energy surface remains a central goal of quantum chemistry. Modern electronic structure theory has arisen from nearly a century of sustained research and development, and the vast majority of modern methods are based on Molecular Orbital (MO) theory,[280–283] either in the form of Kohn–Sham Density Functional Theory[284] or post-Hartree–Fock correlation methods.[285] These approaches came to prominence in no small part due to the natural mapping of the MO representation of the many-electron wavefunction to linear algebra and matrix operations that can be performed with high efficiency on classical digital computers.[286, 287]

Despite the success of MO theory, it has well-documented shortcomings.[218, 288–290] Closed-shell ground states are typically well approximated by a single configuration of electrons in spin-orbitals, but open-shell states, or cases with many competing low-energy configurations, are not. In practice, the accurate description of fundamental chemical processes, such as bond breaking using MO theory, requires a very large number of excited configurations,[291, 292] and sophisticated optimisation techniques.[110, 111, 293, 294] As a consequence, any connection between simple chemical models and data from high-level ab initio calculations is obfuscated.

Looking to the future, it is increasingly probable that quantum computers will become available for scientific computing.[1, 64, 295–297] Many of the constraints on the way that many-electron wavefunctions are represented and optimised in modern MO-based theories do not apply to quantum algorithms and we were therefore motivated to examine the electronic structure of bond formation and breaking afresh.

In this article, we introduce a generalised MO theory that includes spin-coupled radical states. Our theory provides a chemically intuitive picture of bonding that directly maps to a highly compact representation of the many-body wavefunction. Contrary to the current consensus among the electronic structure community, we

find that it is not necessary to use sophisticated multi-reference expansions with very large numbers of variational parameters to describe bond breaking. Instead, we show that the electronic state undergoes transitions between a small number of valence states, each with an easily understood spin-coupled electronic configuration.

Furthermore, we show that our representation encodes strong correlation through prescribed patterns of entanglement, which is ideally suited to quantum computing. We demonstrate that our approach provides a powerful route to construct the sufficiently accurate reference states that are required for the practical application of fault-tolerant algorithms, such as quantum phase estimation[67, 72, 76] and related approaches.[104, 105, 298–300]

3.2 Molecular hydrogen

The prototypical chemical bond between two hydrogen atoms exemplifies the language and conceptual framework within which electronic structure theories operate. In MO theory, the left L and right R 1s atomic orbitals combine to form bonding σ_g and anti-bonding σ_u MOs that transform as the irreducible representations of the $D_{\infty h}$ molecular point group.[282, 283] The wavefunction is written as the closed-shell singlet, doubly occupying the bonding MO, which is the Slater determinant used in Hartree–Fock theory $|\Phi_{\mathrm{RHF}}\rangle = |\sigma_g\bar{\sigma}_g\rangle$. The MO description is accurate around the equilibrium bond length, but deteriorates as the molecule dissociates. This well-known failure arises from the inclusion of ionic terms in the wavefunction, which results from delocalised nature of the MOs[287]

$$|\Phi_{\mathrm{RHF}}\rangle = \frac{1}{2}\Big(\underbrace{|s_L\bar{s}_L\rangle + |s_R\bar{s}_R\rangle}_{\text{ionic}} + \underbrace{|s_L\bar{s}_R\rangle - |\bar{s}_L s_R\rangle}_{\text{covalent}}\Big). \tag{3.1}$$

The strong correlation that localises the electrons on opposite atoms is absent in the MO picture. In MO-based theories, this strong correlation is introduced through mixing with excited configurations, and for H_2 the correct behaviour for the binding curve is obtained using two configurations

$$|\Psi\rangle = c_1\,|\sigma_g\bar{\sigma}_g\rangle + c_2\,|\sigma_u\bar{\sigma}_u\rangle\,. \tag{3.2}$$

In the general case, the number of excited configurations required to recover the correct electronic structure grows exponentially with the number of localised electrons. Modern MO-based methods for strong correlation therefore typically employ sophisticated wavefunction representations to optimise large numbers of variational parameters.

An alternative representation that predates MO theory is the valence bond (VB) approach of Heitler and London.[301] The VB wavefunction is the open-shell singlet spin-coupled state of 1s electrons on each H atom

$$|\Phi_{\mathrm{VB}}\rangle = \mathcal{N}(|\mathrm{s}_L\bar{\mathrm{s}}_R\rangle - |\bar{\mathrm{s}}_L\mathrm{s}_R\rangle). \tag{3.3}$$

where the normalisation constant $\mathcal{N} = (2 + 2S_{LR}^2)^{-1/2}$ depends on the orbital overlap $S_{LR} = \langle \mathrm{s}_L|\mathrm{s}_R\rangle$. The VB description also captures the bonding interaction, which arises through electron delocalisation resulting from the overlap of the L and R 1s orbitals. Approaches based on VB wavefunctions, however, are poorly suited to both classical and quantum computing due to reliance on non-orthogonal orbitals, which introduces cumbersome overlap terms in the working equations.[302–304]

Notwithstanding the differences in the MO and VB representations, the physical process of bond breaking is clear: the electronic state transitions from a singlet spin-coupled bonding σ_{g}^2 configuration in the equilibrium region to a singlet spin-coupled diradical $\mathrm{s}_L^1\mathrm{s}_R^1$ configuration at dissociation. On a classical digital computer, the basis transformations required to combine configurations of different orbitals are costly and often impractical. On a digital quantum computer, basis transformations pose no difficulty. Orbital rotations are encoded through products of exponentials of single excitation and de-excitation operators,[238] and the corresponding quantum circuits have linear depth in orbital number and minimal qubit connectivity.[295]

We therefore choose to generalise MO theory by expressing the wavefunction as a superposition of configurations where electrons occupy both delocalised MOs and localised AOs. This representation directly encodes both the covalent interactions responsible for bonding and the strong correlation that localises electrons

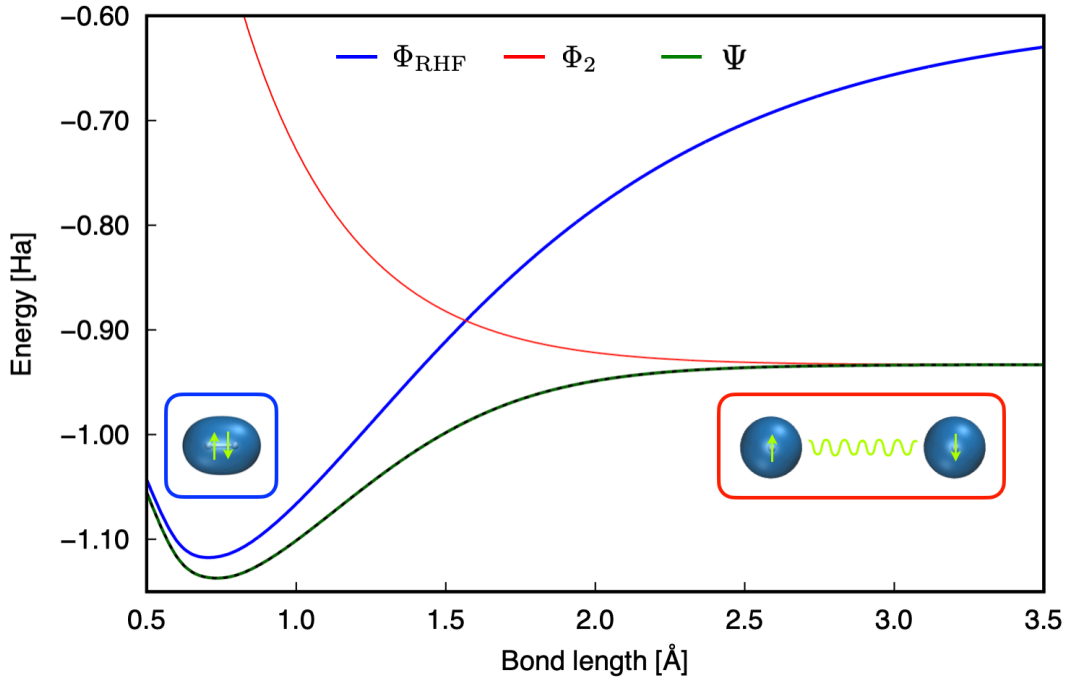


Figure 3.1: Dissociation of H_2 in the STO-3G basis using generalised spin-coupled orbitals.

at different sites, giving a physically accurate trial wavefunction across the whole binding curve without requiring large numbers of variational parameters.

For H_2, the singlet coupled $s_L s_R$ diradical state is

$$|\Phi_2\rangle = \frac{1}{\sqrt{2}}(|s_L\bar{s}_R\rangle - |\bar{s}_L s_R\rangle) = \frac{1}{\sqrt{2}}(|\sigma_g\bar{\sigma}_g\rangle - |\sigma_u\bar{\sigma}_u\rangle). \tag{3.4}$$

where now $s_L = \frac{1}{\sqrt{2}}(\sigma_g + \sigma_u)$ and $s_R = \frac{1}{\sqrt{2}}(\sigma_g - \sigma_u)$ are orthogonal orbitals localised on the left and right atoms. In contrast to the Heitler–London wavefunction, this is a purely repulsive state.[305] The correct binding curve (Figure 3.1) is obtained through the linear combination of the σ_g^2 and $s_L s_R$ states

$$|\Psi\rangle = c_1\,|\Phi_{\mathrm{RHF}}\rangle + c_2\,|\Phi_2\rangle\,. \tag{3.5}$$

The $|\Phi_{\mathrm{RHF}}\rangle$ state is uncorrelated, since a single Slater determinant is an antisymmetrised orbital product corresponding to an independent electron wavefunction. $|\Phi_2\rangle$ is an entangled, or strongly correlated state, since it is a linear combination of determinants that cannot be converted to a single determinant through orbital

rotation. Both states have low complexity since their structures are determined by symmetry considerations. The weights of the entanglement structure in $|\Phi_2\rangle$ are specified through the rules of spin angular momentum coupling to obtain a singlet spin molecular ground state.

In this representation, the bonding interaction arises through increasing character of the delocalised σ_g^2 configuration, which is driven by the favourable kinetic energy of the delocalised state and the attractive electron-nucleus Coulomb interaction. Conversely, the dominant electron correlation process for the H_2 bond is the localisation of the electrons on opposite atoms, introducing partial diradical character, and this correlation increases in strength as the bond length increases. This transition from delocalised to localised electronic states is equivalent to the metal-insulator transition in the Fermi–Hubbard model, where the H_2 bond length plays a similar role to the ratio U/t between the on-site repulsion U and the hopping term t.

In the following, we demonstrate how these simple concepts can be generalised to more complex cases where multiple bonds are broken. We show that the strongly correlated state at dissociation is highly structured and that the structure is determined solely by symmetry and spin-angular momentum coupling considerations. Furthermore, we show that the transition from the simple covalent equilibrium bonding structure to the strongly correlated dissociated state is represented by a small number of intermediate configurations, each with their own spin-coupled structure. We first summarise the configuration state functions that we use to represent spin-coupled open-shell configurations.

3.3 Configuration State Functions

The exact molecular wavefunction is an eigenstate of both the total spin angular momentum operator $\hat{S}^2$ and the spin projection operator $\hat{S}_z$ with quantum numbers S, M. A single Slater determinant, the many-electron basis function in MO-based theories, is an eigenfunction of $\hat{S}_z$, but not of $\hat{S}^2$, except for closed-shell and high-spin open-shell configurations. The physically relevant functions for describing

open-shell states with localised orbitals are configuration state functions (CSF), which are eigenstates of $\hat{S}^2$ with quantum numbers S, M.[306]

There exist several methods for constructing CSFs, as covered in detail by Pauncz.[189, 270]. Examples include the methods by Kotani-Yamanouchi[273], Serber [274] and Rumer,[275] which each generate a different but complete basis of CSFs. More recently, the (graphical) unitary group approach by Paldus [276, 277] and Shavitt [278, 279] has been adopted within the FCI quantum Monte Carlo (FCI-QMC) algorithm.[190]. These methods apply group-theoretic techniques to build CSFs and compute Hamiltonian matrix elements. Their aim is to construct a full basis of spin eigenfunctions for the Hamiltonian, such that pure spin quantum numbers are guaranteed in any approximate wavefunction ansatz, and specific spin sectors can be targetted.

In this work, we are not interested in constructing a full CSF basis. Our aim is to find a compact approximation to the many-electron wavefunction derived from a physically-intuitive picture, and to efficiently exploit the CSFs within electronic structure algorithms on quantum computers. We are interested in identifying a small number of bespoke CSFs, expressed in the Slater determinants basis. These can be constructed with any of the above techniques.

A CSF can be defined by a specification of the open-shell orbitals and of the spin-coupling pattern, which is determined by the order and manner in which the spin-$\frac{1}{2}$ particles in the open-shell orbitals are coupled. Formally, our CSFs take the form

$$|\Phi\rangle = \mathcal{N}\hat{\mathcal{A}}\Big(|\phi_c\bar{\phi}_c\ldots\rangle\,\mathcal{O}^{N,i}_{SM}(\phi_o\bar{\phi}_o\ldots)\Big). \tag{3.6}$$

where $\hat{\mathcal{A}}$ is the antisymmetric permutation operator, c denotes the doubly occupied closed-shell orbitals, o are the open-shell orbitals, and $\mathcal{O}^{N,i}_{SM}(\phi_o\bar{\phi}_o\ldots)$ is an open-shell state with N electrons in orbitals $\{\phi_o, \bar{\phi}_o\}$ coupled with spin quantum numbers S and M in a spin-coupling pattern labelled by i, and $\mathcal{N}$ is a normalisation constant. All occupied spin-orbitals are mutually orthogonal and normalised.

The open-shell state $\mathcal{O}^{N,i}_{SM}$ for coupled angular momenta can be expressed in terms of the uncoupled Slater determinant representation of N electrons in open-shell orbitals $\{\phi_o, \bar{\phi}_o\}$ through the Clebsch–Gordan coefficients for the spin-coupling

pattern. To streamline presentation, we use a simplified notation for $|\Phi\rangle$ in Eq 3.6, listing only the occupancy of the spatial orbitals and suppressing $\hat{\mathcal{A}}$ and $\mathcal{N}$:

$$|\Phi\rangle = |\phi_c^2 \ldots\rangle \, \mathcal{O}_{SM}^{N,i}(\phi_o \ldots). \tag{3.7}$$

A list of CSFs used in this work is provided in Appendix A.1.

Electronic structure methods designed for classical computers that use a CSF basis rather than a Slater determinant basis, such as spin-adapted FCI quantum Monte Carlo,[190] are constrained to use one set of orbitals and one coupling scheme with an orthonormal basis of CSFs. This restriction allows Hamiltonian matrix elements to be evaluated using efficient group theoretical approaches without explicitly building the Slater determinant expansion of each CSF, which has exponential cost.[276, 278] Apart from ensuring symmetry adaption, this is beneficial in situations where the electronic wavefunction is dominated by a single CSF. Given the correct choice of basis, the Hamiltonian and wavefunction can become very sparse and therefore significant compression can be achieved using a CSF basis.[307–309]

Quantum algorithms are not constrained in this way, as quantum computers can efficiently process exponentially large quantum states,[54] e.g. corresponding to very large CI expansions. We are free to use orbitals and coupling patterns that directly encode the physical interactions. We therefore advocate for allowing full freedom in the choice of the CSF and orbital basis, and selecting a set of physically-motivated CSFs, rather than choosing a subset of CSFs from a fixed CSF basis with a common set of orbitals.

This is advantageous over classical algorithms in situations where no single basis yields a very sparse representation of the electronic eigenstate, such as the intermediate region in the binding curve. Here, the wavefunction is truly multi-configurational [224] and requires superpositions of multiple CSFs of very different physical character. On quantum computers, we can combine localised, AO-like configurations for coupled high-spin atomic fragments with MO configurations for delocalised covalent bonds, as well as configurations intermediate between the these

extremes. The symmetric structure of the CSFs allows efficiently loading them onto quantum devices (Section 3.6).

The physical interactions follow established rules connected to the relative strengths of exchange and delocalisation energies, and the CSFs required to represent the electronic state compactly are straightforwardly constructed. For example, the electron configurations of dissociated atomic fragments follow Hund's rules, where electron spins couple to form a local high-spin state, and the net electron spin angular momenta of the atoms couple to form low-spin states. Bonding interactions in the molecule are formed when paired electrons occupy delocalised molecular orbitals. We now illustrate this framework using examples that pose significant challenges for traditional approaches.

3.4 Homolytic bond cleavage

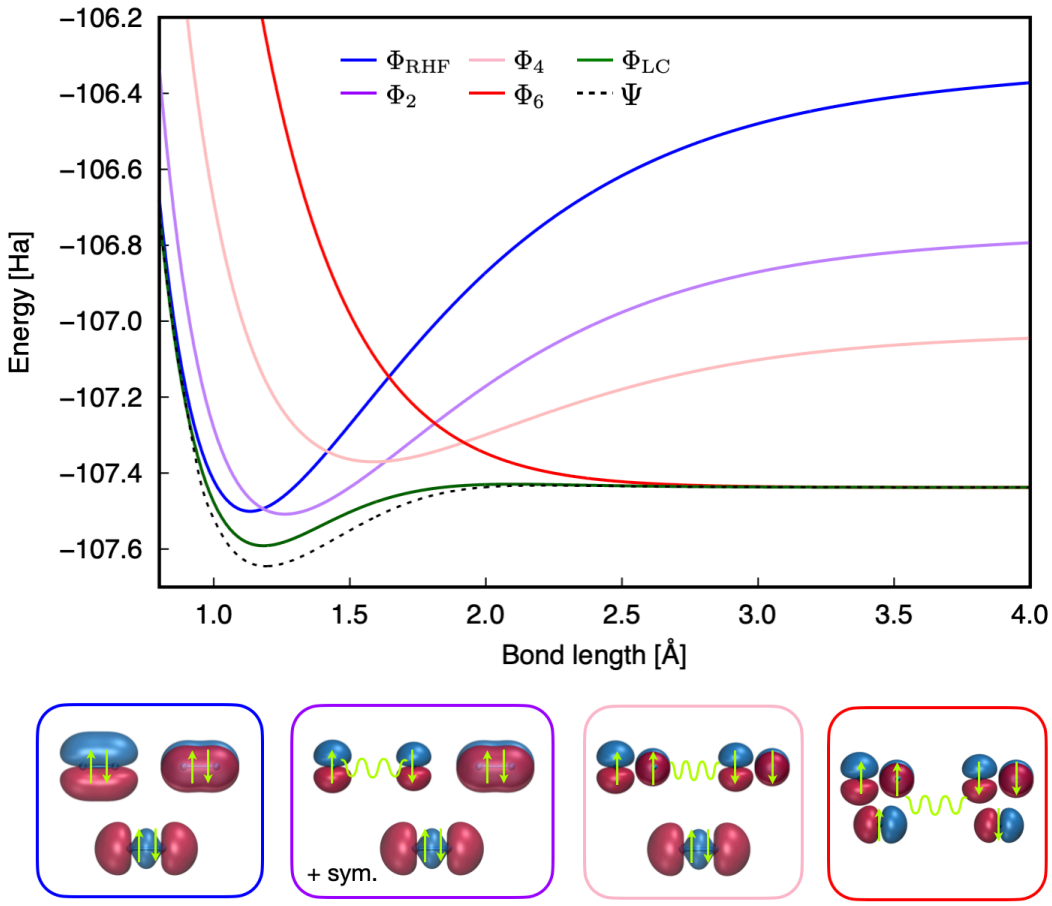
3.4.1 Breaking a triple bond: N_2

Dissociation of the nitrogen molecule is a well-documented challenge for electronic structure theory[163, 310–319] because the six electrons in the triple bond become very strongly correlated when the bond elongates. Even when only considering the six 2p valence orbitals[320] there are 56 Slater determinants with the correct symmetry that contribute to the wavefunction. However, using generalised MO theory, the wavefunction is recovered to better than 92% across the whole binding curve with only four CSFs (Figure 3.2).

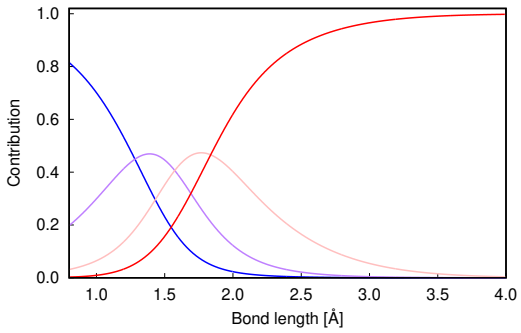
MO theory predicts a $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 3\sigma_g^2 1\pi_u^4$ configuration for the N_2 molecule, with a $^1\Sigma_g^+$ ground state. The corresponding CSF is simply the closed-shell Hartree–Fock wavefunction

$$|\Phi_{\mathrm{RHF}}\rangle = |\mathcal{C}\rangle\, |3\sigma_g^2\, 1\pi_{u,x}^2\, 1\pi_{u,y}^2\rangle \tag{3.8}$$

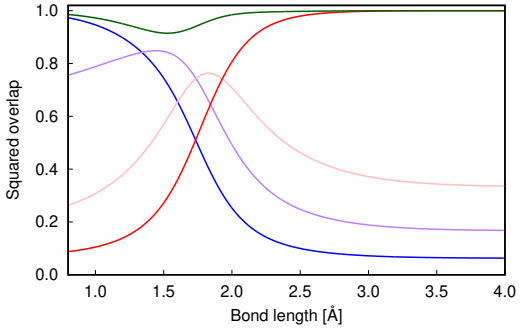
where $|\mathcal{C}\rangle = |1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2\rangle$ is the contribution from the core orbitals that are not involved in bonding. At dissociation, each atom has a $1s^2 2s^2 2p^3$ electronic



(a) Energies and configurations of the valence CSFs.



(b) Coefficients of each CSF in Φ_{LC}.



(c) Squared overlaps $|\langle\Psi|\Phi_{\mathrm{i}}\rangle|^2$.

Figure 3.2: The dissociation of N_2 in a STO-3G basis in the six 2p valence space using generalised MO theory.

configuration with spins locally coupled to a ^4P on each atom, which then couple through space to give a $^1\Sigma_g^+$ molecular state with the CSF structure

$$|\Phi_6\rangle = |\mathcal{C}\rangle\, \mathcal{O}_{00}^{6,1}(\mathrm{x}_L\mathrm{y}_L\mathrm{z}_L\mathrm{x}_R\mathrm{y}_R\mathrm{z}_R) \tag{3.9}$$

where $\mathrm{z}_L = \frac{1}{\sqrt{2}}(3\sigma_g + 3\sigma_u)$, $\mathrm{z}_R = \frac{1}{\sqrt{2}}(3\sigma_g - 3\sigma_u)$, $\mathrm{x}_L = \frac{1}{\sqrt{2}}(1\pi_{u,x} + 1\pi_{g,x})$, $\mathrm{x}_R = \frac{1}{\sqrt{2}}(1\pi_{u,x} - 1\pi_{g,x})$ etc are localised atomic orbitals. The spin function $\mathcal{O}_{00}^{6,1}$ spin-couples the three-electron quartets on each atom into a singlet, and is a specific linear combination of 20 open-shell determinants [Eq. (B.1c)]. The closed-shell part $|\mathcal{C}\rangle$ is invariant to rotations among its constituent orbitals making it equal to $|1\mathrm{s}_L^2 1\mathrm{s}_R^2 2\mathrm{s}_L^2 2\mathrm{s}_R^2\rangle$. As the bond elongates, one π bond breaks first , then both π bonds, and finally the σ bond. The CSFs that describe these intermediate states with one and two broken bonds are

$$|\Phi_4\rangle = |\mathcal{C}\rangle\, |\sigma_\mathrm{g}^2\rangle\, \mathcal{O}_{00}^{4,1}(\mathrm{x}_L\mathrm{y}_L\mathrm{x}_R\mathrm{y}_R) \tag{3.10}$$

$$|\Phi_2\rangle = |\Phi_{2x}\rangle + |\Phi_{2y}\rangle \tag{3.11}$$

$$|\Phi_{2x}\rangle = |\mathcal{C}\rangle\, |\sigma_\mathrm{g}^2\pi_{\mathrm{u,y}}^2\rangle\, \mathcal{O}_{00}^{2,1}(\mathrm{x}_L\mathrm{x}_R) \tag{3.12}$$

$$|\Phi_{2y}\rangle = |\mathcal{C}\rangle\, |\sigma_\mathrm{g}^2\pi_{\mathrm{u,x}}^2\rangle\, \mathcal{O}_{00}^{2,1}(\mathrm{y}_L\mathrm{y}_R) \tag{3.13}$$

where $|\Phi_{2x}\rangle$ and $|\Phi_{2y}\rangle$ are a degenerate pair. The state $\mathcal{O}_{00}^{4,1}$ describes atomic triplet states coupled to form a singlet overall and $\mathcal{O}_{00}^{2,1}$ describes atomic doublet states coupled to a singlet. The representations of $\mathcal{O}_{00}^{6,1}$, $\mathcal{O}_{00}^{4,1}$ and $\mathcal{O}_{00}^{2,1}$ in terms of Slater determinants are collected in Appendix A.1. In this formulation, the localised orbitals $\mathrm{x}_L, \mathrm{x}_R, \mathrm{y}_L, \mathrm{y}_R, \mathrm{z}_L, \mathrm{z}_R$ are mutually orthogonal and are therefore not true atomic orbitals except at infinite separation.

Just as for H_2, the RHF state dissociates to the incorrect limit due to the presence of unphysical ionic terms in the wavefunction, and the exact dissociated state is the CSF where the high-spin atomic states couple to form a singlet (Figure 3.2a). The CSFs with bond orders of two and one behave similarly to the RHF state but with higher energy in the bonding region, lower ionic destabilisation at dissociation, and with minima at progressively longer bond lengths, entirely consistent with their bonding character. The closed-shell orbitals used in these CSFs were obtained from

the RHF wavefunction, but very similar curves are obtained if the orbitals of each CSF are relaxed, for example through RASSCF optimisation (restricted active space self consistent field).[321]

Figure 3.2a also displays the energy of the exact ground state in the (6e,6o) active space of the bonding electrons and the energy obtained from the variationally optimised linear combination of the four valence CSFs. The optimised coefficients and the overlap of each state with the exact ground state are displayed in figures 3.2b and 3.2c.

While the optimised linear combination of these four CSFs is dominated by the RHF configuration at equilibrium, electrons in a bond correlate by localising on opposite atoms, occupying low-energy spin-coupled states, leading to non-negligible contributions from the CSFs with lower bond orders and an equilibrium bond length longer than that of the RHF state. Just as for H_2, the strength of these correlation processes increases as the bond elongates. In this formulation, only four CSFs are required to obtain an accurate binding curve with an overlap with the exact state of at least 92% at all bond lengths. The largest deviations are found in the bonding region where there is significant additional dynamic correlation among the six electrons in the triple bond.

The state at dissociation is strongly correlated in the sense that the electrons are localised on opposite atoms, and is a linear combination of 20 Slater determinants. However, this state is described exactly by a single CSF. The complexity is low; the coefficients in the determinantal representation are determined entirely by symmetry and spin angular momentum considerations.

Our generalised spin-coupled MO theory provides a highly intuitive description of chemical bonding that maps directly onto clearly identifiable contributions to the many-electron wavefunction. Since each contribution is composed of different orbitals (localised and covalent) and different spin-coupling patterns, the states are non-orthogonal and optimisation of the linear combination requires non-orthogonal configuration interaction (CI) theory.[322–326] In the context of fault-tolerant quantum algorithms, such as quantum phase estimation, the increased overlap with the

true ground state obtained by using $|\Phi_{\mathrm{LC}}\rangle$ as the reference wavefunction instead of $|\Phi_{\mathrm{RHF}}\rangle$ translates into more than an order of magnitude reduction in runtime at stretched geometries (see Section 3.6).

3.4.2 First row diatomics

Our spin-coupled generalised MO formalism is straightforwardly generalised to the bonding motifs of all of the first-row diatomic molecules, irrespective of whether the electronic ground state is singlet or triplet spin-coupled. Starting from the HF state given by standard MO theory, CSFs with incrementally lower bond order are generated by localising the electrons in the weakest remaining bond into the corresponding non-bonding atomic orbitals. The spin coupling pattern follows Hund's rules, where electrons first spin-couple with other unpaired electrons on the same atom to give high-spin atomic states, which then spin-couple to give the appropriate overall spin state for the molecule. The $D_{\infty h}$ molecular point group symmetry is satisfied by taking the symmetry adapted linear combinations generated by the symmetry operations.

In Figure 3.3 we plot the potential energy curves for the valence CSFs and the accuracy of the linear combination with respect to full CI in the (Ne, 6o) active space for B_2, C_2, O_2 and F_2. The CSFs are also listed. In all cases accurate binding curves and overlap with the FCI state of at least 88% are obtained, with the largest deviation for the bonding region of O_2 where significant additional dynamic correlation contributes. The triplet state B_2 transitions from the π^2 single-bond configuration to a triplet diradical state at dissociation. For C_2, configurations with either σ or π occupation are both important since the large sp mixing, arising from the small s-p orbital energy gap, destabilises the σ configurations. In the dissociation limit, the ground state tends towards the spin-coupled atomic states, where the electrons are evenly distributed among the p_x,p_y,p_z orbitals. This is also the reason that F_2 requires both the σ-hole and π-hole configurations at dissociation. The triplet state O_2 follows exactly the same physical rules, but the three spin-coupled states contain multiple symmetry-related components.

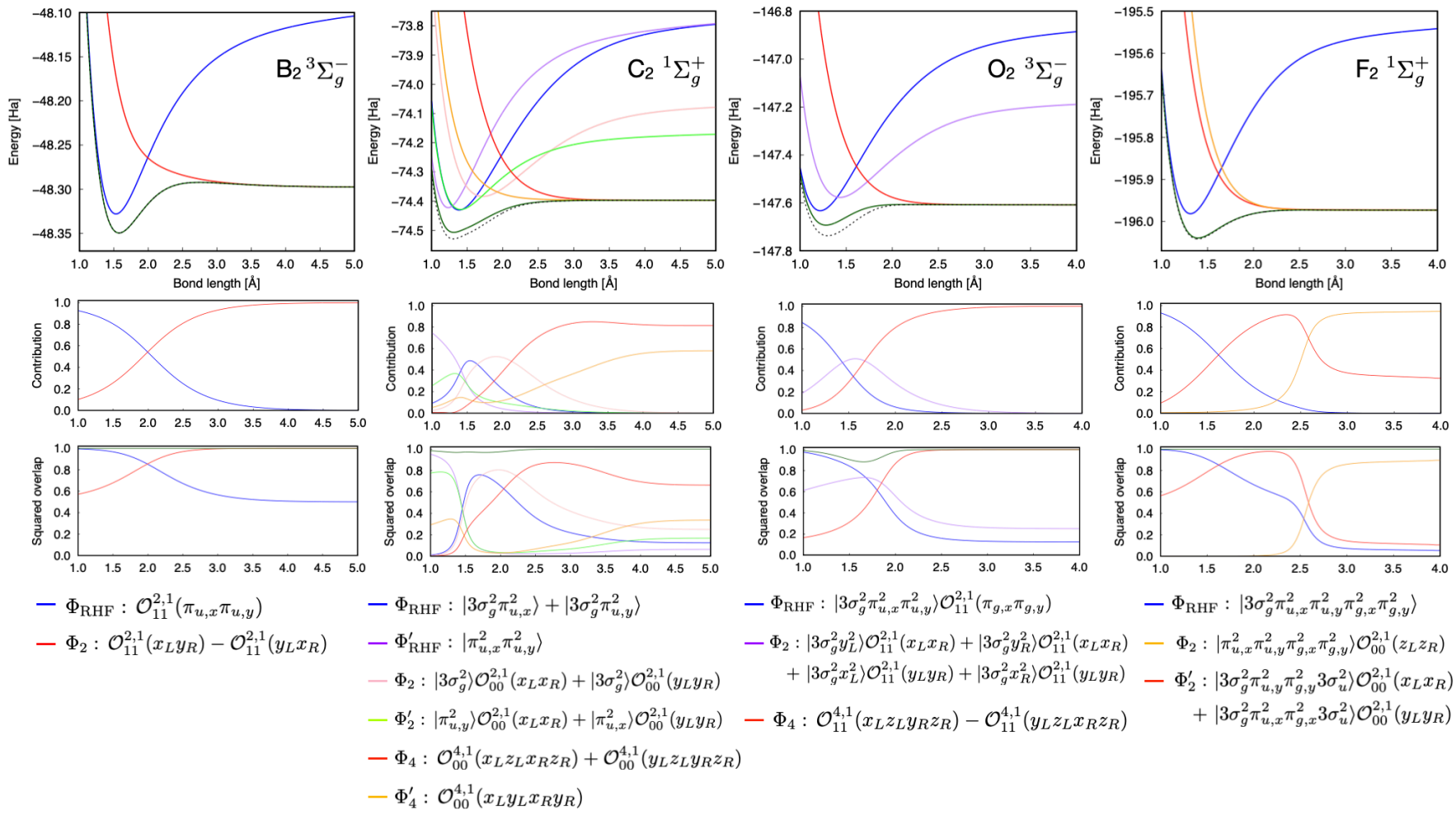


Figure 3.3: CSF energies, configurations and weights in the binding of first row diatomics.

3.4.3 Breaking two single bonds: H_2O

The symmetric dissociation of H_2O is another challenging case for electronic structure theory, involving strong correlation and spin-coupling among the electrons as the bonds break.[327, 328] In contrast to N_2, the dissociation of H_2O involves the localisation of electrons onto more than two sites.

The MO description of water leads to the configuration $1a_1^2 2a_1^2 1b_2^2 3a_1^2 1b_1^2$, with term symbol $^1A_1$, where the x-axis points out of the plane of the molecule. The closed-shell orbitals form the state $|\mathcal{C}\rangle = |1a_1^2 2a_1^2 1b_1^2\rangle$, which contains the 1s, 2s and $2p_x$ orbitals on the oxygen atom, and the RHF wavefunction is

$$|\Phi_{\mathrm{RHF}}\rangle = |\mathcal{C}\rangle\,|1b_2^2 3a_1^2\rangle = |\mathcal{C}\rangle\,|\sigma_L^2\sigma_R^2\rangle \tag{3.14}$$

where σ_L and σ_R are MOs localised on the left and right OH bonds. The CSF for the dissociated state is that where the oxygen electrons from the two cleaved OH bonds spin-couple to form a triplet oxygen state, due to the favourable exchange stabilisation, meaning that the two hydrogen electrons must also spin-couple to form a triplet and these triplets spin-couple to form the singlet molecular state:

$$|\Phi_4\rangle = |\mathcal{C}\rangle\,\mathcal{O}_{00}^{4,1}(o_L o_R h_L h_R) \tag{3.15}$$

where, in a minimal basis, the localised orbitals are obtained through simple rotation of the valence orbitals

$$o_L = (\sigma_L + \sigma_L^*)/\sqrt{2}, \qquad o_R = (\sigma_R + \sigma_R^*)/\sqrt{2}, \tag{3.16}$$

$$h_L = (\sigma_L - \sigma_L^*)/\sqrt{2}, \qquad h_R = (\sigma_R - \sigma_R^*)/\sqrt{2}, \tag{3.17}$$

$$\sigma_L = (1b_2 + 3a_1)/\sqrt{2}, \qquad \sigma_L^* = (2b_2 + 4a_1)/\sqrt{2}, \tag{3.18}$$

$$\sigma_R = (1b_2 - 3a_1)/\sqrt{2}, \qquad \sigma_R^* = (2b_2 - 4a_1)/\sqrt{2}. \tag{3.19}$$

The CSF that describes the intermediate configuration between the fully bonded and fully dissociated states includes one broken OH bond. The molecular point group symmetry demands that this state must be a symmetric superposition of the left and right bonded configurations, which results in

$$|\Phi_2\rangle = |\mathcal{C}\rangle\,|\sigma_L^2\rangle\,\mathcal{O}_{00}^{2,1}(o_R h_R) + |\mathcal{C}\rangle\,|\sigma_R^2\rangle\,\mathcal{O}_{00}^{2,1}(o_L h_L) \tag{3.20}$$

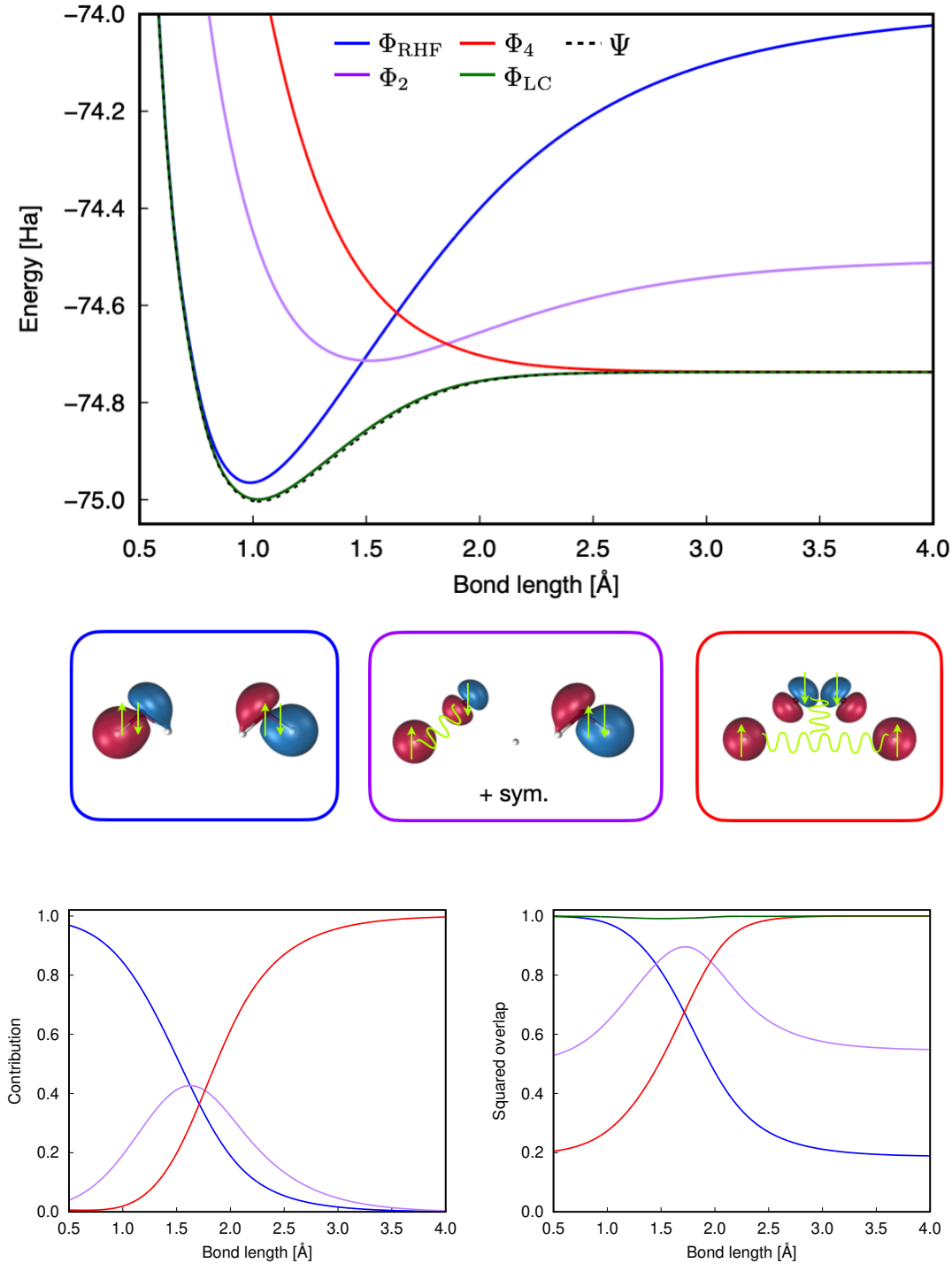


Figure 3.4: Dissociation of H_2O in a STO-3G basis using generalised spin-coupled orbitals.

The relative contributions of each CSF to a variationally optimised linear combination follow the intuitive trend, transitioning from the RHF state at equilibrium to $|\Phi_4\rangle$ at dissociation (Figure 3.4). Remarkably, the electron correlation is almost entirely captured by the three CSFs at all bond lengths, and the squared overlap with the exact wavefunction is at least 99% everywhere. At intermediate bond lengths all three CSFs have approximately equal contributions, which coincides with the region where traditional algorithms encounter the most difficulties.

3.5 Hydrogen chains: H_4

The change in the electronic structure of hydrogen chains as the separation increases is another challenging case for electronic structure theory. This simple system can be used to understand conjugation in π systems and is closely related to the Hubbard model for metal-to-insulator transitions in materials.[329]

We start with the cyclic H_4 structure, which shares electronic structure characteristics with the π system of cyclobutadiene in a square geometry, and has a degenerate LUMO. In a minimal basis, the MOs are uniquely determined by symmetry and are the set $\mathrm{a}_{1g}, \mathrm{e}_{u,x}, \mathrm{e}_{u,y}, \mathrm{b}_{2g}$. We consider the open-shell singlet $^1\mathrm{A}_{1\mathrm{g}}$, which cannot be represented by a Slater determinant and is the CSF

$$|\Phi_0\rangle = |\mathrm{a}_{1g}^2 \mathrm{e}_{u,x}\rangle + |\mathrm{a}_{1g}^2 \mathrm{e}_{u,y}\rangle \tag{3.21}$$

where the x and y axes pass through the sides of the square. Following our reasoning that the next most important configurations are those where spins localise, we introduce bonding and antibonding orbitals localised on each of the four edges, left, right, top and bottom, through transformation of the valence orbitals

$$\begin{aligned}
\sigma_L &= (\mathrm{a}_{1g} + \mathrm{e}_{u,x})/\sqrt{2}, \qquad & \sigma_L^* &= (\mathrm{b}_{2g} + \mathrm{e}_{u,x})/\sqrt{2}, \\
\sigma_R &= (\mathrm{a}_{1g} - \mathrm{e}_{u,x})/\sqrt{2}, \qquad & \sigma_R^* &= (\mathrm{b}_{2g} - \mathrm{e}_{u,x})/\sqrt{2}, \\
\sigma_T &= (\mathrm{a}_{1g} + \mathrm{e}_{u,y})/\sqrt{2}, \qquad & \sigma_T^* &= (\mathrm{b}_{2g} + \mathrm{e}_{u,y})/\sqrt{2}, \\
\sigma_B &= (\mathrm{a}_{1g} - \mathrm{e}_{u,y})/\sqrt{2}, \qquad & \sigma_B^* &= (\mathrm{b}_{2g} - \mathrm{e}_{u,y})/\sqrt{2}.
\end{aligned} \tag{3.22}$$

The atomic 1s orbitals are recovered by taking linear combinations of bonding and antibonding pairs. The CSF that describes the valence state with partial localised bonding is

$$|\Phi_2\rangle = |\sigma_L^2\rangle\, \mathcal{O}_{00}^{2,1}(\mathrm{s}_3, \mathrm{s}_4) + |\sigma_R^2\rangle\, \mathcal{O}_{00}^{2,1}(\mathrm{s}_1, \mathrm{s}_2) + |\sigma_T^2\rangle\, \mathcal{O}_{00}^{2,1}(\mathrm{s}_1, \mathrm{s}_4) + |\sigma_B^2\rangle\, \mathcal{O}_{00}^{2,1}(\mathrm{s}_2, \mathrm{s}_3) \tag{3.23}$$

and the CSF for the fully open-shell state is

$$|\Phi_4\rangle = \mathcal{O}_{00}^{4,2}(\mathrm{s}_1, \mathrm{s}_2, \mathrm{s}_3, \mathrm{s}_4) + \mathcal{O}_{00}^{4,2}(\mathrm{s}_2, \mathrm{s}_3, \mathrm{s}_4, \mathrm{s}_1) \tag{3.24}$$

where $\mathcal{O}_{00}^{4,2}$ is the antisymmetrised product of two open-shell singlets (Eq B.2). In Figure 5.2 we present the binding curves for the symmetric dissociation of H_4. These three valence states are sufficient to reproduce the exact binding curve and represent all of the correlation processes involved.

We now turn to the linear H_4, which shares electronic characteristics with the π system of butadiene. The valence states for linear H_4 are the same as those of the cyclic structure, but where left-right and top-bottom symmetry was present in the square geometry it is now lost, becoming the left-right and centre-edge. The left-right and centre contributions to $|\Phi_2\rangle$ must be included separately, and the edge contributions become unimportant since they are non-bonding. At the dissociation limit, all 2^4 spin states with one electron on each atom are degenerate, but the degeneracy is lifted by the kinetic energy of delocalisation at finite separations. The state $|\Phi_4\rangle$ is the linear combination of $S = 0$ CSFs with maximal nearest neighbour delocalisation energy, and equivalently, minimal nearest neighbour spin coupling $\langle\Phi|\hat{S}_i \cdot \hat{S}_j|\Phi\rangle$ through the isomorphism with the Heisenberg–Dirac spin Hamiltonian.[119, 330, 331]

3.6 Quantum computation using spin-coupled states

We have shown how to recover a significant amount of electron correlation from chemical intuition and symmetry considerations by building a linear combination

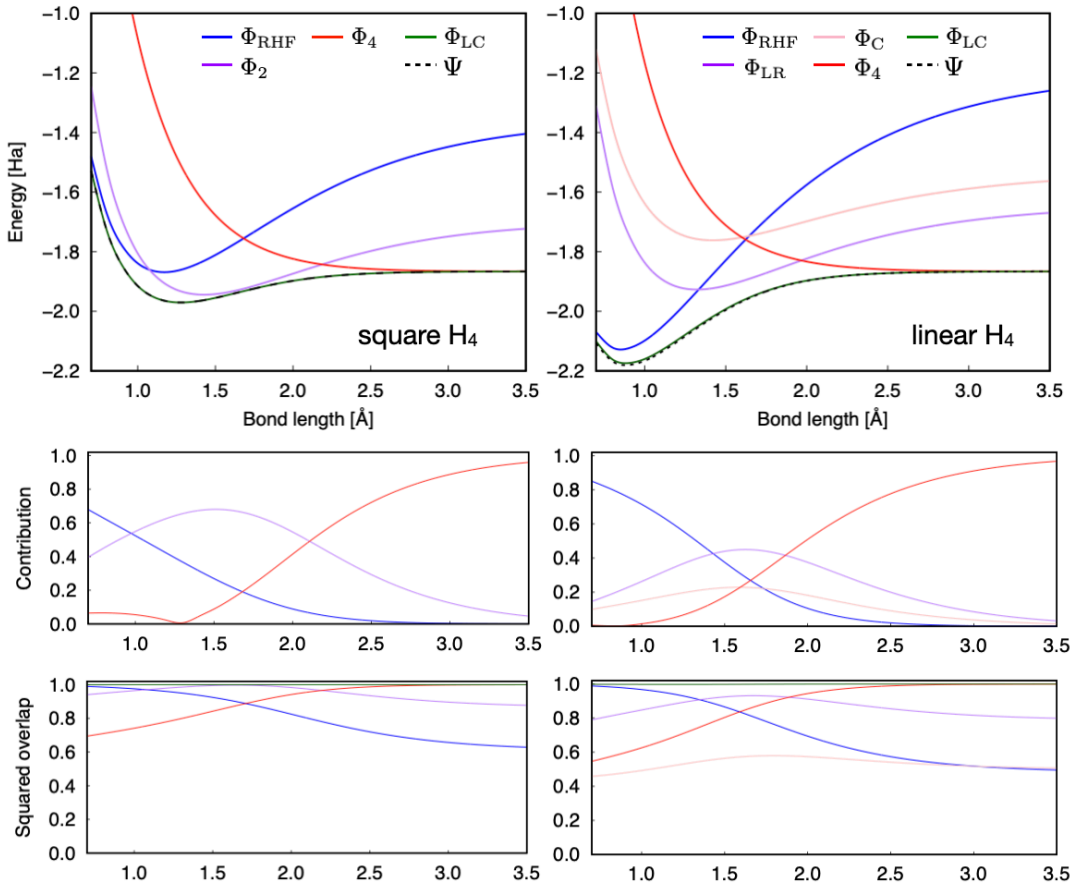


Figure 3.5: Dissociation of square and linear H_4 in a STO-3G basis.

of CSFs. The resulting state has large overlap with the exact ground state wavefunction because the CSFs capture the strong correlation effects between open-shell electrons. Although this provides a qualitatively accurate binding curve, quantitative quantum chemical predictions [306] require powerful computational methods that can simultaneously treat dynamic correlation and orbital relaxation effects, in addition to strong correlation due to spin coupling.

In principle, this could be achieved using classical computational methods based on many-body perturbation theory,[49] but their success strongly depends on the nature of the reference state. There exists a plethora of well-established methods capable of correlating Hartree–Fock reference states, such as truncated coupled cluster.[35, 49] Equation-of-motion-based approaches can yield accurate approximations for systems with few open-shell electrons using diradical or triradical reference

states.[161, 225] However, correlating a reference state with a large number of open-shell electrons remains a challenge.[160]

Quantum processors do not suffer from this limitation. In digital quantum computing, the many-electron wavefunction is encoded by mapping it to a quantum superposition of the qubit states of the quantum device. Quantum algorithms involve preparing the initial qubit state for a known reference wavefunction $|\Phi_{\text{ref}}\rangle$, and either slowly evolving it to ground state of the Hamiltonian $|\Psi\rangle$, for example in quantum imaginary time-evolution[95] and adiabatic state preparation,[76] or collapsing the qubit register to the exact eigenstate through probabilistic projective measurements, for example in quantum phase estimation (QPE).[48, 67, 72, 76] The cost of these approaches depends critically on the overlap of the initial reference state with the exact eigenstate.[41] However, the dimension of the initial state, as given by the number of determinants involved in its expansion in the Slater determinant basis, can be exponential in the number of electrons due the exponential memory scaling with the number of qubits.

QPE involves three steps. First, the initial state $|\Phi_{\text{ref}}\rangle$ must be prepared on the quantum computer. The state is then time-evolved under the action of the Hamiltonian. This applies a phase on the qubit state proportional to the energy of each eigenstate overlapping with the initial state. Finally, the energy of the exact eigenstate is obtained through measurement of the quantum register. Given an accurate initial state, this procedure is guaranteed to yield the exact energy up to an arbitrary numerical precision in a time scaling approximately quadratically in the number of orbitals.[58, 65, 85, 332]

However, the probability of measuring the ground state energy is determined by the squared overlap $\gamma^2 = |\langle\Phi_{\text{ref}}|\Psi\rangle|^2$. Consequently, $\mathcal{O}(|\gamma|^{-2})$ measurements are required to obtain the ground state energy reliably. Although post-QPE approaches require only $\mathcal{O}(|\gamma|^{-1})$ measurements,[104, 105, 298] this overlap problem remains a significant bottleneck.[41]

Most implementations of quantum algorithms employ the Hartree–Fock state as the reference.[32] However, since the overlap between the Hartree–Fock determinant

and the exact ground state decays exponentially with the number of open-shell electrons, the performance of quantum algorithms for strongly-correlated systems rapidly deteriorates.[41] Several works on concrete resource estimates for QPE-based electronic structure for challenging systems predict runtimes for a single-point energy calculation in the order of several days, despite generous assumptions for the quantum hardware requirements.[64, 106, 332] It is clear that reductions in the overall runtime of QPE, as enabled through improved reference states, are of paramount importance to enable practical quantum computation of electronic structure.

One could in principle use classical methods that target the exact wavefunction within the valence space, such as FCI-QMC,[163] to obtain a high-quality reference state. However, loading the reference state onto a quantum computer requires applying a quantum circuit whose time complexity (circuit depth) is linear in the number of determinants.[83, 181] This would be inefficient for strongly correlated reference states with very large CI expansions.

Our spin-coupled formalism provides a systematic approach to constructing reference wavefunctions with good overlaps with the exact state. Furthermore, although the number of determinants in each CSF is equal to the number of fully open-shell determinants with $M_s = 0$, which grows combinatorially with the size of the system as $\binom{N}{N/2}$,[276] the number of distinct coefficients only scales as $\mathcal{O}(N)$. This highly symmetric structure in the wavefunction coefficients means that they can be efficiently prepared on a quantum device at a small cost. In a separate manuscript, we provide quantum circuits for preparation of such CSFs with $\mathcal{O}(N)$ depth.[51] By incorporating the correlation due to symmetry in terms of a few symmetry-derived objects (the CSFs), this approach therefore has the potential to unlock the power of quantum algorithms for challenging chemical problems.[64, 65, 333]

Similarly, the fact that each contributing valence state uses different sets of orthonormal orbitals, and that the valence states are not mutually orthogonal, poses no problem for quantum state preparation.[51] Using our reference wavefunctions

as initial states in QPE directly reduces the overall runtime compared to using a Hartree–Fock reference. In the simple case of N_2, for example, $|\langle\Phi_{\mathrm{RHF}}|\Psi\rangle|^2 \sim 0.06$ at elongated bonds, but $|\langle\Phi_{\mathrm{LC}}|\Psi\rangle|^2 \sim 1$ (Fig. 3.2c), which translates to a sixteen-fold speed-up in QPE execution.

We have presented a selection of CSFs that is well-suited for approximating ground states. One could in principle use a similar framework to build CSFs for excited states. However, excited state energies can also be computed using quantum subspace diagonalisation algorithms, a new class of quantum algorithms with similarities to classical Krylov methods.[252] Here, each reference state is time-evolved on the quantum computer for different times. The resulting (nonorthogonal) states form a low-dimensional subspace for which the Hamiltonian and overlap matrix elements can be evaluated through measurements on the quantum device. The spectrum of the Hamiltonian within this subspace is then obtained by solving the resulting generalised eigenvalue problem on a classical computer. It has been shown that this can lead to accurate ground state energy estimates using a small number of variational parameters.[96, 98, 100–102, 334] In our follow-up work,[51] we demonstrate that using the CSFs derived in this work as the reference states greatly accelerates convergence towards accurate estimates of both ground and excited state energies.

We have used a minimal basis (STO-3G) to discretise the Hilbert space, as this is the simplest chemical model and provides a qualitatively accurate description of the valence electrons. Nevertheless, our approach can be extended straightforwardly to larger basis sets. In a larger basis, one can build the same configurations for those orbitals, since the correlation due to spin coupling is only present in the valence space of open-shell orbitals. The quantum algorithms discussed above can then be applied to compute the remaining correlation effects.

3.7 Conclusions and outlook

Despite the fact that the chemical bond is at the very heart of chemical theory, an accurate representation of the process of bond formation has presented a severe

challenge to quantum chemical methods. Modern approaches[35, 160, 222, 223, 335] are based on complete expansions within a set of active MOs[291] and suffer from high computational and operational complexity, which has encouraged the widespread erroneous belief that the wavefunction is inherently complex. Using the spin-coupled molecular orbital theory presented in this article, we have shown that the wavefunction is in fact highly structured and that bond formation can be represented through a handful of spin-coupled states, each with well-defined spin quantum numbers and physical characteristics.

Our spin-coupled theory exposes a simple interpretation of the chemical bond. The character of the bond at equilibrium bond lengths is predominantly that predicted by MO theory, but electron crowding is relieved by diradical character, where electrons localise on opposite atoms, leading to a reduction in formal bond order and bond elongation. As a bond breaks, the diradical character increases and the MO character reduces. For multiple bonds, the radical states are characterised by high-spin coupling of the electrons localised on each atom. The σ and π bonds break at different bond lengths, with the π bonds typically breaking first, and the bond order reduces sequentially as each pair of electrons in the bonds localise on opposite atoms. Applying this simple theory recovers 90% of the wavefunction resulting from a complete active space calculation, and our simple chemical model therefore has a direct connection to sophisticated wavefunction expansions.

Our model invokes different orbitals for each configuration, and the resulting states are not orthogonal. Moreover, the spin-coupled configurations correspond to different genealogical spin-coupling schemes, with different orbital orderings. The highly compact representation of the wavefunction afforded by our approach is lost when either a single orbital basis, or single CSF basis is used to represent all contributing states. For the archetypical strongly correlated systems considered in this work, the apparent complexity of the wavefunction in traditional approaches arises from the insistence on using a single basis of orthonormal orbitals and a single CSF basis to represent all configurations, so that the Hamiltonian matrix evaluation is straightforward.[276, 278, 285] While matrix elements between nonorthogonal

Slater determinants can be readily computed,[336–338] building the determinant expansion of each CSF has exponential cost using traditional algorithms. To exploit the compact structure revealed by our theory, it is necessary to use methods that can handle non-orthogonal CSF states.

Quantum computing has the potential to overcome many of the remaining challenges in quantum chemistry. Provided that a high-quality reference state can be efficiently parametrised as a quantum circuit, algorithms such as quantum phase estimation and quantum eigenstate filtering can efficiently extract the exact ground state within a given basis set. Our spin-coupled molecular orbital theory provides a framework for constructing highly entangled reference states for strongly correlated molecular ground states using only a basic understanding of a chemical process, a small number of variational parameters, and at the cost of a mean-field calculation. Since the entanglement structure of each configuration is defined through spin and symmetry constraints, they can be prepared and transformed on a quantum register with low-depth quantum circuits,[51] providing a practical route to using quantum algorithms to address challenging problems in quantum chemistry.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Spin-coupled molecular orbitals: chemical intuition meets quantum chemistry
Publication Status	□Published □ Accepted for Publication ☒Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	D. Marti-Dafcik, N. Lee, H. G. A. Burton, D. P. Tew, "Spin-coupled molecular orbitals: chemical intuition meets quantum chemistry", arXiv:2402.08858 (2024)

Student Confirmation

Student Name:	Daniel Marti Dafcik		
Contribution to the Paper	D.M.D. conceived the project and developed the theoretical framework and numerical methodology in collaboration with D.P.T. D.M.D. wrote code to implement the numerical method and carried out part of the computations. N.L. contributed additional code and generated part of the numerical results. D.M.D. and D.P.T. wrote the initial draft of the manuscript. All authors contributed to the revision and final version of the manuscript and provided ideas and guidance throughput the completion of the project.		
Signature		Date	30 September 2025

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof. David P. Tew		
Supervisor comments		
Signature 	Date	1/10/2025

This completed form should be included in the thesis, at the end of the relevant chapter.

4

Spin coupling is all you need: Encoding strong electron correlation in molecules on quantum computers

Contents

4.1 Introduction 104
4.2 Background and conceptual framework 109
4.2.1 Encoding and notation for qubit states 109
4.2.2 Discovering and describing strong electron correlation in molecules through symmetry 110
4.2.3 Configuration State Functions 114
4.2.4 Multireference quantum chemistry 115
4.2.5 Comparison with spin-adapted classical algorithms . . . 117
4.3 This work: overview of results 118
4.4 Quantum circuits for preparation of spin eigenfunctions 120
4.4.1 Spin eigenfunctions for bond dissociation 122
4.4.2 Mapping to Dicke states and symmetric states 123
4.4.3 Preparation of Dicke and symmetric states 125
4.4.4 Notation for operators and circuits 127
4.4.5 Preparation of singlet states with locally ferromagnetic coupling 128
4.4.6 Mapping from spin to Fock space 130
4.4.7 Total circuit cost and examples 131
4.4.8 Other spin eigenfunctions 133
4.4.9 Linear combinations of spin eigenfunctions 135
4.5 Heuristic quantum algorithms 135
4.5.1 Variational quantum eigensolver 136

4.5.2 Quantum subspace diagonalization through real-time evolution 141
4.5.3 ADAPT-QSD: QSD with adaptive ansatz 148
4.5.4 Adiabatic state preparation from the molecular dissociation limit 152
4.6 Avoiding the orthogonality catastrophe in quantum phase estimation 156
4.6.1 Quantum phase estimation 156
4.6.2 Resource estimation for fault-tolerant simulation of FeS systems 159
4.7 Conclusions and discussion 165

Abstract

The performance of quantum algorithms for eigenvalue problems, such as computing Hamiltonian spectra, depends strongly on the overlap of the initial wavefunction and the target eigenvector. In a basis of Slater determinants, the representation of energy eigenstates of systems with N strongly correlated electrons requires a number of determinants that scales exponentially with N. On classical processors, this restricts simulations to systems where N is small. Here, we show that quantum computers can efficiently simulate strongly correlated molecular systems by directly encoding the dominant entanglement structure in the form of spin-coupled initial states. This avoids resorting to expensive classical or quantum state preparation heuristics and instead exploits symmetries in the wavefunction. We provide quantum circuits for deterministic preparation of a family of spin eigenfunctions with $\binom{N}{N/2}$ Slater determinants with depth $\mathcal{O}(N)$ and $\mathcal{O}(N^2)$ local gates. Their use as highly entangled initial states in quantum algorithms reduces the total runtime of quantum phase estimation and related fault-tolerant methods by orders of magnitude. Furthermore, we assess the application of spin-coupled wavefunctions as initial states for a range of heuristic quantum algorithms, namely the variational quantum eigensolver, adiabatic state preparation, and different versions of quantum subspace diagonalization (QSD) including QSD based on real-time-evolved states. We also propose a novel QSD algorithm that exploits states obtained through adaptive quantum eigensolvers. For all algorithms, we demonstrate that using spin-coupled initial states drastically reduces the quantum resources required to simulate strongly correlated ground and excited states. Our work provides a crucial component for enabling scalable quantum simulation of classically challenging electronic systems.

4.1 Introduction

Computing the low-lying eigenstates and energies of electronic Hamiltonians remains a fundamental challenge in the physical sciences and scientific computing. Despite the success of classical computational approaches for quantum chemistry in simulating a wide range of molecules, their application to systems exhibiting strong electron correlation is limited due to the difficulty of efficiently representing highly entangled quantum states.[160]

Quantum computers can generate and transform vectors whose dimension scales exponentially in the number of qubits with remarkable efficiency.[54] This suggests that they might provide a solution to the curse of dimensionality. The reality is more nuanced. Despite the plethora of algorithms developed for computing energy eigenvalues or preparing eigenstates of many-body systems, complexity-theoretic results rule out exponential quantum speedups for worst-case versions of the electronic Schrödinger equation.[39, 40] Although this does not prohibit practical speedups for realistic physical systems, a fundamental problem lies at the core of quantum computing for electronic structure: the performance of nearly all quantum algorithms for eigenvalue problems strongly depends on the accuracy of the initial state.

A random vector in the many-electron Hilbert space is expected to have exponentially small overlap with any exact eigenstate.[339, 340] Therefore, the success of quantum algorithms relies on finding an initial state which approximately contains the structure of the true eigenstate. This can be achieved by computing approximate wavefunctions through classical heuristics, such as Hartree–Fock, configuration interaction,[38] or density matrix renormalization group (DMRG) approaches,[15] and loading such states on quantum hardware.[181] On fault-tolerant quantum hardware, quantum heuristics—such as the variational quantum eigensolver (VQE),[42, 44, 52, 91, 92, 97, 184, 185, 258, 261, 262, 341–347] adiabatic state preparation (ASP),[41, 64, 73, 76, 186–188, 251] and quantum subspace diagonalization (QSD)[93–103, 334, 348, 349] —could be used to prepare initial states

for quantum phase estimation, a method for which the runtime rigorously depends on the overlap between the initial state and the target eigenstate.[48, 67, 72]

The challenge is that, for systems where such heuristics are accurate, the problem can often be solved entirely using classical algorithms (to sufficient accuracy). This casts doubts on the advantage of using quantum computers over classical machines for quantum chemistry.[41] To obtain quantum speedups, the approximate initial state found through heuristics must be of insufficient accuracy for chemistry applications. At the same time, its structure must be such that further refining the wavefunction through classical algorithms is hard, while encoding it as an initial state for quantum algorithms is easy. In this work, we tackle the strong correlation problem in quantum chemistry by leveraging a recently-developed classical heuristic that satisfies these requirements and is ideally suited for initial state preparation on quantum computers.[50]

To understand what features make an initial state preparation method advantageous for quantum computation, consider the success and limitations of scalable classical algorithms. All polynomially-scaling methods for quantum chemistry rely on identifying qualitatively accurate, yet approximate, initial states with a simple, compact wavefunction description. Restricted Hartree–Fock theory provides good initial states for systems with weak electron correlation, because it accurately encodes the mean-field character of delocalized molecular wavefunctions.[218] The many-body Hartree–Fock state is simply an antisymmetrized product of delocalized molecular orbitals, which corresponds to a single Slater determinant. Since this state has a compact representation in the single-particle basis of Hartree–Fock orbitals, one can expand around it in a controlled manner, e.g. with methods based on many-body perturbation theory, to improve the state without incurring exponential cost.[35]

For strongly correlated systems, the Hartree–Fock state is inaccurate because the molecular orbital picture underpinning it breaks down. Even a qualitatively accurate wavefunction for such systems requires a superposition of Slater determinants that scales exponentially with the number of strongly correlated electrons,

regardless of the choice of single-particle basis. In such scenarios, no efficient classical heuristics exist, and one resorts to brute-force algorithms which in general scale exponentially as they require a linear combination of all the relevant Slater determinants.[163, 226, 227]

The high dimension of such states also poses a challenge for quantum algorithms due to the initial state dependency. Suppose that an approximate wavefunction is found through a state-of-the-art classical algorithm, such as selective configuration interaction or DMRG. Even then, its use in quantum algorithms is inefficient, because preparing a superposition with arbitrary coefficients through a quantum circuit requires a number of steps (gates) proportional to the number of basis states (determinants).[181, 192] Quantum heuristics could in principle offer an alternative, but often also suffer in strongly correlated regimes.[184] Therefore, there appears to be a trade-off between the accuracy of the initial state used in quantum algorithms and the cost of state preparation.[350–352] Fundamentally, the challenge is that quantum computers need to exploit structure in the problem, and while such structure is present in Hartree–Fock states, it is washed away in strongly correlated wavefunctions due to the brute-force nature of algorithms that are used for their approximation.

In this work, we present a heuristic state preparation method that solves the initial state problem for a range of strongly correlated molecules. This relies on our recently-developed generalized molecular orbital theory,[50] where we showed that molecular electronic wavefunctions that arise during the stretching of chemical bonds are often highly structured and can be approximated to high accuracy through few spin-coupled states. These spin-coupled states can be deduced from chemical intuition and spatial-spin symmetry arguments, and their representation is given directly from the standard Clebsch–Gordan coefficients.[189] This challenges the common notion that strongly correlated states are inevitably complex. While they might be highly-dimensional when expressed as vectors expanded in a single-particle basis, they encode a relatively small amount of information.

Here, we enable their efficient application in quantum algorithms by providing quantum circuits that prepare a family of spin-coupled states with $\binom{N}{N/2}$ determinants in depth $\mathcal{O}(N)$ and using $\mathcal{O}(N^2)$ gates. We achieve this by exploiting the symmetry structure in spin-coupled wavefunctions and connecting them to Dicke states, a well-known family of entangled states.[353] This approach avoids the exponential scaling of generic, black-box state preparation methods.[181, 192] This is possible because our spin-coupled molecular orbital theory assigns a unique set of orbitals to each reference state, ensuring each wavefunction has a highly symmetric and compact representation. The states can then be rotated to a common basis using standard, linear-depth quantum circuits.[238]

We numerically assess the application of this state preparation method in VQE, ASP, and QSD, where the quantum algorithms are initialized with spin-coupled states. We also propose a new QSD algorithm, ADAPT-QSD, which is of interest in its own right. It builds a subspace from states obtained through adaptive quantum eigensolvers such as ADAPT-VQE.[261] For all algorithms, we demonstrate that the use of spin-coupled initial states greatly reduces the quantum resources (circuit depth and gate counts) and the number of degrees of freedom (variational parameters) required to achieve a given accuracy, compared to using the Hartree–Fock reference. This confirms our spin-coupled framework as a low-cost approach for improving the performance of heuristic quantum algorithms.

We highlight QSD as a class of algorithms that can most strongly benefit from access to different reference states, in particular for *multiconfigurational* eigenstates i.e. those for which multiple dominant configurations (e.g. as provided through multiple initial states) are required for an accurate wavefunction description. These are generally hard to tackle using classical algorithms and therefore a promising target for quantum computation. Combining QSD with spin-coupled states also allows computation of excited state energies at low cost.

Finally, we analyze the state preparation question in the context of fault-tolerant computation of electronic structure based on quantum phase estimation. We consider the number of non-Clifford gates required to prepare initial states with

high ground state overlap for systems with many (up to 35) spin-coupled electrons such as FeMoCo [41, 64, 65] Our extrapolated gate count estimates suggest that one could greatly reduce the state preparation cost by directly encoding the entanglement due to spin coupling through a circuit similar to the ones presented here, compared to preparation of states obtained from classical black-box algorithms such as DMRG or selected CI. This would enable efficient fault-tolerant quantum simulation of such strongly correlated molecules—precisely the systems for which classical methods are most likely to remain insufficient.

Overall, our work provides a scalable framework with the necessary ingredients required to unlock the unique power of quantum computers for challenging chemical systems.

We note that we have only studied strongly correlated electrons in molecules, and it is unclear how one would apply similar concepts to solid state systems. While the state preparation circuits can in principle be used for any electronic or spin-$\frac{1}{2}$ system, our choice of a few bespoke CSFs is unlikely to provide a solution to the strong correlation problem in solid state materials. Molecules consist of a relatively small number of electrons, which makes it possible to identify the spin coupling within and across different subsystems, as we show below.

In Sec. 4.2, we introduce our wavefunction notation and encoding and the background on spin-coupled states. Most of the content appeared in our previous work, Ref. 50, in a language tailored to chemists; here, we present these concepts to a general quantum science audience. In Sec. 4.3 we present a brief overview of our main results. In Sec. 4.4, we provide the quantum circuits for preparing spin eigenfunctions. In Sec. 4.5, we discuss the numerical results obtained from classical simulations of quantum algorithms, and introduce the ADAPT-QSD algorithm. In Sec. 4.6, we analyze the initial state preparation task within the longer-term, fault-tolerant quantum computing context.

4.2 Background and conceptual framework

4.2.1 Encoding and notation for qubit states

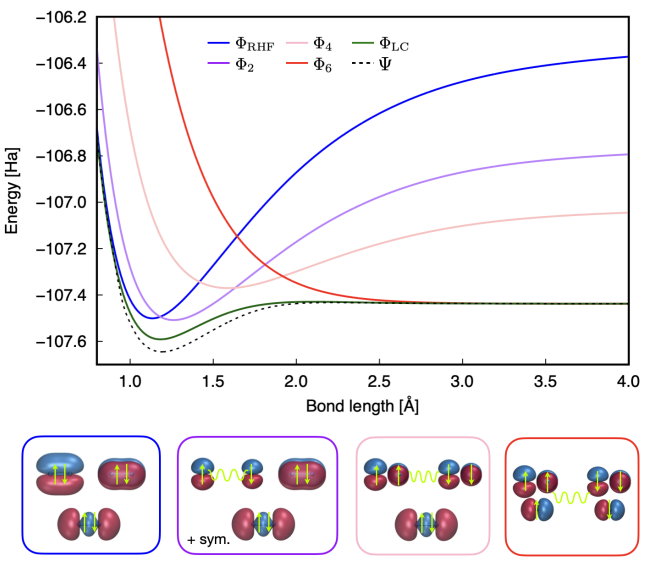
We work with the second-quantized representation, where basis states are antisymmetric product states (Slater determinants) in a fermionic Fock space and the antisymmetry is manifest through anticommuting operator algebra. Using the Jordan-Wigner mapping,[70] an electronic wavefunction of $2M$ spin-orbitals (M spatial orbitals) can be mapped to a system of $2M$ qubits, where each qubit represents the occupation of a spin-orbital (0 means unoccupied, 1 means occupied). Every computational basis state $|j\rangle$ uniquely encodes a Slater determinant represented as

$$|j\rangle = \bigotimes_{i=1}^{M} \Big(|f_{i_\alpha}\rangle \otimes |f_{i_\beta}\rangle \Big), \tag{4.1}$$

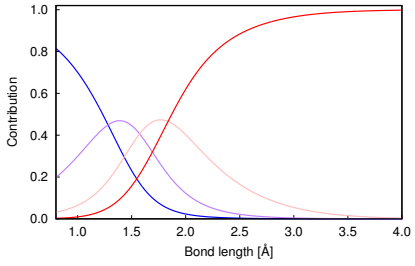
where $f_{i_\alpha}, f_{i_\beta} \in \{0, 1\}$ represent the occupation of the α (spin-up) and β (spin-down) spin-orbital corresponding to the i-th spatial orbital, and the Hamming weight (number of 1s) is the number of electrons. We define the qubit ordering such that a spin-up orbital is followed by the corresponding spin-down orbital. This convention allows us to introduce a more compact notation:

$$|j\rangle = \bigotimes_{i=1}^{M} |g_i\rangle \,, \tag{4.2}$$

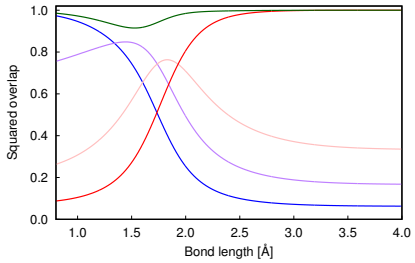
where $g_i \in \{0, 2, \alpha, \beta\}$ is the occupation of a spatial orbital: 0 if unoccupied, 2 if doubly occupied, α or β if singly occupied by a spin-up or spin-down electron. We will typically drop the tensor product i.e. $|g_i\rangle \otimes |g_j\rangle =: |g_i g_j\rangle$, and often write products of K identical single-qubit states as $|g_i\rangle^K$. The compact notation (Eq. (4.2)) relates to the standard notation (Eq. (4.1)) as follows: $|0\rangle \to |00\rangle$, $|\alpha\rangle \to |10\rangle$, $|\beta\rangle \to |01\rangle$, $|2\rangle \to |11\rangle$. Since this notation only specifies the occupation of each orbital, we must specify the nature and ordering of the orbitals separately. We will do so through an ordered list which we define by curly brackets $\{\phi_1, \phi_2, ...\}$.



(a) Energies and configurations of the valence CSFs.



(b) Coefficients of each CSF in Φ_{LC}.



(c) Squared overlaps $|\langle\Psi|\Phi_i\rangle|^2$.

Figure 4.1: The dissociation of N_2 in a STO-3G basis in the valence space of six electrons in six orbitals (twelve qubits) using four configuration state functions (CSFs) Φ_i, $i = 0, 2, 4, 6$, as predicted by generalized MO theory. Here, Φ_0 is the RHF mean-field state, Φ_2, Φ_4, Φ_6 are CSFs with two, four, and six spin-coupled (strongly correlated) electrons, and Ψ is the exact ground state. Panel 4.1a shows the energy of each CSF; panel 4.1b displays the coefficients of each CSF in the optimal linear combination of CSFs; panel 4.1c shows the squared overlap between each CSF and the exact ground state. The figure is reproduced with permission from Ref. 50.

4.2.2 Discovering and describing strong electron correlation in molecules through symmetry

Common examples of strongly correlated electronic wavefunctions in molecules include electronic eigenstates at dissociation. Here, we consider the nitrogen molecule, which illustrates the challenges of both classical[163, 319] and quantum[184, 187, 354] algorithms.

In Ref. 50, we showed that the ground state wavefunction can be represented in terms of a few states with both mean-field and strongly-correlated character. Each state encodes a unique entanglement pattern that corresponds to a clear physical motif. Bonding is manifest by electrons that doubly occupy delocalized molecular orbitals, and strong correlation corresponds to electrons occupying open-shell, localized orbitals that couple through their spin angular momentum. A superposition of four reference states yields an accurate wavefunction approximation at any nuclear

geometry (Fig. 4.1), without requiring a large number of variational parameters. Below, we discuss the derivation of the four reference states.

In a minimal STO-3G basis, the Restricted Hartree–Fock method yields the set of delocalized molecular orbitals $\{1\sigma_\mathrm{g}, 1\sigma_\mathrm{u}, 2\sigma_\mathrm{g}, 2\sigma_\mathrm{u}, 3\sigma_\mathrm{g}, 1\pi_{\mathrm{u},x}, 1\pi_{\mathrm{u},y}, 1\pi_{\mathrm{g},x}, 1\pi_{\mathrm{g},y}, 3\sigma_\mathrm{u}\}$. We use a Hamiltonian that *freezes* the lowest four orbitals and therefore only work with the 1π-orbitals and the 3σ valence orbitals.[1] This yields a system of six electrons in six orbitals that can be mapped to twelve qubits. We drop the shell indices from the specification of the orbitals and therefore the delocalized orbitals considered here are: $\{\sigma_\mathrm{g}, \pi_{\mathrm{u,x}}, \pi_{\mathrm{u,y}}, \pi_{\mathrm{g,x}}, \pi_{\mathrm{g,y}}, \sigma_\mathrm{u}\}$. In this basis, the RHF state doubly occupies the three lowest orbitals (Fig. 4.1a, blue), given by

$$|\Phi_\mathrm{RHF}\rangle = |222000\rangle \tag{4.3}$$

using the compact notation introduced in Eq. (4.2).

Although a rigorous definition of electron correlation remains an open research question,[355] one natural definition is to take the view that any product state in second quantization is uncorrelated. In this sense, the RHF state in Eq. (4.3) is uncorrelated. This is consistent with it being a fermionic mean-field state, and with the widely adopted definition of the correlation energy $E_\mathrm{corr} = E - E_\mathrm{RHF}$. In this paradigm, correlation is therefore a measure proportional to the deviation of an entangled quantum state from a product state, using the Hartree–Fock orbitals as the prescribed single-particle basis.

Around the equilibrium bond length, the RHF state is a relatively accurate approximation of the true ground state (Fig. 4.1, blue). This reflects the success of molecular orbital theory in chemistry, which describes bonding as consisting of electrons occupying delocalized orbitals, a useful model for many molecules at equilibrium.

However, this model breaks down at dissociation. The electronic wavefunction exhibits strong correlation effects, as reflected in the poor energy of the RHF state (Fig. 4.1a) and its low overlap with the exact ground state (Fig. 4.1c). In the

[1] The state of the frozen core orbitals was explicitly represented as $|\mathcal{C}\rangle$ in Ref. 50, but we drop it in this paper for compactness.

basis of RHF orbitals, the exact singlet ground state is a superposition of 44 Slater determinants of similar weights. Classical methods encounter challenges in such situations since the wavefunction is not compact in the RHF basis. Quantum algorithms inherit these challenges if the RHF state is used as the initial state, as all the entanglement needs to be *discovered* throughout the quantum computation, e.g. by optimizing the ansatz formed by a parametrized quantum circuit in a variational quantum algorithm.

Fortunately, many of the low-energy eigenstates that appear in nature are highly structured, even if they can be highly entangled. Approximations thereof can be obtained from simple symmetry considerations.

Let us now consider the wavefunction for N_2 at dissociation. The wavefunction for the dissociated state is not the direct product of the atomic states, because the local spins of the dissociated atoms are coupled and the wavefunction is entangled across the two atomic subsystems. The ground state of each isolated nitrogen atom is well-understood from basic chemical principles: the three valence electrons repel each other and therefore form the maximally antisymmetric spatial wavefunction, which maximizes the average distance between electrons. Due to the overall antisymmetry requirement for fermionic wavefunctions, the spin wavefunction must be fully symmetric. The wavefunction for the atom can therefore be expressed as a quartet (total spin $s = 3/2$) that ferromagnetically couples three spin-1/2 particles (Fig. 4.1a, red). This spin-coupled state can be determined directly from the Clebsch–Gordan coefficients for angular momentum coupling:

$$|3, \frac{3}{2}\rangle = \begin{cases} |\frac{3}{2}, \frac{3}{2}\rangle = |\alpha\alpha\alpha\rangle \\ |\frac{3}{2}, \frac{1}{2}\rangle = \frac{1}{\sqrt{3}}\Big(|\alpha\alpha\beta\rangle + |\alpha\beta\alpha\rangle + |\beta\alpha\alpha\rangle \Big) \\ |\frac{3}{2}, -\frac{1}{2}\rangle = \frac{1}{\sqrt{3}}\Big(|\alpha\beta\beta\rangle + |\beta\alpha\beta\rangle + |\beta\beta\alpha\rangle \Big) \\ |\frac{3}{2}, -\frac{3}{2}\rangle = |\beta\beta\beta\rangle \,, \end{cases} \tag{4.4}$$

where on the left we have used the notation $|n, s\rangle$ for the particle and spin quantum number, in the center we have used $|s, m_s\rangle$ for the total spin and spin projection quantum numbers s and m_s, and on the right we use α and β to represent the state of each site as either spin-up or spin-down.

The representation on the right hand side of Eq. 4.4 is equivalent to electronic Slater determinants with singly-occupied spatial orbitals using the compact notation for fermions/qubits from Eq. (4.2), and assuming a single-particle basis of localized orbitals. These localized orbitals can be obtained from the RHF orbitals for N_2 through a trivial unitary transformation of the molecular orbitals: $\mathrm{z}_L = \frac{1}{\sqrt{2}}(\sigma_\mathrm{g} + \sigma_\mathrm{u})$, $\mathrm{z}_R = \frac{1}{\sqrt{2}}(\sigma_\mathrm{g} - \sigma_\mathrm{u})$, $\mathrm{x}_L = \frac{1}{\sqrt{2}}(\pi_{\mathrm{u,x}} + \pi_{\mathrm{g,x}})$, $\mathrm{x}_R = \frac{1}{\sqrt{2}}(\pi_{\mathrm{u,x}} - \pi_{\mathrm{g,x}})$, $\mathrm{y}_L = \frac{1}{\sqrt{2}}(\pi_{\mathrm{u,y}} + \pi_{\mathrm{g,y}})$, $\mathrm{y}_R = \frac{1}{\sqrt{2}}(\pi_{\mathrm{g,y}} - \pi_{\mathrm{g,y}})$. Here, the delocalized RHF orbitals have been transformed to a set of atomic-like, mutually orthogonal molecular orbitals x, y, z localized on the *left* (L) or *right* (R) atom.

In the dissociation limit, the Coulomb force between electrons localized on different atoms is zero. Thus, the N_2 ground state can be any state that couples the two atomic quartets with proper quantum numbers. However, as the interatomic distance R decreases, singlet-coupled electrons can delocalize across the bond, which reduces the kinetic energy. This one-body delocalization lifts the degeneracy between eigenstates and the ground state requires that the two three-electron subsystems couple to a six-electron singlet. This state, which we will refer to as $|\Phi_6\rangle = |\mathcal{O}^{6,1}_{0,0}\rangle$, can also be determined from the Clebsch–Gordan coefficients and corresponds to a superposition of 20 Slater determinants in the localized orbital basis (Eq. (B.1c) in Appendix B.1). In contrast to the mean-field RHF state, this spin-coupled wavefunction is exact at dissociation but very inaccurate at shorter bond lengths (Fig. 4.1). Precisely due to its symmetry origin, the state at dissociation is predictable despite being strongly correlated. It has relatively low Kolmogorov complexity in the sense that few steps are required to specify it: one must simply define the orbitals involved and their spin coupling pattern. From this, the coefficients of each basis state can be computed through the standard formula for Clebsch–Gordan coefficients.

As shown in Ref. 50, the same is true for the dissociation of a range of molecules including multiple diatomics, water, and hydrogen clusters: in the limit of long bond length, they can all be described by a single spin-coupled wavefunction, whose

general form we will define below. For N strongly correlated electrons, the spin-coupled state $|\mathcal{O}^{N,1}_{0,0}\rangle$ is a superposition of $\binom{N}{N/2}$ Slater determinants (where $N = 6$ for N_2). This exponential scaling in the number of electrons means that it is hard to exploit such wavefunctions as reference states in classical algorithms, as it is difficult to refine highly-entangled reference states e.g. through pertubation theory or coupled cluster methods.[160]

At intermediate bond lengths, neither the RHF state nor the fully spin-coupled state ($|\mathcal{O}^{6,1}_{0,0}\rangle$ for N_2) are an accurate description of the ground state. Instead, multiple reference states have significant weights in the wavefunction expansion. Such wavefunctions are referred to as multiconfigurational in quantum chemistry,[34, 356] and methods that approximate these by using different references states are known as *multireference*.

In the following, we introduce the notation used for the general configuration state functions that describe each of the contributing states relevant to strongly correlated regimes where bonds are partially broken.

4.2.3 Configuration State Functions

The strong correlation corresponding to localized spin-coupled orbitals can be described through entangled states whose coefficients are determined from symmetry. The Hamiltonian operator commutes with both the total spin operator, $\hat{S}^2$, and the spin projection operator, $\hat{S}_z$. Therefore, energy eigenstates must simultaneously be eigenstates of the spin operators with corresponding total spin and spin projection quantum numbers S, M_S. A Slater determinant is in general an eigenfunction of $\hat{S}_z$ but not of $\hat{S}^2$. Exceptions include fully closed-shell determinants with $S, M_S = 0, 0$, such as the state in Eq. (4.3).

Configuration state functions (CSFs) are eigenfunctions of both $\hat{S}^2$ and $\hat{S}_z$ with quantum numbers formed by a symmetry-adapted linear combination of determinants.[38] All the reference states considered in this work are CSFs. These are products of closed-shell delocalized orbitals as well as open-shell spin-coupled orbitals. The spin-coupled part is defined by the pattern in which the single

electrons/orbitals are coupled to form the many-body wavefunction. We define the following first-quantized representation for the CSFs used in this work (Eq. 6 in Ref. 50):

$$|\Phi\rangle = \mathcal{N}\hat{\mathcal{A}}\,|\phi_c^2\cdots\rangle\,\mathcal{O}^{N,i}_{S,M_S}(\phi_o\cdots), \tag{4.5}$$

where S, M_S are the spin quantum numbers and N is the number of spin-coupled electrons (and therefore the number of spin-coupled, localized spatial orbitals). Here, we have included the antisymmetrization operator $\hat{\mathcal{A}}$ and normalization constant $\mathcal{N}$, but we omit these in the remaining text. The term $|\phi_c^2\cdots\rangle$ is a product state with doubly occupied (closed-shell) orbitals $\{\phi_c, ...\}$. The term $\mathcal{O}^{N,i}_{S,M_S}(\phi_o\cdots)$ is a state with singly-occupied (open-shell) orbitals $\{\phi_o, ...\}$ that are spin-coupled. The index i specifies the coupling pattern and the amplitudes of the expansion of the CSF in a product basis are simply the Clebsch–Gordan coefficients; a list of relevant spin-coupled states is provided in Appendix B.1.

In the compact, second-quantized qubit representation (Eq. (4.2)), the closed-shell state maps as $|\phi_c^2\cdots\rangle \to |2\cdots\rangle$, the open-shell state is $|\mathcal{O}^{N,i}_{S,M_S}(\phi_o\cdots)\rangle$, and we explicitly include the occupation of the unoccupied (virtual) orbitals $\{\phi_v, ...\}$ as $|0\cdots\rangle$. Therefore, the qubit representation of any CSF with orbitals $\{\phi_c, ..., \phi_o, ..., \phi_v, ...\}$ takes the form

$$|\Phi\rangle = |2\cdots\rangle\,|\mathcal{O}^{N,i}_{S,M_S}(\phi_o\cdots)\rangle\,|0\cdots\rangle\,, \tag{4.6}$$

where the indices c, o and v run over the closed-shell, open-shell, and virtual orbitals.

4.2.4 Multireference quantum chemistry

To describe situations where multiple correlation effects occur simultaneously, such as bond breaking in the intermediate region of the binding curve, we form hybrid, partially entangled states. By localizing only the RHF orbitals involved in the π_{x}-bond to obtain the basis $\{\sigma_{\mathrm{g}}, \pi_{\mathrm{u,y}}, \mathrm{x}_L, \mathrm{x}_R, \pi_{\mathrm{g,y}}, \sigma_{\mathrm{u}}\}$, we form a state with one stretched bond (two spin-coupled electrons):

$$|\Phi_{2_x}\rangle = |\sigma_{\mathrm{g}}^2\pi_{\mathrm{u,y}}^2\rangle\,\mathcal{O}^{2,1}_{0,0}\,|\pi_{\mathrm{g,y}}^2\sigma_{\mathrm{u}}^2\rangle \to |22\rangle\,\frac{1}{\sqrt{2}}(|\alpha\beta\rangle - |\beta\alpha\rangle)\,|00\rangle\,. \tag{4.7}$$

This is traditionally referred to as a diradical state. Due to the cylindrical symmetry of the molecule, whose bond points along the z-axis, the state has an exactly degenerate counterpart with spin coupling along the y direction. Using orbitals $\{\sigma_g, \pi_{u,x}, y_L, y_R, \pi_{g,x}, \sigma_u\}$, this state has the same qubit representation

$$|\Phi_{2_y}\rangle = |\sigma_g^2 \pi_{u,x}^2\rangle \, \mathcal{O}_{0,0}^{2,1} \, |\pi_{g,x}^2 \sigma_u^2\rangle \rightarrow |22\rangle \, \frac{1}{\sqrt{2}}(|\alpha\beta\rangle - |\beta\alpha\rangle) \, |00\rangle \, . \tag{4.8}$$

In calculations, we will form a fixed, symmetric linear combination of the two (non-orthogonal) states

$$|\Phi_2\rangle = \frac{1}{\sqrt{3}}(|\Phi_{2_x}\rangle + |\Phi_{2_y}\rangle) \tag{4.9}$$

and treat this as a single reference state without any loss in circuit/state expressivity.

We also form a state describing two stretched bonds by localizing four of the orbitals. With orbitals $\{\sigma_g, x_L, y_L, x_R, y_R, \sigma_u\}$, the state is

$$|\Phi_4\rangle = |\sigma_g^2\rangle \, \mathcal{O}_{0,0}^{4,1}(x_L, y_L, x_R, y_R) \, |\sigma_u^2\rangle \rightarrow |2\rangle \, |\mathcal{O}_{0,0}^{4,1}\rangle \, |0\rangle \, , \tag{4.10}$$

where $|\mathcal{O}_{0,0}^{4,1}\rangle$ is given in the Appendix (Eq. (B.1b)). At arbitrary bond lengths, the wavefunction is accurately described by the optimal linear combination of all CSFs,

$$|\Phi_{\mathrm{LC}}\rangle = c_1 \, |\Phi_{\mathrm{RHF}}\rangle + c_2 \, |\Phi_2\rangle + c_3 \, |\Phi_4\rangle + c_4 \, |\Phi_6\rangle \, , \tag{4.11}$$

which can be found by diagonalizing the Hamiltonian in the subspace built from the four (nonorthogonal) states.

The binding curve (Fig. 4.1a) clearly displays the expected behaviour: the RHF state is accurate around the equilibrium distance but poor at dissociation. The spin-coupled states $|\Phi_2\rangle$, $|\Phi_4\rangle$ have increasingly longer minima and improved energies at dissociation. The fully spin-coupled state $|\Phi_6\rangle$ is exact at dissociation.

The trend seen in the energies is fully matched by the trend in the coefficients and squared overlaps (Figs. 4.1b and 4.1c). At short bond lengths, the RHF state dominates the linear combination with high overlap with the exact ground state. As the bond is stretched, the intermediate CSFs $|\Phi_2\rangle$ and $|\Phi_4\rangle$ become increasingly

important, and at long bond lengths the fully spin-coupled state $|\Phi_6\rangle$ dominates, with 100% overlap in the infinite limit. The linear combination is least accurate at the intermediate region; nevertheless the minimum squared overlap is 92%. Using the linear combination state $|\Phi_{\mathrm{LC}}\rangle$ instead of $|\Phi_{\mathrm{RHF}}\rangle$ directly leads to a reduction in the runtime of QPE of more than one order of magnitude at stretched geometries (see Section 4.6 for an extended discussion).

4.2.5 Comparison with spin-adapted classical algorithms

We note that the techniques for construction of CSFs have long been established,[37, 189] and that one can choose to construct different CSF bases with different features.[273–275] Such CSF bases enable implementations of spin-adapted versions of classical electronic structure algorithms.[16, 190] The insight that working in a particular CSF and orbital basis can enable a compression of many-electron wavefunctions for systems where spin coupling dominates has recently also been exploited in works on quantum chemistry on classical computers. For example, Li Manni *et al.* showed that one can greatly increase the sparsity in the Hamiltonian matrix if one expresses it in a CSF basis with a set of localized orbitals,[307, 308] in line with what we saw for diatomic molecules at dissociation where a single CSF is exact (Section 4.2.2). For such systems in which a single spin-coupled CSF dominates the true eigenstate, this speeds up the classical FCI-QMC algorithm which uses stochastic sampling of Hamiltonian terms and therefore benefits from increased sparsity.[163] However, for multiconfigurational wavefunctions [356] in which multiple configurations of different character contribute significantly to the exact eigenstate, no many-body basis consisting of orthogonal CSFs optimally compresses the wavefunction, and therefore denser Hamiltonians and eigenvectors cannot be avoided. This is why exponentially vanishing CSF coefficients were recently observed for more challenging molecular systems in a fixed CSF basis.[41]

In contrast, our approach allows efficiently simulating challenging molecular systems by simply preparing each of the dominant (mutually nonorthogonal) CSFs

as a linear combination of Slater determinants, where each CSF is initially represented in a different orbital basis but subsequently transformed to one common basis through simple orbital rotation circuits. Our method does not bypass the exponential memory complexity, as this is unavoidable: by choosing a Slater determinant basis (as opposed to a CSF basis) for the many-body Hilbert space, the RHF state has a compact description (one computational basis state) but we are forced to pay an exponential memory cost in the dimensionality of the localized, spin-coupled initial states. For example, the spin-coupled CSF $|\mathcal{O}_{0,0}^{N,1}\rangle$ requires an exponential amount of Slater determinants, whereas once could construct a CSF basis where $|\mathcal{O}_{0,0}^{N,1}\rangle$ would be just one basis vector. However, exponential memory growth is not an issue when using quantum computers, and since preparation of any of the CSFs considered here is efficient through the circuits that we provide, the time complexity of our CSF-based quantum algorithms is polynomial (Section 4.4).

4.3 This work: overview of results

The ability to build a qualitatively accurate reference state for weakly correlated electronic systems through mean-field theory is what has driven the success of most practical approaches to computational chemistry. The key result from Ref. 50 is that a qualitatively accurate wavefunction can be built for systems with strong electron correlation by parametrizing strong correlation effects due to spin coupling. Just as MO theory predicts the Hartree–Fock state to be the dominant electron configuration, our generalized, spin-coupled MO theory predicts the dominant entangled states for strongly spin-coupled systems from chemical intuition and symmetry considerations. This type of correlation is typically referred to as strong or *static* correlation in quantum chemistry.

To obtain quantitatively accurate energy estimates, the remaining (*dynamic* or *weak*) correlation needs to be added to the spin-coupled reference states.[211] This requires powerful computational methods to generate a large amount of additional Slater determinants with small weights. Classical computational methods generally struggle to provide such accuracy due to the unavoidable memory/time complexity

of simultaneously storing and processing spin-coupled as well as delocalized mean-field states, as this requires exponential wavefunction expansions regardless of the choice of single-particle basis.

On the other hand, quantum information processors can efficiently store and transform exponentially large quantum states and therefore this limitation is not present in quantum models of computation.[54] Within the digital quantum circuit model, the key requirement to leverage such reference states is the ability to efficiently prepare them by applying elementary gates on the initial product state of the qubits, typically $|0\rangle \cdots |0\rangle$. In Section 4.4, we show how this can be achieved through quantum circuits of depth $\mathcal{O}(N)$ and using $\mathcal{O}(N^2)$ gates for systems of N spin-coupled electrons, by mapping such states to the well-studied family of Dicke states. This confirms that, in addition to their low Kolmogorov complexity, the quantum circuit complexity of such states is also low owing to their high degree of symmetry.

This approach unlocks the power of quantum algorithms for strongly correlated electronic structure. Through extensive classical simulations of quantum algorithms, we demonstrate how the use of spin-coupled initial states, in particular when represented in different bases, drastically enhances the performance of quantum algorithms for challenging electronic systems. In VQE, it reduces the number of variational parameters, and therefore the number of parametrized quantum gates, required to achieve chemical accuracy (Section 4.5.1). In quantum subspace approaches, it greatly lowers the circuit depth and the number of measurements (Sections 4.5.2 and 4.5.3). In adiabatic state preparation, it increases the accuracy that can be achieved through a given number of steps, effectively speeding up the adiabatic evolution process required to reach a wavefunction of high accuracy (Section 4.5.4).

Finally, spin-coupled reference states reduce the cost of fault-tolerant algorithms based on QPE by orders of magnitude (Section 4.6). Although alternative methods for initial state preparation have been proposed by other authors,[181, 192, 253] fault-tolerant applications relying on QPE will greatly benefit from our approach.

This stems from the fact that, by directly encoding the entanglement due to spin coupling through bespoke quantum circuits, our state preparation circuits are much more efficient than generic, black-box state preparation techniques.

4.4 Quantum circuits for preparation of spin eigenfunctions

To exploit spin-coupled initial states in quantum algorithms, they must be efficiently loaded onto quantum registers. Despite their exponentially-scaling support in the computational basis, the spin-coupled states considered in this work have a highly symmetric and well-defined structure, and the number of distinct coefficients scales linearly with the number of electrons. Crucially, we can exploit this structure to derive quantum circuits that prepare the spin eigenfunctions $|\mathcal{O}^{N,1}_{0,0}\rangle$, which correspond to a superposition of $\binom{N}{N/2}$ determinants, with $\mathcal{O}(N^2)$ rotation and CNOT gates and depth $\mathcal{O}(N)$. We achieve this by connecting the spin eigenfunctions $|\mathcal{O}^{N,1}_{0,0}\rangle$ to Dicke states, which form a different family of entangled wavefunctions.[353] The recursive circuit structure also reduces the cost of preparing linear combinations of spin-coupled states. Furthermore, we consider their implementation on a fault-tolerant architecture, where there is an additional $\log(N)$ overhead in the non-Clifford gate count due to the need for synthesizing rotation gates through Clifford + T gates (Section 4.4.7 and Appendix B.3.4).

We compute the exact CNOT and Toffoli gate counts and find that both are very low in practice for a range of systems sizes relevant to quantum chemistry (Table 4.1 in Section 4.4.7). For example, for $N = 34$, only $\sim 10^3$ gates are required to prepare the spin-coupled CSF $|\mathcal{O}^{34,1}_{0,0}\rangle$, which is a superposition of $L \sim 10^9$ Slater determinants.[2]

In contrast, there exist general algorithms for preparing arbitrary linear combinations of computational basis states (Slater determinants).[56, 83, 181, 192, 253, 357] Such methods do not exploit any particular features of the states that are

[2]This assumes a basis of localized orbitals, and L would be significantly higher using delocalized basis functions such as Hartree–Fock orbitals.

being prepared and their cost scales at best linearly with the number of Slater determinants. Therefore, the scaling is exponential for strongly correlated systems. For the example of $N = 34$, our method is at least eight orders of magnitude more efficient than preparing the same linear combination of computational basis states one determinant at a time using the algorithm presented in Ref. 253 (see Section 4.6.2 for more details on the comparison between our method and others).

Other works have focused on the preparation of certain spin eigenfunctions.[333, 358, 359]. However, these either scale exponentially[359] or only prepare trivial geminal products that couple two-electron singlets[333, 358] (Eqn. 4.31) and therefore do not have the structure necessary to encode the electronic correlation present in many molecular systems such as the ones considered in this work (Section 4.2 and Ref. 50).

The work in Ref. 360 is similar in spirit to ours in that it considers superpositions of states with a varying number of spin-coupled electrons. However, it defines the states in terms of delocalized orbitals and relies on symmetry-broken mean-field solutions. By choosing appropriate localized orbitals and obtaining the spin-coupled states directly from the Clebsch–Gordan coefficients, our work uses spin-pure wavefunctions that are derived from a clear physical picture, and provides a clean mathematical framework for systematic construction of the relevant CSFs.

In Section 4.4.1, we analyze the general form of a family of spin eigenfunctions that captures the entanglement structure of molecular bonds. We discuss the connection to Dicke states in Section 4.4.2. In Section 4.4.3, we review the circuits for Dicke state preparation from Ref. 353. We introduce a more explicit notation for operators and circuits in Section 4.4.4, discuss the circuits for preparation of the states $|\mathcal{O}_{0,0}^{N,1}\rangle$ in Section 4.4.5, the mapping from the representation in the spin Hilbert space to the Fock space in Section 4.4.6, the total state preparation cost in Section 4.4.7, the preparation of linear combinations of CSFs in Section 4.4.9, as well as generalizations to other spin eigenfunctions in Section 4.4.8.

4.4.1 Spin eigenfunctions for bond dissociation

Each spin eigenfunction presented in Section 4.2 includes a set of doubly occupied, delocalized orbitals (typically, Hartree–Fock orbitals), and a set of singly occupied, localized orbitals which capture the strong electron correlation (Eqs. (4.5) and (4.6)). The weakly correlated component can be trivially prepared by applying Pauli-X gates on each of the qubits corresponding to the closed-shell Hartree–Fock orbitals. This is the circuit structure used in most quantum algorithm implementations for preparing the Hartree–Fock state. The entanglement due to spin coupling occurs in the remaining singly-occupied orbitals, which are localized in this work. The state preparation task thus reduces to preparing a particular type of spin eigenfunction $|\mathcal{O}_{S,M}^{N,i}\rangle$ for N strongly correlated electrons in $2N$ spin orbitals (qubits).

The physical model that we propose requires a partitioning of the molecular electronic system into subsystems, and considers the coupling within the subsystem and across the subsystems. For the systems considered in Ref. 50, these subsystems always contain the same number of electrons. Therefore, for a total number of N strongly correlated electrons in N_S subsystems, each subsystem has $n = N/N_S$ electrons.

In the case of diatomic bond breaking at dissociation, N strongly correlated electrons localize onto $n = N/2$ subsystems, where $N = 2$ for H_2, and $N = 6$ for N_2 (Section 4.2). The local coupling within a molecular fragment is ferromagnetic, which corresponds to a state of maximum spin multiplicity for each subsystem. By considering the local coupling first, we can build the global wavefunction from coupling subsystem wavefunctions. We continue to use the example of stretching a triple bond as in N_2 for illustration.

The subsystem state with spin $s = 3/2$ state that ferromagnetically couples three electrons is the quartet in Eq. (4.4). The state $|\mathcal{O}_{0,0}^{6,1}\rangle$ that describes the dissociation of a triple molecular bond into two three-electron quartets is the

product of linear combinations of the $2s+1$ components of each quartet subsystem that forms an overall N-electron state of spin $S = 0, M_S = 0$:

$$\begin{aligned}
|\mathcal{O}^{6,1}_{0,0}\rangle &= \frac{1}{2}\Big[|\tfrac{3}{2}\rangle\,|-\tfrac{3}{2}\rangle - |-\tfrac{3}{2}\rangle\,|\tfrac{3}{2}\rangle - |\tfrac{1}{2}\rangle\,|-\tfrac{1}{2}\rangle + |-\tfrac{1}{2}\rangle\,|\tfrac{1}{2}\rangle \Big] \\
&= \frac{1}{2}\Big[|\alpha\alpha\alpha\beta\beta\beta\rangle - |\beta\beta\beta\alpha\alpha\alpha\rangle \Big] \\
&+ \frac{1}{6}\Big[|\alpha\beta\beta\alpha\alpha\beta\rangle + |\alpha\beta\beta\alpha\beta\alpha\rangle + |\alpha\beta\beta\beta\alpha\alpha\rangle + |\beta\beta\alpha\beta\alpha\alpha\rangle \\
&+ |\beta\beta\alpha\alpha\beta\alpha\rangle + |\beta\beta\alpha\alpha\alpha\beta\rangle + |\beta\alpha\beta\beta\alpha\alpha\rangle + |\beta\alpha\beta\alpha\beta\alpha\rangle \\
&+ |\beta\alpha\beta\alpha\alpha\beta\rangle - |\beta\alpha\alpha\beta\beta\alpha\rangle - |\beta\alpha\alpha\beta\alpha\beta\rangle - |\beta\alpha\alpha\alpha\beta\beta\rangle \\
&- |\alpha\beta\alpha\beta\beta\alpha\rangle - |\alpha\beta\alpha\beta\alpha\beta\rangle - |\alpha\beta\alpha\alpha\beta\beta\rangle - |\alpha\alpha\beta\beta\beta\alpha\rangle \\
&- |\alpha\alpha\beta\beta\alpha\beta\rangle - |\alpha\alpha\beta\alpha\beta\beta\rangle \Big].
\end{aligned} \tag{4.12}$$

In the first row, we have used the notation $|m_{sL}\rangle\,|m_{sR}\rangle$ for compactness to refer to the spin projection quantum number of the *left* and *right* spin subsystems, i.e., we have dropped the spin quantum number $s = 3/2$ and only specify the spin projection quantum numbers of the subsystems. The general form reads:

$$|\mathcal{O}^{N,1}_{0,0}\rangle = \frac{1}{\mathcal{N}} \sum_m \Big[a_m\,|m\rangle\,|-m\rangle + b_m\,|-m\rangle\,|m\rangle\,(1-\delta_{m,0}) \Big], \tag{4.13}$$

where $s = n/2 = N/4$ is the spin of each subsystem, $m = s, s-1, ... \geq 0$ is the absolute magnitude of the spin projection quantum number, $m = |m_s|$, $\mathcal{N} = \sqrt{n+1}$ is a normalizing constant, and the relative signs $a_m, b_m \in \{1, -1\}$ are given from the Clebsch–Gordan coefficients.

4.4.2 Mapping to Dicke states and symmetric states

Although Eq. (4.13) contains $\binom{N}{n} = \binom{N}{N/2}$ Slater determinants, we can efficiently prepare this expansion on a digital quantum computer by mapping it to Dicke states, for which efficient quantum circuits are known. A Dicke state $|D^n_k\rangle$ is defined as an equal-weight superposition of all possible states of an n-qubit system with Hamming weight k[361]

$$|D^n_k\rangle = \binom{n}{k}^{-1/2} \sum_{\mathcal{P}} \mathcal{P}\Big(|1\rangle^k\,|0\rangle^{n-k} \Big), \tag{4.14}$$

where $\mathcal{P}$ denotes any permutation of qubits and the sum runs over all $\binom{n}{k}$ possible permutations. The open-shell spin eigenfunctions discussed in this work can naturally be expressed in terms of Dicke states by rewriting $\alpha \to 1$ and $\beta \to 0$. For example $|D_2^3\rangle = \frac{1}{\sqrt{3}}\Big(|110\rangle + |101\rangle + |011\rangle \Big)$ corresponds to $|\frac{3}{2}, \frac{1}{2}\rangle$ in Eq. (4.4).

This encoding is natural for spin systems, where the N-qubit states exist in the 2^N-dimensional Hilbert space of N spin-1/2 sites, as might be found in site-based model Hamiltonians. However, in quantum chemistry we consider spin-1/2 fermions, and each *site* is a localized spatial orbital that can have four possible occupations. In this context, the appropriate encoding requires one qubit per spin-orbital (Section 4.2.1). Therefore, we first only work with the M qubits corresponding to the spin-up orbitals to prepare the $|\mathcal{O}_{0,0}^{N,1}\rangle$ spin eigenfunction, i.e., we map spin-coupled CSFs $|\mathcal{O}_{0,0}^{N,1}\rangle$ to a superposition of M-qubit basis states

$$|j_\alpha\rangle = \bigotimes_{i=1}^{M} |f_{i_\alpha}\rangle \,. \tag{4.15}$$

Afterwards, we map the occupation from this spin space to the Fock space of $2M$ spin-orbitals (qubits) by populating the β orbitals to recover the correct encoding (Eqs. (4.1) and (4.2)), as detailed in Section 4.4.6.

A symmetric state of n qubits $|\psi_{\mathcal{S}}^n\rangle$ is a state that remains invariant under permutations $\mathcal{P}$ of the symmetry group of n qubits, $\mathcal{S}_n$:

$$|\psi_{\mathcal{S}}^n\rangle = \mathcal{P}\,|\psi_{\mathcal{S}}^n\rangle \,. \tag{4.16}$$

Furthermore, the $n+1$ Dicke states $|D_k^n\rangle$ of an n qubit-system ($k \in \{0, 1, ..., n\}$) are orthonormal and form a complete basis for the $n+1$ symmetric states of n qubits.[362] Therefore, any symmetric state can be expressed as a linear combination of Dicke states:

$$|\psi_{\mathcal{S}}^n\rangle = \sum_{k=0}^{n} c_k\,|D_k^n\rangle \,. \tag{4.17}$$

The states $|\mathcal{O}_{0,0}^{N,1}\rangle$ are constructed from the products of states of highest multiplicity, which are symmetric (Eqs. (4.4), (4.12), and (4.13)). Thus, the construction of $|\mathcal{O}_{0,0}^{N,1}\rangle$ is equivalent to constructing linear combinations of products of Dicke

states with different Hamming weights. For example, the state in Eq. (4.12) can be rewritten as

$$|\mathcal{O}^{6,1}_{0,0}\rangle = \frac{1}{2}\Big(|D^3_3\rangle\,|D^3_0\rangle - |D^3_0\rangle\,|D^3_3\rangle - |D^3_2\rangle\,|D^3_1\rangle + |D^3_1\rangle\,|D^3_2\rangle \Big). \tag{4.18}$$

The general N-electron case (Eq. (4.13)) expressed in the Dicke basis reads:

$$|\mathcal{O}^{N,1}_{0,0}\rangle = \frac{1}{\mathcal{N}} \sum_m \Big[a_m\,|D^n_{s+m}\rangle\,|D^n_{s-m}\rangle + b_m\,|D^n_{s-m}\rangle\,|D^n_{s+m}\rangle\,(1-\delta_{m,0}) \Big]. \tag{4.19}$$

We now discuss the preparation of Dicke states and symmetric states following the approach proposed in Ref. 353.

4.4.3 Preparation of Dicke and symmetric states

Dicke states have long been created on experimental platforms such as ion traps,[363] but their efficient preparation through quantum circuits was a longstanding challenge.[364] Recently, Bärtschi and Eidenbenz provided a protocol for efficiently preparing any Dicke state $|D^n_k\rangle$ or any symmetric state using shallow circuits and without requiring any ancilla qubits.[353] Dicke circuits have been benchmarked in experimental implementations on quantum hardware.[365, 366] Dicke states have direct applications in combinatorial optimization, where they can be used as initial states of the Quantum Alternative Operator Ansatz algorithm,[367] but their connection to spin eigenfunctions has previously not been explored. In quantum simulation, Dicke circuits have been considered for preparing antisymmetrized geminal power states.[368] Here, we use the Dicke state preparation method from Ref. 353 as a circuit primitive to prepare spin eigenfunctions. We summarize the relevant results below.

1. Efficient preparation of Dicke states: The Dicke state $|D^n_k\rangle$, for $k \leq n$, can be prepared deterministically by applying a unitary $U_{n,k}$ on the input state $|0\rangle^{n-k}\,|1\rangle^k$. A decomposition of $U_{n,k}$ in terms of single-qubit and CNOT gates is provided. The resulting quantum circuit has depth $\mathcal{O}(n)$ and uses $\mathcal{O}(kn)$ gates.

2. Efficient preparation of symmetric states: Given any input state of the same number of qubits n but with a smaller Hamming weight, $|0\rangle^{n-l}|1\rangle^{l}$ with $l \leq k$, applying the operator $U_{n,k}$ onto said state outputs the Dicke state $|D_l^n\rangle$. Consequently, any superposition of Dicke states of equal n and different l can be generated by applying the unitary $U_{n,n}$ on the appropriate superposition of input states:

$$|\psi\rangle = \sum_l c_l |D_l^n\rangle = U_{n,n} \sum_l c_l |0\rangle^{n-l} |1\rangle^{l}. \tag{4.20}$$

Since the $n+1$ Dicke states $|D_l^n\rangle$ form an orthonormal basis of the fully symmetric subspace of all n-qubit states (Eq. (4.17)), the circuits for unitaries $U_{n,n}$ can be used to prepare any symmetric state $|\psi_S^n\rangle$, provided that they act on the appropriate input states. The input state $\sum_l c_l |\psi_l^n\rangle$ can be prepared by a quantum circuit of depth and gate count $\mathcal{O}(n)$. The overall complexity for preparation of any symmetric state is thus depth $\mathcal{O}(n)$ and $\mathcal{O}(n^2)$ gates. These circuits can be implemented directly on a qubit architecture with linear connectivity, which is beneficial e.g. for applications on superconducting quantum hardware.

The key insight that enables this efficient protocol is the following recursive formula:

$$|D_k^n\rangle = \sqrt{\frac{k}{n}} |D_{n-1}^{k-1}\rangle |1\rangle + \sqrt{\frac{n-k}{n}} |D_k^{n-1}\rangle |0\rangle. \tag{4.21}$$

Given a unitary operator $U_{n,k}$ that prepares the Dicke state $|D_k^n\rangle$ when acting on the input state $|0\rangle^{n-k}|1\rangle^{k}$, one can recursively decompose the unitary operator $U_{n,k}$ into products of two types of elementary operators, $M_{l,l-1}$ and $M_{l,k}$ (Lemma 2 in Ref. 353):

$$U_{n,k} = \prod_{l=2}^{k} M_{l,l-1} \prod_{l=k+1}^{n} M_{l,k}. \tag{4.22}$$

For preparation of symmetric states we set $k = n$ and the unitary reduces to

$$U_{n,n} = \prod_{l=2}^{n} M_{l,l-1} := S_n. \tag{4.23}$$

Here, the operators $M_{l,l-1}$ can be constructed from products of two primitive building blocks: one two-qubit gate block (Fig. 4.2b) followed by $l-2$ three-qubit gate blocks (Fig. 4.2c). This enables the preparation of an n-qubit Dicke

state in terms of subcircuits acting on $l < n$ qubits. An example circuit for S_4 is shown in Fig. 4.2. We discuss the decomposition of such circuits into gates consisting of only CNOT and single-qubit rotations in Appendix B.3, to show that the number of CNOT gates required is,

$$C_{S_n} = \frac{5}{2}n^2 + \frac{9}{2}n + 2 \tag{4.24}$$

as derived in Eq. (B.11).

One attractive feature of this protocol is the fact that only one unitary S_n needs to be applied to generate a linear combination of multiple Dicke states, and therefore no special techniques such as Linear Combination of Unitaries [56, 174] are required for state preparation. Instead, the desired superposition of Dicke states is achieved by preparing a linear combination with the appropriate weights in the relatively simple input state. This is much more efficient than applying state preparation unitaries controlled by ancilla qubits.[56, 192, 226] By slightly modifying the circuit for input state preparation, we can efficiently use this circuit structure to prepare the CSFs relevant for chemical bond breaking. For reference, the cost for up to $n = 6$ is shown in Table B.1 in Appendix B.3.

4.4.4 Notation for operators and circuits

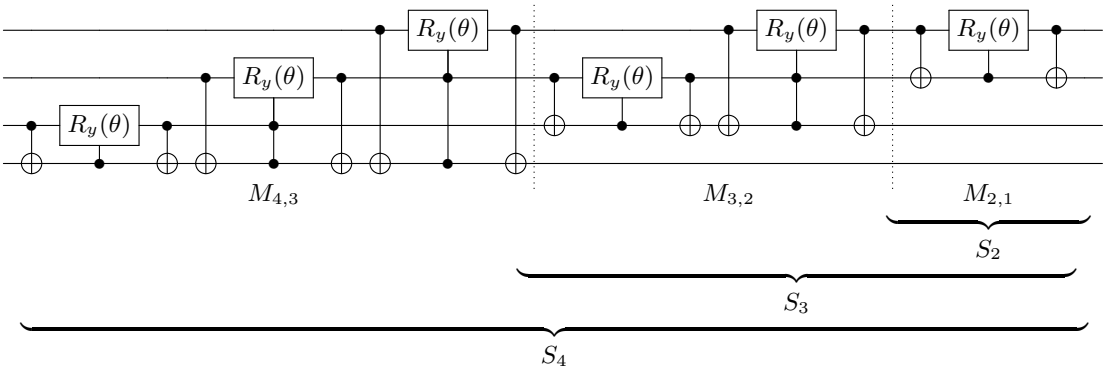
In the above, the qubits on which $M_{l,k}$ acts are specified in the circuit diagrams (Fig. 4.2). For further clarity, we define the following more explicit notation for all quantum circuits below.

Superscripts denote the qubits on which gates act, e.g. X^i is a Pauli-X gate on qubit i, where 1 is the top qubit in the circuit diagram and its state is specified by the leftmost bit in the ket: $|j\rangle = |j_1, j_2, ...\rangle$. We use $U^{i,j}_{j-i+1}$ to denote an operator acting on $(j - i + 1)$ qubits with indices $i, i + 1, ..., j - 1, j$, e.g. $S_n^{1,n}$ means S_n acting on qubits 1 through n. For two unitary operators U and V, the product is UV (unlike for states where products of kets are tensor products, see Section 4.2.1), therefore the tensor product of two unitaries must be written

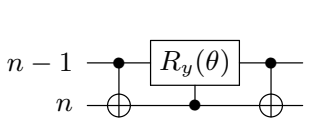
explicitly as $U \otimes V$. Tensor products with identity are implied whenever the unitaries act on a subset of the qubits.

For two-qubit controlled gates, we use the following notation: $CU^{c,t}$ applies the operator U on target qubit t controlled by qubit c, i.e. $CU^{c,t} = P_0^c \otimes I^t + P_1^c \otimes U^t$, where the projectors are $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$. On circuit diagrams, this corresponds to a full black dot on qubit i. We use $\overline{CU}^{c,t}$ for a gate that controls application of the operator U on target qubit t such that U is applied if c is in the state $|0\rangle$, i.e. $\overline{CU}^{c,t} = P_0^c \otimes U^t + P_1^c \otimes I^t$. This corresponds to a white dot on the control qubit c in the circuit diagram. We often use the CNOT gate with $\overline{CX}^{c,t} = X^c CX^{c,t} X^c$.

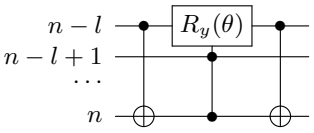
4.4.5 Preparation of singlet states with locally ferromagnetic coupling



(a) Circuit for $S_4 = S_3 M_{4,3} = (S_2 M_{3,2}) M_{4,3} = M_{2,1} M_{3,2} M_{4,3}$.



(b) Two-qubit gate.



(c) Three-qubit gate.

Figure 4.2: (a) Quantum circuit proposed in Ref. 353 to implement the symmetric state preparation unitary $S_4 = M_{2,1} M_{3,2} M_{4,3}$ (see Eq. (4.23)). (b), (c) Gate blocks in Dicke-state preparation circuits such as (a), also from Ref. 353. The angles are $\theta = 2\arccos\left(\sqrt{1/n}\right)$ for the two-qubit gate blocks and $\theta = 2\arccos\left(\sqrt{l/n}\right)$ for the three-qubit gate blocks.

Equipped with the circuits for Dicke state preparation, we can rewrite the CSF for bond breaking (Eq. (4.19)) in terms of the corresponding unitaries acting on

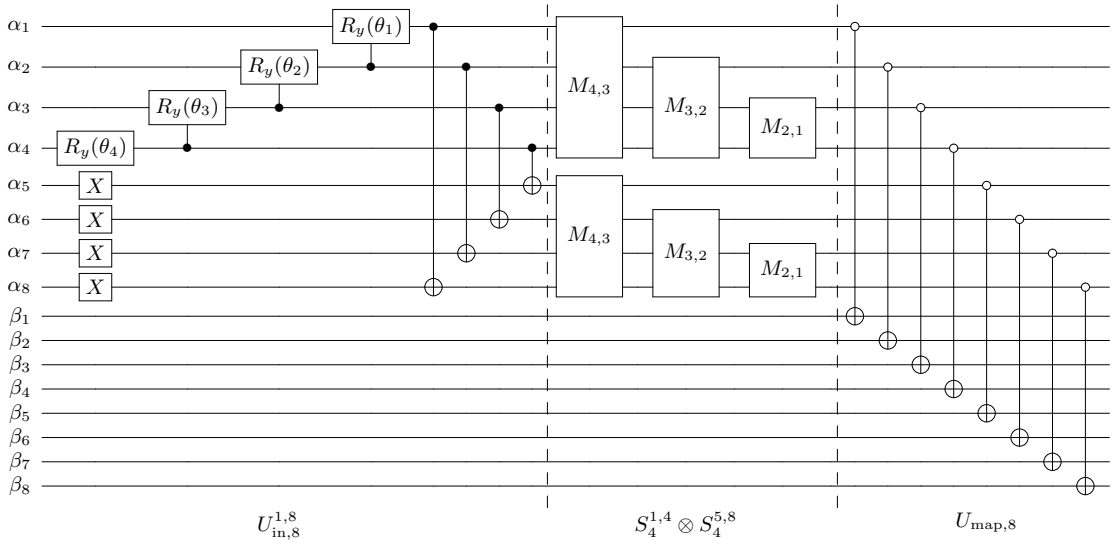


Figure 4.3: Quantum circuit that implements the unitary V_8 to prepare the spin-coupled state $|\mathcal{O}_{0,0}^{8,1}\rangle = V_8 |00\rangle^8$ (Eq. (4.25)). It consists of three parts, delimited by the dashed vertical lines: preparation of the input state (left), preparation of the spin eigenfunction through Dicke state preparation circuits (center, see Fig. 4.2), and mapping from spin to Fock space occupation (right). Qubits labelled α_i and β_i correspond to the up and down spin-orbital with the same spatial orbital ϕ_i.

carefully chosen input states:

For example, to prepare the spin-coupled state $|\mathcal{O}_{0,0}^{8,1}\rangle$, we only need to apply the operator S_4 twice, in parallel, on an input state which is a linear combination of five states:

$$\begin{aligned}|\mathcal{O}_{0,0}^{8,1}\rangle =& \frac{1}{\sqrt{5}}\Big(|D_4^4\rangle\,|D_0^4\rangle + |D_0^4\rangle\,|D_4^4\rangle - |D_3^4\rangle\,|D_1^4\rangle - |D_1^4\rangle\,|D_3^4\rangle + |D_2^4\rangle\,|D_2^4\rangle\Big) \\ = S_4^{1,4} \otimes S_4^{5,8} & \frac{1}{\sqrt{5}}\Big(|1111000\rangle + |00001111\rangle - |01110001\rangle - |00010111\rangle + |00110011\rangle\Big).\end{aligned} \tag{4.25}$$

To prepare input states of the form

$$\sum_l c_l\,|0\rangle^{n-l}\,|1\rangle^l\,, \tag{4.26}$$

one could simply use the circuits in Fig. 4 of Ref. 353, which are composed of ladders of controlled-R_y gates applied on the initial state $|0\rangle^n$. However, the spin eigenfunctions considered in this work are not symmetric and require parallel application of two circuits for S_n, with input states of the form in Eq. (4.25).

We define the input state $|v_N\rangle$ for preparation of the CSF $|\mathcal{O}_{0,0}^{N,1}\rangle = S_n^{1,n} \otimes S_n^{n+1,2n}\,|v_N\rangle$ as $|v_N\rangle = |L\rangle\,|R\rangle = \sum_i c_i\,|L_i\rangle\,|R_i\rangle$, where the *left* and *right* states $|L\rangle$

and $|R\rangle$ correspond to qubits $\{1, 2, ..n\}$ and qubits $\{n+1, n+2, ..., 2n\}$. With $n = N/2$, the left and right qubits correspond to dividing the system into two molecular subsystems (Section 4.2). The following circuit produces the appropriate initial state:

1. Flip the last n qubits by applying X-gates to obtain the state $\otimes_n^{2n} X^n |0\rangle^{2n} = |0\rangle^n |1\rangle^n$.
2. Apply $\left(\prod_{i=1}^{n-1} CR_y^{i+1,i}(\theta_i)\right) R_y^n(\theta_n)$ on the first n qubits to produce the state $\sum_i c_i |L_i\rangle |1\rangle^n$.
3. Apply a CNOT *accordion* $\prod_{i=0}^{N-1} CX^{1+i,N-i}$ that flips the state of the *right* qubits controlled by the occupation of the *left* qubits to obtain the input state $|v_N\rangle = |L\rangle |R\rangle = \sum_i c_i |L_i\rangle |R_i\rangle$.

We define $U_{\text{in},N}^{1,N}$ as the unitary for preparation of the input state for N spin-$\frac{1}{2}$ particles (spin-coupled electrons): $|v_N\rangle = U_{\text{in},N}^{1,N} |0\rangle^N$. An example for $N = 8$ is shown in Fig. 4.3 (left panel).

4.4.6 Mapping from spin to Fock space

If one desires to prepare spin eigenfunctions for pure spin systems, e.g. for quantum simulation of Heisenberg or Ising spin models, this circuit above completes the entire state preparation procedure. However, here we are interested in systems of electrons (spin-$\frac{1}{2}$ fermions). In this context, the $|\alpha\rangle$ and $|\beta\rangle$ states in the expansion of the spin eigenfunctions (Appendix B.1) represent the spin projection of an electron occupying a localized spatial orbital (rather than of a spin orbital or qubit), where we have restricted each orbital to be occupied by exactly one electron. For a general many-electron wavefunction, any spatial orbital can be singly occupied, unoccupied or doubly occupied (Eq. (4.2)). We must therefore map the state from the spin space of N spins, with dimension 2^N, to the Fock space of N electrons occupying M spatial orbitals, with dimension $\binom{2M}{N}$. For a minimal basis set and N spin-coupled valence electrons in the fully dissociated limit, we have $M = N$ spatial orbitals (Section 4.2).

To account for this, we have so far chosen to work entirely on the qubits corresponding to α spin orbitals to prepare the superposition of Dicke states, leaving the qubits corresponding to β spin-orbitals in the $|0\rangle$ state. After the spin eigenfunction is prepared through Dicke circuits, we map the occupation as $|\alpha\rangle = |10\rangle \rightarrow |10\rangle$ and $|\beta\rangle = |00\rangle \rightarrow |01\rangle$, where the first qubit corresponds to an α spin-orbital and the second qubit represents a β spin-orbital. This can be achieved by a single layer of CNOT gates that flips the state of the i-th β spin-orbital if the state of the i-th α spin-orbital is $|0\rangle$: $U_{\mathrm{map},N} = \bigotimes_{i=1}^{N} \overline{CX}^{i_\alpha, i_\beta}$.

The full operator V_N for preparing the N-electron CSF in two molecular subsystems ($n = N/2$) is thus

$$|\mathcal{O}_{0,0}^{N,1}\rangle = V_N \, |00\rangle^N = U_{\mathrm{map},N}\Big(S_n^{1,n} \otimes S_n^{n,n+1}\Big) U_{\mathrm{in},N}^{1,N} \, |00\rangle^N . \tag{4.27}$$

An example for $N = 8$ is shown in Fig. 4.3.

4.4.7 Total circuit cost and examples

The total number of CNOTs assuming all-to-all qubit connectivity is

$$C_{\mathrm{tot}}^{\mathrm{all}} = \frac{5}{4}N^2 + 2N + 2 \tag{4.28}$$

(see Eq. (B.14)). In Appendix B.3, we provide a detailed derivation and discuss the case with linear or planar qubit connectivity. The latter might be of interest for experiments on near-term superconducting hardware.[295] The depth of the circuits with all-to-all and planar connectivity is $\mathcal{O}(N)$. For linear connectivity, we obtain a depth scaling as $\mathcal{O}(N^2)$.

The CNOT counts are the most important metric for implementations on noisy quantum hardware without error correction because two-qubit gates are usually the source of the largest errors.[256] In Appendix B.3.4, we discuss a fault-tolerant implementation of the state preparation circuit V_N using a Clifford + Toffoli gate set. In short, the number of non-Clifford (Toffoli) gates depends on the number of rotations, R, and the accuracy required for synthesizing each continuous rotation gate through a discrete gate set. To bound the error in the state preparation

unitary as $||V_N - \tilde{V}_N|| \leq \epsilon$, where ϵ is the target accuracy, the number of Toffoli gates required is

$$T = R\Big[0.2875\lceil \log(R/\epsilon)\rceil + 4.6\Big]. \tag{4.29}$$

Here, $R = \frac{1}{4}N^2$ is the total number of rotation gates in the circuit that implements $V_N : |\mathcal{O}^{N,1}_{0,0}\rangle = V_N |00\rangle^N$ (including R_y, CR_y, and CCR_y gates, see Eq. (B.25)). We use log for logarithm to the base 2 throughout this manuscript.

Table 4.1 shows the total cost of preparing a spin eigenfunction $|\mathcal{O}^{N,1}_{0,0}\rangle$, and the number of determinants (computational basis states) in the expansion of $|\mathcal{O}^{N,1}_{0,0}\rangle$ in a product basis. We have chosen $\epsilon = 10^{-7}$ for the accuracy of initial state preparation in the fault-tolerant setting. It is evident that the preparation of these states is very efficient in practice, due to the $\mathcal{O}(N^2)$ scaling and the relatively small constant factors.

H_2, H_2O, N_2, Cr_2 are examples of molecules for which, in a minimal basis of N spatial orbitals, a single spin eigenfunction describes the exact ground state at dissociation. For $N = 2$, $|\mathcal{O}^{2,1}_{0,0}\rangle = |\mathcal{O}^{2,2}_{0,0}\rangle$ and we can therefore directly use the more compact circuit for preparation of $|\mathcal{O}^{N,2}_{0,0}\rangle$, which only contains Clifford gates (see Section 4.4.8). For the FeS systems, we hypothesize that the gate counts are within the correct order of magnitude because the number of open-shell electrons N is known from existing literature,[17, 18, 112] and the number of determinants as well as the cost of CSF preparation depends mainly on N. However, the estimates are by no means exact since the correct spin eigenfunctions are currently unknown (see Section 4.6.2 for details).

Finally, depending on the application, the total cost of the reference state preparation might require additional gates to rotate the single-particle basis, which can be done in depth $\mathcal{O}(N)$ and at worst $\mathcal{O}(N^2)$ gates, although often $\mathcal{O}(N)$ gates suffice. We exclude these from the analysis in this section since they depend on the system considered, the choice of basis, as well as the quantum algorithm in which they are being used. In Appendix B.2, we present the circuits for basis rotations, their cost, and optimal implementations for the molecules considered here.

Table 4.1: Number of gates required for preparation of the state $|\mathcal{O}_{0,0}^{N,1}\rangle$ with N spin-coupled sites through our state preparation circuits. In the quantum chemistry context discussed here and in Ref. 50, N is the number of open-shell electrons occupying N localized orbitals. The total CNOT cost of preparing each spin eigenfunction is given by $C_{\text{tot}} = C_{\text{in}} + 2C_{S_{N/2}} + C_{\text{map}}$, where the superscript indicates different hardware connectivities. The number of determinants (computational basis states) with non-zero coefficients in the state $|\mathcal{O}_{0,0}^{N,1}\rangle$ is L. The number of Toffoli gates in a fault-tolerant implementation is given by T (see Appendix B.3.4). We also report b, the number of digits required for the binary representation of the rotation angle.

N	$C_{\text{tot}}^{\text{all}}$	$C_{\text{tot}}^{\text{pla}}$	$C_{\text{tot}}^{\text{lin}}$	b	T	L	System
2	3	3	5	-	0	2	H_2
4	14	19	63	13	49	6	H_2O
6	35	49	163	14	114	20	N_2
8	66	93	309	14	203	70	
10	107	151	501	14	317	252	Fe_2S_2 (est.)
12	158	223	739	15	477	924	Cr_2
18	371	523	1729	15	1072	48620	Fe_4S_4 (est.)
34	1379	1939	6393	16	3989	2.3×10^9	FeMoCo (est.)

4.4.8 Other spin eigenfunctions

We have so far discussed the preparation of CSFs for the coupling of two subsystems with maximum spin, corresponding to fully symmetric spin functions, into an overall singlet. While this spin coupling pattern represents the dominant entanglement structure of electronic eigenstates in typical covalent bonds,[50] which are ubiquitous in chemistry, other systems might require different types of spin eigenfunctions. For completeness, here we present circuits required to prepare the remaining CSFs in Ref. 50.

The state of linear and square (cyclic) H_4 at stretched geometries can be expressed as a superposition of two states,[50] each of the form in Eq. (B.2):

$$|\mathcal{O}_{0,0}^{4,2}\rangle = \left[\frac{1}{\sqrt{2}}(|\alpha\beta\rangle - |\beta\alpha\rangle)\right]^2. \tag{4.30}$$

The general form is a tensor product of singlets (antisymmetric functions) of $N/2$ two-spin systems

$$|\mathcal{O}_{0,0}^{N,2}\rangle = \left[\frac{1}{\sqrt{2}}(|\alpha\beta\rangle - |\beta\alpha\rangle)\right]^{\frac{N}{2}}. \tag{4.31}$$

This state, also considered in Refs. 358 and 333 is a product of Bell states and can be implemented trivially in constant depth and using only one CNOT gate per two-electron singlet $\left[\frac{1}{\sqrt{2}}(|10\rangle - |01\rangle)\right]^{N/2} = \bigotimes_{i=1}^{N/2} \overline{CX}^{2i,2i-1} R_y^{2i}(-\pi/2)\,|00\rangle^{N/2}$. The total CNOT cost for preparing $|\mathcal{O}_{0,0}^{N,2}\rangle$, including the cost of the spin-to-Fock-space mapping U_{map}, is given by (Appendix B.3.6):

$$C_{\text{tot}}^{\text{all}} = \frac{3}{2}N. \tag{4.32}$$

Since $R_y(-\pi/2) = HX$, this rotation can be implemented without any non-Clifford (e.g., Toffoli) gates. On a fault-tolerant architecture, the preparation of such states has negligible cost (Appendix B.3.4).

One might wonder if the circuits can be extended to prepare any type of CSF. The efficient preparation of arbitrary spin eigenfunctions remains an open problem, as existing approaches with rigorous performance guarantees require exponentially many gates.[359] In general, highly symmetric states allow efficient state preparation procedures, and one might expect that the complexity of the state preparation circuit increases with the number of distinct coefficients in the CSF. We hypothesize that most chemical systems contain enough structure that their spin coupling is given by paths that have a significant amount of symmetry.

Crucially, for applications in electronic structure, the choice of CSFs greatly depends on the choice of the single-particle basis as well as the method used to construct the many-body CSFs.[189] In particular, rather than working in a fixed orthonormal basis, it is more efficient to use a bespoke basis for each physically-relevant CSF. The choice of basis should appropriately reflect the physical spin coupling pattern, e.g., by partitioning the system into molecular fragments and separately parametrizing the correlations due to local vs. global couplings, as is done in this work. The extension of our state preparation protocols to more systems is an interesting direction for future work.

4.4.9 Linear combinations of spin eigenfunctions

When multiconfigurational initial states are required for VQE or QPE, such as for molecular bonds at intermediate bond lengths or for transition metal clusters, it is useful to prepare a linear combination of CSFs $|\Phi_{\mathrm{LC}}\rangle = \sum_j c_j |\Phi_j\rangle$ for $j \in \{0, 2, ..., N\}$. For example, $N = 0, 2, 4, 6$ for N_2 (Eq. (4.11)). This can be done by controlling the state preparation unitaries $V_j : |\mathcal{O}_{0,0}^{j,1}\rangle = V_j |00\rangle^N$ from an ancilla register, for each j.

The encoding used for the ancilla register can take various forms that trade off circuit depth with space (qubits). One could either work with one ancilla qubit with a higher gate cost, as shown in Ref. 181, use a one-hot encoding with L ancilla qubits for L CSFs, or use a compressed register of $\log(L)$ ancilla qubits as is common in general Linear Combination of Unitaries approaches.[56] The optimal encoding of the ancilla register depends on the particular hardware and molecule considered.

The gate overhead for controlling the unitaries V_j scales linearly with N, which is lower than the cost of the circuits for V_j themselves (Appendix B.3.5). Exploiting the recursive structure of the Dicke circuits enables an additional reduction of the cost of implementing $|\Phi_{\mathrm{LC}}\rangle$. Specifically, from Eq. (4.23) it follows that

$$S_n = M_{l,l-1} S_{n-1}. \tag{4.33}$$

This implies that the circuit for V_{j-2} can be reused as a subcircuit for V_j rather than being applied twice.

To summarize, the cost for preparing linear combinations of CSFs is only slightly higher than the cost of preparing a single CSF, where the exact overhead depends on the particular case considered.

4.5 Heuristic quantum algorithms

While the algorithmic error of fault-tolerant quantum algorithms can be bound analytically,[67, 72, 104, 105, 298] the factors determining the performance of heuristic quantum algorithms such as VQE, quantum subspace diagonalization,

and adiabatic state preparation are less well understood. Regardless, the numerical evidence accumulated so far suggests that the accuracy and efficiency of heuristic algorithms strongly depends on the quality of the initial state. In this section, we show that spin-coupled reference states have the correct features to provide this increase in accuracy and efficiency.

4.5.1 Variational quantum eigensolver

We first consider the VQE approach, which avoids deep quantum circuits by combining quantum and classical processors in a hybrid algorithm. A quantum circuit with parametrized gates prepares the electronic wavefunction on a quantum computer, and a classical computer variationally optimizes the parameters. This is repeated in an iterative loop until convergence is reached.

At a fundamental level, VQE exploits the variational principle to rotate the initial qubit state $|0\cdots0\rangle$ to the true electronic ground state. Naturally, the more complex the eigenstate, the more parametrized quantum gates are required to accurately approximate it. In practice, finding a compact and accurate ansatz is difficult.[52] In particular, strong electron correlation exacerbates the challenges of VQE, which reduces the potential of this quantum approach to outperform classical algorithms.[184, 185] For example, benchmark calculations show that many *ansätze* accurately describe weakly-correlated wavefunctions of molecular bonds at equilibrium geometries, but their accuracy deteriorates significantly at strongly-correlated stretched geometries.[184]

To reduce the computational cost, VQE implementations use a classical heuristic to obtain an approximate initial state, $|\Phi\rangle$, usually the Hartree–Fock state. Here, we show that using a spin-coupled initial state can greatly reduce the cost required to achieve high-accuracy approximations of strongly correlated ground states.

Formally, the quantum computer is first initialized to the reference state

$$|\Phi\rangle = U_{\text{ref}}\,|0\rangle^{2M}\,. \tag{4.34}$$

Here, $2M$ is the number of spin-orbitals (qubits) and U_{ref} is the unitary transformation required to prepare the reference state. To further correlate the reference

state, a unitary transformation $U_{\text{ansatz}}(\boldsymbol{\theta})$ is applied to it. This is implemented as a sequence of parametrized quantum gates:

$$|\Psi(\boldsymbol{\theta})\rangle = U_{\text{ansatz}}(\boldsymbol{\theta})\,|\Phi\rangle = \prod_i U_i(\theta_i)\,|\Phi\rangle\,. \tag{4.35}$$

The particular choice and ordering of elementary unitary transformations U_i defines a variational *ansatz*, which determines the wavefunctions that can be generated by the quantum circuit through variation of the rotation parameters $\boldsymbol{\theta}$, starting from the reference state $|\Phi\rangle$. The variational energy

$$E(\boldsymbol{\theta}) = \langle\Phi|U^{\dagger}_{\text{ansatz}}(\boldsymbol{\theta})\,\hat{H}\,U_{\text{ansatz}}(\boldsymbol{\theta})|\Phi\rangle \tag{4.36}$$

is estimated by measuring Hamiltonian expectation values on each qubit followed by operator averaging.[92] These estimates are then passed to a classical computer, which uses traditional optimization algorithms to proposes updates to $\boldsymbol{\theta}$ that minimise $E(\boldsymbol{\theta})$.

Practical VQE implementations face substantial challenges. Firstly, although compact and accurate electronic states can in principle be prepared with short circuit depths,[52, 261, 262, 347, 369] finding these *ansätze* currently relies on computationally expensive strategies that optimize the gate sequences.[52, 261] Secondly, the numerical optimization appears to be fundamentally challenging. The non-linearity of the *ansatz* means that $E(\boldsymbol{\theta})$ is highly non-convex and the optimization is prone to get stuck in local minima.[258] Finally, estimating the energy may incur prohibitive measurement costs.[42, 342]

Strong electron correlation greatly increases the difficulties of VQE.[184, 185] This poor performance can be attributed to the use of a mean-field RHF initial state, which is uncorrelated and therefore burdens the ansatz circuit U_{ansatz} to recover a large amount of electron correlation. Instead, we propose to use a spin-coupled wavefunction as the initial state, which has a greater overlap with the true ground state. This greatly reduces the number of parametrized ansatz operators required to reach quantitative accuracy.

We demonstrate this improvement by performing classical numerical simulations of VQE with the quantum-number-preserving (QNP) *ansatz*,[344] which uses a layered circuit structure that can be systematically improved to the exact result by increasing the number of repeating layers k. Each layer includes a series of spin-adapted one-body rotations and paired two-body rotation operators acting between neighboring spatial orbitals, with the arrangement of these operators depicted in Fig. 4.4. This choice of operators preserves the spin quantum numbers of the initial state, which ensures that approximate wavefunctions are exact spin eigenfunctions. Ref. 52 proves that this type of *ansatz* is universal, i.e. it can yield exact wavefunctions within the spin- and particle-number-preserving subspace. In practice, it is unclear how many layers are required to achieve sufficient accuracy.

We consider two versions of VQE: a *single reference* approach in which a single unitary transformation is applied on one sole reference state composed of a linear combination of CSFs, and a *multireference* approach that applies a different unitary transformation on each individual CSF.

Single-reference VQE with variable linear combination

Our first approach is to apply the QNP *ansatz* to an initial state defined as the variable linear combination

$$|\Phi\rangle = \sum_{I=1}^{L} C_I \, |\Phi_I\rangle \, . \tag{4.37}$$

We optimize both the QNP rotation angles as well as the linear coefficients C_I in the initial state definition, since the latter are also parametrized in terms of single-qubit rotations through the circuits for preparing linear combinations of CSFs (Section 4.4.9). This allows the linear combination of spin-coupled states to relax in the presence of the correlation generated by the QNP circuit.

For the strongly correlated N_2 binding curve, the variable linear combination of CSFs significantly reduces the energetic error compared to the RHF initial state at large bond lengths (Fig. 4.5), reaching chemical accuracy at $R > 2.5\,\text{Å}$ even with a single layer (i.e. $k = 1$). This result is expected since the spin-coupled wavefunction

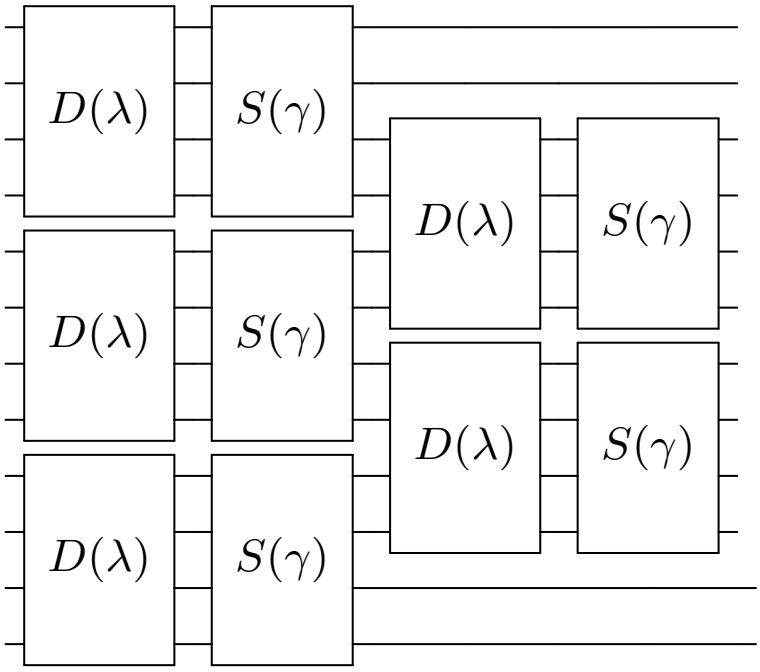


Figure 4.4: One layer of the QNP *ansatz* ($k = 1$) includes an alternating series of paired two-body operators $D(\lambda)$ and spin-preserving one-body operators $S(\gamma)$. These are applied between neighboring spatial orbitals, each acting on four qubits, as described in Ref. 344. The fermionic excitation operators $S(\gamma)$ and $D(\lambda)$ can be decomposed into Givens rotations that act non-trivially on a two-qubit subspace.[344, 345]

becomes exact in the dissociation limit (Section 4.2.2). Each layer corresponds to ten variational parameters (five for each operator in Fig. 4.4), and there is a fixed constant number of three parameters required to prepare the linear combination of four CSFs (one is fixed through normalization). Even around the equilibrium region, where the RHF configuration dominates the linear combination of CSFs (Fig. 4.1b), the spin-coupled wavefunction improves the accuracy by around half an order of magnitude for each value of k. These results demonstrate that using a linear combination of CSFs to define the initial state reduces the number of parameters and unitary operators required to reach quantitative accuracy with a VQE *ansatz*.

Multireference ansatz with a nonorthogonal variational quantum eigensolver

An alternative strategy is to uniquely correlate each reference state. This is particularly useful for multiconfigurational systems where more than one CSF dominates

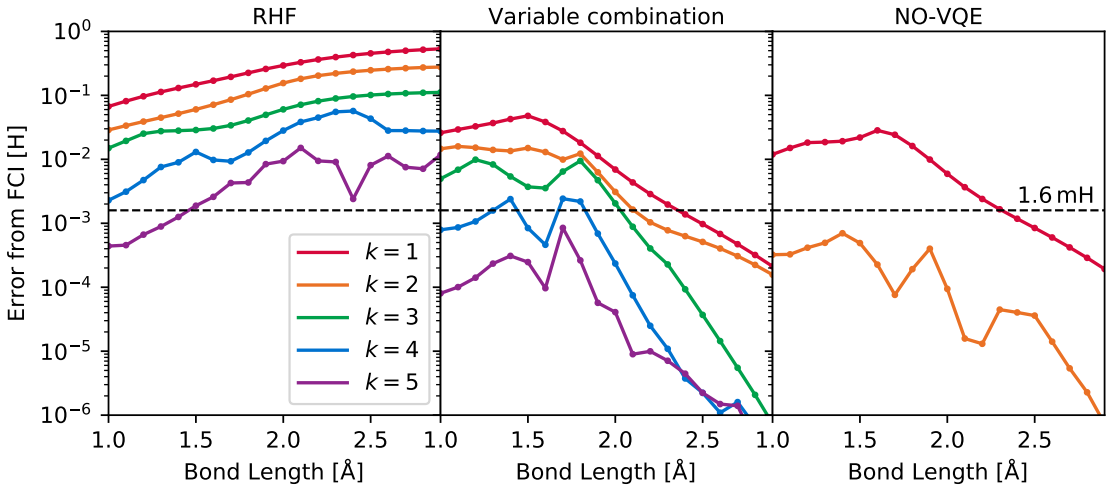


Figure 4.5: Energy errors from the QNP *ansatz* for N_2 (STO-3G) using either a fixed RHF reference state or a variable linear combination of CSFs. The linear combination of CSFs significantly improves the accuracy in the strongly-correlated dissociation limit. The NO-VQE approach applies a bespoke QNP circuit to each reference state prior to constructing the linear combination and reaches chemical accuracy using shallower circuits with $k = 2$.

the wavefunction. Since each CSF is a unique vector in the Hilbert space, separately rotating each vector enables multiple regions of the Hilbert space to be explored in a targetted manner. Although this idea has been explored in other works e.g. using unrestricted Hartree–Fock determinants,[346] our spin-coupled wavefunction approach has the advantage that significant entanglement is already included in the reference state. Achieving the same result through single-determinant reference states would require exponentially many initial states. Furthermore, since we strictly use CSFs, any spin-preserving ansatz is guaranteed to produce spin-pure wavefunctions, without relying on approximately recovering the correct spin state after symmetry breaking.

On a quantum circuit implementation, the multireference approach avoids the preparation of a linear combination of different CSFs on the quantum register and instead relies on the measurement of Hamiltonian and overlap matrix elements between the individually correlated reference states. Furthermore, the increased number of variational parameters provides a very flexible ansatz using limited circuit depth.

Formally, a unique set of unitary operations is applied to the each reference state before constructing the linear combination, giving

$$|\Psi(\boldsymbol{\Theta}, \boldsymbol{C})\rangle = \sum_{I=1}^{L} C_I \left(\prod_i U_{Ii}(\theta_{Ii}) \, |\Phi_I\rangle \right), \tag{4.38}$$

where $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_L)$. Each reference state is correlated with a bespoke set of quantum gates that each have a unique variational parameter, and the gate parameters $\boldsymbol{\Theta}$ are optimized simultaneously with the linear combination $\boldsymbol{C}$. The correlated basis states in this expansion will generally not be orthogonal, and we must therefore obtain the overlap matrix elements to guide the optimization as in any nonorthogonal quantum eigensolver.[97, 346] Here, we run classical simulations of this nonorthogonal VQE (NO-VQE) algorithm. We describe details of our numerical implementation in Appendix B.4.

The multireference NO-VQE provides chemically accurate energies across the full binding curve of N_2 (STO-3G) using only two layers of the QNP *ansatz* for each reference state (Fig. 4.5, right panel). In contrast, five layers are required to reach an equivalent accuracy when applied on a single reference state (Fig. 4.5, central panel), as per the approach described in Section 4.5.1. This corresponds to a reduction in circuit depth by a factor of 2 to 3, which is highly desirable on noisy quantum hardware. Therefore, although the NO-VQE approach introduces more parameters, it achieves greater accuracy with a shallower circuit than applying a single unitary to the variable linear combination of CSFs. The efficient exploration of the Hilbert space achieved by starting from different spin-coupled wavefunctions makes this a natural application of our work in multiconfigurational systems.

4.5.2 Quantum subspace diagonalization through real-time evolution

Although VQE might be suited for noisy hardware because typical ansatz circuits have low gate counts and depths, severe implementation challenges in the optimization and measurement make its practical use questionable.[42, 44, 258]

In this section, we consider an approach that avoids non-linear optimization and instead requires real-time-evolution of the reference states under the Hamiltonian.

This falls within the broader category of quantum subspace diagonalization (QSD) methods (one could also consider the nonorthogonal VQE from Section 4.5.1 as a QSD approach). QSD methods are uniquely well-suited for exploiting different reference states like the ones presented in this work, as they can explore the Hilbert space from multiple directions in parallel, through transformations applied separately on each reference state. The ability to rotate each reference state individually relaxes the accuracy requirement for each ansatz circuit, as the classical diagonalization step gives increased variational freedom. Here, we will show how using spin-coupled reference states can greatly improve the performance of QSD methods for both ground and excited state calculations.

Background

Quantum subspace diagonalization methods have emerged as a promising class of hybrid quantum-classical algorithms for computing Hamiltonian eigenvalues. Starting from a reference state $|\Phi\rangle$, the quantum device is used to generate a set of basis states $\{|\Psi_j\rangle\}$ using various unitary transformations U_j, where $|\Psi_j\rangle = U_j |\Phi\rangle$. The Hamiltonian is then diagonalized within the corresponding subspace to give the optimal linear combination

$$|\psi(\boldsymbol{v})\rangle = \sum_j v_j |\Psi_j\rangle . \tag{4.39}$$

In general, the basis states are mutually nonorthogonal and the subspace expansion takes the form of a nonorthogonal configuration interaction with the generalized eigenvalue problem

$$\boldsymbol{H}\boldsymbol{v} = E\boldsymbol{S}\boldsymbol{v}, \tag{4.40}$$

where $H_{jk} = \langle\Psi_j|\hat{H}|\Psi_k\rangle$ and $S_{jk} = \langle\Psi_j|\Psi_k\rangle$. This eigenvalue problem is solved on a classical computer. An advantage of this subspace diagonalization is that the quantum device is only used to measure the matrix elements H_{jk} and S_{jk}, avoiding the significant overhead in gate count and circuit depth associated with explicitly preparing the linear combination in Eq. (4.39) as a quantum circuit. Furthermore,

QSD also gives direct access to excited state energies. Various strategies for generating the basis states have been proposed, including:

1. fermionic excitation operators applied to a ground state approximation to target excited states,[93, 94]

2. application of parametrized quantum circuits $U_j(\boldsymbol{\theta_j})$ defined through optimization[97] or perturbation theory,[346]

3. application of the imaginary time-evolution operator $U_j = e^{-H\tau_j}$, where $\tau_j = j\,\Delta t$ is total duration of the evolution in imaginary time defined by increasing integers $j = 0, 1, 2, ...$,[95]

4. real-time evolution operators $U_j = e^{-iH\tau_j}$, where i is the imaginary unit.[98–102, 334, 348, 349]

Here, we focus on the real-time evolution approach, also referred to as variational quantum phase estimation (VQPE)[100] or quantum Krylov.[98, 334, 349] This method only requires transforming the reference state through unitary time-evolution, which is a natural operation for quantum processors whose quantum circuit implementation has greatly been optimized over the past decades.[65, 72, 76, 88, 102] Previous work has shown that the real-time-evolved states provide a very compact subspace for diagonalizing the Hamiltonian, in terms of the number of variational parameters needed.[98, 100, 102, 348] Furthermore, numerical[100] and theoretical analysis[101, 349] suggest that this algorithm is robust to noise. Starting from the RHF reference state, the subspace is typically built using a linear grid of N_T equally-spaced time points $\tau_j = j\Delta t$ $(j = 0, 1, ..., N_T)$, where the fixed time step Δt must be chosen prior to the calculation. Naively, a total of $2M^2$ matrix elements would need to be evaluated on the quantum device in order to set up the eigenvalue problem (Eq. (4.40)), where M is the number of expansion states. This computation can be reduced to $2M$ unique matrix elements if the time-evolution operator is implemented exactly, since the Hamiltonian and overlap matrices then have a Toeplitz structure, $H_{j,k} = H_{j+1,k+1}$ and $S_{j,k} = S_{j+1,k+1}$.[100] However, if

the time-evolution operator is approximated through Trotterization,[76] then the Toeplitz structure of the Hamiltonian matrix is lost. In this scenario, and for linear time grids, the Toeplitz structure can still be recovered if the eigenvalue problem is reformulated using the time-evolution operator instead of the Hamiltonian (see Appendix B.5 and Refs. 96100).

The main disadvantage of VQPE is that relatively deep quantum circuits are required to implement the time-evolution operator. Specifically, the depth and gate count scale linearly with N_T. The performance of VQPE worsens in the presence of strong correlation, meaning that more time steps are required.[98] The convergence with respect to N_T has also been shown to be slower for excited-state energies.[100, 102] Furthermore, numerical results suggest that the accuracy of ground-state calculations deteriorates when the time-evolution is implemented with low-order Trotter approximations.[98] Higher-order Trotter formulas could be applied, but this comes at great cost since the circuit depth for the k-th order Trotter formula scales exponentially as $\mathcal{O}(5^k)$. Below, we show how combining QSD with spin-coupled reference states can significantly mitigate these limitations.

QSD with spin-coupled reference states

Rather than using more time steps, the deep circuits associated with the time-evolution operator can be reduced by building the subspace using multiple reference states.[98, 99] For example, Stair *et al.* used a set of reference determinants that are identified using iteratively-grown subspaces.[98] Here, we show that using the physically-inspired spin-coupled wavefunctions introduced in Ref. 50 (see Sections 4.2, 4.4) to define the reference states allows for the accurate computation of ground- and excited-state energies at significantly reduced circuit depth compared to a single RHF reference state. Starting from N_R spin-coupled reference states $\{|\Phi_1\rangle, \ldots, |\Phi_{N_R}\rangle\}$, we construct a subspace $\{e^{-iH\tau_j}|\Phi_1\rangle, \ldots, e^{-iH\tau_j}|\Phi_{N_R}\rangle\}$ for $j = 1, \ldots, N_T$, which has $N_R \times (1 + N_T)$ states. Note that each reference state is time-evolved independently.

We test this approach for N_2 at $R = 1.5$ Å using noiseless simulations on classical hardware. We use a maximum of $N_T = 120$ steps in a linear time grid $t_j = j\Delta t$ for three different time-step values $\Delta t \in \{0.1, 1.0, 2.0\}$. We set thresholds of 10^{-6}, 10^{-4}, and 10^{-2} for the singular values of the overlap matrix to remove the null space in the generalized eigenvalue problem. Higher thresholds might be more suited for implementation on noisy devices, whereas using lower thresholds can help to retain more expansion states and achieve faster convergence. We observed no qualitative difference for the three different thresholds and therefore limit our discussion to results with the threshold 10^{-6}. We apply a first-order Trotter approximation of the time-evolution operator for a Hamiltonian given by a sum of terms $H = \sum_k h_k$ in the Pauli basis, given by

$$e^{-iHt_j} \approx \prod_k e^{-ih_k t_j}. \tag{4.41}$$

Compared to a single RHF reference state, using multiple spin-coupled reference states provides significantly faster convergence with respect to the number of time steps (and thus circuit depth) for every Δt considered (Fig. 4.6, top left). This implies that the multireference approach should be faster and less susceptible to hardware noise, which becomes worse for deeper quantum circuits. One might wonder whether this comes at the cost of an increased number of measurements, since the increased number of reference states results in an additional $N_R - 1$ subspace states per time step. Encouragingly, we find that the multireference expansion also converges more rapidly with respect to the total number of expansion states, meaning that fewer measurements would be required overall (Fig. 4.6, top right). The faster convergence with the number of expansion states indicates that starting from different reference states and independently time-evolving each of them might help to explore the Hilbert space more efficiently than if the time-evolution is applied onto a single same initial state.

The accuracy of VQPE strongly depends on the choice of Δt. We can understand this dependence through the phase cancellation picture discussed in Ref. 100, since Δt controls the phase applied to each eigenstate contained in the reference

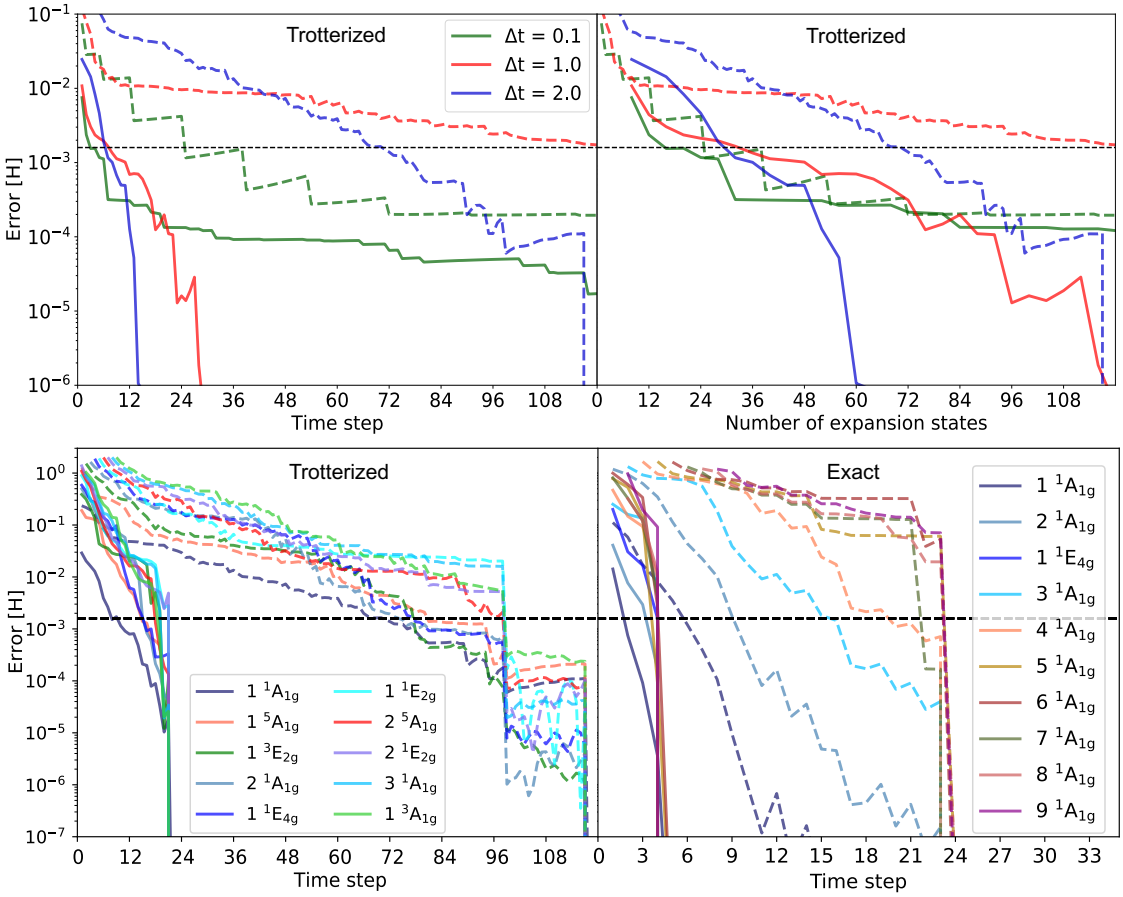


Figure 4.6: When applied on spin-coupled reference states, real-time evolution accurately yields ground and excited state energies after subspace diagonalization, requiring very few time steps. Top row: Ground state energy error as a function of the number of time steps (left) and number of expansion states (right), where the time-evolved states are obtained through first-order Trotter evolution (Eq. (4.41)). Bottom row: Energy error for the 10 lowest eigenstates obtained from first-order Trotter evolution (left), and for the 10 lowest singlet eigenstates obtained from exact time-evolution (right), with time-evolution step size $\Delta t = 2.0$. Dashed lines correspond to subspaces formed from time-evolving the RHF state; solid lines correspond to time-evolving four spin-coupled reference states. The labels indicate the spin (superscript) and spatial symmetry of each eigenstate, where the first number is the index of the eigenstate of a particular symmetry, in increasing energy and starting at 1 for the ground state. All results are for N_2 in the STO-3G basis at bond length $R = 1.5\,\text{Å}$.

space. Our results suggest that larger values of Δt enable faster cancellations and avoid the step-like plateaus that occur for $\Delta t = 0.1$, which are associated with near-linear dependencies in the expansion subspace.

VQPE can also provide access to excited state energies, a key benefit that has received surprisingly limited attention so far. When the time-evolution is Trotterized at first order, VQPE with both the RHF or multireference expansion can compute the lowest eigenstates with spatial symmetry corresponding to A_{1g}, E_{2g}, and E_{4g}, with no restriction on the spin symmetry (Fig. 4.6, bottom left). Like for the ground state calculation, the multireference expansion requires an order of magnitude fewer time steps to reach the exact excited state energies compared to the single RHF reference state. The presence of excited states with different spin symmetry to the reference state arises because the Trotter approximation of the time-evolution operator breaks the spin symmetry of two-body operators in the Hamiltonian. Similarly, non-abelian spatial symmetries in the $D_{\infty h}$ point group are also partially broken since the Trotterized time-evolution operator only conserves symmetries within the D_{2h} abelian subgroup. Since the initial states are totally symmetric, this means that we obtain excited states that reduce to A_g in D_{2h} (which correspond to the $D_{\infty h}$ irreps A_{1g} and E_{ng} for even n). Consequently, although approximate energies can be obtained with high accuracy, the corresponding eigenstates might not have pure spin or spatial symmetry unless they are exact (within numerical accuracy).

Contrast this with the version where we apply the exact time-evolution operator (Fig. 4.6, bottom right). Since the exact time-evolution operator commutes with all the symmetry operators of the Hamiltonian, the time-evolved subspace states are guaranteed to preserve the total spin and the spatial symmetry of the reference state. This allows us to systematically target states of a particular symmetry. Here, we consider the lowest $^1A_{1g}$ states in N_2. QSD with exact time evolution applied on multiple CSFs yields all ten eigenenergies with only 5 time steps, whereas QSD applied on the RHF state requires 24 time steps. (Note that our multireference approach also yields states of $^1E_{4g}$ symmetry since the CSF $|\Phi_2\rangle$, defined in

Eq. (4.9), contains spatial contaminants with azimuthal orbital angular momentum $L_z = 4, 8, \ldots$.) It is remarkable that our multireference approach provides such rapid convergence despite the fact that the spin-coupled configurations are tailored explicitly for the ground state.[50] This can be understood because the excited states of N_2 correspond to configurations where electrons are promoted from bonding to antibonding orbitals, which are exactly the configurations included in our CSF reference states.

In summary, defining a set of multiple reference states using spin-coupled configurations significantly reduces the number of time steps required to converge ground- and excited-state VQPE energies compared to a single RHF reference state. This faster convergence results in shallower quantum circuits, making the VQPE algorithm more suitable for implementation on quantum devices. We believe that this success arises because the multireference expansion is conducive to exploring distant sectors of the Hilbert space. Any QSD method requires the generation of linearly dependent basis states,[100] but this is generally hard to achieve due to the redundancies that arise when applying the time-evolution operator to a single reference state. By independently time-evolving each reference state, we generate a basis that has fewer linear dependencies and is thus more efficient at phase cancellation.

The *perturb-then-diagonalize* form of QSD is naturally well-suited for multireference quantum chemistry. Just as for the nonorthogonal VQE algorithm explored in 4.5.1, QSD with real-time-evolved states is greatly enhanced by using spin-coupled initial states that contain a large part of the relevant electron correlation.

4.5.3 ADAPT-QSD: QSD with adaptive ansatz

Despite its simplicity, the real-time evolution approach to QSD requires circuits that are significantly deeper than is feasible with near-term quantum hardware, because each time-evolution step requires the application of every term in the Hamiltonian. Although the circuit depth can be reduced through randomized Hamiltonian simulation,[334] typically such circuits are still deeper than those for

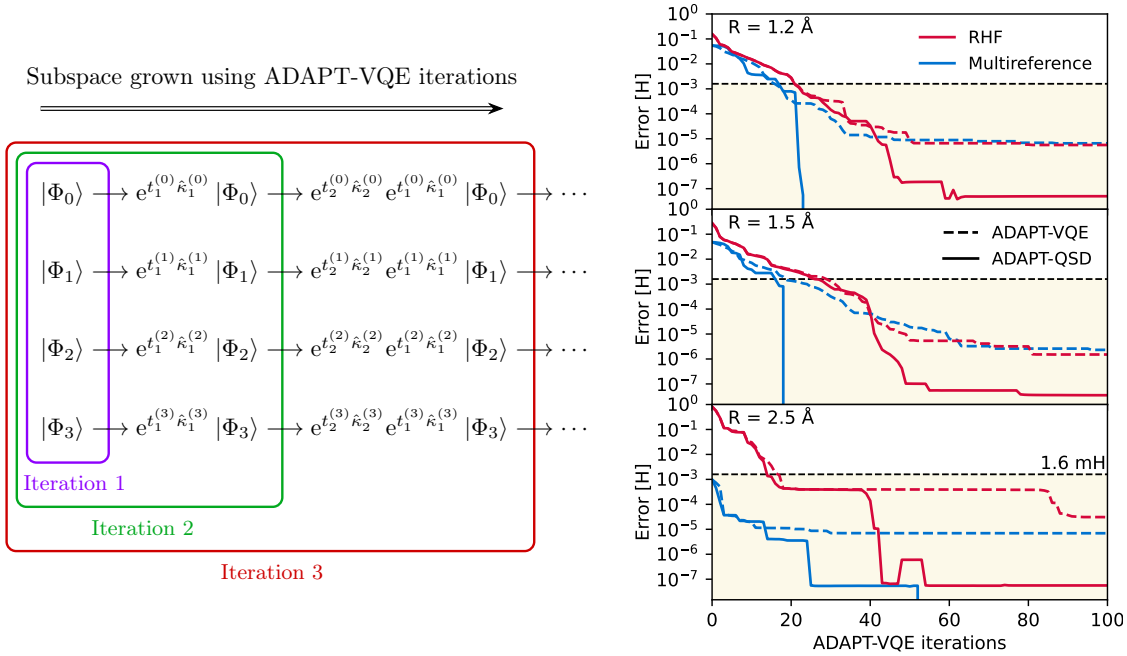


Figure 4.7: ADAPT-QSD: Performing QSD in the basis of sequential ADAPT-VQE states can significantly accelerate the convergence with respect to the number of operators and avoid stagnation, as demonstrated here for N_2 (STO-3G) at $R = 1.5$ Å.

a VQE ansatz. To mitigate this, we propose an alternative QSD approach using adaptive quantum eigensolvers.

Rather than using a fixed quantum circuit, the ADAPT-VQE algorithm dynamically grows an ansatz circuit tailored to the system of interest.[261] At each iteration, the gradient of the energy estimate with respect to a pool of candidate operators is measured. The operator with the largest gradient is appended to the ansatz circuit, and all variational parameters of the ansatz unitary are reoptimized following the usual quantum-classical VQE loop. The ADAPT-VQE algorithm has been shown to provide very compact representations of electronic wavefunctions.[261, 262] However, as the ansatz growth is only guided by a local gradient criterion, its convergence can sometimes be slow and it can get stuck in a local minimum within the discrete operator space.[52] Furthermore, as with any VQE algorithm, the parameter optimization can converge to one of the many high-energy local minima in the continuous space.[258]

Here, we propose a QSD approach where a subspace is defined using the sequence of wavefunctions obtained on each iteration of the ADAPT-VQE algorithm

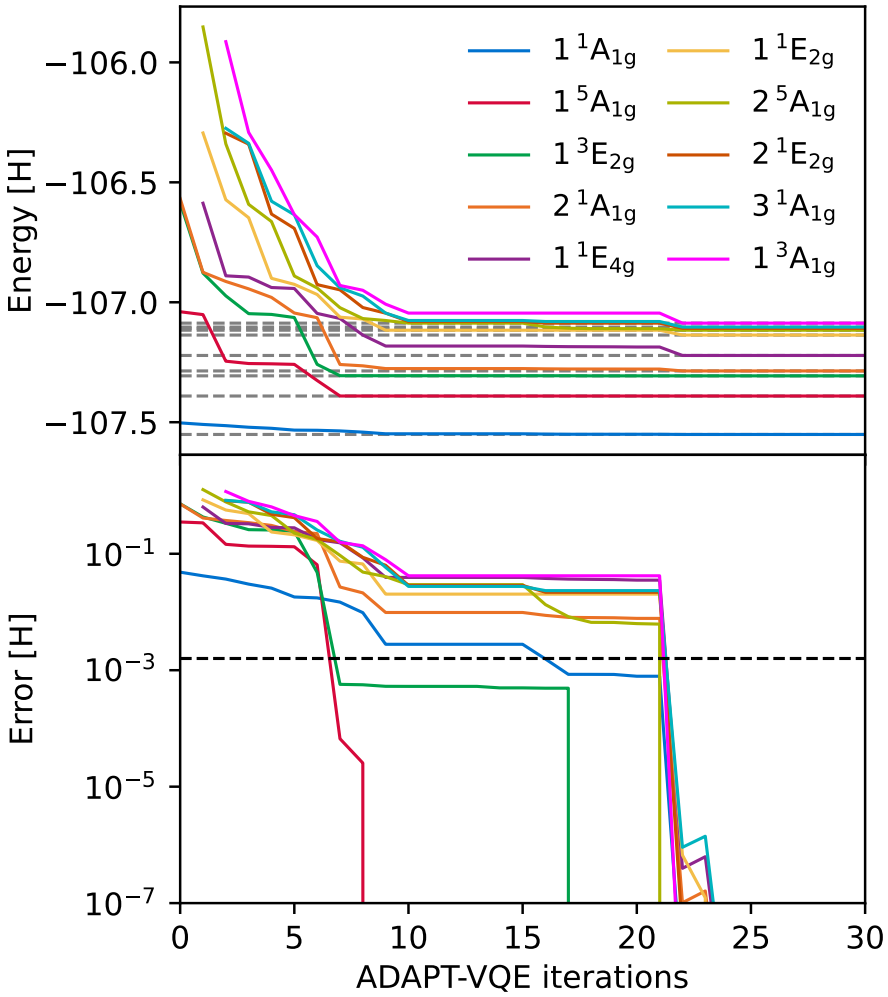


Figure 4.8: Multireference ADAPT-QSD captures the low-lying excited states with the same spatial symmetry as the input CSFs, as shown for N_2 at $R = 1.5\,\text{Å}$. The lower energy states converge more rapidly as the subspace expands, and the states considered here are converged below chemical accuracy using 22 ADAPT-VQE iterations.

(Fig 4.7; left). This can be done either using a single reference wavefunction such as the restricted Hartree–Fock (RHF) state, or multiple reference states including different spin-coupled wavefunctions. ADAPT-QSD performs individual ADAPT-VQE calculations on each reference state, and uses the combined set of correlated states to define the subspace within which the Hamiltonian is diagonalized.

Our numerical simulations on N_2 demonstrate that, even in the single reference case, this approach can mitigate two of the main difficulties of the ADAPT-VQE algorithm. Firstly, ADAPT-QSD significantly accelerates the convergence of the energy estimates with respect to the number of operators, and therefore reduces

the circuit depth (Fig 4.7, right). Secondly, when the ADAPT-VQE wavefunctions stagnate at a local energy mininum, subspace diagonalization in ADAPT-QSD can escape that minimum, yielding an energy estimate that is orders of magnitude lower than ADAPT-VQE at comparable circuit depth.

Comparing the role of the reference states, our results again confirm the advantage of using multiple spin-coupled reference states in QSD approaches. We find that the multireference approach to ADAPT-QSD requires significantly fewer ADAPT-VQE iterations compared to the approach that employs a single RHF reference state. We achieve convergence to the exact ground state energy using multireference ADAPT-QSD with roughly half the number of operators per circuit needed in the single reference case, further reducing the circuit depth requirements.

ADAPT-QSD also systematically yields the low-lying excited states in the molecule, as shown in Fig. 4.8. This enables excited state energies to be computed using the ADAPT-VQE formalism without requiring constrained optimization.[341] For the results presented here, we chose to work with the operators presented in the original ADAPT-VQE paper,[261] which are not spin-symmetry-preserving. Furthermore, these operators also do not conserve all spatial symmetries for non-abelian point groups. Therefore, like the Trotterized RT-QSD approach, the ADAPT-QSD expansion yields excited states that transform as the A_g irreducible representation of the D_{2h} abelian subgroup. One could use symmetry-preserving operators if states of a particular symmetry were desired.

To summarize, we have considered three algorithms that perform subspace diagonalization: the nonorthogonal quantum eigensolver, QSD with real-time-evolved states (also known as VQPE or quantum Krylov), and our newly-proposed ADAPT-QSD. We have seen that the performance of all three algorithms significantly improves when combined with spin-coupled reference states. This is particularly useful for multiconfigurational wavefunctions, including excited states, which are precisely the situations for which classical algorithms are known to struggle. QSD appears particularly promising because it exploits the quantum device to represent highly entangled states in different bases, as well as to yield matrix elements

in nonorthogonal bases. For comparison, the ability to efficiently provide such information through measurements is also the core argument for potential quantum advantage with other quantum algorithms.[97, 346, 370, 371]

There is an additional reason to prefer QSD over VQE methods. Often, VQE implementations struggle to resolve the last few digits of accuracy in the energy estimate. Typically, this occurs either due to a lack of expressiveness in the ansatz or because the optimization gets stuck in local traps.[258, 343] This becomes particularly difficult when hardware noise is present in a real quantum device implementation, because the gate errors are usually higher than the $10^{-3}\,\mathrm{E_h}$ accuracy needed for the energy estimate in chemistry applications.[44] We expect that QSD methods can mitigate this problem since even if hardware noise corrupts the measured overlap and Hamiltonian matrix elements, high accuracy might still be achieved through the linear QSD expansion, in particular when the diagonalization step is regularized through thresholding the eigenvalue problem, as done in our simulations. There is theoretical[101, 349] and numerical[100, 103] evidence that the quantum Krylov approach discussed in Section 4.5.2 is indeed noise resilient when combined with thresholding. It is likely that this noise-resilience is also a feature of ADAPT-QSD, but further work is needed to establish this.

4.5.4 Adiabatic state preparation from the molecular dissociation limit

All the algorithms we have considered so far are hybrid quantum-classical algorithms that exploit the variational principle to minimize the energy through iterative updates of some trial ansatz parameters. Adiabatic state preparation (ASP) offers an interesting alternative. It is a purely-quantum algorithm for preparing Hamiltonian ground states on digital quantum computers that does not rely on any parametrized *ansatz*.[73, 251] Instead, it exploits the adiabatic theorem, which requires using some Hamiltonian whose ground state can be easily prepared on a quantum computer.

Typically, the initial state of choice is the Hartree–Fock state.[76] As usual, this works well for weakly correlated problems, but it is inefficient for strongly correlated systems. The CSFs considered in this work are the exact ground state of a family of molecular Hamiltonians at dissociation (Section 4.2 and Ref. 50). Here, we show that by starting from the dissociation limit with a spin-coupled state, we can speed up adiabatic state preparation of strongly correlated eigenstates.

Background

The ASP algorithm was proposed as a method to prepare initial states with high overlap with the ground state for quantum phase estimation.[41, 64, 76] Starting from some initial Hamiltonian $\hat{H}_0$ whose ground state can be easily prepared on a quantum computer, the ground state of the full Hamiltonian $\hat{H}_\mathrm{F}$ can be generated by slowly evolving the state along a pathway that interpolates between $\hat{H}_0$ and $\hat{H}_\mathrm{F}$, defined as

$$\hat{H}\Big(s(t)\Big) = \Big(1 - s(t)\Big)\hat{H}_0 + s(t)\hat{H}_\mathrm{F}. \tag{4.42}$$

Here, $s(t)$ is some continuous function with $s(0) = 0$ and $s(\tau) = 1$, where τ is the total evolution time. The required total evolution time τ can be estimated as[251]

$$\tau \gg \max_{s\in[0,1]} \frac{|\langle\Psi_1(s)|\,\partial_s H(s)\,|\Psi_0(s)\rangle|}{|E_1(s) - E_0(s)|^2}, \tag{4.43}$$

where $|\Psi_0\rangle$ and $|\Psi_1\rangle$ are the ground and first excited states of $H(s)$, respectively. On a digital quantum device, the path variable $s(t)$ can be discretized using constant time steps Δt. Therefore, slower evolution, which increases the likelihood of remaining in the ground state, requires a larger number of time steps and deeper quantum circuits.

Since the ground state of $\hat{H}_0$ must be easily obtained and prepared, the obvious choice for electronic problems is the mean-field Fock operator, for which the (restricted) Hartree–Fock wavefunction is the ground state.[186] The accuracy of ASP is typically quantified using the squared overlap $|\langle\Psi(\tau)|\Psi_0\rangle|^2$ between the final state $|\Psi(\tau)\rangle$ and the physical ground state $|\Psi_0\rangle$. Numerical studies have shown

that the accuracy strongly depends on the initial state,[41, 186–188] meaning that a mean-field reference is unlikely to be sufficient for strong electron correlation. Indeed, simulations of chemical bond breaking have demonstrated that much larger τ values are required at long bond lengths due to both the inadequacy of the Hartree–Fock state and the small energy gap between eigenstates that leads to an exact degeneracy for $R \to \infty$.[186, 188]

To overcome the limitations of the mean-field Fock operator, alternative initial states and reference Hamiltonians have been explored, including Hamiltonians involving a subset of the molecular orbitals with complete active space configuration interaction (CASCI) wavefunctions,[186, 187] active space Hamiltonians obtained from n-electron valence perturbation theory (NEVPT)[41], and unrestricted Slater determinants that break $\hat{S}^2$ symmetry.[188] The CASCI approach reduces the adiabatic evolution time, but it requires the brute-force computation of the ground state wavefunction within the active space, which is difficult to predict *a priori*. Thus, active-space methods are unlikely to be applicable for systems with many strongly correlated electrons. Furthermore, while the symmetry-broken states analyzed in Ref. 188 can be easily prepared on a quantum circuit, the spin symmetry must be restored by adding a penalty term to constrain the $\langle \hat{S}^2 \rangle$ expectation value, which increases the circuit cost of implementing time-evolution.

Interpolation from the exact dissociation limit

Here, we exploit the fact that the ground state of dissociated molecules can often be described using only one spin eigenfunction of the form $|\mathcal{O}^{N,1}_{0,0}\rangle$.[50] This suggests an alternative approach to apply ASP in the regime of strong correlation. We propose using the ground state and the fully-interacting Hamiltonian at molecular dissociation as the starting point, and applying ASP to interpolate between the dissociation limit and the target geometry to compute the entire binding curve. We numerically investigate this approach through the N_2 molecule.

The ground state for N_2 at dissociation is exactly represented by one open-shell CSF $|\mathcal{O}^{6,1}_{0,0}\rangle$ which spin-couples the six valence p-electrons and can be predicted *a pri-*

ori, from symmetry arguments (see Sections 4.2). We define the initial Hamiltonian $\hat{H}_0$ as the full interacting Hamiltonian within the (6, 6) active space at $R = 4.5\,\text{Å}$, which is representative of the dissociation limit since $1 - \left|\langle \mathcal{O}^{6,1}_{0,0} | \Psi_0 \rangle\right|^2 \approx 4 \times 10^{-7}$, where $|\Psi_0\rangle$ is the exact ground state. For simplicity, we only consider the linear interpolation $s(t) = t/\tau$. We use exact time evolution with a constant timestep of $\Delta t = 0.1\,\text{E}_\text{h}^{-1}$ and different total evolution times up to $\tau = 300\,\text{E}_\text{h}^{-1}$. We compare CSF-based simulations with standard ASP calculations that start from the RHF reference state with $\hat{H}_0$ being the Fock operator.

The accuracy of the adiabatically-evolved RHF state for a given τ decreases as the bond length increases (Fig. 4.9). For example, a very long evolution time of $\tau = 300\,\text{E}_\text{h}^{-1}$ is required to achieve over 90 % fidelity (squared overlap) with the true ground state in the dissociation limit, and the corresponding energetic error is still much larger than chemical accuracy. In contrast, ASP starting from the fully open-shell CSF $|\mathcal{O}^{6,1}_{0,0}\rangle$ achieves fidelities of $> 90\%$ for all geometries above $R > 1.4\,\text{Å}$ with a short evolution time $\tau \geq 30\,\text{E}_\text{h}^{-1}$. This demonstrates that the spin-coupled CSF provides a much better starting point in the strongly correlated dissociation regime. For a given τ, simulations starting from the open-shell CSF perform better than those starting from the Hartree–Fock state for $R > 1.7\,\text{Å}$, which corresponds to the bond length at which the open-shell CSF switches with the RHF determinant as the dominant contribution to the ground wavefunction (Fig. 4.1c). This observation provides further numerical evidence for the direct correlation between the accuracy of the initial state and the cost of ASP, in line with previous research.[41, 186–188]

Since the circuit depth and gate cost of implementing time-evolution on a quantum computer scale polynomially with the total time,[254] state preparation based on open-shell CSFs provides orders of magnitude more efficiency for strong correlation compared to RHF-based ASP. Furthermore, since the CSF is explicitly an eigenfunction of the $\hat{S}^2$ operator, we do not require additional modifications of the time-evolution circuits to enforce spin-pure wavefunction, in contrast to the symmetry-broken approach in Ref. 188. Finally, because the cost of preparing

the open-shell CSFs is negligible compared to the cost of implementing the time-evolution operator and we avoid the classical complexity of identifying CASCI initial wavefunctions,[186, 226] our approach can be extended to many strongly correlated electrons.

Looking forward, although we have only considered a minimal basis representation of the N_2 dissociation, the same principles can be readily applied to larger basis sets. For example, one possible protocol would be to start with a minimal basis representation incorporating the dominant spin coupling and continuously evolving the state into the larger basis set at the target geometry using ASP by *turning on* the matrix elements coupling the orbitals in the minimal and large bases. A similar approach could be applied to other strongly-correlated systems, such as FeS clusters,[64, 112] where the fully-interacting ground state could be prepared starting from a suitable model Hamiltonian with a single CSF ground state.

4.6 Avoiding the orthogonality catastrophe in quantum phase estimation

4.6.1 Quantum phase estimation

Large-scale error-corrected quantum hardware remains a distant prospect and near-term applications are likely to rely on heuristic algorithms. However, in the long term, the most anticipated application of quantum computing in chemistry is the simulation of challenging systems through quantum phase estimation.[41, 64, 65, 76, 88]

QPE can be used to project some initial state $|\Phi_{\text{ref}}\rangle = \sum_i \gamma_i |E_i\rangle$ onto one of the energy eigenstates $|E_i\rangle$ of the Hamiltonian.[48, 67, 72] Although the algorithm is highly promising for quantum advantage with fault-tolerant hardware architectures, its scaling depends inversely on the initial state overlap. This is problematic for strongly correlated electrons where the Hartree–Fock wavefunction has poor overlap with the exact state, which is also where classical methods are most limited. For example, for N_2, $|\gamma_0|^2 = |\langle \Phi_{\text{RHF}}|E_0\rangle|^2 \approx 0.92$ at equilibrium but decreases to $|\gamma_0|^2 \approx 0.06$ at long bond lengths (Fig. 4.1c).

Here, we propose using an initial state state that consists of a linear combination of the fully spin-coupled state $|\mathcal{O}_{0,0}^{N,1}\rangle$ and intermediate states with increasingly fewer correlated electrons ($N-2, N-4, ...0$), including the fully uncorrelated RHF state (see Eq. (4.11) for N_2, where $N = 0, 2, 4, 6$). This allows us to prepare states with high overlap with the exact ground state using $\mathcal{O}(N^2)$ Toffoli gates (Eq. (4.29)), mitigating the initial state problem for systems where the accuracy of the Hartree–Fock state diminishes rapidly with the number of electrons.

The QPE algorithm involves three steps:

1. The main register is initialized to some reference state through application of a state preparation circuit $U_{\text{ref}} |0\rangle = |\Phi_{\text{ref}}\rangle$.

2. An invertible function of the Hamiltonian is applied in the form of a unitary operator. This is often the time-evolution operator $U = e^{-iHt}$,[64, 72, 76] but it can also be a qubitized quantum walk $U = e^{i \arcsin(H/\lambda)}$,[58, 65, 85] where λ is a rescaling parameter proportional to the norm of the Hamiltonian. Evolution of the initial state under U controlled by the ancilla register accumulates a phase on the ancilla qubits corresponding to the eigenvalues of the operator.

3. After application of the Quantum Fourier Transform on the ancilla register, the ancilla qubits are measured in the computational basis. The measurement results encode the energy eigenvalue as a binary bitstring that is read out classically.

If $|\Phi_{\text{ref}}\rangle = |E_i\rangle$, the measurement yields the energy corresponding to the i-th eigenstate with probability 1 (exactly, up to a desired numerical precision). However, if the initial state is a superposition of energy eigenstates, the probability of projecting to the desired eigenstate and measuring the corresponding energy eigenvalue is given by the squared overlap $|\gamma_i|^2 = |\langle \Phi_{\text{ref}} | E_i \rangle|^2$ between the initial state and the desired eigenstate. Therefore, to obtain a sufficiently accurate energy estimate, the entire procedure must be repeated several times, with the number of repeats scaling as $\mathcal{O}(|\gamma_i|^{-2})$. Post-QPE techniques can improve this scaling

to $\mathcal{O}(|\gamma_i|^{-1})$.[104, 105, 298] However, the overlap dependency is a fundamental feature of all projective algorithms, and indeed the $\mathcal{O}(|\gamma_i|^{-1})$ scaling has been found to be near-optimal.[104]

Hamiltonian simulation is remarkably efficient, and the latest qubitization-based algorithms achieve quadratic scaling with the basis set size (number of orbitals).[65, 85] However, the QMA-hardness of electronic structure manifests through the initial state dependency of QPE.[40] The reliance on the global overlap of an approximate wavefunction means that in principle, the algorithm will suffer from the orthogonality catastrophe,[339, 340] requiring a number of repetitions that scales exponentially in the system size. In practice, this worst-case scenario can be avoided by identifying approximate wavefunctions that contain the dominant entanglement structure of the exact wavefunction.

For the N_2 example studied here, a linear combination of four CSFs (Eq. (4.11)) can enhance the squared overlap of the reference state and the exact ground state by a factor of 16 when compared to the Hartree–Fock state at stretched geometries (Fig. 4.1c). This implies an equivalent reduction in the total runtime of QPE since the cost of preparing the initial state is negligible compared to the cost of Hamiltonian simulation. In Ref. 50, we presented analogous results for more systems including most second-row diatomics and the water molecule, where the overlaps of linear combinations of CSFs are always above 85%. The accuracy of this linear combination state lies in the ability to capture spin coupling effects between localized electrons in open-shell orbitals.

This improved initial state removes the bias of Hartree–Fock theory towards delocalized states and instead encodes the main physical effect that causes strong correlation in such systems: ferromagnetic coupling within a local subsystem of electrons and global antiferromagnetic coupling across subsystems. We expect that diatomics with even more strongly correlated electrons, such as Cr_2, can also be treated in this way.

Nevertheless, all systems mentioned so far, including the Cr_2,[108] are tractable with classical methods, at least if only the ground state is considered. This can

be attributed to the fact that a relatively small number of electrons are strongly spin-coupled ($N = 12$ for Cr_2), and therefore compressed classical representations of the wavefunction can still be processed on classical devices without a big loss in accuracy. In other words, the exponential asymptotic scaling in N is not prohibitive when $N = 12$.

4.6.2 Resource estimation for fault-tolerant simulation of FeS systems

The long-term hope is that fault-tolerant quantum devices with hundreds of logical qubits will be able to compute energies and properties of systems that are out of reach for all classical methods.[64, 65, 88]

An often-cited class of molecules with potential technological applications, whose simulation is classically challenging, are transition metal clusters such as systems composed of multiple iron-sulfur centers.[17, 18] In particular, the electronic structure of FeMoCo is considered a benchmark for QPE-based quantum advantage.[41, 64, 65, 112] It is strongly correlated, containing up to 35 open-shell electrons.[18] By increasing the number of FeS centers, it has been shown that both the Hartree–Fock state and a naive single CSF have an overlap with the best-available matrix product state that diminishes exponentially with increasing system size.[41] Although the poor Hartree–Fock overlap is expected, we hypothesize that it is possible to obtain a much higher overlap using a linear combination of a small number of CSFs, provided a bespoke, compact representation is chosen for each CSF. This requires identifying the relevant spin coupling patterns through a basic understanding of the electronic structure, and mapping these to their corresponding CSFs. Crucially, to obtain the desired wavefunction compression, one must choose an appropriate set of orbitals (single-particle basis) as well as CSFs (many-body basis). A detailed study on CSFs for FeS systems is beyond the scope of this work. Nevertheless, one can already see the potential of this CSF-based approach as a state preparation method for QPE, under reasonable assumptions.

To get a rough estimate for the cost of preparing an accurate initial state, first assume that a CSF like the state $|\mathcal{O}_{0,0}^{N,1}\rangle$ is an accurate approximation of the ground state, where N is the number of open-shell, spin-coupled electrons. We expect the number of determinants to be determined mostly by the number of open-shell electrons. Therefore, the open-shell CSF $|\mathcal{O}_{0,0}^{N,1}\rangle$, whose structure and circuit representation we have studied in this work, serves as a proxy for the cost of preparing a high-overlap approximation of the eigenstates of FeS systems.

From previous studies,[17] it is known that $N = 10$ and $N = 18$ for Fe_2S_2 and Fe_4S_4. FeMoCo is believed to have up to 35 open-shell electrons,[18] and we thus set $N = 34$. The smaller clusters have been tackled with success using state-of-the-art DMRG programs,[17] and satisfactory approximations have also been obtained for the P-cluster,[18] a system believed to facilitate the nitrogen reduction process between the $[Fe_4S_4]$ Fe cluster and FeMoCo. However, accurately resolving the electronic structure of FeMoCo remains an open problem.[41, 112]

Table 4.1 contains the number of determinants $\binom{N}{N/2}$ required in the expansion of the state $|\mathcal{O}_{0,0}^{N,1}\rangle$. For $N = 34$, this is 2.3×10^9. In this situation, it is clear that one would need post-Hartree–Fock initial states for QPE as the overlap between the Hartree–Fock state and the ground state is very small. Indeed, Lee *et al.* [41] found an overlap of $\sim 10^{-7}$ between the Hartree–Fock state and a matrix product state obtained through the DMRG algorithm. Recent work by Ollitrault *et al.* shows that the overlap of a single CSF and the ground state can be boosted by optimizing the orbitals for the $[Fe_2S_2]$ and $[Fe_4S_4]$ Fe clusters.[372] This provides a useful improvement but it does not solve the fundamental issue of exponential decay. Therefore, more sophisticated state preparation methods are required for larger systems.

As an alternative to preparing the Hartree–Fock state or an open-shell CSF, one could consider using the best possible selected CI state,[253] e.g., as found through FCIQMC,[163, 373] heat-bath CI,[227], or adaptive sampling CI.[181, 226] However, quantum circuits that prepare a generic linear combination of L determinants typically require $\mathcal{O}(L)$ gates,[192, 374] usually with an additional $\mathcal{O}(\log L)$ [192]

or $\mathcal{O}(N)$ factor for N qubits,[181] as well as the use of an ancilla register. At best, the cost can be reduced to $\tilde{\mathcal{O}}(\sqrt{L})$, but this requires trading off Toffoli gates for $\mathcal{O}(\sqrt{L})$ qubits.[192]

The state-of-the-art algorithm for computing the ground state energy of FeMoCo through phase estimation requires 3.2×10^{10} Toffoli gates.[65] Using the clean qubit approach from Low *et al.*,[192] recent work gave an algorithm for preparing a CI state of L determinants with cost of

$$(2 \log L - 2)L + 2^{\log L + 1} + L \tag{4.44}$$

Toffoli gates.[253] Therefore, preparing a superposition of the $L = 2.3 \times 10^9$ determinants that form the state $|\mathcal{O}_{0,0}^{34,1}\rangle$ would cost 1.43×10^{11} Toffoli gates. This would make the cost of initial state preparation higher than the cost of QPE, which is problematic given the already-long expected runtimes for evaluation of a single energy, which are in the order of multiple days on future superconducting quantum hardware.[65]

To reduce the state preparation cost, one could consider only preparing a few of the most important determinants that dominate the wavefunction. Examining the CSF $|\mathcal{O}_{0,0}^{N,1}\rangle$, while the number of determinants scales exponentially as $\binom{N}{N/2}$, the determinants $|\alpha\rangle^n |\beta\rangle^n$ and $|\beta\rangle^n |\alpha\rangle^n$, where $n = N/2$, have significantly more weight that the rest. Specifically, their amplitude is $1/\sqrt{n+1}$ (Eq. (4.13)). One could therefore build a state

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|\alpha\rangle^n |\beta\rangle^n - |\beta\rangle^n |\alpha\rangle^n) \tag{4.45}$$

that would have squared overlap with the $|\mathcal{O}_{0,0}^{N,1}\rangle$ CSF scaling inverse linearly in the number of electrons: $|\gamma_N|^2 = |\langle\phi|\mathcal{O}_{0,0}^{N,1}\rangle|^2 = 4/(N+2)$. For example, for N_2, this would give an initial state with squared overlap $|\gamma_6|^2 = 0.5$ with the $|\mathcal{O}_{0,0}^{6,1}\rangle$ CSF (and therefore with the ground state near dissociation) at a very cheap cost.

However, this has several downsides. First, the state $|\phi\rangle$ is not a spin eigenfunction. Furthermore, $|\phi\rangle$ has significant overlap with eigenstates other than the ground state, which is problematic for application in QPE, as it makes it highly

probable to measure an undesired eigenstate. In the N_2 example, at $R = 3.0\,\text{Å}$ the squared overlap of $|\phi\rangle$ with the singlet ($S = 0$) ground state and a quintet ($S = 2$) excited state is equal at 50%, and therefore it is equally likely to measure either state. This situation would arise in a much more pronounced way in more challenging systems such as transition metal clusters, where the strong correlation due to spin coupling leads to many near-degenerate eigenstates of different spin, or eigenstates of the same spin but different spin coupling. In such situations, it is essential to use accurate initial states to separate the different eigenstates, which requires a much larger number of determinants.

A further alternative would be to use a classical DMRG calculation to find a good initial state.[234] The tensor product structure of the MPS can be exploited to directly encode the DMRG state in a quantum circuit, circumventing any explicit CI expansion. To our knowledge, the only method for MPS preparation with guaranteed success is the sequential algorithm proposed by Schön *et al.*[375] The number of gates scales as $\mathcal{O}(M\chi^2)$ for a MPS with M sites and bond dimension χ.[253] For FeMoCo, we can take $M = N = 34$, i.e. the number of sites of the MPS is the number of open-shell orbitals/spin-coupled electrons.

Despite the polynomial scaling in N and χ of MPS preparation circuits, the cost of gate-level implementations have large constant factors as they require implementing non-trivial quantum arithmetic operations and the use of $\mathcal{O}(\chi)$ ancilla qubits.[253] For example, Formichev *et al.* [253] gave a circuit decomposition of the algorithm by Schön *et al.* [375] with approximate cost

$$T_{\text{MPS}} \approx (M-1)\chi[32\chi + (b+1)\log(4\chi)] \tag{4.46}$$

(Appendix B.6). More fundamentally, the entanglement captured by a MPS of bond dimension χ scales as $\mathcal{O}(\log(\chi))$. Therefore, for systems whose entanglement grows with the system size, as opposed to Hamiltonians with area-law ground states,[235] the bond dimension can scale very steeply, at worst exponentially in N to achieve high accuracy through a MPS.

In practice, the success of the DMRG algorithm when applied to strongly correlated systems, at least for smaller FeS clusters,[17, 18] indicates that the worst-case exponential scaling might not be strong enough to prohibit classical simulations for the system sizes considered. Another explanation could be that locality dampens electron-electron interactions, e.g. due to charge screening, reducing the strength of long-range entanglement between different FeS clusters. It remains to be seen what bond dimensions would be required to accurately resolve larger systems like FeMoCo, and what would be the exact cost of the quantum circuits for sufficiently-accurate state preparation in the context of phase estimation. State-of-the-art DMRG calculations for different FeS systems typically involve bond dimensions of up to several thousands $(2000 - 8000)$ to achieve accurate energy estimates.[17, 18, 41, 372] Table 4.2 compares the Toffoli cost of MPS state preparation for different FeS systems as a function of the bond dimension, as per Eq. (4.46). Clearly, preparation of a MPS of high ($\chi > 1000$) bond dimension would constitute a significant fraction of the overall cost of one run of QPE.

Formichev *et al.* also proposed reducing the cost of MPS state preparation by taking a MPS with high bond dimension (e.g., $\chi \approx 2000$) obtained from a DMRG calculation, and compressing it to a MPS with lower bond dimension ($\chi \approx 5 - 10$).[253] For some systems, including Fe_4S_4, they found that while the energy of that MPS is very inaccurate, compression does not significantly affect the wavefunction quality, in the sense that there remains a strong overlap between the compressed MPS and the MPS of higher bond dimension. Assuming that such compressed states with bond dimension in the order of $\chi = 10$ indeed provide accurate ground state approximations, the state preparation cost would be relatively low compared to the cost of QPE (Table 4.2). Although this compression approach could be promising for cheaper state preparation, it requires significant classical cost to initially compute the MPS with high bond dimension. It also remains to be seen how accurate the compression is for larger and more complex systems.

In contrast, our CSF preparation circuits directly encode the relevant entanglement structure into bespoke quantum circuits, which we expect to maximize the

Table 4.2: Estimated number of Toffoli gates required for preparation of matrix product states for N electrons in $M = N$ spatial orbitals for different choices of bond dimension χ. Here, b is the number of digits (ancilla qubits) that encode each rotation gate in binary representation, as required to guarantee a state preparation error of at most $\epsilon = 10^{-7}$. For comparison, the number of Toffoli gates required for running qubitization-based quantum phase estimation circuits is 3.2×10^{10} for FeMoCo.[65]

N	System	χ	b	T_{MPS}
10	Fe_2S_2	10	34	4.56×10^4
		50	37	8.51×10^5
		2000	43	1.16×10^9
18	Fe_4S_4	10	35	8.7×10^5
		50	37	1.61×10^6
		2000	44	2.2×10^9
34	FeMoCo	10	36	1.71×10^5
		50	38	3.13×10^6
		2000	44	4.26×10^9

state preparation efficiency nearly optimally. Preparation of $|\mathcal{O}_{0,0}^{N,1}\rangle$ for $N = 34$ has a low cost of $\approx 4 \times 10^3$ Toffoli gates (Table 4.1), and does not require any ancillas, significantly reducing the space-time volume and hardware connectivity requirements compared to the aforementioned black-box state preparation approaches.

Our current understanding of the spin coupling structure of FeMoCo is not as precise as that of the bond breaking examples in Ref. 50 (Section 4.2). However, it is well-known that the nature of electron correlation in stretched bonds and transition metal clusters is similar.[356, 376] Moreover, data from state-of-the-art electronic structure simulations reveals analogous phenomena in these systems:[18, 112] atoms form clusters such as Fe_4S_4 cubanes, where the $3d$ orbitals in each FeS cluster contribute a number of unpaired electrons (spins). These spins interact through shorter-range (local/intracubane) coupling within the cluster as well as longer-range (global/intercubane) entanglement between the cubanes. Typically, the local coupling is nearly maximally ferromagnetic whereas the coupling between the cubanes is antiferromagnetic. This locally ferromagnetic and globally antiferromagnetic alignment mirrors the mechanisms occurring in the bond breaking examples of Ref. 50, including N_2 (Section 4.2.2).

FeS systems are more complex than the organic molecules we studied here and in Ref. 50 due to the presence of sulfur ligands, charge hopping effects, and different oxidation states. These complicate the picture, as they introduce a larger number of spin coupling patterns that are energetically close, necessitating the inclusion of a larger number of near-degenerate CSFs in the wavefunction ansatz to account for all dominant effects. Nevertheless, the main many-body effect remains the entanglement/correlation induced by spin coupling, and the number of dominant couplings is greatly limited by symmetries, typically in the order of the number of electrons.[18, 376] By employing a bespoke and appropriate choice of basis for each CSF or for sets of CSFs corresponding to similar electronic configurations, this additional complexity should only increase the number of CSFs by a moderate factor.

In conclusion, we anticipate that the proposed approach could be extended to FeS systems or similar transition metal clusters, and that the number of relevant states required for high overlap would scale moderately (rather than exponentially) with the number of open-shell electrons. Preparing linear combinations of CSFs would add a small overhead to the $\sim 10^3$ gate counts for state preparation in Table 4.1 (see Section 4.4.9), but it is unlikely to be significant enough to affect our conclusions.

4.7 Conclusions and discussion

While quantum algorithms such as QPE can in principle compute the exact eigenvalues of arbitrary electronic Hamiltonians, in practice their performance depends critically on the overlap of the target eigenstate with the initial state. Since the overlap between two random vectors is inversely proportional to the size of the Hilbert space, strategies for preparing initial states with the principal features of the Hamiltonian eigenstates are an essential component of quantum simulation. In quantum chemistry, the mean-field Hartree–Fock method provides an initial state with a high overlap with the exact ground state for weakly correlated electronic systems, which pushes the orthogonality catastrophe to systems with a very large

number of electrons. Classically challenging problems, however, involve strongly correlated states. Here, the overlap of the RHF wavefunction with the exact eigenstate shrinks exponentially with the number of strongly correlated electrons, which presents a severe initial state problem.

In previous work, we demonstrated that the structure of strongly correlated eigenstates in chemical systems can be predicted from symmetry arguments using generalized spin-coupled molecular orbital theory.[50] Replacing the HF wavefunction with more general spin-coupled wavefunctions, linear combinations of CSFs, provides high-quality initial states across the whole range of correlation regimes found in molecules. We have shown that the CSF-based initial states improve the performance of a wide range of quantum algorithms for the simulation of quantum chemistry on quantum computers, which can be applied both on fault-tolerant quantum hardware as well as on near-term devices. The CSFs directly encode the strong correlation in molecules and constitute specific patterns of entanglement. We have presented circuits for building CSFs with depth linear in the number of electrons N, which generate a linear combination of states of dimension exponential in N.

For VQE with a fixed-depth quantum number preserving ansatz, our numerical simulations for N_2 show that the CSF-based initial state returns significantly more accurate energies than a HF initial state, with orders of magnitude improvement at stretched geometries. The multireference non-orthogonal variant of this VQE algorithm produces chemically accurate energies across the whole binding curve at much reduced gate depths. Quantum subspace diagonalization through real-time evolution also benefits greatly from using CSF-based initial states. We find that the subspace formed by independently evolving each of the chemically relevant CSF states rapidly spans the relevant low-energy regions of Hilbert space. This provides a powerful method for computing accurate excited states, as well as accelerating convergence to the ground state. Since the CSFs are eigenfunctions of space and spin symmetry operators, this has the advantage that the space and spin symmetry sectors can be treated separately, which makes it possible to target specific molecular excited states. For ASP, beginning with a single spin-coupled

wavefunction at dissociation enables preparation of the ground state at stretched geometries with a much-reduced evolution time and therefore lower circuit depth.

In the language of quantum chemistry, the initial state is constructed to contain the static, or strong correlation contributions, and the subsequent refinement through the quantum algorithms captures the remaining dynamic correlation and orbital relaxation. Although we have restricted the analysis to a minimal basis set, quantitatively accurate energy estimates require a more refined discretization of the Hilbert space in the form of a larger basis set.[211] This does not present a problem since the CSFs in a minimal basis can be projected into the larger basis, and orbital relaxation effects captured e.g. through variational quantum algorithms.

A key benefit of the QSD algorithm based on real-time evolution is that it can be used to compute excited states and has less stringent requirements on the accuracy of each ansatz circuit when compared e.g. to VQE which relies on a single wave function ansatz and corresponding unitary. However, implementing real-time evolution according to the Hamiltonian typically requires higher gate counts and depth per circuit compared to the parametrised ansätze used in VQE. The dissociation of N_2 in a STO-3G basis in the valence space of six electrons in six orbitals (twelve qubits) using four configuration state functions (CSFs) Φ_i, $i = 0, 2, 4, 6$, as predicted by generalized MO theory. Here, Φ_0 is the RHF mean-field state, Φ_2, Φ_4, Φ_6 are CSFs with two, four, and six spin-coupled (strongly correlated) electrons, and Ψ is the exact ground state. The figure is reproduced with permission from Ref. 50. We have therefore introduced a new quantum algorithm, ADAPT-QSD, that combines the advantages of QSD with the benefit of the short-depth circuits of VQE. We have demonstrated that performing QSD in the basis of sequential ADAPT-VQE states can significantly accelerate the convergence with respect to the number of operators and avoid stagnation, and that ADAPT-QSD yields an energy estimate that is orders of magnitude lower than ADAPT-VQE at comparable circuit depth. In view of the high popularity of ADAPT-VQE for ground state computation, we believe that ADAPT-QSD is a promising algorithm for excited state calculations.

Our state preparation heuristic avoids brute-force computation and instead exploits physical insight and symmetries to identify and prepare bespoke initial states. Although it remains to be seen how effectively one can apply the same techniques to other systems, our framework provides a basis to tackle strongly correlated molecules where spin coupling dominates. Previous work showed how, analogously to the N_2 molecule studied here, one can identify a small set of accurate CSFs for a range of other systems including H_2O, B_2, C_2, O_2, F_2, and clusters of hydrogen atoms.[50] Here, we have used the example of N_2 in all our numerical simulations of quantum algorithms, but we expect a similar advantage when enhancing quantum algorithms through CSFs for other molecules.

Looking farther ahead, the hope is that future fault-tolerant quantum devices with hundreds of logical qubits will be able to solve the electronic Schrödinger equation for systems that are out of reach for all classical methods. For strongly correlated molecules, such as FeMoCo, which is considered an exemplary target for quantum simulation, our CSF-based formalism provides a systematic approach to constructing initial states with high overlap with the exact state. From approximate resource estimates, we conclude that our spin-coupling-based approach to initial state preparation will be greatly beneficial in solving challenging chemical problems through quantum phase estimation. This is because the CSF states can be identified at the cost of mean-field classical computation and prepared with much lower numbers of gates than expensive approximate classical wavefunctions such as DMRG, for a similar level of initial state overlap.

We anticipate that a possible fault-tolerant workflow would be the following: 1) prepare the relevant CSFs in the minimal basis, 2) optimize the orbitals, given the CSF state, as proposed in Ref. 372, 3) (optionally) apply a heuristic quantum algorithm to compute additional correlation arising in a larger basis, e.g. VQE or a QSD algorithm, and/or (4) use QPE or post-QPE algorithms to obtain the final energy.

The ultimate goal of computational chemistry is to increase our understanding of complex chemical systems. Chemical properties are often the result of finely

balanced competing effects and large parts of modern development have focused on methods for obtaining tightly converged energy estimates. Powerful classical algorithms such as DMRG and FCI-QMC fall into this category, as do the prevailing quantum algorithms. While these methods can accurately compute properties directly from quantum mechanics, the brute-force nature of these algorithms obscures the interpretation of the results. In contrast, our spin-coupled molecular orbital theory reveals configurations that correspond to the dominant contributions to the ground and low-lying excited eigenstates. The CSFs correspond to clear bonding motifs, providing direct insight into the underlying electronic structure as well as a simplified, informationally-compact description of many-body wavefunctions that otherwise appear complicated and unintelligible.

The central role of spin coupling in strongly correlated molecular states [17, 18] as well as the machinery for construction and application of CSFs [189, 273–275] has long been known in the quantum chemistry community. However, the degree to which one can exploit this insight in classical algorithms has remained limited because it is fundamentally challenging to achieve a significant compression of truly multiconfigurational wavefunctions [356] using classical computers. To maximize the compactness of such wavefunctions, one must work in different orbital bases, and possibly different CSF bases, for each or at least some of the CSFs. The valence bond literature emphasizes the limitations of canonical molecular orbitals and the power of CSFs.[377] However, valence-bond-based ansätze involve superpositions over all possible CSFs [378–381] and are therefore constrained to small systems by the curse of dimensionality; furthermore, they impose a common set of (potentially nonorthogonal) orbitals for each CSF. Molecular-orbital-based spin-adapted algorithms on classical computers,[16, 190] also recently proposed for state preparation on quantum computers,[41, 376] involve the use of a fixed single-particle basis of orthogonal orbitals, as well as a fixed many-body basis of CSFs, e.g. as enumerated through a geneological coupling scheme.[189, 376] These limitations have led to the view that CSF-based initial state preparation for quantum algorithms will also require expansions with exponentially many variational coefficients.[41, 376]

Here, we take the view that, in the context of quantum computation, these limitations are unnecessary and the standard spin-adapted approach is unsuitable. Instead, we allow for full freedom in the choice of both the single-particle basis as well as CSF basis. This ensures that a compact representation of each CSF is possible in terms of its Slater determinant expansion, i.e. the basis is chosen such that the determinants' coefficients exhibit a highly symmetric structure, even if many Slater determinant are required for some CSFs. As a consequence, this enables efficient state preparation of each CSF in its corresponding orbital basis. Our approach combines the unique strength of quantum computers (exponential memory scaling, ease of transforming between different bases and obtaining corresponding matrix elements) with the wavefunction/circuit compression and conceptual benefits (increased chemical insight) of working with CSFs. Note that a similar line of argument was recently taken by Leimkuhler and Whaley to motivate the potential advantage of using quantum computers in the context of a tensor network ansatz.[371]

In summary, the ability to distil strongly correlated wavefunctions into a small number of symmetry-derived components (the dominant CSFs) greatly enhances the performance of quantum simulation methods. It enables designing efficient state preparation circuits to bias quantum algorithms towards relevant, highly entangled regions of the Hilbert space at a reduced computational cost. We have here shown how to achieve this for a range of systems and quantum algorithms and presented a blueprint for how to extend it to more challenging Hamiltonians. As such, our work provides a conceptual framework and quantum algorithmic techniques with the necessary ingredients for scalable and interpretable quantum simulation of classically challenging molecular electronic systems.

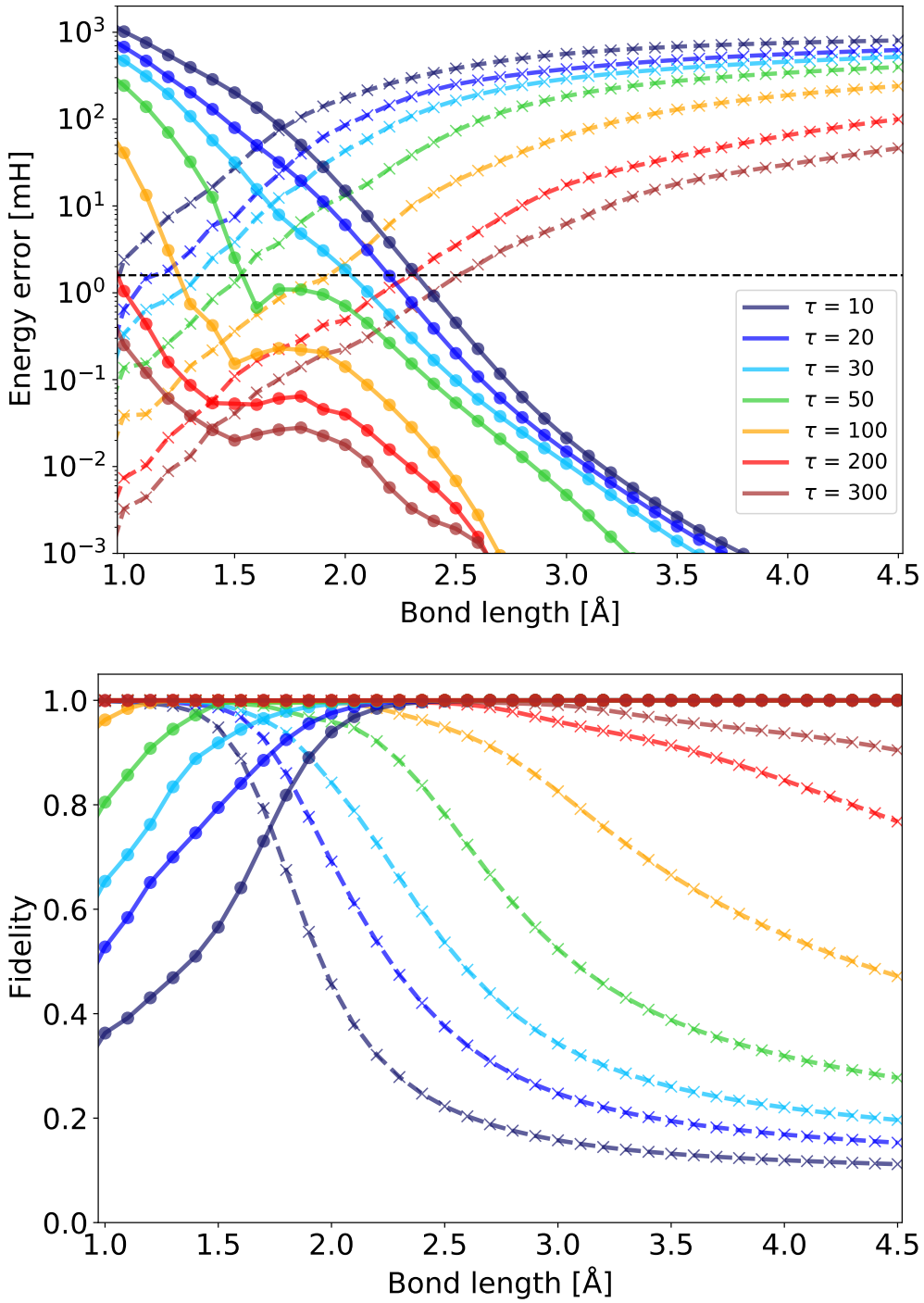


Figure 4.9: Adiabatic state preparation is greatly accelerated by using a spin-coupled wavefunction in the strong correlation limit. (Top) Error of energy estimates and (bottom) squared overlap between the exact ground state for N_2 and the adiabatically evolved states after full evolution ($s(\tau) = 1$). Solid lines correspond to the version of ASP that starts with the state $|\mathcal{O}^{6,1}_{0,0}\rangle$ and the Hamiltonian at dissociation, whereas dashed lines refer to ASP starting with the RHF state and the Fock operator. The total evolution time is τ, and the horizontal black line denotes chemical accuracy in the energy estimate.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Spin coupling is all you need: Encoding strong electron correlation in molecules on quantum computers
Publication Status	☒Published □ Accepted for Publication □Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	D. Marti-Dafcik, H. G. A. Burton, D. Tew, "Spin coupling is all you need: Encoding strong electron correlation in molecules on quantum computers", Physical Review Research **7**, 013191 (2025)

Student Confirmation

Student Name:	Daniel Marti Dafcik		
Contribution to the Paper	D.M.D. and D.P.T. conceived the project. D.M.D. developed the theoretical framework and algorithms. D.M.D. and H.G.A.B. wrote code to implement the numerical methods and carried out the computations. D.M.D. wrote the initial draft of the manuscript. All authors contributed to the revision and final version of the manuscript and provided ideas and guidance throughput the completion of the project.		
Signature 		Date	30 September 2025

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof. David P. Tew		
Supervisor comments		
Signature 	Date	1/10/2025

This completed form should be included in the thesis, at the end of the relevant chapter.

5

Exact electronic states with shallow quantum circuits from global optimisation

Contents

5.1 Introduction 178
5.2 Results 181
5.2.1 Symmetry-preserving unitary product states 181
5.2.2 Discretely optimised variational quantum eigensolver . . 182
5.2.3 Analysis of computational scaling 185
5.2.4 Exact wave functions using DISCO-VQE 186
5.2.5 Applicability to open-shell molecules 188
5.2.6 Accuracy of truncated approximations 189
5.2.7 Maximising the efficiency of quantum circuits 192
5.2.8 Balancing weak and strong electron correlation 193
5.3 Discussion 195

Abstract

Quantum computers promise to revolutionise molecular electronic simulations by overcoming the exponential memory scaling. While electronic wave functions can be represented using a product of fermionic unitary operators, the best *ansatz* for strongly correlated electronic systems is far from clear. In this contribution, we construct universal wave functions from gate-efficient, spin symmetry-preserving fermionic operators by introducing an algorithm that globally optimises the wave function in the discrete ansatz design and continuous parameter spaces. Our approach maximises the accuracy that can be obtained with near-term quantum circuits and provides a practical route for designing *ansätze* in the future. Numerical simulations for strongly correlated molecules, including water and molecular nitrogen, and the condensed-matter Hubbard model, set a new standard for the accuracy of gate-efficient quantum circuits for simulating strongly correlated chemistry.

5.1 Introduction

Computing molecular properties and processes relies on solving the Schrödinger equation to find the many-electron wave function. An exact solution to this problem formally scales exponentially with the number of electrons. Therefore, current high-accuracy algorithms with polynomial scaling, including coupled-cluster and many-body perturbation theory, fail when the electronic entanglement and correlation cannot be easily approximated,[285] for example, in open-shell systems with strong spin coupling, such as stretched molecular bonds and transition metal complexes.[160] For this reason, universally accurate simulations of complex processes, such as metalloenzyme catalysis and high-temperature superconductivity, remain difficult.[34] Gate-based quantum computation has the potential to overcome this challenge by representing the electronic wave function using polynomially scaling quantum resources, providing a route for solving strongly correlated electronic structure problems.[382]

Current and near-term quantum devices are impaired by noise and are limited to shallow quantum circuits.[32] The most promising near-term approaches optimise the parameters of an ansatz for the quantum state, and algorithms such as the variational quantum eigensolver[91] (VQE) are expected to outperform classical computational methods with modest quantum resources. However, the best ansatz for strongly correlated wave functions is far from clear, raising the question of how near-term quantum computing can usefully address unsolved electronic structure problems in chemistry. These questions are further complicated by the lack of numerical insight into the mathematical limits of different ansätze, such as how the accuracy scales with the number of operators or circuit depth.

We present a method for discovering highly compact ansätze by combining products of spin symmetry-preserving fermionic operators with a new global optimisation algorithm, the *Discretely Optimised Variational Quantum Eigensolver* (DISCO-VQE), which optimises both the discrete sequence of operators and their continuous variational parameters. Numerical simulations demonstrate that accurate wave functions for strongly correlated electronic systems can be parametrised

with efficient quantum circuits. These advances facilitate the design of high-accuracy quantum ansätze for solving challenging electronic structure problems on quantum devices, which will provide a deeper theoretical understanding of how these mathematical approximations encode strong electron correlation.

Ansatz-based approaches apply a unitary transformation $\hat{U}$ to an initial state $|\Phi_0\rangle$ in the form of a quantum circuit, constructed through a series of parametrised building blocks as $\hat{U} = \hat{U}_1(\theta_1)\cdots\hat{U}_N(\theta_N)$. The parameters θ_i are optimised on a classical computer by extremising an objective function evaluated using the quantum computer. This product structure defines a family of wave functions that we term unitary product states (UPS), and encompasses every quantum computational chemistry ansatz that has been proposed to date.

The ideal unitary product state should satisfy:

- Universality — describe any electronic eigenstate;
- Practicality — minimise circuit depth with few entangling gates;
- Symmetry — preserve good quantum numbers for operators that commute with the Hamiltonian.

Currently, the most popular ansätze use either "hardware-efficient" unitary operators constructed from qubit operations that minimise the depth of a specific quantum circuit,[260] or "physically-motivated" fermionic unitary operations[237] based on the second-quantised representation of electronic systems.[285] Second-quantised fermionic operators conserve particle number and Pauli antisymmetry, ensuring that truncated approximations predict accurate properties beyond the energy, and providing suitable initial states for fault-tolerant quantum phase estimation, where efficiency depends on the initial overlap with exact states. However, fermionic operators require multiple gate operations that lead to deep quantum circuits, creating a trade-off between symmetry-preservation and gate efficiency.[237]

Initial developments based on fermionic operators included the unitary coupled cluster approach[91, 237] (UCC), inspired by the success of coupled cluster theory for classical computing.[35] The UCC unitary transformation is an exponential

of a sum of anti-Hermitian fermionic operators, usually restricted to one- and two-body excitations.[237] However, since fermionic operators generally do not commute, representing this unitary transformation as a quantum circuit requires a Trotter approximation, which is only exact in the infinite limit: Instead, focus has shifted to a particular UPS form called disentangled UCC.[383] While it has been shown theoretically that exact states can be represented using a suitably-ordered infinite product of one- and two-body fermionic operators,[383] finite expansions are sensitive to the operator order, fail for strong correlation,[384, 385] and do not conserve the $\langle \hat{S}^2 \rangle$ quantum number.[386] In our view, refering to these states as UCC is no longer appropriate since the connection to the many-body cluster operator is not retained in practice. The ADAPT-VQE optimisation algorithm,[261] and its extensions,[262, 387–389] iteratively build a UPS from a pool of operators, selecting the operator with the largest energy improvement at each step. Although promising, variants of ADAPT-VQE that use qubit operators may not converge to the exact ground state,[390] and current qubit and fermionic implementations are not guaranteed to conserve spin symmetry.[389, 391] Furthermore, as we show in this work, the ADAPT-VQE optimisation strategy will generally not identify the best representation with the fewest number of operators from a given pool.

Here we describe an approach that fulfils the ansatz requirements using a symmetry-preserving unitary product state (s-UPS) that conserves Pauli antisymmetry, particle number symmetry, and the $\langle \hat{S}_z \rangle$ and $\langle \hat{S}^2 \rangle$ quantum numbers using a fermionic operator pool. We prove that this ansatz is universal and demonstrate that it is gate-efficient. Identifying the appropriate choice of the operators is system-dependent and challenging. We solve this problem by introducing the DISCO-VQE algorithm, which simultaneously optimises the operator sequence and their variational parameters. Our approach discovers the most accurate s-UPS representation with a minimal number of parameters, demonstrating that high-accuracy gate-efficient quantum circuits can be achieved, and providing a practical route for designing highly accurate ansätze in the future.

5.2 Results

5.2.1 Symmetry-preserving unitary product states

Fermionic unitary operators correspond to an exponential of anti-Hermitian second-quantised operators $\hat{\kappa}^{rs\cdots}_{pq\cdots} = a^\dagger_r a^\dagger_s \cdots a_q a_p - a^\dagger_p a^\dagger_q \cdots a_s a_r$. These operators ensure Pauli antisymmetry, conserve particle number, and can be mapped onto qubit operators using fermion-to-qubit encodings.[70, 239] We define a symmetry-preserving unitary product state (s-UPS) of length M as

$$|\Psi(\boldsymbol{t}, \boldsymbol{\mu})\rangle = \prod_{i=1}^{M} e^{t_i \hat{\kappa}_{\mu_i}} |\Phi_0\rangle \,, \tag{5.1}$$

where $|\Phi_0\rangle$ is an arbitrary initial state, and only spin-adapted generalised one-body and paired two-body operators are allowed, that is $\hat{\kappa}_{\mu_i} \in \{\hat{\kappa}^q_p + \hat{\kappa}^{\bar{q}}_{\bar{p}}, \hat{\kappa}^{q\bar{q}}_{p\bar{p}}\}$. The indices p, q (where $p < q$) denote arbitrary spin orbitals and the absence (presence) of an overbar indicates a high-spin (low-spin) orbital. Individual operators may appear multiple times with different amplitudes, defined by the continuous coordinates $\boldsymbol{t} = (t_1, \ldots, t_M)$. Because the $\hat{\kappa}_{\mu_i}$ need not commute, the ansatz depends on the sequence of operators, denoted by the ordered indices $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_M)$.

Using generalised fermionic operators, where p and q represent arbitrary spin orbitals, removes any dependence on the separation of occupied or virtual orbitals in a mean-field reference. In contrast to many-body perturbation theory, these operators represent Lie algebra generators[392] for unitary transformations in the D-dimensional Hilbert space and can rotate any initial state into any exact state. Universality of the generalised (unpaired) one- and two-body fermionic operators has been proved by realising that their nested commutators can generate the full Lie algebra.[383] Since a product of unitary operators corresponds to an exponential of a sum of operators and their commutators, $e^{\hat{A}} e^{\hat{B}} = e^{\hat{A} + \hat{B} + \frac{1}{2}[\hat{A}, \hat{B}] + \cdots}$, combining enough operators in a suitable order can represent any unitary transformation in the Hilbert space.[272] The universality of the s-UPS ansatz can be proved using the additional observation that generalised two-body operators can be expressed as nested commutators of one-body and paired two-body operators, as we show in

Appendix C.1. This proof requires the reference state to have at least one pair of doubly occupied and doubly unoccupied orbitals (Appendix C.1.1). We postulate that $M \leq (D-1)$ will allow any initial state to be rotated into an exact state without redundant parameters, and provide numerical examples that support this hypothesis. In practice, we expect sufficient accuracy with smaller M values.

Since these elementary operators are spin-adapted (the commutators $[\hat{S}^2, \hat{\kappa}]$ and $[\hat{S}_z, \hat{\kappa}]$ vanish) the s-UPS ansatz conserves the spin symmetry of the initial state for every truncation M, avoiding spin-constrained optimisation[386] or spin projection.[389] In contrast, spin-adapted unpaired two-body operators require the exponentiation of a sum of non-commuting operators, and their application in a quantum circuit requires a Trotter approximation, which destroys the spin adaptation.[386] Furthermore, paired two-body operators $\hat{\kappa}_{p\bar{p}}^{q\bar{q}}$ can be implemented with a constant number of two-qubit controlled-NOT (CNOT) gates, while the CNOT count for unpaired two-body operators grows linearly with the number of orbitals (Appendix C.2).

The s-UPS ansatz can be viewed as a generalisation of the quantum-number-preserving (QNP) approach[344] and the swap-network k-UpGGCSD framework,[169] where spin-adapted "orbital rotation" gates (generalised spin-adapted one-body operators) and "pair exchange" gates (paired two-body operators) between nearest-neighbour qubits are used to define low-depth quantum circuits with local connectivity. Since we include operators connecting any pair of orbitals, and allow a variable operator ordering, we expect greater accuracy with fewer parameters. Furthermore, while universality of the QNP ansatz was only postulated and demonstrated numerically,[344] our fermionic representation has allowed a rigorous proof.

5.2.2 Discretely optimised variational quantum eigensolver

Fixed operator orderings are typically poor approximations to the optimal unitary product state and very little is understood about what constitutes a good ordering.[261, 393] Without *a priori* information, the optimal operator ordering can be approached as a coupled discrete and continuous optimisation problem

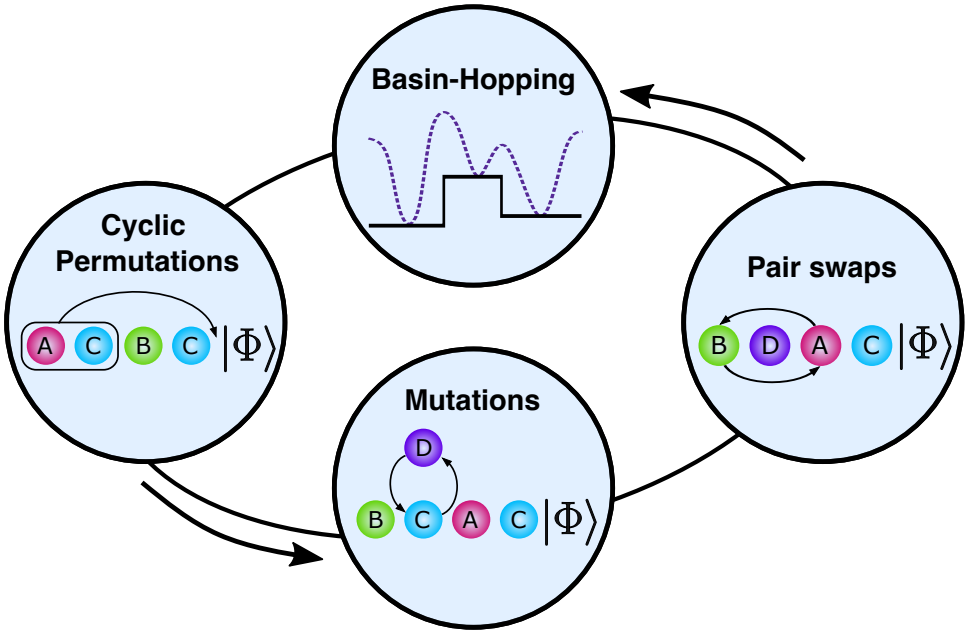


Figure 5.1: DISCO-VQE is a generalised basin-hopping approach[399, 400] that optimises the continuous coordinates $\boldsymbol{t}$ and the discrete space of operator orders $\boldsymbol{\mu}$ to identify the best s-UPS wave function using cyclic permutations, mutations, and pair swaps of the unitary operators.

for the continuous operator amplitudes $\boldsymbol{t}$ and the discrete ordered operator set $\boldsymbol{\mu}$. Exploiting the variational principle, the optimal s-UPS representation can be identified by minimising the energy

$$E(\boldsymbol{t},\boldsymbol{\mu}) = \langle\Phi_0|\left(\prod_{i=M}^{1} e^{-t_i\hat{\kappa}_{\mu_i}}\right)\hat{H}\left(\prod_{i=1}^{M} e^{t_i\hat{\kappa}_{\mu_i}}\right)|\Phi_0\rangle\,. \tag{5.2}$$

The non-linearity of $E(\boldsymbol{t},\boldsymbol{\mu})$ with respect to $\boldsymbol{t}$ creates local minima that hinder global optimisation.[385, 394] Previous discrete optimisation strategies include mapping the problem to a probabilistic framework,[395] genetic algorithms,[396, 397] or an iterative gate-by-gate optimisation algorithm.[398] However, stochastic approaches are not guaranteed to find the global minimum, while the gate-by-gate approach cannot escape local discrete minima, and the strong interdependence of the operators means that crossover steps in genetic algorithms are less likely to lower the energy.[396] Instead, we have developed the *Discretely Optimised Variational Quantum Eigensolver* (DISCO-VQE), which facilitates global optimisation by taking steps in both the continuous and discrete spaces for the s-UPS ansatz.

For a given $\boldsymbol{\mu}$, we define a local discrete neighbourhood as the operator sequences connected by:

- cyclically permuting the ordered set;
- mutating one operator into another from the pool;
- swapping the position of two operators.

Following the generalised basin-hopping (GBH) approach,[399–401] we define a *biminimum*[400] as a configuration $(\boldsymbol{t}, \boldsymbol{\mu})$ that is a minimum in the continuous parameters and lies lower in energy than neighbouring minima after a step in the discrete space, including relaxation of the continuous parameters. DISCO-VQE alternates between a series of basin-hopping[402, 403] steps that explore the continuous energy landscape for a fixed set of operators, and discrete steps in the space of operator orderings (Fig. 5.1). After each round of basin-hopping steps (Appendix C.3.1), we identify the lowest-energy configuration that can be reached by cyclically permuting the operators and reoptimisation of the continuous coordinates. Similarly, we identify the lowest energy mutation and pair swap for each operator and accept this step if it lowers the energy (Appendix C.3.2).

By testing mutations and pair swaps for every operator individually, DISCO-VQE can completely change the discrete parameters in each macrocycle and, in principle, the optimal operator sequence could be discovered after one macro iteration. In practice, the combined space of continuous and discrete parameters exhibits many local biminima and uphill moves are required to escape these traps (Appendix C.3.3). The full algorithm is detailed in Appendix C.3. To avoid a bias towards any particular region of the discrete space, the search is initialised with identity operators that represent the initial state $|\Phi_0\rangle$. In the future, the use of 'informed' guesses based on intuition into the best operator order may improve the optimisation efficiency.

5.2.3 Analysis of computational scaling

Obtaining high-accuracy wave functions using adaptive algorithms, such as DISCO-VQE and ADAPT-VQE, increases the computational cost relative to a fixed ansatz approach. For M operators chosen from a pool of size N, the computational scaling of DISCO-VQE is determined by the cost of the operator mutations and pair swaps. One cycle of operator mutations tests whether the energy is lowered when each of the M operators is replaced with another operator from the pool, giving a scaling of $\mathcal{O}(MN)$. Similarly, one cycle of pair swaps considers every possible way to interchange the current operators, with a scaling of $\mathcal{O}(M^2)$. The overall scaling of the discrete search is therefore $\mathcal{O}(\max(M^2, MN))$, allowing the combinatorial space of operator sequences to be searched with polynomial scaling in M and N.

Global optimisation allows highly accurate wavefunctions to be discovered with smaller operator pools and fewer variational parameters than methods such as ADAPT-VQE. To avoid stagnation, the standard fermionic ADAPT-VQE implementation[261] must consider all unpaired two-body operators with a pool that scales as $N \sim \mathcal{O}(n_\mathrm{q}^4)$, where n_q is the number of qubits. The discrete gradients needed to iteratively grow the ansatz scale as $\mathcal{O}(N)$, giving fermionic ADAPT-VQE an overall scaling of $\mathcal{O}(n_\mathrm{q}^4)$. In contrast, our spin-adapted one-body and paired two-body operator pool scales as $N \sim \mathcal{O}(n_\mathrm{q}^2)$. Assuming that chemical accuracy can be achieved with $M \leq N$, which is true for the examples we consider, the overall scaling of DISCO-VQE becomes $\mathcal{O}(Mn_\mathrm{q}^2)$, which is at worst the same as fermionic ADAPT-VQE with unpaired two-body operators. Further reductions may be achieved with additional restrictions, such as constraining the operators to act between nearest-neighbour qubits to give $N \sim \mathcal{O}(n_\mathrm{q})$. The current state of knowledge does not provide a prediction for how the number of operators M required to achieve a given accuracy will scale with the system size.

The computational cost of DISCO-VQE scales linearly with the number of basin-hopping steps in the continuous space and the number of discrete macrocycles, which are primarily determined by the amount of computing resources available. In this work, our aim is to calibrate the best achievable accuracy, so we use a large

number of cycles, as detailed in Appendix C. However, any number of basin-hopping steps must always be at least as accurate as the current best optimisation with the same initial guess, while the reoptimisation in the discrete search means that the continuous space can still be explored without any basin-hopping steps (Appendix C.4.2). Furthermore, because DISCO-VQE starts by selecting mutations for the initial empty set of operators, which is essentially the same as the ADAPT-VQE procedure, even a single discrete macrocycle will ensure that DISCO-VQE is at least as accurate as ADAPT-VQE with the same operator pool (Appendix C.4.2). Consequently, even for a single macrocycle, DISCO-VQE will always match the accuracy of current state-of-the-art VQE calculations, making it the best optimisation algorithm for any problem accessible to the VQE procedure.

5.2.4 Exact wave functions using DISCO-VQE

The universality of the s-UPS ansatz means that wave functions with an arbitrarily small energy error can be obtained using a finite number of spin-adapted one-body and paired two-body unitary operators. DISCO-VQE calculations on linear and tetrahedral H_4 show that chemical accuracy ($1.59\,\mathrm{m}E_\mathrm{h}$) requires only nine and five operators, respectively (Fig. 5.2, **A**–**B**). Furthermore, exact ground states can be identified using thirteen or eight operators, respectively, compared to a Hilbert space of 36 configurations. These results support our postulate that exact s-UPS representations require a finite number of operators M that is less than the Hilbert space size, in contrast to previous theoretical work that could only demonstrate universality in the limit $M \to \infty$.[383]

Practical applications generally require consistent accuracy for every relevant geometry of a molecule. DISCO-VQE achieves this goal for linear H_4 by identifying exact wave functions with thirteen operators at all geometries (Fig. 5.2 **C**). Fixing the ansatz to the ordering identified at equilibrium gives near-exact energies across the full binding curve (Fig. C.4), demonstrating that accurate and transferable fixed ansätze can be defined. The operator sequence and optimised parameters for this fixed ansatz are included in Appendix C.4.4, providing a high-accuracy

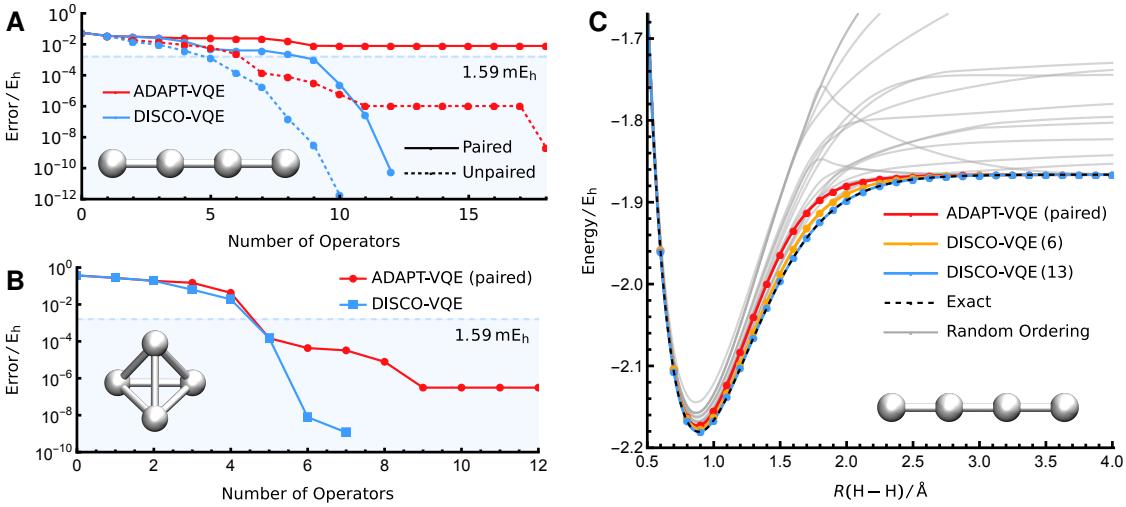


Figure 5.2: DISCO-VQE produces exact wave functions using global optimisation. The panels illustrate the convergence of DISCO-VQE and ADAPT-VQE with respect to the number of operators for **(A)** linear H_4, R(H−H) = 0.90 Å, and **(B)** tetrahedral H_4, R(H−H) = 1.98 Å, using the STO-3G basis set.[210] **(C)** DISCO-VQE identifies the exact wave function throughout the linear H_4 binding curve, while ADAPT-VQE and random ordered sets of thirteen operators exhibit a large variation in accuracy. DISCO-VQE surpasses the accuracy of ADAPT-VQE using only six operators, while chemically accurate binding curves require ten operators (Fig. C.5, Appendix C). Computational details for the DISCO-VQE and ADAPT-VQE calculations are provided in Appendix C.4.

quantum circuit for use in practical VQE experiments. In comparison, random fixed ansätze are relatively accurate near equilibrium, but exhibit a large energy variation towards dissociation (Fig. 5.2**C**). Furthermore, optimisation using ADAPT-VQE with the paired operator pool accurately describes equilibrium and dissociation, but gives appreciable errors at intermediate geometries, where electron correlation is most challenging (Fig. 5.2**C**).

Global optimisation of the continuous coordinates and discrete operator ordering is essential for realising the universality of the s-UPS ansatz. The improved algorithmic optimisation provided by DISCO-VQE is clear from comparison with ADAPT-VQE calculations using the same pool of operators. By adding one operator at a time, ADAPT-VQE performs a local optimisation, which may not converge to the exact ground-state wave function, as demonstrated in Fig. 5.2**A**–**B**. DISCO-VQE can escape these local minima and implements global optimisation for any sequence of operators.

Stagnation of the ADAPT-VQE optimisation was previously described in Ref. 390 and arises from combining the gradient-based selection criterion with particular operator pools. The restriction to paired two-body operators means that the ADAPT-VQE gradient[261] no longer coincides with the residual of the Anti-Hermitian Contracted Schrodinger Equation[404] and converged solutions are not necessarily eigenstates.[405] A near-exact ADAPT-VQE result for linear H_4 can be obtained by including unpaired two-body operators (Fig. 5.2**A**), although DISCO-VQE is still more accurate with fewer operators, and truncated wave functions exhibit spin contamination (Fig. C.3).

5.2.5 Applicability to open-shell molecules

High-spin molecules are an important target for quantum algorithms because conserving open-shell spin symmetry in standard methods, such as coupled-cluster theory, introduces significant computational complexity.[225] Rubin *et al.* have recently shown that ADAPT-VQE using the full unpaired operator pool struggles to converge to the high-spin triplet ground state of O_2 using the restricted open-shell HF (ROHF) reference state, while starting from a more accurate ROHF-

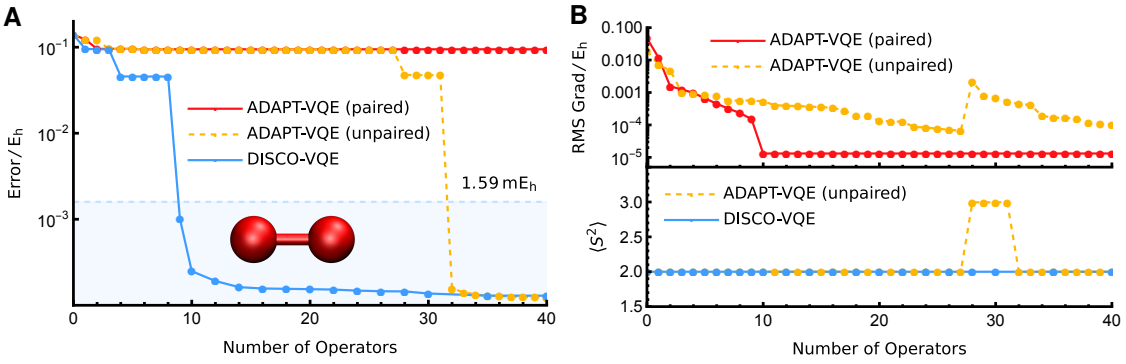


Figure 5.3: The s-UPS ansatz accurately predicts the high-spin triplet state of O_2 (STO-3G) at $R(\mathrm{O{-}O) = 2.55}$ Å while preserving the $\langle S^2 \rangle$ quantum number. Two core molecular orbitals are frozen. (**A**) DISCO-VQE optimisation achieves chemical accuracy with only nine operators from the ROHF reference state, while ADAPT-VQE with the paired operator pool fails to converge to the correct state. (**B**) ADAPT-VQE using the full unpaired operator pool can become spin contaminated and exhibits a plateau in the energy with respect to the number of operators, which creates a minimum in the magnitude of the ADAPT-VQE discrete gradient. Computational details for DISCO-VQE and ADAPT-VQE calculations are provided in Appendix C.5.

CCSD state introduces spin contamination.[391] These issues arise from a plateau in the energy with respect to the number of operators (Fig. 5.3**A**), which causes the magnitude of the discrete gradient to temporarily fall below the convergence threshold used in Ref. 391 (Fig. 5.3**B**: top panel). Adding more operators eventually causes the energy to be decrease, although 32 operators are ultimately required to reach chemical accuracy.

In contrast, the s-UPS ansatz provides a chemically accurate representation of the O_2 ground state using only nine operators from a Hilbert space of 448 configurations. Remarkably, this highly accurate open-shell s-UPS approximation is achieved despite only including spin-adapted one-body and paired two-body operators in the pool, with no modifications to the DISCO-VQE procedure. As for the H_4 results, the importance of global optimisation in the discrete space is demonstrated by the stagnation of ADAPT-VQE calculations using the paired operator pool. Furthermore, enforcing spin symmetry at the operator level ensures that every s-UPS truncation preserves the $\langle S^2 \rangle = 2$ property of the ROHF reference state, guaranteeing a spin-pure representation (Fig. 5.3**B**: bottom panel). Hence, universal wave functions for any spin state can be achieved using the s-UPS approach.

5.2.6 Accuracy of truncated approximations

Practical quantum computations on large molecules will still require truncated wave function approximations. Global optimisation using DISCO-VQE provides the most accurate s-UPS for a given number of operators, while conserving spin symmetry at every truncation. For example, DISCO-VQE calculations on linear H_6, with a Hilbert space of 400 configurations, achieve chemical accuracy with 30 operators (Fig. 5.4**A**) while ADAPT-VQE optimisation stagnates between 14 and 40 operators (Fig. 5.4**A**). DISCO-VQE requires 10 operators to surpass the accuracy of ADAPT-VQE using the paired operator pool (Fig. 5.4**B**) and significantly improves the energy using fewer parameters compared to previous fermionic

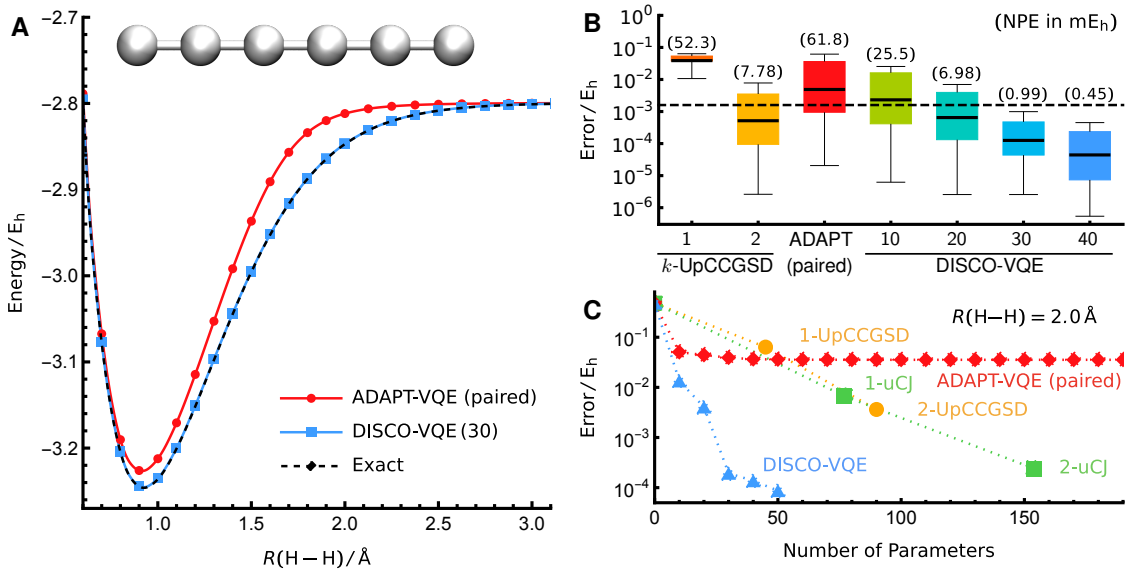


Figure 5.4: (**A**) DISCO-VQE with 30 operators provides chemically accurate energies for the linear H_6 binding curve (STO-3G). This number of operators is chosen to match the size of the operator pool. (**B**) Comparison of DISCO-VQE errors and the previous state-of-the-art methods across the full binding curve. ADAPT-VQE calculations use the same paired operator pool. The un-Trotterised 1-UpCCGSD and 2-UpCCGSD results are taken from Ref. 385. The non-parallelity error (NPE) is the difference between the maximum and minimum error in the energy. Boxes illustrate the upper and lower quartiles, and whiskers represent the maximum and minimum values. (**C**) DISCO-VQE produces highly accurate energies at R(H−H) = 2.0 Å with fewer continuous parameters than previous algorithms. The k-uCJ results are taken from Ref. 406. The accuracy of the 2-UpCCGSD[385] or 2-uCJ[406] approaches is surpassed with 30 operators, while ADAPT-VQE with the full unpaired operator pool required at least 78 operators to obtain comparable results.[261] Computational details for the DISCO-VQE and ADAPT-VQE calculations are provided in Appendix C.6.

operator methods (Fig. 5.4**C**). Hence, DISCO-VQE defines a new standard for accuracy with a specified number of fermionic operators in linear H_6.

Symmetrically stretched H_2O and N_2 dissociation represent strongly correlated molecular tests for wave function approximations, with Hilbert space dimensions of 441 and 400, respectively, when the four lowest molecular orbitals of N_2 are frozen. DISCO-VQE calculations with a truncated s-UPS ansatz produce highly accurate energies across the H_2O binding curve (Fig. 5.5**A**). In contrast, ADAPT-VQE optimisation using the paired operator pool fails to identify chemically accurate wave functions at any bond length (Fig. 5.5**A**), while previous studies that include unpaired two-body operators are much less accurate than our approach.[407] The

un-Trotterised k-UpCCGSD ansatz required 126 variational parameters to reach an accuracy of 0.07 mE_h,[393] while Trotterised calculations with over 250 operators have residual errors of 0.02–200 mE_h.[183] Our global optimisation strategy significantly increases the accuracy with fewer operators than the previous state-of-the-art algorithms.

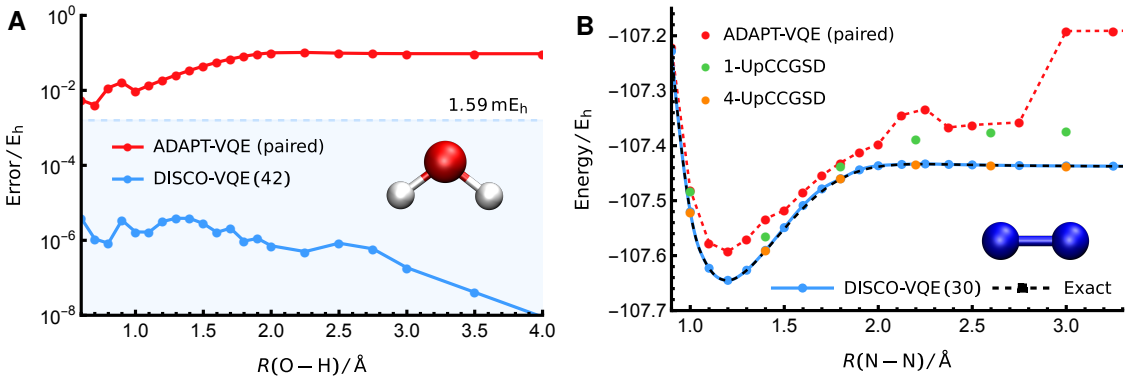


Figure 5.5: Wave functions built using DISCO-VQE accurately describe the potential energy surface for strongly correlated molecules. The number of operators in these DISCO-VQE calculations is chosen to match the size of the operator pool in the STO-3G basis set, with **(A)** 42 operators for the symmetric stretch of H_2O and **(B)** 30 operators for the dissociation of N_2. The four lowest energy molecular orbitals are frozen in N_2. Un-Trotterised k-UpCCGSD results for N_2 are taken from Ref. 393. Computational details for the DISCO-VQE and ADAPT-VQE calculations are provided in Appendix C.7 and C.8.

DISCO-VQE achieves chemical accuracy at nearly every N_2 bond length using 30 operators, giving a very accurate binding curve (Fig. 5.5**B**). In comparison, UPS representations built from randomly ordered sets that include all two-body operators have errors of 16–64 mE_h,[385] demonstrating the importance of global optimisation that treats both the discrete operator space and the continuous space of operator amplitudes. The parameter reduction is highlighted by comparing to the un-Trotterised k-UpCCGSD results, which require over 180 parameters to obtain a better mean-average error[393] (Table C.2). In contrast, ADAPT-VQE optimisation with the paired operator pool does not produce a smooth binding curve (Fig. 5.5**B**), while including unpaired two-body operators breaks the total spin symmetry for truncated approximations.[389] The s-UPS ansatz conserves the initial $\langle \hat{S}^2 \rangle$ value at every truncation, and DISCO-VQE global optimisation

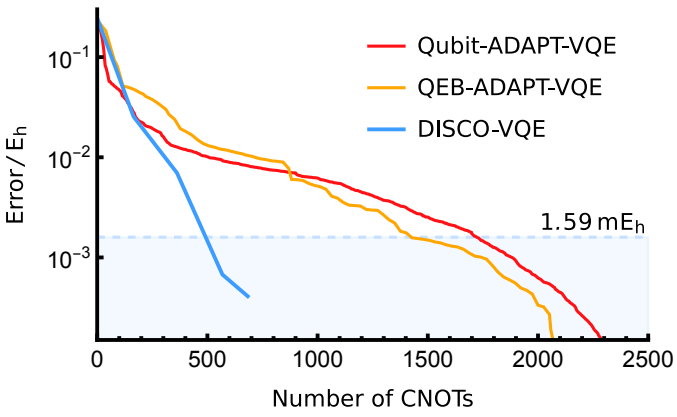


Figure 5.6: DISCO-VQE locates accurate and efficient quantum circuits for linear H_6 at $R(\mathrm{H}{-}\mathrm{H}) = 1.5\,\text{Å}$ with significantly fewer CNOT gates than the previous state-of-the-art. Data for the qubit-ADAPT-VQE and QEB-ADAPT-VQE methods are taken from Ref. 262. CNOT counts are evaluated using the efficient circuit design introduced in Ref. 408 (Appendix, section S2.3).

achieves a highly accurate and operator efficient description of this strongly correlated electronic system.

5.2.7 Maximising the efficiency of quantum circuits

Applications on near-term quantum devices must balance accuracy against the number of two-qubit CNOT gates in the quantum circuit, which contribute the largest source of hardware noise. This challenge has motivated hardware-efficient ansätze where the operator pool requires few CNOT gates, such as qubit-ADAPT-VQE[387] and qubit-excitation-based (QEB) ADAPT-VQE.[262] However, these methods can break fermionic antisymmetry, particle number, or spin symmetry, giving unphysical electronic states.

Despite employing fermionic operators, DISCO-VQE achieves chemical accuracy for linear H_6 using less than a third of the CNOT gates required for hardware-efficient ansätze (Fig. 5.6; Appendix C.2.3). These favourable CNOT counts are achieved by removing redundancies from the quantum circuit using discrete optimisation and through the CNOT efficiency of paired two-body operators (Appendix C.2.2). Therefore, efficient quantum circuits can be achieved using symmetry-preserving fermionic operators without relying on a hardware-efficient ansatz, creating a new paradigm for quantum-compatible electronic algorithms.

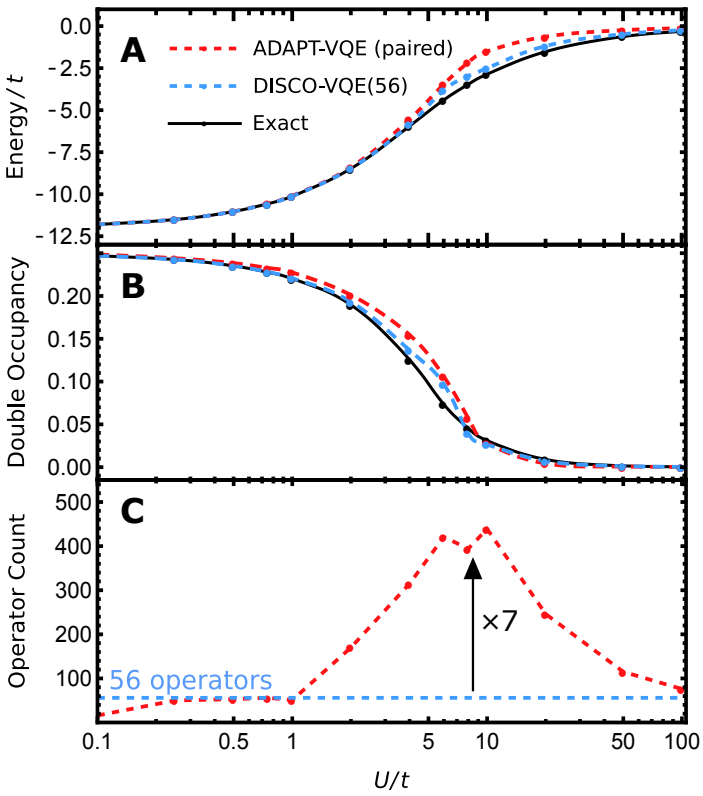


Figure 5.7: DISCO-VQE captures the correct physics for a strongly correlated two-dimensional Hubbard lattice (4×2) using 56 operators. The corresponding Hilbert space contains 4900 configurations, and the number of operators is chosen to match the size of the operator pool. The DISCO-VQE energy (**A**) is very accurate in both the weak correlation ($U \ll t$) and strong correlation ($U \gg t$) regimes. Similarly, DISCO-VQE accurately predicts the double occupancy (**B**), which becomes exact in the weak and strong correlation limits. Analogous ADAPT-VQE calculations using the paired operator pool require seven times more operators at intermediate correlation but give less accurate results (**C**). Computational details for DISCO-VQE and ADAPT-VQE calculations are provided in Appendix C.9.

5.2.8 Balancing weak and strong electron correlation

Balancing weak and strong electron correlation is essential for quantum chemistry, but has proved difficult to achieve in practice. Many-body perturbation techniques can be quantitatively accurate for weak correlation, but fail for strong correlation where the molecular orbital picture breaks down. The accuracy of the DISCO-VQE results across the H_6, H_2O, and N_2 potential energy surfaces demonstrate that the s-UPS ansatz can quantitatively describe both correlation regimes.

The strongly correlated half-filled Hubbard lattice represents a challenging condensed matter physics problem for many-body approximations.[409] The corre-

lation strength is determined by the balance of the on-site repulsion U and the one-electron adjacent site hopping term t. DISCO-VQE calculations with a truncated s-UPS wave function provide excellent accuracy in both the weak ($U \ll t$) and strong ($U \gg t$) correlation limits (Fig. 5.7**A**) using a consistent number of operators and parameters. ADAPT-VQE optimisation with the paired operator pool provides an accurate representation of weak correlation, but is less accurate than DISCO-VQE as the correlation strength increases. The converged ADAPT-VQE wave functions include significantly more operators than DISCO-VQE for intermediate correlation ($U/t \approx 10$), but are less accurate (Fig. 5.7**C**). We see that global optimisation using DISCO-VQE is essential for obtaining quantitatively accurate s-UPS wave functions for both weak and strong correlation with a consistent quantum circuit cost.

Physical properties beyond the energy are also important for understanding quantum phase behaviour. The strongly correlated half-filled Hubbard model is characterised by an antiferromagnetic phase, where the electrons are unpaired and localised on individual sites. The double occupancy $\langle \hat{n}_p \hat{n}_{\bar{p}} \rangle$ measures the probability of simultaneously finding two electrons on the same site, which tends to zero for $U/t \to \infty$ and 0.25 for $U/t \to 0$. Truncated s-UPS representations identified using DISCO-VQE accurately predict this change across different U/t values (Fig. 5.7**B**) and exhibit the correct limits. Hence both the wave function and the energy can be accurately predicted, providing a physically complete description of strong electron correlation.

Finally, the 1/8-filling fraction is more challenging for standard quantum Monte–Carlo techniques due to the fermionic sign problem.[410] DISCO-VQE calculations for the 4×2 lattice (Appendix C.9.1) achieve exact wave functions using only seven operators (Fig. 5.8). This number is significantly lower than the 56 operators used for the inexact half-filled result, suggesting that the performance of the s-UPS ansatz for a fixed number of operators is primarily related to the size of the relevant number- and S_z-preserving Hilbert space sector, which is 64-dimensional

in this case. These results provide further evidence for the applicability of our approach across various different electron correlation regimes.

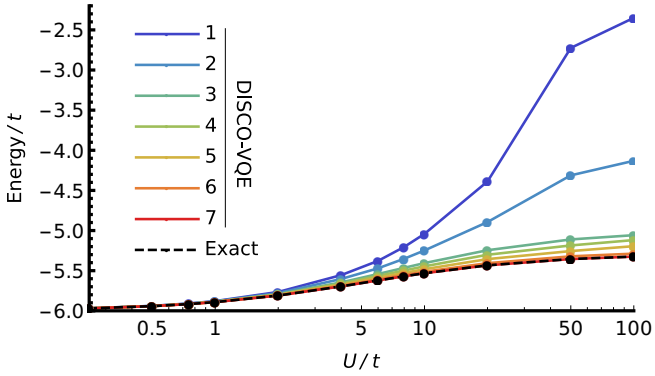


Figure 5.8: DISCO-VQE can identify a numerically exact representation of the 1/8-filled 4×2 Hubbard model using seven operators, demonstrating the applicability across various different electron correlation regimes.

5.3 Discussion

We have shown that an arbitrary electronic state can be parametrised with the s-UPS ansatz built from spin-adapted one-body and paired two-body fermionic operators. This ansatz conserves fermionic antisymmetry, particle number, and the $\langle \hat{S}^2 \rangle$ and $\langle \hat{S}_z \rangle$ quantum numbers without constrained optimisation or spin projection, and can parametrise any spin state. The flexibility and accuracy of these wave functions comes from identifying the best sequence of unitary operators. DISCO-VQE is the first algorithm that can universally address this challenge by performing a coupled discrete and continuous global optimisation of the wave function. Numerical simulations demonstrate that chemically accurate energies can be achieved for weakly and strongly correlated molecules using significantly fewer operators and CNOT gates than previous state-of-the-art techniques.

The s-UPS ansatz offers key advantages for practical quantum algorithms: it can parametrise arbitrarily accurate wave functions; preserves the Hamiltonian spin symmetries at every truncation; and provides low CNOT counts suitable for noisy quantum hardware. Conserving physical symmetries is essential for subsequent fault-tolerant quantum phase estimation, where efficiency depends on the overlap

of the initial and exact states. Furthermore, formulating generalised fermionic operators as Lie algebra generators for unitary rotations moves beyond the many-body unitary coupled cluster framework and embraces the natural capabilities of qubit rotations in quantum computation, bridging the gap between fermionic operators and hardware-efficient approaches such as the quantum-number-preserving ansatz.[344] These advances set a new standard for quantitative, physically accurate, and gate-efficient quantum circuits for simulating strongly correlated chemistry using near-term quantum computers.

We have not explicitly considered spatial symmetry in the present calculations, meaning that the spatial symmetry of the initial state may not be preserved, because the unitary operators are applied sequentially. In principle, for point groups with non-degenerate irreducible representations, conservation of spatial symmetry can be trivially enforced by restricting the one-body operators to act between orbitals with the same spatial symmetry.[1] However, the effect of this restriction on the convergence of DISCO-VQE is beyond the scope of the current work, and we leave this for future investigations.

The success of many proposed quantum algorithms relies on the accuracy of the initial state, which must be prepared at a polynomial cost. Whether this accuracy is achievable using current UPS ansätze remains an open question and depends on: how the ansatz expressibility grows with the number of operators and parameters; whether strong correlation can be parametrised with low circuit depth; and how the number of operators scales with the system size for a given level accuracy. Definitive answers to these questions will require further work. Here we have shown for the first time how global optimisation using DISCO-VQE numerically provides reliable lower bounds on the minimum number of operators and circuit depth required for a given operator pool, providing new insight into ansatz design. Our results clearly demonstrate how a combination of theory and numerical simulations can produce high-accuracy ansätze to treat strong correlation on quantum devices.

[1] Paired two-body operators automatically conserve spatial symmetry.

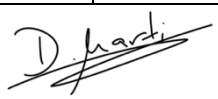
Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Exact electronic states with shallow quantum circuits from global optimisation
Publication Status	☒Published □ Accepted for Publication □Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	H. G. A. Burton, D. Marti-Dafcik, D. P. Tew, D. J. Wales, "Exact electronic states with shallow quantum circuits from global optimisation", npj Quantum Information **9**, 75 (2023)

Student Confirmation

Student Name:	Daniel Marti Dafcik		
Contribution to the Paper	H.G.A.B., D.M.D., and D.J.W. co-designed the discrete optimisation algorithm. D.M.D., H.G.A.B. and D.P.T. co-developed and analysed the wavefunction ansatz. H.G.A.B. and D.J.W. wrote the DISCO-VQE code and H.G.A.B. carried out the simulations. D.M.D. and H.G.A.B. performed data analysis and wrote code to generate the quantum circuit representation of the ansatz. All authors contributing to writing the manuscript.		
Signature		Date	30 September 2025

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof. David P. Tew		
Supervisor comments		
Signature 	Date	1/10/2025

This completed form should be included in the thesis, at the end of the relevant chapter.

6
Conclusions

Quantum computers promise to solve some of the most demanding electronic structure problems, particularly those involving strongly correlated electrons. Accurate simulation of such systems could enable the design of highly-entangled quantum matter such as room-temperature superconductors [34] and catalysts or drugs based on transition metal clusters.[64, 65, 106] Currently, resolving the electronic structure of some of these systems remains out of reach for the most powerful classical computers.[112] Despite the exponential memory advantage of quantum computers over classical computers, the performance of quantum algorithms also remains limited for strongly correlated systems due to the curse of dimensionality.[39–41]

Classical electronic structure methods mitigate the exponential growth of the Hilbert space by introducing approximations and heuristics that focus on physically relevant regimes, biasing the algorithms towards small corners of the Hilbert space. Although this manifests differently, quantum algorithms face the same underlying need: without guidance from such heuristics, they inherit the inefficiencies of brute-force search. Yet, for strongly correlated systems, suitable heuristics have been scarce, perhaps precisely because classical approaches struggle in this regime. Developing them has therefore become essential for both classical and quantum algorithmic progress.

In Chapter 3, we introduced a heuristic framework for identifying and parametrising the wavefunctions of strongly correlated electrons. The resulting spin-coupled molecular orbital theory approximates molecular eigenstates through a linear combination of few configuration state functions that explicitly encode spin coupling effects. In Chapter 4, we demonstrated how to efficiently prepare these states on a quantum computer and showed that they significantly enhance the performance of existing quantum algorithms. Together, these methods establish a promising route towards tackling the strong electron correlation problem on quantum computers.

In Chapter 5, we addressed the problem of running quantum algorithms on NISQ devices with very limited coherence times. To maximise the efficiency of variational wavefunction parametrisations, we proposed a symmetry-preserving ansatz and DISCO-VQE, a coupled continuous-discrete optimisation algorithm that aims to find the optimal quantum circuit for preparing molecular ground states through products of parametrised unitary operators. Besides serving as a numerical method for the time-independent Schrödinger equation, this approach reveals qualitative insights into quantum ansätze. Notably, we found that, given a flexible choice of ansatz, quantum parametrisations can describe both strongly correlated and weakly correlated electrons using a similar amount of degrees of freedom, a feature which is generally not true for classical wavefunction parametrisations.[160] Although expectations of near-term quantum advantage via VQE have diminished since the time when we began this line of research,[44] the algorithm could remain valuable as a state preparation routine for fault-tolerant quantum computations.

The CSF-based framework presented in Chapters 3 and 4 remains highly promising but requires extensions to address more complex systems. While our numerical demonstrations focused on stretched bonds, we expect these ideas to generalise to other strongly correlated molecular systems including transition metal clusters.[51] We are actively pursuing this direction and hope that our work will encourage both the classical and quantum communities to develop improved heuristics for strong electron correlation. We also emphasise the importance of identifying additional

systems that are both strongly correlated and of clear scientific or technological relevance, since these represent the most plausible regimes where quantum algorithms could provide a meaningful advantage and impact over classical methods.[41]

Another direction that we have begun exploring is the application of this methodology to finer-grained wavefunction discretisations, i.e., larger basis sets. The set-up is relatively straightforward: the same CSFs identified in the minimal basis are projected onto the larger basis. For state preparation, we can employ the same quantum circuits we derived for the minimal basis and realise the projection onto the larger basis through basis rotation circuits. While these rotations are unitary operators, they effectively act like projectors when embedded within a larger qubit space (corresponding to the orbitals outside of the active space). We believe that the quantum Krylov algorithm could be a particularly useful method to combine with CSFs in this setting: it would enable computing dynamic correlation effects within the larger basis set using a low-order Trotter formula while parametrising most of the static correlation through the initial CSFs that spin-couple the valence electrons. Given that the dominant correlation is already encoded in the initial CSFs, this approach could yield a lower-cost alternative to quantum phase estimation while maintaining high accuracy.

We have also begun investigating a quantum-inspired extension of this idea, replacing the quantum Krylov procedure with a classical analogue based on tensor networks which maps CSFs to matrix-product states and the time-evolution operator to a matrix-product operator. The aim is to construct a classical counterpart of the quantum Krylov algorithm that can serve both to benchmark future quantum computers and, potentially, as a practical electronic-structure method in its own right, complementing existing post-HartreeFock approaches.

Ultimately, the goal is a future where the quantum algorithms discussed in this thesis are implemented on large-scale, fault-tolerant quantum computers. Considerable theoretical progress remains necessary to close the gap with classical methods, especially given the substantial overhead of quantum error correction. I hope that continued advances in both quantum algorithms and hardware will one day enable

routine simulation of strongly correlated electrons. Achieving this would extend the practical reach of quantum mechanics to the vast majority of chemistry, thereby overcoming, at least in practice, the seemingly intractable problem that Dirac identified in 1929.[119] This prospect alone remains an enduring and compelling motivation for the continued development of quantum computing.

Appendices

A

Appendix for Chapter 3: Spin-coupled molecular orbital theory

Contents

A.1 Spin eigenfunctions 205

A.1 Spin eigenfunctions

A CSF $\mathcal{O}^{N,i}_{SM}(\phi_1 \dots \phi_N)$ has N electrons in spatial orbitals $\{\phi\}$ and is a spin eigenfunction with quantum numbers S and M. Each CSF is a linear combination of determinants, here represented through the α or β occupation of the orbitals ϕ.

$$|\mathcal{O}^{2,1}_{0,0}\rangle = \frac{1}{\sqrt{2}}(|\alpha\beta\rangle - |\beta\alpha\rangle), \tag{A.1}$$

$$|\mathcal{O}^{2,1}_{1,1}\rangle = |\alpha\alpha\rangle\,, \tag{A.2}$$

$$\begin{aligned} |\mathcal{O}^{4,1}_{0,0}\rangle = \frac{1}{\sqrt{3}}(|\alpha\alpha\beta\beta\rangle + |\beta\beta\alpha\alpha\rangle) \\ -\frac{1}{2\sqrt{3}}(|\alpha\beta\alpha\beta\rangle + |\beta\alpha\beta\alpha\rangle + |\alpha\beta\beta\alpha\rangle + |\beta\alpha\alpha\beta\rangle), \end{aligned} \tag{A.3}$$

$$|\mathcal{O}^{4,2}_{0,0}\rangle = \frac{1}{2}(|\alpha\beta\alpha\beta\rangle - |\alpha\beta\beta\alpha\rangle - |\beta\alpha\alpha\beta\rangle + |\beta\alpha\beta\alpha\rangle), \tag{A.4}$$

$$|\mathcal{O}_{1,1}^{4,1}\rangle = \frac{1}{2}(|\alpha\alpha\alpha\beta\rangle - |\alpha\alpha\beta\alpha\rangle - |\alpha\beta\alpha\alpha\rangle + |\beta\alpha\alpha\alpha\rangle) \tag{A.5}$$

$$\begin{aligned} |\mathcal{O}_{0,0}^{6,1}\rangle = \frac{1}{2}(|\alpha\alpha\alpha\beta\beta\beta\rangle) - |\beta\beta\beta\alpha\alpha\alpha\rangle) \\ +\frac{1}{6}(|\alpha\beta\beta\alpha\alpha\beta\rangle + |\alpha\beta\beta\alpha\beta\alpha\rangle + |\alpha\beta\beta\beta\alpha\alpha\rangle + |\beta\beta\alpha\beta\alpha\alpha\rangle \\ + |\beta\beta\alpha\alpha\beta\alpha\rangle + |\beta\beta\alpha\alpha\alpha\beta\rangle + |\beta\alpha\beta\beta\alpha\alpha\rangle + |\beta\alpha\beta\alpha\beta\alpha\rangle \\ + |\beta\alpha\beta\alpha\alpha\beta\rangle - |\beta\alpha\alpha\beta\beta\alpha\rangle - |\beta\alpha\alpha\beta\alpha\beta\rangle - |\beta\alpha\alpha\alpha\beta\beta\rangle \\ - |\alpha\beta\alpha\beta\beta\alpha\rangle - |\alpha\beta\alpha\beta\alpha\beta\rangle - |\alpha\beta\alpha\alpha\beta\beta\rangle - |\alpha\alpha\beta\beta\beta\alpha\rangle \\ - |\alpha\alpha\beta\beta\alpha\beta\rangle - |\alpha\alpha\beta\alpha\beta\beta\rangle), \end{aligned} \tag{A.6}$$

B

Appendix for Chapter 4: Spin coupling is all you need: Encoding strong electron correlation in molecules on quantum computers

Contents

B.1 Spin eigenfunctions 208
B.2 Quantum circuits for basis rotations 208
B.2.1 General fermionic basis rotations 208
B.2.2 Basis rotations for diatomic bond breaking 210
B.3 Quantum circuits for spin eigenfunctions: explicit decompositions and cost 211
B.3.1 All-to-all connectivity 212
B.3.2 Linear connectivity 213
B.3.3 Planar connectivity 215
B.3.4 Fault-tolerant circuits and non-Clifford cost 217
B.3.5 Controlling state preparation circuits 220
B.3.6 Other spin eigenfunctions 221
B.4 Energy and gradients in Nonorthogonal VQE 222
B.5 Matrix elements for QSD based on real-time evolution 224
B.6 Quantum circuits for preparation of matrix product states 225

B.1 Spin eigenfunctions

Below are some examples of the spin-coupled state $|\mathcal{O}^{N,1}_{0,0}\rangle$. The general form is given in Eq. (4.13).

$$|\mathcal{O}^{2,1}_{0,0}\rangle = \frac{1}{\sqrt{2}}(|\alpha\beta\rangle - |\beta\alpha\rangle), \tag{B.1a}$$

$$\begin{aligned} |\mathcal{O}^{4,1}_{0,0}\rangle &= \frac{1}{\sqrt{3}}(|\alpha\alpha\beta\beta\rangle + |\beta\beta\alpha\alpha\rangle) \\ &-\frac{1}{2\sqrt{3}}(|\alpha\beta\alpha\beta\rangle + |\beta\alpha\beta\alpha\rangle + |\alpha\beta\beta\alpha\rangle + |\beta\alpha\alpha\beta\rangle), \end{aligned} \tag{B.1b}$$

$$\begin{aligned} |\mathcal{O}^{6,1}_{0,0}\rangle &= \frac{1}{2}(|\alpha\alpha\alpha\beta\beta\beta\rangle) - |\beta\beta\beta\alpha\alpha\alpha\rangle) \\ &+\frac{1}{6}(|\alpha\beta\beta\alpha\alpha\beta\rangle + |\alpha\beta\beta\alpha\beta\alpha\rangle + |\alpha\beta\beta\beta\alpha\alpha\rangle + |\beta\beta\alpha\beta\alpha\alpha\rangle \\ &+ |\beta\beta\alpha\alpha\beta\alpha\rangle + |\beta\beta\alpha\alpha\alpha\beta\rangle + |\beta\alpha\beta\beta\alpha\alpha\rangle + |\beta\alpha\beta\alpha\beta\alpha\rangle \\ &+ |\beta\alpha\beta\alpha\alpha\beta\rangle - |\beta\alpha\alpha\beta\beta\alpha\rangle - |\beta\alpha\alpha\beta\alpha\beta\rangle - |\beta\alpha\alpha\alpha\beta\beta\rangle \\ &- |\alpha\beta\alpha\beta\beta\alpha\rangle - |\alpha\beta\alpha\beta\alpha\beta\rangle - |\alpha\beta\alpha\alpha\beta\beta\rangle - |\alpha\alpha\beta\beta\beta\alpha\rangle \\ &- |\alpha\alpha\beta\beta\alpha\beta\rangle - |\alpha\alpha\beta\alpha\beta\beta\rangle) \end{aligned} \tag{B.1c}$$

As shown in Ref. 50, other systems such as hydrogen clusters are best described by a different type of CSF which is also a singlet but corresponds to a different spin coupling pattern. This CSF is just a product of identical Bell states, whose general form is given by Eq. 4.31. The $N = 4$ example used in Ref. 50 is:

$$|\mathcal{O}^{4,2}_{0,0}\rangle = \frac{1}{2}(|\alpha\beta\alpha\beta\rangle - |\alpha\beta\beta\alpha\rangle - |\beta\alpha\alpha\beta\rangle + |\beta\alpha\beta\alpha\rangle). \tag{B.2}$$

B.2 Quantum circuits for basis rotations

B.2.1 General fermionic basis rotations

We can rotate the single-particle basis of any many-particle state by applying exponential unitary transformations that in general act on the entire many-body Hilbert space. A basis transformation of K basis functions (here, spin-orbitals or qubits) can be represented by a $K \times K$ unitary matrix with entries u_{pq}:

$$\tilde{\phi}_p = \sum_q \phi_q u_{pq}, \tag{B.3}$$

where the sets $\{\phi_p\}$ and $\{\tilde{\phi}_p\}$ are the original and the rotated basis set. Given a many-body state, this single-particle rotation is equivalent to applying a linear transformation on the second-quantized operators:

$$\tilde{a}_p^\dagger = \sum_q u_{pq} a_{pq}^\dagger \tag{B.4a}$$

$$\tilde{a}_p = \sum_q u_{pq}^* a_q \tag{B.4b}$$

Following Thouless' theorem,[215] the action of the single-particle rotation on a many-body wavefunction $|\psi\rangle$ can be expressed by the following operator:

$$U(\boldsymbol{u}) = \exp\Big(\sum_{pq} [\log(\boldsymbol{u})]_{pq} (a_p^\dagger a_q - a_q^\dagger a_p)\Big). \tag{B.5}$$

While this operator in general acts on the entire Hilbert space (its dimension is $D \times D$, where $D = 2^K$ for an K-qubit computational basis or $\binom{K}{N}$ for a particle-number-conserving basis of N fermions in K single-particle states), it can be implemented efficiently on a quantum circuit as follows.[238]

Following the approach outlined in Ref. 238, the unitary operator $U(\boldsymbol{u})$ can be decomposed into a series of Givens rotations with the form

$$U(\boldsymbol{\theta}) = \prod_{pq} \exp\Big(\theta_{pq}(a_p^\dagger a_q - a_q^\dagger a_p)\Big) \tag{B.6}$$

without any trotterization error. The corresponding rotational angles θ_{pq} can be identified by performing a QR decomposition of the orbital transformation matrix $\log(\boldsymbol{u})$, which can be solved classically. Each exponentiated one-body operator can be implemented individually using efficient quantum circuits.[32, 238] In general, the indices run over all elements $p, q \in \{1, 2, .., K\}$ with $p > q$ and thus there are $\binom{K}{2}$ such one-body operators. However, for many of the systems that we consider in this work, the matrix $\boldsymbol{u}$ has many zero entries since only few orbitals in $\{\phi_q\}$ contribute to each $\tilde{\phi}_p$, and therefore the cost of the basis transformations is lower than $\binom{K}{2}$.

The CNOT cost of implementing the exponential of a one-body operator

$$\exp[\theta_{pq}(a_p^\dagger a_q - a_q^\dagger a_p)] \tag{B.7}$$

that rotates between spin-orbitals p and q on a quantum circuit is [344, 411]

$$C_{pq} = \begin{cases} 2(p-q)+1, & p-q > 2 \\ 2, & p-q \in \{1,2\}. \end{cases} \tag{B.8}$$

In this work, we use restricted orbitals and therefore the transformations are the same for both spin-up and spin-down orbitals that share the same spatial orbital. The cost for a transformation between two spatial orbitals is therefore $2\,C_{pq}$. The linear scaling with the orbital indices $p-q$ stems from the requirement of implementing the Pauli-Z-strings that arise when mapping the fermionic operators onto qubit operators via the Jordan-Wigner encoding (see SI in Ref. 52). To avoid this overhead, we can geometrically arrange the ordering of qubits such that the orbitals that contribute to the same orbital transformation neighbor each other.

B.2.2 Basis rotations for diatomic bond breaking

For the diatomic systems discussed in Section 4.2 and Ref. 50, transforming a delocalized orbital to a localized one requires building a linear combination of two spatial orbitals. For N spin-coupled electrons, we have N such transformations.

Consider the spin-coupled states $|\mathcal{O}^{N,1}_{0,0}\rangle$ (Appendix B.1). To obtain the localized orbitals from the Hartree–Fock orbitals, we need rotations of the form $\tilde{\phi}_L = \frac{1}{\sqrt{2}}(\phi_1 + \phi_2)$, $\tilde{\phi}_R = \frac{1}{\sqrt{2}}(\phi_1 - \phi_2)$ for each pair of spin-coupled electrons. Each rotation involves two spatial orbitals, or four spin-orbitals. In total, it requires four one-body operators of the form in Eq. (B.8), each with cost C_{pq}, to localize two electrons/spatial orbitals. The total cost for localizing or delocalizing N electrons/orbitals is

$$C_{\mathrm{rot},N} = 2NC_{pq}. \tag{B.9}$$

For example, for N_2, we have

$$\begin{aligned} \mathrm{z}_L &= \frac{1}{\sqrt{2}}(\sigma_\mathrm{g} + \sigma_\mathrm{u}), & \mathrm{z}_R &= \frac{1}{\sqrt{2}}(\sigma_\mathrm{g} - \sigma_\mathrm{u}), \\ \mathrm{x}_L &= \frac{1}{\sqrt{2}}(\pi_{\mathrm{u,x}} + \pi_{\mathrm{g,x}}), & \mathrm{x}_R &= \frac{1}{\sqrt{2}}(\pi_{\mathrm{u,x}} - \pi_{\mathrm{g,x}}), \\ \mathrm{y}_L &= \frac{1}{\sqrt{2}}(\pi_{\mathrm{u,y}} + \pi_{\mathrm{g,y}}), & \mathrm{y}_R &= \frac{1}{\sqrt{2}}(\pi_{\mathrm{g,y}} - \pi_{\mathrm{g,y}}). \end{aligned} \tag{B.10}$$

as described in Section 4.2.2. Using the qubit (spin-orbital) ordering $\{\sigma_{\mathrm{g}}, \bar{\sigma}_{\mathrm{g}}, \sigma_{\mathrm{u}}, \bar{\sigma}_{\mathrm{u}}, \pi_{\mathrm{g},x}, \bar{\pi}_{\mathrm{g},x}, \pi_{\mathrm{g},y}, \bar{\pi}_{\mathrm{g},y}\}$, where the absence (presence) of an overbar indicates an α (β) spin-orbital, respectively, we always have $p - q = 2$ and thus $C_{pq} = 2$ (Eq. (B.8)). Therefore, to localize 2, 4, and 6 spatial orbitals, as required to apply the circuits for preparation of ($|\phi_{2x}\rangle$, $|\phi_{2y}\rangle$), $|\phi_4\rangle$, and $|\phi_6\rangle$ (Section 4.2), the cost is $4N$ i.e. 8, 16, and 24 CNOTs, respectively.

B.3 Quantum circuits for spin eigenfunctions: explicit decompositions and cost

Below we analyze the exact cost of the CSF preparation circuit in Section 4.4. We focus on counting CNOT gates because entangling gates are typically the biggest sources of noise on devices without error correction.[256] In Section B.3.4, we also compute the Toffoli gate counts for longer-term fault-tolerant implementations of these circuits, since Toffoli gates dominate the cost when applying error-correcting codes.[83] We separately consider qubit architectures with all-to-all interactions between qubits, as well as devices with restricted nearest-neighbor-only connectivity; specifically, linear, and planar (grid-like) connectivity. The latter require decomposing gates between distant qubits into nearest-neigbor gates, which introduces a gate overhead. This detailed analysis might be relevant in particular when considering implementations on near-term quantum hardware. For this, we made some circuit design choices to get concrete numbers for the gate counts, but these are not necessarily optimal and could possibly be improved if tailored for a particular architecture or for the simulated system.

The circuit for preparing a single spin eigenfunction of the form $|\mathcal{O}^{N,1}_{0,0}\rangle$ contains three parts (see Sections 4.4.5, 4.4.6 and the circuit diagram in Fig. 4.3):

1. Preparation of the input state
2. Application of the two S_n unitaries
3. Mapping from spin to Fock space.

Below we discuss their cost in terms of the number of CNOT gates, assuming all-to-all connectivity. While the unitaries S_n can be implemented using circuits with only linear (nearest-neighbor) connectivity, this is not the case for circuits that implement the input state and the mapping from spin to Fock space. We therefore compare the cost assuming all-to-all connectivity and the cost with connectivity restrictions (linear and planar connectivity).

B.3.1 All-to-all connectivity

Let C_2 and C_3 denote the number of CNOT gates required to implement the two- and three-qubit gate blocks required for the Dicke state preparation circuits (Figs. 4.2b, 4.2c). Each two-qubit gate block consists of a controlled R_y-gate conjugated by two CNOTS. This can be decomposed into 3 CNOT gates: $C_2 = 3$ (Fig. 3 in [411]). The three-qubit gate consists of a R_y-gate controlled by two qubits, which can be decomposed into 5 CNOT gates: $C_3 = 5$ (Fig. 3 in [353]). Since each unitary $M_{l,l-1}$ consists of one two-qubit gate block and $l-2$ three-qubit gate blocks (Section 4.4.3), the overall CNOT cost for implementation of the operator S_n is

$$\begin{aligned} C_{S_n} &= \sum_{l=2}^{n} C_{M_{l,l-1}} = \sum_{l=2}^{n} \Big(C_2 + C_3(l-2)) \Big) \\ &= \sum_{l=2}^{n} \Big(3 + 5(l-2) \Big) = \frac{5}{2} n^2 - \frac{9}{2} n + 2, \end{aligned} \tag{B.11}$$

where $C_{M_{l,l-1}}$ denotes the CNOT cost of implementing the unitary $M_{l,l-1}$. Table B.1 shows the cost up to $n = 6$.

Table B.1: Number of two-qubit and three-qubit gate blocks (Figs. 4.2b, 4.2c) required for application of symmetric state preparation unitaries S_n, and cost C_{S_n} (number of CNOT gates) in the circuit for S_n after decomposition into elementary gates (Eq. (B.11)).

	2-qubit gates	3-qubit gates	C_{S_n}
S_2	1	0	3
S_3	2	1	11
S_4	3	3	24
S_5	4	6	42
S_6	5	10	65

The cost of the circuit for input state preparation is linear in the number of CNOT gates (Section 4.4.5). The first step only contains X gates; the second step requires controlled rotation gates $CR_y(\theta_i)$ which can be decomposed into single-qubit rotations and two CNOTs (Fig. B.1) to give CNOT cost of $2(n-1) = N-2$; the third step requires $n = N/2$ CNOT gates. The total CNOT count for input state preparation is therefore

$$C_{\text{in}}^{\text{all}} = \frac{3N}{2} - 2. \tag{B.12}$$

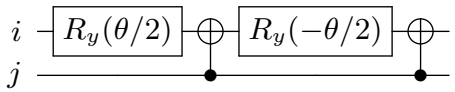


Figure B.1: Decomposition of $CR_y^{j,i}(\theta)$.

The mapping from spin to Fock space costs

$$C_{\text{map}}^{\text{all}} = N \tag{B.13}$$

for hardware with all-to-all connectivity, giving a total cost of

$$\begin{aligned} C_{\text{tot}}^{\text{all}} &= 2C_{S_n} + C_{\text{in}}^{\text{all}} + C_{\text{map}}^{\text{all}} \\ &= \left(\frac{5}{4}N^2 - \frac{9}{2}N + 4\right) + \left(\frac{3}{2}N - 2\right) + N \\ &= \frac{5}{4}N^2 - 2N + 2. \end{aligned} \tag{B.14}$$

B.3.2 Linear connectivity

For hardware with linear connectivity, the circuit S_n can be implemented directly without any modification.[353] However, in the last step of the input start preparation circuits, some CNOT gates act between distant qubits and therefore cannot be implemented as such on a linear topology. The circuit can be adapted for hardware with nearest-neighbor connectivity via a recompilation of the CNOT accordion into nearest-neighbor CNOTs (see Fig. B.2). Evidently, some of the CNOT gates cancel out after the decomposition. The CNOT count for the accordion with linear connectivity is

$$C_{\text{acc}}(N) = N - 1 + 2\sum_{i=0}^{\frac{N}{2}-2}(N - 2i - 2) = \frac{N^2}{2} - 1. \tag{B.15}$$

The overall CNOT cost for preparation of the input state with linear connectivity is

$$C_{\text{in}}^{\text{lin}} = (N-2) + \left(\frac{N^2}{2} - 1\right) = \frac{N^2}{2} + N - 3. \tag{B.16}$$

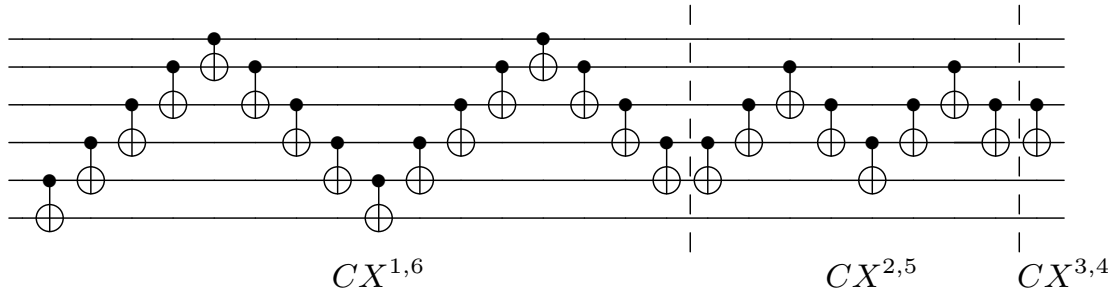


Figure B.2: Recompilation of CNOT accordion for $N = 6$, namely $CX^{1,6}CX^{2,5}CX^{3,4}$, into gates restricted to linear connectivity. This corresponds to the last part of the circuit for preparation of the input state (see Fig. 4.3 for $N = 8$ example).

To implement U_{map}, one could apply the same naive decomposition of the CNOTs into linear connectivity as in Fig. B.2, where in this case it is not directly obvious if any CNOTs cancel out. This has a large overhead as what was previously $C_{\text{map}}^{\text{all}} = N$ CNOTs now becomes

$$C_{\text{map}}^{\text{lin}} = N\Big[N + 2(N-1) + (N-2)\Big] = 4N^2 - 4N. \tag{B.17}$$

The total CNOT cost for hardware with linear connectivity would thus be

$$\begin{aligned} C_{\text{tot}}^{\text{lin}} &= 2C_{S_n} + C_{\text{in}}^{\text{lin}} + C_{\text{map}}^{\text{lin}} \\ &= \left(\frac{5}{4}N^2 - \frac{9}{2}N + 4\right) + \left(\frac{1}{2}N^2 + N - 3\right) + \left(4N^2 - 4N\right) \\ &= \frac{23}{4}N^2 - \frac{15}{2}N + 1. \end{aligned} \tag{B.18}$$

Due to the decomposition of the CNOT accordion into nearest-neigbor gates, the depths of the circuits for input state preparation $U_{\text{in},N}$ and for spin-to-Fock mapping U_{map} scale as $\mathcal{O}(N^2)$, which is worse than the $\mathcal{O}(N)$ scaling of the depth for the Dicke circuits S_n. Furthermore, $C_{\text{map}}^{\text{lin}}$ dominates the gate count of CSF preparation with linear connectivity (Eq. (B.18)). Below, we show that the overhead can be removed in the case of planar connectivity.

B.3.3 Planar connectivity

Many state-of-the-art quantum devices based on superconducting qubits, such as Google's Sycamore processor, consist of planar (two-dimensional) structures with nearest-neighbor connectivity.[295, 296] In cases with such a hardware topology, one can carefully align the spin-orbitals to avoid the overhead required when going from all-to-all to linear connectivity associated with recompiling the two-qubit gates required by the spin-to-Fock space mapping U_{map} (Appendix B.3.2).

Starting from a basis of localized orbitals where $\{i_L\}$ and $\{i_R\}$ are spin-up orbitals localized on the left and right atom, and $\{\bar{i}_L\}$ and $\{\bar{i}_R\}$ are the corresponding spin-down orbitals, the circuits for preparation of a single CSF $|\mathcal{O}_{0,0}^{N,1}\rangle$ require the following pairs of qubits/spin-orbitals to be directly connected to avoid any gate overhead due to recompiling two-qubit (CNOT) gates between non-neigboring qubits:

1. Spin-up orbitals i_L and $(i+1)_L$ for all $i \in \{1, 2, ...N-1\}$, must be connected to each other. This is required for the circuits for input state preparation $U_{\text{in},N}$ and the circuits $S_{N/2}$ (left and middle section in Fig. 4.3). The ordering within the left and right subsystem is irrelevant due to the symmetry of the state within that subspace.

2. Every spin-up orbital $i_{L/R}$ must be connected with their corresponding spin-down orbital $\bar{i}_{L/R}$. This is required for the spin-to-Fock-space mapping U_{map} (right section in Fig. 4.3).

3. For the basis rotation circuits (Appendix B.2.2), each *left* orbital i_L or $\bar{i}_L$ must be connected to its corresponding *right* orbital i_R or $\bar{i}_R$.

Assuming a rectangular grid, there is no ordering that satisfies all three conditions. If we only wish to prepare a single CSF without performing orbital rotations, we can place all spin-up qubits $i_{L/R}$ in one line, and all the corresponding spin-down qubits $\bar{i}_{L/R}$ in a line below or above it. With this ordering, the first two conditions are met and therefore there is no overhead in implementing the input and Dicke state preparation circuits. However, basis rotations would require applying

SWAP gates to place the relevant qubit pairs next to each other, with an overhead scaling as $\mathcal{O}(N^2)$.

If we wish to use multiple CSFs (including localized and delocalized states) in a quantum algorithm, we can prepare the CSF in a basis of localized orbitals and then rotate the basis as in Appendix B.2.2. This is preferably done using the following arrangement to exploit planar connectivity without a large gate cost: set all left spin-up orbitals $\{i_L\}$, $i = \{1, 2, ...n\}$ in a horizontal line, and all right spin-up orbitals $\{i_R\}$, $i = \{1, 2, ...n\}$ in parallel, below the spin-up line. Then, set the left spin-down orbitals $\{\bar{i}_L\}$, $i = \{1, 2, ...n\}$ in a line on top, and the right spin-down orbitals $\{\bar{i}_R\}$, $i = \{1, 2, ...n\}$ in a line below. This enables applying the circuits for preparation of a single CSF with nearly the same cost as if we had all-to-all connectivity. The overhead using this grid structure comes from the need of decomposing the CNOT accordion for preparation of the input state to a linear array (Appendix B.3.2). Since the spin-up orbitals i_L and i_R are already connected $\forall i \in \{1, 2, ..., n\}$, those rotations can be implemented directly. However, to connect the spin-down orbitals $\bar{i}_L$ and $\bar{i}_R$, we must swap i_L and $\bar{i}_L$ and i_R and $\bar{i}_R$ for all i. Then, we can implement the input state preparation circuit. The overhead for going from all-to-all to planar connectivity is therefore $2n = N$ SWAP gates, or $3N$ CNOT gates. This linear overhead is small compared to the total $\mathcal{O}(N^2)$ cost of CSF preparation (Eq. (B.14)).

Using this strategy, the total CNOT count for state preparation of a single CSF with N spin-coupled electrons in the localized basis on a device with planar connectivity is:

$$C_{\text{tot}}^{\text{pl}} = C_{\text{tot}}^{\text{all}} + 3N = \frac{5}{4}N^2 + N + 2. \tag{B.19}$$

Preparing that state and then rotating N spin-orbitals orbitals (occupied by N electrons) to a delocalized basis has an additional cost of $2NC_{p,q} = 4N$ (Eq. (B.9)). This might not be needed depending on the application (e.g. when preparing linear combinations of CSFs, we must rotate only a subset of the orbitals for each CSF).

B.3.4 Fault-tolerant circuits and non-Clifford cost

Universal gate sets for fault-tolerant quantum computation include Clifford + Toffoli or Clifford + T gates. Since Clifford gates including the set of Clifford group generators $\{H, S, CX\}$ are efficiently classically simulable,[48] most fault-tolerant resource estimates typically assume that their cost is negligible compared to the overall cost.[64, 65] On the other hand, expensive magic-state distillation protocols are required to apply noisy non-Clifford gates with high fidelity within an error-correcting code, resulting in a large space-time overhead orders of magnitude larger than the cost of fault-tolerant Clifford gates.[118] We therefore focus on minimizing and counting non-Clifford gates. This allows direct comparison with the literature on Hamiltonian simulation for electronic structure (see [64, 65, 85, 88] and Table III in Ref. 65 for an overview).

The circuit V_N for preparation of the CSF $|\mathcal{O}^{N,1}_{0,0}\rangle$, presented in Section 4.4.5, requires converting the continuous rotation gates $R_y(\theta)$ to a discrete gate set of Clifford + Toffoli gates. Due to the finite precision in the binary representation of continuous numbers, this introduces an error ϵ_r per rotation gate that can be exponentially suppressed since the gate and qubit counts scale as $\mathcal{O}(\log(1/\epsilon_r))$. The number of rotations in V_N scales as $\mathcal{O}(N^2)$ and therefore the asymptotic Toffoli complexity remains $\tilde{\mathcal{O}}(N^2)$ after suppression of polylogarithmic factors. Below, we derive the exact cost.

Rotation gate synthesis

A commonly-used approach is to synthesize rotations in terms of Clifford + T gates. The dominant factor affecting the number of T-gates per rotation gate is

$$N_{TR} = c\lceil \log(1/\epsilon_r) \rceil, \tag{B.20}$$

where c is a constant that depends on the implementation. For deterministic algorithms, c is lower-bounded by 3,[412, 413] but this can be improved using randomized methods such as the repeat-until-success (RUS) approach in Ref. 414,

which finds an empirical average of $c = 1.15$. Here, we consider the probabilistic RUS technique, where the cost includes an extra additive factor:

$$N_{TR} \approx c\lceil \log(1/\epsilon_r) \rceil + 9.2. \tag{B.21}$$

Since we later compare the cost of initial state preparation with the cost of implementing quantum phase estimation, we translate the T-gate counts into Toffoli counts to compare with the Toffoli counts in state-of-the-art qubitization-based algorithms for electronic structure.[65, 88] The corresponding number of Toffoli gates is roughly $T_R \approx \frac{1}{2}N_{TR}$ in a surface code implementation, since the cost of applying a Toffoli gate is approximately twice the cost of applying a T-gate using the magic state factories introduced in Ref. 118. Thus, the number of T-gates per rotation is $\frac{1}{2}(cR\lceil \log(1/\epsilon_r) \rceil + 4.6)$.

The circuits for preparation of $|\mathcal{O}^{N,1}_{0,0}\rangle$ also require application of controlled R_y gates CR_y and $CCRy$ (Section 4.4). Controlling rotation gates does not have any non-Clifford overhead if the rotation angle is given by a classical register, as we can simply use CNOT gates to change the direction of the rotation $\theta \leftrightarrow -\theta$.[266] This is indeed the case in the circuits for preparation of the spin eigenfunction $|\mathcal{O}^{N,1}_{0,0}\rangle$, since the entire state preparation circuit (including the angles) can be specified classically before any quantum computation. To determine the number of Toffoli gates, we must therefore simply count the number of rotations, irrespective of whether these are controlled or not.

Finally, consider a circuit with R rotations with overall error ϵ. A naive error analysis based on a triangle inequality $\epsilon \leq \sum_{r=1}^{R} \epsilon_r$ would suggest allocating an error $\epsilon_r = \epsilon/R$ for each rotation, where the number of bits required to represent each rotation angle is $b = \log(1/\epsilon_r)$. Ref. 415 suggested that one can reduce this pessimistic bound if one regards errors as random rather than coherent. The corresponding random walk complexity analysis gives an error bound that is tighter by a quadratic factor $\epsilon \to \sqrt{\epsilon}$, or equivalently halves the number of bits b. With this, we obtain a reduced average of

$$T_R = \frac{1}{4}cR\lceil \log(1/\epsilon_r) \rceil + 4.6 \tag{B.22}$$

Toffoli gates per rotation, where $b = \frac{1}{2}\lceil \log(1/\epsilon_r) \rceil$ and $c = 1.15$.

Toffoli cost for preparation of spin eigenfunctions

As discussed in Section 4.4.3, each subcircuit $M_{l,l-1}$ that implements the symmetric state preparation unitary S_n (Eq. (4.22)) consists of one two-qubit gate gate block (Fig. 4.2b) and $l-2$ three-qubit gate blocks. The total number of rotations for implementation of S_n is therefore

$$R_{S_n} = \sum_{l=2}^{n} (l-1) = \frac{1}{2} n(n-1). \tag{B.23}$$

This must be implemented twice to implement $V_N : |\mathcal{O}_{0,0}^{N,1}\rangle = V_N |00\rangle^N$. The number of rotation gates required for preparation of the input state is n (Section 4.4.5). Thus, the total number of rotation gates for V_N, where $n = N/2$ is

$$R = 2R_{S_{N/2}} + R_{\mathrm{in},N} = n(n-1) + n = n^2 = \frac{1}{4} N^2. \tag{B.24}$$

The total (average) Toffoli cost thus becomes

$$T = R \times T_R = R\Big[0.2875\lceil \log(R/\epsilon) \rceil + 4.6\Big]. \tag{B.25}$$

The method requires only a single ancilla qubit to check if each rotation was implemented successfully.[414]

Our goal is to use the spin eigenfunctions as initial states in fault-tolerant quantum algorithms. Thus, we must consider two sources of errors: the error due to state preparation ϵ_{SP}, and the error due to implementation of the quantum algorithm itself, e.g. quantum phase estimation, ϵ_{QPE}. The total error will at worst be

$$\epsilon_{\mathrm{tot}} \leq \epsilon_{\mathrm{SP}} + \epsilon_{\mathrm{QPE}}. \tag{B.26}$$

Since our circuits for initial state preparation are very efficient, a reasonable strategy would be to allocate most of the error to the quantum algorithm itself, rather than the state preparation task. For example, in the context of quantum phase estimation, one could choose $\epsilon_{\mathrm{tot}} = 0.0016$ (in Hartree atomic units) to achieve chemical accuracy in the energy estimation. We divide this into $\epsilon_{\mathrm{SP}} = 10^{-7}$

and $\epsilon_{\mathrm{QPE}} = \epsilon_{\mathrm{tot}} - 10^{-7}$. Inserting these values and Eq. (B.24) into Eq. (B.25), we get the Toffoli counts for preparation of $|\mathcal{O}^{N,1}_{0,0}\rangle$ states reported in Table 4.1. Even for the largest systems with $N = 34$, the cost is only $T = 3989 \approx 4 \times 10^3$. This is indeed very low compared to the typical cost of quantum phase estimation, which is in the order of 10^{10} using state-of-the-art techniques.[65, 88]

Note that we have also considered the elegant phase gradient technique from Ref. 416 as an alternative to rotation gate synthesis. While we found that the Toffoli cost per rotation can be slightly lower depending on the choice of ϵ, the advantage is washed away due to the one-off cost of preparing the phase gradient state (which in itself requires rotation gate synthesis), even for the largest system we consider, where $N = 34$.

B.3.5 Controlling state preparation circuits

In some quantum algorithms, e.g. when preparing initial states for VQE or phase estimation at intermediate bond lengths, it is necessary to prepare linear combinations of spin eigenfunctions (Section 4.4.9). This can be achieved by controlling the circuits for preparation of a single CSF of the form $|\mathcal{O}^{N,1}_{0,0}\rangle$. It does not affect the asymptotic scaling with N but introduces a constant factor overhead. The overhead is small because one only needs to control a small part of the circuit for input state preparation, $U_{\mathrm{in},N}$, as well as the circuit for U_{map}.

To control the circuit for input state preparation and the Dicke circuits $S_{N/2}$, it is sufficient to simply control the single-qubit Pauli X-gates at the beginning of $U_{\mathrm{in},N}$, as well as the first single-qubit rotation gate $R_y(\theta_{N/2})$. If the control qubit is in 0, the remaining part of the circuit $U_{\mathrm{in},N}$ and the entire circuit $S_{N/2} \otimes S_{N/2}$ act like the identity, because the CNOT and rotation gates therein have no effect, and thus this effectively controls the entire circuit. Therefore the only overhead for controlling this is $n = N/2$ CNOT gates, as well as the cost of converting the $R_y(\theta_n)$ to a $CR_y(\theta_n)$ gate. The $CR_y(\theta_n)$ can be implemented with 2 CNOTs and two R_y gates (Figure B.1) and therefore has an overhead of 2 CNOT gates and one R_y gate.

Finally, consider the circuit U_{map}, consisting of N CNOT gates of the form $\overline{CX}^{i,j} = X_i CX^{i,j} X_i$. This can be controlled by simply controlling the two X gates and therefore requires $2N$ CNOTs gates. The overhead (cost to add on top of state preparation cost without controlled qubits) is $3N - N = 2N$. The total CNOT overhead becomes:

$$C_{\text{ctrl}} = \left(\frac{N}{2} + 2\right) + 2N = \frac{5}{2}N + 2. \tag{B.27}$$

In fault-tolerant hardware, the dominant cost comes from implementing the non-Clifford (Toffoli) gates, not CNOT gates. As discussed in Appendix B.3.4, rotations controlled by arbitrary qubits only have non-Clifford overhead compared to uncontrolled rotations.[266] Therefore, there is no Toffoli overhead for controlling any part of the state preparation circuit V_N.

B.3.6 Other spin eigenfunctions

We briefly discuss the cost for preparing the state

$$|\mathcal{O}_{0,0}^{N,2}\rangle = \left[\frac{1}{\sqrt{2}}(|\alpha\beta\rangle - |\beta\alpha\rangle)\right]^{N/2}. \tag{B.28}$$

(Section 4.4.8, Eq. 4.31). Working entirely in the spin space, where this maps to the qubit state

$$\left[\frac{1}{\sqrt{2}}(|1001\rangle - |0110\rangle)\right]^{N/2}, \tag{B.29}$$

state preparation only requires one CNOT per *two*-electron singlet, therefore $N/2$ CNOTs. Two additional CNOT gates are required to implement the spin-to-Fock space mapping U_{map}. Therefore the total CNOT cost assuming all-to-all connectivity is $C_{\text{tot}}^{\text{all}} = N/2 + N = 3/2N$.

This is the same for planar connectivity, $C_{\text{tot}}^{\text{pla}} = C_{\text{tot}}^{\text{all}}$ if we carefully align the spin-orbitals on the qubit grid using the strategy in Appendix B.3.3. For linear connectivity, we can choose the qubit ordering $\{\alpha_1, \beta_1, \alpha_2, \beta_2, ..., \alpha_N, \beta_N, \}$ to minimize the cost of U_{map} to $2 \times (N/2)$. The CNOT between orbitals α_i and α_{i+1},

which are separated by the β_i qubit, can be decomposed into 3 CNOTs using the CNOT accordion (Fig. B.2). The total cost for linear connectivity thus becomes

$$C_{\text{tot}}^{\text{lin}} = \frac{5}{2}N. \tag{B.30}$$

Controlling these circuits only requires controlling the $R_y(-\theta/2)$ rotation for each two-electron singlet, which costs two CNOTs (Fig. B.1). Adding the cost of controlling U_{map}, the CNOT overhead for controlling the preparation of $|\mathcal{O}_{0,0}^{N,2}\rangle$ is $N/2 + N = 3/2N$.

On a fault-tolerant device, $R_y(-\theta/2)$ can be factorized into HX, therefore the state preparation and its controlled version can be implemented entirely with Clifford gates.

B.4 Energy and gradients in Nonorthogonal VQE

The nonorthogonal VQE (NO-VQE) algorithm variationally optimizes the energy of a wavefunction corresponding to the linear combination defined in Eq. (4.38). A simultaneous optimization of the gate parameters $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M)$ and linear coefficients $\boldsymbol{C}$ requires the gradient of the energy with respect to each variable. In what follows, we define the correlated basis states as

$$|\Psi_I(\boldsymbol{\theta}_I)\rangle = \prod_i U_{Ii}(\theta_{Ii})\,|\Phi_I\rangle\,, \tag{B.31}$$

such that the full wavefunction is given by

$$|\Psi(\boldsymbol{\Theta}, \boldsymbol{C})\rangle = \sum_{I=1}^{L} C_I\,|\Psi_I(\boldsymbol{\theta}_I)\rangle\,. \tag{B.32}$$

Since the states $\{|\Psi_I(\boldsymbol{\theta}_I)\rangle\}$ are not mutually orthogonal, the VQE optimization requires gradients of the energy expectation value

$$E(\boldsymbol{\Theta}, \boldsymbol{C}) = \frac{\langle\Psi(\boldsymbol{\Theta}, \boldsymbol{C})|\hat{H}|\Psi(\boldsymbol{\Theta}, \boldsymbol{C})\rangle}{\langle\Psi(\boldsymbol{\Theta}, \boldsymbol{C})|\Psi(\boldsymbol{\Theta}, \boldsymbol{C})\rangle}, \tag{B.33}$$

which are obtained through the quotient rule as

$$\frac{\partial E}{\partial \boldsymbol{\Theta}} = 2\Big[\,\langle\partial_{\boldsymbol{\Theta}}\Psi|\hat{H}|\Psi\rangle - E\,\langle\partial_{\boldsymbol{\Theta}}\Psi|\Psi\rangle\Big], \tag{B.34}$$

and likewise for $\frac{\partial E}{\partial \boldsymbol{C}}$. Here, $\partial_{\boldsymbol{\Theta}} \equiv \frac{\partial}{\partial \boldsymbol{\Theta}}$ and we have exploited the Hermitian symmetry, e.g. $\langle \Psi | \hat{H} | \partial_{\boldsymbol{\Theta}} \Psi \rangle = \langle \partial_{\boldsymbol{\Theta}} \Psi | \hat{H} | \Psi \rangle$.

Using Eq. (B.32), explicit expressions for the derivatives with respect to the linear coefficients can be obtained as

$$\begin{aligned} \frac{\partial E}{\partial C_I} &= 2\Big[\langle \Psi_I(\boldsymbol{\theta}_I) | \hat{H} | \Psi(\boldsymbol{\Theta}, \boldsymbol{C}) \rangle - \langle \Psi_I(\boldsymbol{\theta}_I) | \Psi(\boldsymbol{\Theta}, \boldsymbol{C}) \rangle\Big], \\ &= 2 \sum_{J=1}^{M} [H_{IJ} - S_{IJ}] C_J \end{aligned} \tag{B.35}$$

where the Hamiltonian and overlap coupling elements are

$$H_{IJ} = \langle \Psi_I(\boldsymbol{\theta}_I) | \hat{H} | \Psi_J(\boldsymbol{\theta}_J) \rangle \,, \tag{B.36a}$$

$$S_{IJ} = \langle \Psi_I(\boldsymbol{\theta}_I) | \Psi_J(\boldsymbol{\theta}_J) \rangle \,, \tag{B.36b}$$

and we assume that the linear coefficients $\boldsymbol{C}$ are real valued. These derivatives may be computed using the established circuits for measuring nonorthogonal coupling terms outlined in Ref. 97, which are also used in subspace diagonalization approaches. Similarly, using Eq. (B.31), the derivatives with respect to the gate parameters are

$$\begin{aligned} \frac{\partial E}{\partial \theta_{Ii}} = 2\, C_I \sum_J \Big[& \langle \partial_{\theta_{Ii}} \Psi_I(\boldsymbol{\theta}_I) | \hat{H} | \Psi_J(\boldsymbol{\theta}_J) \rangle \\ & - \langle \partial_{\theta_{Ii}} \Psi_I(\boldsymbol{\theta}_I) | \Psi_J(\boldsymbol{\theta}_J) \rangle \Big] C_J, \end{aligned} \tag{B.37}$$

where the partial derivative of the wavefunction is

$$|\partial_{\theta_{Ii}} \Psi_I(\boldsymbol{\theta}_I)\rangle = \prod_{j<i} U_{Ij}(\theta_{Ij}) \, \frac{\partial U_{Ii}(\theta_{Ii})}{\partial \theta_{Ii}} \prod_{k>i} U_{Ij}(\theta_{Ij}) \, |\Phi_I\rangle \,. \tag{B.38}$$

For the QNP *ansatz*, the circuits for these partial derivatives can be constructed using the parameter shift rules detailed in Ref. 344. Therefore, the coupling terms $\langle \partial_{\theta_{Ii}} \Psi_I(\boldsymbol{\theta}_I) | \hat{H} | \Psi_J(\boldsymbol{\theta}_J) \rangle$ and $\langle \partial_{\theta_{Ii}} \Psi_I(\boldsymbol{\theta}_I) | \Psi_J(\boldsymbol{\theta}_J) \rangle$ can be evaluated with the same circuit architecture used to evaluate the Hamiltonian and overlap terms in Eq. (B.36) with a constant prefactor.

With these gradient expressions, the NO-VQE algorithm proceeds using the standard L-BFGS optimization approach. The initial linear coefficients $\boldsymbol{C}$ are

obtained by solving the generalized eigenvalue problem with $\boldsymbol{\Theta} = 0$. Although optimizing the expectation value of the energy means that the linear expansion [Eq. (B.32)] does not need to be normalized, we obtain more stable optimization by normalizing $\boldsymbol{C}$ on each iteration.

B.5 Matrix elements for QSD based on real-time evolution

For simplicity, we restrict the analysis below to the case where we only time-evolve a single reference state $|\Psi_0\rangle$, and note that the conclusions remain unchanged for cases with multiple reference states. Choosing a linear time grid, $t_j = j\Delta t$ with $j = 0, 1, ..., N_T$, we form a subspace of $M = N_T + 1$ states. The overlap matrix elements between the time-evolved states forming the subspace are

$$S_{j,k} = \langle \Psi_j | \Psi_k \rangle = \langle \Phi_0 | e^{-iH\Delta t(k-j)} | \Phi_0 \rangle \,. \tag{B.39}$$

Replacing the Hamiltonian operator with the time-evolution operator $U(\Delta t) := e^{-iH\Delta t}$, we can write $S_{j,k} = \langle \Phi_0 | [U(\Delta t)]^{k-j} | \Phi_0 \rangle$. The matrix elements of $U(\Delta t)$ in the basis of expansion states are:

$$\begin{aligned} U_{j,k} &= \langle \Phi_j | U(\Delta t) | \Phi_k \rangle = \langle \Phi_0 | e^{-iH\Delta t(k-j+1)} | \Phi_0 \rangle \\ &= S_{j,k+1} = S_{j-1,k}. \end{aligned} \tag{B.40}$$

These expressions show that, if we reformulate the eigenvalue problem in Eq. (4.40) to use the time-evolution operator $U(\Delta t)$ rather than the Hamiltonian, the matrix elements $U(\Delta t)_{j,k}$ become equivalent to the overlap matrix elements (shifted by one row or column), which has a Toeplitz structure. This equivalence is advantageous for implementations on quantum hardware, as it only requires measuring the overlap matrix rather than separately measuring the Hamiltonian and overlap (intuitively, the matrix elements in Eq. (B.40) correspond to an autocorrelation function).[96, 100] The number of matrix elements is reduced from $2M^2$ to $M + 1$. This result holds even when the time-evolution operator is Trotterized, as long as the time grid is linear.[100]

B.6 Quantum circuits for preparation of matrix product states

Here, we discuss the cost of known quantum circuits for preparation of matrix product states (MPS). An MPS of M sites has the form

$$|\Psi\rangle = \sum_{\boldsymbol{m}} A_1^{[m_1]} A_2^{[m_2]} \cdots A_M^{[m_M]} \prod_{i=1}^{M} |m_i\rangle \,. \tag{B.41}$$

Here, $m_i \in \{0, 1, ..., d-1\}$ is the physical index that runs over the possible occupations of the Hilbert space of a single site (of *local* dimension d), and the vector $\boldsymbol{m} = (m_1, m_2, ..., m_M)$ defines the occupation numbers for each site. In the case of quantum chemistry, each site can be mapped to a spatial orbital, therefore the site's Hilbert space dimension is $d = 4$, and M is the number of spatial orbitals.[15]. Each $A_i^{[m_i]}$ is a tensor of order χ_i, where χ_i is the bond dimension, and each bond dimension is bounded by the maximal bond dimension, $\chi_i \leq \chi$.

The preparation of an MPS on a quantum computer can be achieved using the sequential method by Schön *et al.*[375] This scales as $\mathcal{O}(M)$ in the number of gates and has depth $\mathcal{O}(M)$. Although an improved scaling of depth $\mathcal{O}(\log(M))$ can be reached through the recent technique by Malz *et al.* [417], which is provably optimal,[417] this only applies to MPS with short-range correlations. For long-range correlated states such as GHZ states, it has been proven that state preparation circuits of depth M are optimal based on Lieb-Robinson bounds.[417, 418] Since we do not expect any area-laws to apply to the entanglement of most molecular eigenstates,[234] we must consider the method for general MPS.[375]

The scaling of the deterministic method in Ref. 375 is also polynomial in the bond dimension. Specifically, for generic MPS, each circuit contains M gate blocks, one for each of the tensors $A_i^{[m_i]}$. Each gate block is a multi-qubit unitary acting on $n_a = \log(\chi_i)$ ancilla qubits (corresponding to the virtual Hilbert space) and one physical qubit. Decomposing a block into single and two-qubit gates requires a circuit with depth exponential in n_a, or $\mathcal{O}(\chi_i)$.[417, 419, 420] Therefore, the total gate complexity is $\mathcal{O}(M\chi^2)$, which is proportional to the scaling of the dimension of the overall tensor.[15] Although approximate methods might ameliorate the

scaling,[419–421] it is unclear if they can retain sufficient accuracy and their error cannot be theoretically bound.[419]

Formichev *et al.* provided a detailed derivation of the cost and a more explicit circuit implementation of the MPS state preparation circuits from Ref. 375 using modern quantum linear algebra techniques.[253] This implementation has Toffoli cost of $\chi_{i-1}[8\chi_i d + b\log(\chi_i d) + \log(\chi_i d)]$ for each block. Assuming $\chi_i = \chi$ for all i, we obtain the following approximate cost for the entire MPS preparation circuit:

$$\begin{aligned} T_{\mathrm{MPS}} &= \sum_{i=2}^{M} \chi_{i-1}[8\chi_i d + b\log(\chi_i d) + \log(\chi_i d)] \\ &\approx (M-1)\chi[32\chi + (b+1)\log(4\chi)]. \end{aligned} \tag{B.42}$$

C

Appendix for Chapter 5: Exact electronic states with shallow quantum circuits from global optimisation

Contents

C.1 Universality of one-body and paired two-body operator pool 228
C.1.1 Edge cases for the universality proof 230
C.2 Qubit operators and quantum circuit implementation 231
C.2.1 General operators 231
C.2.2 Operators in s-UPS ansatz 232
C.2.3 Circuit costs 233
C.3 Summary of the DISCO-VQE algorithm 233
C.3.1 Continuous optimisation using basin-hopping 233
C.3.2 Performing discrete steps: generalised basin-hopping . . 235
C.3.3 Taking uphill discrete steps 237
C.3.4 Summary of hyperparameters 243
C.3.5 Computational implementation 243
C.4 Linear and tetrahedral H_4 (STO-3G) 244
C.4.1 Computational Details 244
C.4.2 Variants of DISCO-VQE with minimal computational cost 245
C.4.3 Comparison to ADAPT-VQE using unpaired two-body operators 246
C.4.4 Binding curve using a fixed ansatz 247
C.4.5 Comparison of DISCO-VQE and ADAPT-VQE binding curves 248
C.5 Triplet ground-state of O_2 (STO-3G) 249
C.5.1 Computational Details 249

C.6 Linear H_6 binding curve (STO-3G) 249
C.6.1 Computational Details 249
C.7 Symmetric H_2O binding curve (STO-3G) 250
C.7.1 Computational Details 250
C.8 N_2 binding curve (STO-3G) 251
C.8.1 Computational Details 251
C.8.2 Comparison of binding curve accuracy 252
C.9 Two-dimensional Hubbard lattice 252
C.9.1 Investigating the 1/8-filling fraction 253

C.1 Universality of one-body and paired two-body operator pool

In the main text, we state that an arbitrary wave function can be represented as a unitary product state using only one- and paired two-body fermionic operators. The universality of this ansatz arises as an extension to the proof presented for Disentangled Unitary Coupled Cluster in Ref. 383. In that work, it was shown that any exact state can be built as a unitary product of generalised one-body and two-body operators. This construction is achieved by decomposing higher-body operators, which appear in the disentangled exact wave function, into nested commutators of one- and two-body operators, which can then be expanded as a

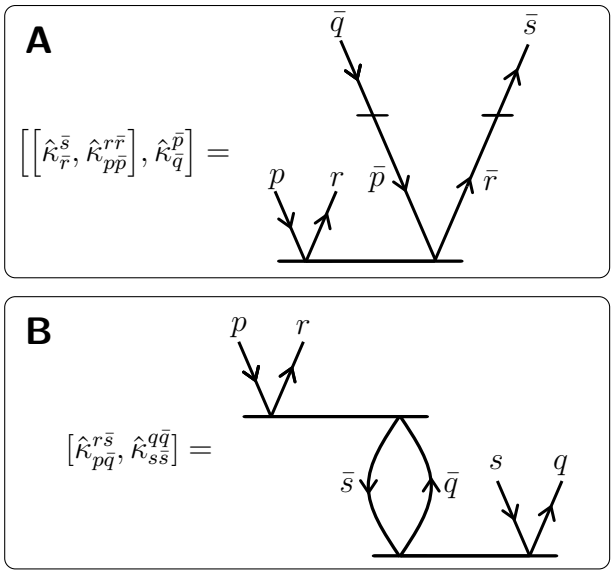


Figure C.1: Goldstone diagrams (**A**) and (**B**) represent the commutators in Eqs. (C.1a) and (C.1b), respectively.

product of individual unitary operators.[383] Here, we show that this derivation can be extended to restrict the operator pool to only paired two-body operators. In particular, we show that unpaired two-body operators can be decomposed into commutators that include only one-body and paired two-body operators as

$$\Big[\Big[\hat{\kappa}^{\bar{s}}_{\bar{r}}, \hat{\kappa}^{r\bar{r}}_{p\bar{p}}\Big], \hat{\kappa}^{\bar{p}}_{\bar{q}}\Big] = \hat{\kappa}^{r\bar{s}}_{p\bar{q}}, \tag{C.1a}$$

$$[\hat{\kappa}^{r\bar{s}}_{p\bar{q}}, \hat{\kappa}^{q\bar{q}}_{s\bar{s}}] = \hat{\kappa}^{rq}_{ps}(\hat{n}_{\bar{s}} - \hat{n}_{\bar{q}}). \tag{C.1b}$$

Goldstone diagrams[49] for these commutator combinations are illustrated in Fig. C.1. The additional factor $(\hat{n}_{\bar{s}} - \hat{n}_{\bar{q}})$ introduces number operators (e.g. $\hat{n}_{\bar{q}} = \hat{a}^\dagger_{\bar{q}}\hat{a}_{\bar{q}}$) that will create a stronger dependence on the ordering of unitary operators. However, since these commute with all (de)excitation operators, the commutator expansions for higher-order operators are unaffected and can be derived using the process described in Ref. 383.

To prove the commutator expansions in Eqs. (C.1a) and (C.1b), we exploit the general representation of fermionic excitation operators $\hat{\tau}^q_p = \hat{a}^\dagger_q\hat{a}_p$[285] and define $p \neq q \neq r \neq s$. The derivation for the commutator expansion in Eq. (C.1a) starts with

$$\begin{aligned}
\Big[\hat{\kappa}^{\bar{s}}_{\bar{r}}, \hat{\kappa}^{r\bar{r}}_{p\bar{p}}\Big] &= \Big(\hat{\tau}^{\bar{s}}_{\bar{r}} - \hat{\tau}^{\bar{r}}_{\bar{s}}\Big)\Big(\hat{\tau}^r_p\hat{\tau}^{\bar{r}}_{\bar{p}} - \hat{\tau}^p_r\hat{\tau}^{\bar{p}}_{\bar{r}}\Big) - \Big(\hat{\tau}^r_p\hat{\tau}^{\bar{r}}_{\bar{p}} - \hat{\tau}^p_r\hat{\tau}^{\bar{p}}_{\bar{r}}\Big)\Big(\hat{\tau}^{\bar{s}}_{\bar{r}} - \hat{\tau}^{\bar{r}}_{\bar{s}}\Big) \\
&= \hat{\tau}^r_p\Big[\hat{\tau}^{\bar{s}}_{\bar{r}}, \hat{\tau}^{\bar{r}}_{\bar{p}}\Big] + \hat{\tau}^p_r\Big[\hat{\tau}^{\bar{r}}_{\bar{s}}, \hat{\tau}^{\bar{p}}_{\bar{r}}\Big] \\
&= \hat{\tau}^r_p\hat{\tau}^{\bar{s}}_{\bar{p}} - \hat{\tau}^p_r\hat{\tau}^{\bar{p}}_{\bar{s}} \\
&= \hat{\kappa}^{r\bar{s}}_{p\bar{p}},
\end{aligned} \tag{C.2}$$

where we have exploited the general identities[384] $\Big[\hat{\tau}^q_p, \hat{\tau}^s_r\Big] = \hat{\tau}^q_r\,\delta_{ps} - \hat{\tau}^s_p\,\delta_{rq}$, and $\hat{\tau}^r_p\hat{\tau}^r_q = \hat{\tau}^p_r\hat{\tau}^q_r = 0$.

The full commutator expansion then follows by taking:

$$\begin{aligned}
\Big[\hat{\kappa}^{r\bar{s}}_{p\bar{p}}, \hat{\kappa}^{\bar{p}}_{\bar{q}}\Big] &= \Big(\hat{\tau}^r_p\hat{\tau}^{\bar{s}}_{\bar{p}} - \hat{\tau}^p_r\hat{\tau}^{\bar{p}}_{\bar{s}}\Big)\Big(\hat{\tau}^{\bar{p}}_{\bar{q}} - \hat{\tau}^{\bar{q}}_{\bar{p}}\Big) - \Big(\hat{\tau}^{\bar{p}}_{\bar{q}} - \hat{\tau}^{\bar{q}}_{\bar{p}}\Big)\Big(\hat{\tau}^r_p\hat{\tau}^{\bar{s}}_{\bar{p}} - \hat{\tau}^p_r\hat{\tau}^{\bar{p}}_{\bar{s}}\Big) \\
&= \hat{\tau}^r_p\Big[\hat{\tau}^{\bar{s}}_{\bar{p}}, \hat{\tau}^{\bar{p}}_{\bar{q}}\Big] + \hat{\tau}^p_r\Big[\hat{\tau}^{\bar{p}}_{\bar{s}}, \hat{\tau}^{\bar{q}}_{\bar{p}}\Big] \\
&= \hat{\tau}^r_p\hat{\tau}^{\bar{s}}_{\bar{q}} - \hat{\tau}^p_r\hat{\tau}^{\bar{q}}_{\bar{s}} \\
&= \hat{\kappa}^{r\bar{s}}_{p\bar{q}}.
\end{aligned} \tag{C.3}$$

Continuing this approach allows Eq. (C.1b) to be derived using the commutator expansion:

$$\begin{aligned}\left[\hat{\kappa}^{r\bar{s}}_{p\bar{q}}, \hat{\kappa}^{q\bar{q}}_{s\bar{s}}\right] &= \left(\hat{\tau}^r_p\hat{\tau}^{\bar{s}}_{\bar{q}} - \hat{\tau}^p_r\hat{\tau}^{\bar{q}}_{\bar{s}}\right)\left(\hat{\tau}^q_s\hat{\tau}^{\bar{q}}_{\bar{s}} - \hat{\tau}^s_q\hat{\tau}^{\bar{s}}_{\bar{q}}\right) - \left(\hat{\tau}^q_s\hat{\tau}^{\bar{q}}_{\bar{s}} - \hat{\tau}^s_q\hat{\tau}^{\bar{s}}_{\bar{q}}\right)\left(\hat{\tau}^r_p\hat{\tau}^{\bar{s}}_{\bar{q}} - \hat{\tau}^p_r\hat{\tau}^{\bar{q}}_{\bar{s}}\right) \\ &= \left(\hat{\tau}^r_p\hat{\tau}^q_s - \hat{\tau}^p_r\hat{\tau}^s_q\right)\left[\hat{\tau}^{\bar{s}}_{\bar{q}}, \hat{\tau}^{\bar{q}}_{\bar{s}}\right] \\ &= \left(\hat{\tau}^r_p\hat{\tau}^q_s - \hat{\tau}^p_r\hat{\tau}^s_q\right)\left(\hat{\tau}^{\bar{q}}_{\bar{q}} - \hat{\tau}^{\bar{s}}_{\bar{s}}\right) \\ &= \hat{\kappa}^{rq}_{ps}(\hat{n}_{\bar{s}} - \hat{n}_{\bar{q}}).\end{aligned} \tag{C.4}$$

On the last line, we have identified:

$$\hat{\tau}^{\bar{q}}_{\bar{q}} = \hat{a}^\dagger_{\bar{q}}\hat{a}_{\bar{q}} = \hat{n}_{\bar{q}}$$

as the number operator for spin-orbital $\bar{q}$. An equivalent derivation can be used to obtain the low-spin two-body operators. Through these expansions, we can expand any n-body fermionic operator as a series of nested commutators involving only one-body or paired two-body operators. Therefore, the unitary product states built with these operators are universal.

C.1.1 Edge cases for the universality proof

As for the quantum-number-preserving ansatz,[344] this operator pool is not universal when the high-spin or low-spin orbitals are either fully occupied or fully unoccupied. In these cases, there are no orbital pairs where one is doubly occupied and the other is fully unoccupied. Therefore, none of the paired two-body operators have any action on the wave function and the optimisation reduces to Hartree–Fock theory. These edge cases can be derived from the commutator expansions because the same-spin two-body operators required for universality cannot be accessed due to the number operators $\hat{n}_{\bar{s}}$ and $\hat{n}_{\bar{q}}$ in Eq. (C.4), which cancel out when all the low-spin orbitals are fully occupied or unoccupied. The fully occupied edge case can be resolved by introducing an additional unoccupied orbital into the system.

C.2 Qubit operators and quantum circuit implementation

C.2.1 General operators

General fermionic one-body and two-body operators are given by:

$$\hat{\kappa}_p^q = (\hat{a}_q^\dagger \hat{a}_p - \hat{a}_p^\dagger \hat{a}_q), \tag{C.5a}$$

$$\hat{\kappa}_{pq}^{rs} = (\hat{a}_r^\dagger \hat{a}_s^\dagger \hat{a}_q \hat{a} - \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r), \tag{C.5b}$$

where the operators $\hat{a}_p^\dagger$ and $\hat{a}_p$ are creation and annihilation operators for spin-orbital p, respectively, and satisfy fermionic anticommutation relations $[\hat{a}_p, \hat{a}_q^\dagger]_+ = \delta_{pq}$, and $[\hat{a}_p, \hat{a}_q]_+ = [\hat{a}_p^\dagger, \hat{a}_q^\dagger]_+ = 0$. In the following derivations, we define the ordering of qubits such that spin-up and spin-down orbitals that share the same spatial wave function are adjacent, i.e. the indices map from fermionic to qubit operators as $p \to 2p$ and $\bar{p} \to 2p+1$. The fermionic operators can then be mapped to qubit operators using the Jordan-Wigner transformation[32, 70]

$$\hat{a}_p = \frac{1}{2}(X_{2p} + iY_{2p}) \prod_{k=1}^{2p-1} Z_k \qquad \text{and} \qquad \hat{a}_{\bar{p}} = \frac{1}{2}(X_{2p+1} + iY_{2p+1}) \prod_{k=1}^{2p} Z_k \tag{C.6a}$$

$$\hat{a}_p^\dagger = \frac{1}{2}(X_{2p} - iY_{2p}) \prod_{k=1}^{2p-1} Z_k \qquad \text{and} \qquad \hat{a}_{\bar{p}}^\dagger = \frac{1}{2}(X_{2p+1} - iY_{2p+1}) \prod_{k=1}^{2p} Z_k \tag{C.6b}$$

where X_p, Y_p, Z_p are the Pauli matrices acting on qubit p. In this transformation, the strings of Pauli-Z matrices account for the parity of the state, as required for fermionic antisymmetry. Assuming $r > s > q > p$ without loss of generality, the qubit representations of fermionic one-body and two-body operators between qubits p, q, r and s are

$$F_p^q = \frac{i}{2}\Big(Y_p X_q - X_p Y_q\Big) \prod_{k=p+1}^{q-1} Z_k, \tag{C.7a}$$

$$\begin{aligned} F_{pq}^{rs} = \frac{i}{8}\Big(&X_p X_q Y_s X_r + Y_p X_q Y_s Y_r + X_p Y_q Y_s Y_r + X_p X_q X_s Y_r \\ &- Y_p X_q X_s X_r - X_p Y_q X_s X_r - Y_p Y_q Y_s X_r - Y_p Y_q X_s Y_r\Big) \prod_{k=p+1}^{q-1} Z_k \prod_{l=r+1}^{s-1} Z_l, \end{aligned} \tag{C.7b}$$

where the $F_{p\ldots}^{q\cdots}$ notation distinguishes these as qubit operators.

To prepare a wave function with N spin-orbitals on an N-qubit register, the unitary transformations $\exp\Big(t_p^q F_p^q\Big)$ and $\exp\Big(t_{pq}^{rs} F_{pq}^{rs}\Big)$ need to be decomposed into elementary one- and two-qubit gates.[237] Such circuits involve applying a series of CNOT gates, known as a CNOT cascade, that implements the parity-computing action of the Pauli-Z strings in Eq. (C.7). These Z-strings are highly non-local as they act on $\mathcal{O}(N)$ spin-orbitals. Consequently, imposing the fermionic anticommutation relations on qubit operators requires a number of CNOT gates that scales linearly with the size of the single-particle basis, dominating the CNOT cost of fermionic ansätze. The development of qubit-excitation operators[262, 408] has been proposed to remove the Z-strings in Eq. (C.7) and reduce the number of CNOT gates, but this formalism does not satisfy fermionic antisymmetry.

C.2.2 Operators in s-UPS ansatz

The s-UPS ansatz introduced in the main text uses an operator pool of spin-adapted one-body operators $\hat{\kappa}_p^q + \hat{\kappa}_{\bar{p}}^{\bar{q}}$ and paired two-body operators $\hat{\kappa}_{p\bar{p}}^{q\bar{q}}$. The one-body operators map to the qubit representation as $\hat{\kappa}_p^q + \hat{\kappa}_{\bar{p}}^{\bar{q}} \rightarrow F_{2p}^{2q} + F_{2p+1}^{2q+1}$ and, due to the commutation $[\hat{\kappa}_p^q, \hat{\kappa}_{\bar{p}}^{\bar{q}}] = 0$, the exponential of the spin-adapted qubit operators can be written as $\exp\Big(t(F_{2p}^{2q} + F_{2p+1}^{2q+1})\Big) = \exp\Big(t\, F_{2p}^{2q}\Big) \exp\Big(t\, F_{2p+1}^{2q+1}\Big)$. We can therefore concatenate the circuits for each one-body spin-up and spin-down operator in a series and assign the same variational parameter to both.

Paired two-body operators are more gate-efficient compared to general two-body operators because the CNOT-overhead due to the parity computation vanishes if we define the ordering of qubits such that spin-up and spin-down orbitals that share the same spatial wave function are adjacent. Inserting these index definitions into Eq. (C.7), we see that the products of Z-operators vanish for the qubit operators corresponding to fermionic paired two-body operators $\kappa_{p\bar{p}}^{q\bar{q}} \rightarrow F_{2p,2p+1}^{2q,2q+1}$, giving

$$\begin{aligned} F_{2p,2p+1}^{2q,2q+1} = \frac{i}{8}\Big(& X_{2p}X_{2p+1}Y_{2q}X_{2q+1} + Y_{2p}X_{2p+1}Y_{2q}Y_{2q+1} + X_{2p}Y_{2p+1}Y_{2q}Y_{2q+1} \\ & + X_{2p}X_{2p+1}X_{2q}Y_{2q+1} - Y_{2p}X_{2p+1}X_{2q}X_{2q+1} - X_{2p}Y_{2p+1}X_{2q}X_{2q+1} \\ & - Y_{2p}Y_{2p+1}Y_{2q}X_{2q+1} - Y_{2p}Y_{2p+1}X_{2q}Y_{2q+1}\Big). \end{aligned} \tag{C.8}$$

Consequently, there is no parity-induced CNOT overhead for paired two-body fermionic operators and the number of CNOT gates required for this implementation is constant, as opposed to the $\mathcal{O}(N)$ CNOT scaling with respect to the number of spin orbitals and the excitation indices for general (unpaired) two-body operators. In fact, the operators are equivalent to the qubit excitations in Ref. 408. Removing the scaling of the CNOT count with respect to the number of spin orbitals will be especially important for larger systems since the implementation of two-qubit entangling gates is the main source of decoherence on noisy quantum hardware.

C.2.3 Circuit costs

Different circuit primitives exist for implementing fermionic operators (see Ref. 408 for a detailed discussion). The CNOT counts for our simulations (Fig. 5 in the main text) are based on the quantum circuits presented in Ref. 408 because these implementations were also used in Ref. 262, which contains the data that we compare our results against. In this approach, the number of CNOT gates required for each spin-adapted one-body fermionic operator $\exp\Big(t(\hat{\kappa}_p^q + \hat{\kappa}_{\bar{p}}^{\bar{q}})\Big)$ in the s-UPS ansatz is $8(p-q)+2 \geq 10$ (where p and q index the spatial orbitals), while paired two-body operators $\exp\Big(t\,\hat{\kappa}_{p\bar{p}}^{q\bar{q}}\Big)$ have a constant CNOT cost of 13.

C.3 Summary of the DISCO-VQE algorithm

This section summarises the implementation of the DISCO-VQE algorithm that is outlined in Fig. 1 in the main text.

C.3.1 Continuous optimisation using basin-hopping

For a given ordered set of M operators $\boldsymbol{\mu}$ with the wave function

$$|\Psi(\boldsymbol{t},\boldsymbol{\mu})\rangle = \prod_{i=1}^{M} e^{t_i\,\hat{\kappa}_{\mu_i}}\,|\Phi_0\rangle\,, \tag{C.9}$$

the electronic energy is defined by the expectation value

$$E(\boldsymbol{t},\boldsymbol{\mu}) = \langle\Phi_0|e^{-t_M\hat{\kappa}_{\mu_M}}\cdots e^{t_1\hat{\kappa}_{\mu_1}}\hat{H}e^{t_1\hat{\kappa}_{\mu_1}}\cdots e^{t_M\hat{\kappa}_{\mu_M}}|\Phi_0\rangle\,. \tag{C.10}$$

We use the lowest-energy restricted Hartree–Fock solution[269] to define the initial state $|\Phi_0\rangle$, although more advanced initial states could be considered. The corresponding energy landscape is highly non-linear with respect to the continuous coordinates $\boldsymbol{t}$, and we expect to find a number of local minima. Minimisation of the continuous coordinates is performed using a customised L-BFGS[422] (limited memory Broyden,[423] Fletcher,[424] Goldfarb,[425] Shanno[426]) routine implemented in `GMIN`.[427] Since we consider only real-valued operators, Hamiltonians, and coordinates, the analytic energy gradients are given by

$$\frac{\partial E(\boldsymbol{t},\boldsymbol{\mu})}{\partial t_j} = 2\,\langle\Psi(\boldsymbol{t},\boldsymbol{\mu})|\left(\hat{H}\left(\prod_{i=1}^{j} e^{t_i\hat{\kappa}_{\mu_i}}\right)\hat{\kappa}_{\mu_j}\left(\prod_{i=j+1}^{M} e^{t_i\hat{\kappa}_{\mu_i}}\right)\right)|\Phi_0\rangle\,. \tag{C.11}$$

These analytic gradients can be efficiently evaluated using the algorithm described in the supplementary information for Ref. 261.

We use the basin-hopping[402, 403] procedure to systematically search for the global minimum on this continuous energy landscape with a fixed set of operators $\boldsymbol{\mu}$. Basin-hopping uses random perturbations followed by minimisation to step between local minima on the continuous energy landscape, as described in detail elsewhere [see Ref. 428]. In summary, random steps in the continuous coordinate space are taken from the current minimum, with energy E_{old}, and the coordinates are reoptimised to obtain a new minimum with energy E_{new}. A Metropolis condition is used to accept steps if $E_{\text{new}} < E_{\text{old}}$, or if $E_{\text{old}} > E_{\text{new}}$ and $\exp((E_{\text{old}} - E_{\text{new}})/T_{\text{BH}})$ is greater than a random number in the range $[0, 1]$. Here, T_{BH} is an effective temperature that controls the magnitude of the energy increase for accepted uphill steps. This approach allows basin-hopping to systematically overcome the barriers between different minima on a continuous energy landscape. Periodic boundary conditions are used for the continuous coordinates, with the corresponding periodicity pre-computed using the eigenvalues of the (de)excitation operator matrix representations.

C.3.2 Performing discrete steps: generalised basin-hopping

Following a series of continuous basin-hopping steps, the lowest-energy continuous coordinates are saved and a series of discrete steps are tested. To extend the global optimisation procedure to include the discrete operator space we exploit the generalised basin-hopping approach,[399, 400] which was developed to treat multicomponent materials, and has since been applied to proteins and nucleic acids when the sequence of amino acids or nucleotides is mutated.[401] First, we consider the operator sets that can be reached by performing a cyclic permutation of the ordered set, before then considering mutations of each operator, and finally considering every possible swap between the position of two operators. The algorithm for each process is described below.

Cyclic permutations provide a way to quickly reorganise the ordered operator set without changing the relative position of certain operators. This discrete step is particularly useful in the early stages of the optimisation process where operators may be wrongly placed too close to the reference state. However, later in the optimisation we find that cyclic permutations rarely lower the energy and can have a detrimental impact on convergence. Therefore, a cyclic permutation is only accepted if it lowers the energy. These discrete steps are tested by starting from the best set of coefficients $(\boldsymbol{t}_0, \boldsymbol{\mu}_0)$ with energy E_0 identified in the most recent set of basin-hopping steps. For each possible cyclic permutation j, we permute both $\boldsymbol{t}$ and $\boldsymbol{\mu}$ and perform a single minimisation to obtain the optimised energy E_j. We then identify the lowest energy permutation E_k and accept this discrete step if $E_k < E_0$ (Algorithm 1). If the step is accepted, then the best coordinates and ordered operator set $(\boldsymbol{t}_0, \boldsymbol{\mu}_0)$ and the energy E_0 are updated with the new optimised variables.

Algorithm 1: Procedure for testing discrete cyclic permutations.

Current energy, continuous coordinates, and discrete operators
$E_0, \boldsymbol{t}_0, \boldsymbol{\mu}_0$

Loop over possible permutations
for $j = 1, M-1$ **do**
  Permute continuous coordinates and operators by j positions
  $\boldsymbol{t}_j, \boldsymbol{\mu}_j = \text{permute}(\boldsymbol{t}_0, \boldsymbol{\mu}_0, j)$

  Minimise energy and save result
  $E_j = \text{minimise}(\boldsymbol{t}_j, \boldsymbol{\mu}_j)$

Find index for the lowest energy permutation
$k = \text{find_min}(E_1, \ldots, E_{M-1})$

Return best result if energy is lower than initial state
if $E_k < E_0$ **then**
  return $E_k, \boldsymbol{t}_k, \boldsymbol{\mu}_k$
else
  return $E_0, \boldsymbol{t}_0, \boldsymbol{\mu}_0$

Operator mutations are the primary mechanism for changing the type of operators in the ordered set $\boldsymbol{\mu}$ and are considered once the best cyclic permutation has been selected, with variables $(\boldsymbol{t}_0, \boldsymbol{\mu}_0)$ and energy E_0. The mutation of each operator is considered individually, starting with the operator closest to the reference state (position M). The discrete search proceeds by testing every possible mutation of each operator i with an operator j selected from the pool with size N_{pool}. Following the mutation of an operator, a line search is performed to identify the best corresponding coordinate while the remaining coordinates are fixed. The result is then used as the initial guess for a minimisation of all the continuous coordinates, and the optimal energy is stored as E_j. Once all mutations for the operator i have been tested, a low-energy discrete step is selected based on the criterion described in Algorithm 4 and acceptance of this step is determined using a Metropolis criterion detailed in Algorithm 5. This process is then repeated for the next operator in the ordered set. A taboo criterion is used to ensure that uphill steps do not lead to cycling between a subset of operator orderings. This algorithm is outlined in

Algorithm 2.

Swapping the position of two operators allows the relative ordering to be changed while retaining the overall composition of the ordered operator set. These discrete steps are considered after the operator mutations have been tested and start with the optimal variables $(\boldsymbol{t}_0, \boldsymbol{\mu}_0)$ with energy E_0. As for the operator mutations, the pair swaps for each operator are considered individually, starting with the operator closest to the reference state (position M). The discrete search proceeds by testing every possible swap between this operator i and the another j in the ordered operator set. Once the operators and the corresponding continuous coordinates have been swapped, the configuration is optimised to identify the energy E_j. A low-energy discrete step is then selected using the algorithm described in Algorithm 4 and the acceptance of this step is determined using a Metropolis criterion detailed in Algorithm 5. This process is then repeated for the next operator in the ordered set. A taboo criterion is used to ensure that uphill steps do not lead to cycling between operator orderings. This algorithm is outlined in Algorithm 3.

C.3.3 Taking uphill discrete steps

The discrete landscape of ordered operator sets can have many local minima, and efficient global optimisation requires uphill moves to escape these traps. Therefore, we include a further Metropolis-Hastings condition[429, 430] with a second fictitious temperature parameter T_d to accept uphill steps for biminima,[399, 400] analogous to standard basin-hopping between local minima in the continuous parameter space.[402, 403] To ensure that we do not compromise the optimisation towards a given biminimum, these uphill steps are only considered when a full macro-cycle of discrete steps and continuous basin-hopping does not lower the energy and the algorithm is stuck in a biminimum. The energy after the last downhill step is recorded as E_best.

Algorithm 2: Procedure for taking discrete operator mutation steps. Routines for selecting a discrete step (select_step) and testing whether to accept it (test_step) are described in Algorithms 4 and 5 respectively.

Current energy, continuous coordinates, and discrete operators
$E_0, \boldsymbol{t}_0, \boldsymbol{\mu}_0$

Save initial energy for testing uphill steps
$E_{\text{best}} \leftarrow E_0$

Loop over operator positions, starting closest to the reference
for $i = M, 1$ **do**

  Loop over operators in the pool
  for $j = 1, N_{\text{pool}}$ **do**

    Reset coordinates and discrete operators
    $\boldsymbol{t}_j \leftarrow \boldsymbol{t}_0; \boldsymbol{\mu}_j \leftarrow \boldsymbol{\mu}_0$

    Mutate operator at position i into operator j
    $\boldsymbol{\mu}_j(i) \leftarrow j$

    Check if new operator ordering is not taboo
    $allow \leftarrow$ check_not_taboo$(\boldsymbol{\mu}_j)$

    if *allow* **then**

      Optimise i-th continuous coordinate using a line search
      $\boldsymbol{t}_j(i) \leftarrow$ line_search$(\boldsymbol{t}_j, \boldsymbol{\mu}_j, i)$

      Minimise energy and save result
      $E_j =$ minimise$(\boldsymbol{t}_j, \boldsymbol{\mu}_j)$

    Select mutation to test using fictitious temperature T_{d}
    $k \leftarrow$ select_step(E_j, E_0, T_{d})

    Determine whether to accept the step using fictitious temperature T_{d}
    $accept \leftarrow$ test_step$(E_k, E_0, E_{\text{best}}, T_{\text{d}})$

    if *accept* **then**

      Update current energy, continuous coordinates, and discrete operators
      $E_0 \leftarrow E_k$
      $\boldsymbol{t}_0 \leftarrow \boldsymbol{t}_k$
      $\boldsymbol{\mu}_0 \leftarrow \boldsymbol{\mu}_k$

Return best result
return $E_0, \boldsymbol{t}_0, \boldsymbol{\mu}_0$

Algorithm 3: Procedure for taking discrete operator swap steps. Routines for selecting a discrete step (select_step) and testing whether to accept it (test_step) are described in Algorithms 4 and 5 respectively.

Current energy, continuous coordinates, and discrete operators
$E_0, \boldsymbol{t}_0, \boldsymbol{\mu}_0$

Save initial energy for testing uphill steps
$E_{\text{best}} \leftarrow E_0$

Loop over operator positions, starting closest to reference
for $i = M, 1$ **do**

  Loop over position in the ordered operator set
  for $j = 1, M$ **do**

    Reset coordinates and discrete operators
    $\boldsymbol{t}_j \leftarrow \boldsymbol{t}_0; \boldsymbol{\mu}_j \leftarrow \boldsymbol{\mu}_0$

    Swap operators and coordinates at positions i and j
    $\boldsymbol{\mu}_j(i, j) \leftarrow \boldsymbol{\mu}_j(j, i)$
    $\boldsymbol{t}_j(i, j) \leftarrow \boldsymbol{t}_j(j, i)$

    Check if new operator ordering is not taboo
    allow $\leftarrow$ check_not_taboo($\boldsymbol{\mu}_j$)

    if *allow* **then**

      Minimise energy and save result
      $E_j = \text{minimise}(\boldsymbol{t}_j, \boldsymbol{\mu}_j)$

    Select mutation to test using fictitious temperature T_{d}
    $k \leftarrow \text{select_step}(E_j, E_{\text{best}}, T_{\text{d}})$

    Determine whether to accept the step using fictitious temperature T_{d}
    accept $\leftarrow \text{test_step}(E_k, E_0, E_{\text{best}}, T_{\text{d}})$

    if *accept* **then**

      Update current energy, continuous coordinates, and discrete operators
      $E_0 \leftarrow E_k$
      $\boldsymbol{t}_0 \leftarrow \boldsymbol{t}_k$
      $\boldsymbol{\mu}_0 \leftarrow \boldsymbol{\mu}_k$

Return best result
return $E_0, \boldsymbol{t}_0, \boldsymbol{\mu}_0$

When moving uphill, there is often a choice between many discrete steps with a similar energy. A test step is selected with a certain probability based on the energy difference relative to the lowest-energy discrete step under consideration. Specifically, if the lowest discrete step E_{min} increases the energy relative to the last downhill step (i.e. $E_{\text{min}} > E_{\text{best}}$), then the probability of selecting a step with energy E_j is given by $\exp(-(E_j - E_{\text{min}})/T_{\text{d}})$. Note that T_{d} does not have to be the same as the temperature used for controlling continuous basin-hopping steps, T_{BH}. Any step that decreases the energy relative to the last downhill step ($E_{\text{min}} < E_{\text{best}}$) is always selected. This approach, summarised in Algorithm 4, avoids any artificial bias in the algorithm towards a particular nearly-degenerate step. In addition, when uphill steps are not being considered, degenerate downhill discrete steps are selected with an equal probability.

Algorithm 4: Procedure for selecting temperature dependent steps.

Energies for n discrete steps under consideration, and last downhill step
E_j, E_{best}

Control variable for uphill steps
uphill

Identify lowest energy after discrete steps
$E_{\text{min}} = \text{min}_{j=1,n}(E_j)$

Compute probabilities $P(j)$ for taking step j
for *j=1,n* **do**
  if *uphill* **or** $E_{\text{min}} > E_{\text{best}}$ **then**
    $P(j) = \exp(-(E_j - E_{\text{min}})/T_{\text{d}})$
  else
    if $E_j = E_{\text{min}}$ **then**
      $P(j) = 1$
    else
      $P(j) = 0$

Normalise the probabilities
$P(j) = P(j) \,/\, \text{sum}_{j=1,n}(P(k))$

Randomly select a discrete step using computed probabilities
$k = \text{random_select}(P(1:n))$

Return selected step index
return k

Once a step with energy E_k has been selected using Algorithm 4, it is either accepted or rejected based on a a Metropolis criteria using the same fictitious temperature T_{d}. When uphill steps are not being considered (*uphill* = False), discrete steps are only accepted if $E_k < E_0$. When uphill steps are being considered (*uphill* = True), a discrete step with energy E_k is accepted if it lowers the energy relative to the last step, $E_k < E_0$, or with probability $\exp(-(E_k - E_0)/T_{\text{d}})$ if $E_k > E_0$. If any discrete step lowers the energy below the last downhill step, $E_k < E_{\text{best}}$, then no further uphill steps are considered in the current macro-cycle (i.e. *uphill* $\rightarrow$ False). This condition ensures that we do not take any more uphill steps than are needed to escape a biminimum. The algorithm for accepting or

rejecting proposed discrete steps is detailed in Algorithm 5.

Algorithm 5: Procedure for testing temperature-dependent steps.

Energies of proposed step k, last step, and last downhill step
$E_k, E_0, E_{\text{best}}$

Control variable for uphill steps
uphill

Turn off uphill steps if energy goes below last downhill step
if $E_k < E_{best}$ **then**
  uphill $\leftarrow$ False

Compute probability P for accepting step
if $E_k < E_0$ **then**
  Always accept a downhill step
  $P \leftarrow 1$
else
  Otherwise accept based on Metropolis criterion
  if *uphill* **then**
    $P = \exp(-(E_k - E_0)/T_{\text{d}})$
  else
    $P = 0$

Test if step is accepted using a random number in range $[0, 1]$
$accept = P >$ random$(0, 1)$

return *accept*

We also employ the basin-hopping parallel tempering (BHPT) scheme[431] to improve efficiency. Multiple generalised basin-hopping runs, comprising the continuous and discrete optimisation steps, are performed in parallel using different temperature parameters for accepting moves between biminima, with exchange between replicas analogous to the parallel tempering formalism.[432] The different temperatures apply to the Metropolis criteria used in both the continuous basin-hopping and discrete uphill steps (i.e., $T_{\text{BH}} = T_{\text{d}}$). In this approach, the high-temperature replicas can efficiently move over high barriers on the energy landscape while low-temperature replicas fine-tune the optimisation in each basin

of attraction. Finally, we employ a taboo criterion to encourage the discrete search to climb uphill out of biminima and prevent cycling between a local neighbourhood of ordered operator sets. In this approach, a taboo is placed on ordered sets of operators that prevent a given configuration from being revisited within a certain number of discrete steps N_t.

C.3.4 Summary of hyperparameters

In addition to the standard basin-hopping methodology, the DISCO-VQE algorithm introduces the following hyperparameters:

- **Number of Basin-Hopping steps (N_BH)**
  The number of basin-hopping steps performed on each macrocycle DISCO-VQE;
- **Taboo delay (N_t)**
  A particular ordered set of operators cannot be revisited within N_t discrete steps;
- **Discrete temperature (T_d)**
  The fictitious temperature associated with selecting and accepting discrete uphill steps.

Preliminary numerical studies identified suitable values as $N_\mathrm{BH} = 100$, $N_\mathrm{t} = 20$, and $T_\mathrm{d} = 0.0001$–$0.01\,\mathrm{E_h}$.

C.3.5 Computational implementation

We have implemented DISCO-VQE in the `GMIN` program,[427] developed by some of us for finding global minima on continuous energy landscapes. We use `PySCF`[433] to generate reference Hartree–Fock solutions and the corresponding one- and two-electron Hamiltonian integrals. We then use `OpenFermion`[434] to generate Hilbert-space matrix representations of the Hamiltonian operator and the anti-Hermitian spin-adapted one-body and paired two-body (de)excitation operators. These matrices provide the input for DISCO-VQE calculations in `GMIN`, where all subsequent

computations are performed. All calculations are performed with the restricted Hartree–Fock[269] reference as the initial state, corresponding to an empty set of unitary operators and $\boldsymbol{t} = \mathbf{0}$.

C.4 Linear and tetrahedral H_4 (STO-3G)

C.4.1 Computational Details

The same calculation setup was used for the convergence with respect to the number of operators and the full DISCO-VQE ground-state binding curves in both linear and tetrahedral H_4:

- 8 BHPT replicas at temperatures distributed exponentially from 0.0001 to 0.01 E_h. Replica exchange considered after every 10 continuous basin-hopping steps. Same fictitious temperature for basin-hopping steps and discrete uphill steps ($T_\mathrm{BH} = T_\mathrm{d}$);
- Maximum continuous step size of 0.5, updated dynamically to give an average 50 % acceptance;
- Total of 100 000 basin-hopping steps;
- $N_\mathrm{BH} = 100$ basin-hopping steps for each DISCO-VQE macrocycle;
- Taboo delay of $N_\mathrm{t} = 20$.

For the linear H_4 binding curves computed using twenty fixed random orderings of 13 operators and the the best operator ordering identified at $R(\mathrm{H{-}H}) = 0.9\,\text{Å}$, the optimisation was only performed for the continuous coordinates using the BHPT approach. The following input parameters were used:

- 8 BHPT replicas at temperatures distributed exponentially from 0.0001 to 0.01 E_h. Replica exchange considered after every 10 continuous basin-hopping steps;
- Maximum continuous step size of 0.5, updated dynamically to give an average 50 % acceptance;
- Total of 100 000 basin-hopping steps.

The use of 100 000 steps is chosen, since numerical testing suggests that this value allows the global minimum to be identified while significant further sampling does not produce any further improvement.

C.4.2 Variants of DISCO-VQE with minimal computational cost

The computational cost of DISCO-VQE scales linearly with the number of basin-hopping steps and discrete macrocycles. In the main text, we argue that the accuracy of DISCO-VQE must always be at least as good as ADAPT-VQE (with a paired operator pool) even for a single macrocycle, while the discrete search can still explore the continuous space without any continuous basin-hopping steps. To illustrate these properties, we explore two variants of DISCO-VQE for the convergence of linear H_4 (STO-3G) with the number of operators. The first variant employs a single discrete macrocycle of operator mutations and pair swaps starting from the empty operator pool (Fig. C.2; orange line). This approach is at least as accurate as ADAPT-VQE using the paired operator pool (Fig. C.2; red line), although it also stagnates at the same energy once the ansatz contains at least 10 operators. The second variant employs 1000 macrocycles of the discrete search with no continuous basin-hopping steps, with 8 BHPT replicas at temperatures distributed exponentially from 0.0001 to 0.01 E_h for the discrete search, and a taboo delay of $N_\mathrm{t} = 20$ (Fig. C.2; green line). This approach reaches chemical accuracy at 9 operators and is almost as accurate as the full DISCO-VQE calculations reported in the main text (Fig. C.2; blue line), demonstrating the ability of the discrete search to explore the discrete space without continuous basin-hopping. We expect that including the efficiency gained by using continuous basin-hopping will be increasingly important for large numbers of operators, where the continuous energy landscape will be more non-convex.

C.4.3 Comparison to ADAPT-VQE using unpaired two-body operators

In the ADAPT-VQE algorithm, convergence is assessed using the magnitude of the discrete energy gradient with respect to adding an operator $\hat{\kappa}_\mu$ from the pool, defined by the commutators[261]

$$\frac{\partial E}{\partial \hat{\kappa}_\mu} = 2 \langle \Psi(\boldsymbol{t}, \boldsymbol{\mu}) | [\hat{H}, \hat{\kappa}_\mu] | \Psi(\boldsymbol{t}, \boldsymbol{\mu}) \rangle \,. \tag{C.12}$$

When the operator pool $\{\hat{\kappa}_\mu\}$ includes all unpaired generalised one- and two-body operators, these gradient elements coincide with the residual of the Anti-Hermitian Schrödinger equation (ACSE),[404, 435] suggesting that converged ADAPT-VQE solutions should correspond to exact eigenstates.[405] In contrast, using ADAPT-VQE with the generalised spin-adapted one-body and paired two-body operators no longer satisfies this relationship. Therefore, using the paired operator pool, the ACSE residual components for unpaired two-body operators can grow while ADAPT-VQE cannot directly optimise them to zero (i.e. operator combinations are required).

While ADAPT-VQE with unpaired two-body operators should be able to converge to an exact eigenstate, the global optimisation using DISCO-VQE should give the same accuracy with an equivalent, or smaller, number of operators. To assess the advantage of the DISCO-VQE optimisation algorithm in this context, calculations were performed for linear H_4 using the operator pool defined in the original ADAPT-VQE implementation, corresponding to generalised spin-adapted one-body and spin-complemented two-body operators, $\{\hat{\kappa}^q_p + \hat{\kappa}^{\bar{q}}_{\bar{p}},\ \hat{\kappa}^{rs}_{pq} + \hat{\kappa}^{\bar{r}\bar{s}}_{\bar{p}\bar{q}},\ \hat{\kappa}^{r\bar{s}}_{p\bar{q}} + \hat{\kappa}^{s\bar{r}}_{q\bar{p}}\}$, with $p < q$ and $r < s$.[261] Note that these spin-complemented two-body operators cannot be implemented on a quantum device without an additional Trotter expansion, since they require the exponentiation of a sum of non-commuting operators. Furthermore, while they conserve the $\langle \hat{S}_z \rangle$ value of the initial state, they can break the total spin symmetry associated with the $\hat{S}^2$ operator.

As expected, using ADAPT-VQE to optimise the wave function with these unpaired two-body operators avoids the stagnation observed for the paired operator pool, reaching an energetic error of $10^{-8}\,\mathrm{E_h}$ with 18 operators (Fig. 5.2A). In contrast, DISCO-VQE can obtain a numerically exact result using only 11 operators, providing much faster and more robust convergence than ADAPT-VQE. This DISCO-VQE representation of the exact wave function requires fewer operators than the paired operator pool because the action of unpaired operators no longer needs to be constructed through the combination of multiple operators. However, the truncated wave functions using unpaired two-body operators are spin-contaminated (Fig. C.3), with unphysical values of $\langle \hat{S}^2 \rangle$, while using only paired two-body operators conserves spin-symmetry at every truncation level.

C.4.4 Binding curve using a fixed ansatz

The optimal operator ordering identified with thirteen operators at a bond length of 0.90 Å was used to define a fixed ansatz with the operator sequence

$$\{\hat{\kappa}^{2\bar{2}}_{1\bar{1}},\ \hat{\kappa}^{4}_{2}+\hat{\kappa}^{\bar{4}}_{\bar{2}},\ \hat{\kappa}^{4\bar{4}}_{1\bar{1}},\ \hat{\kappa}^{4\bar{4}}_{2\bar{2}},\ \hat{\kappa}^{4}_{2}+\hat{\kappa}^{\bar{4}}_{\bar{2}},\ \hat{\kappa}^{4\bar{4}}_{1\bar{1}},\ \hat{\kappa}^{3}_{1}+\hat{\kappa}^{\bar{3}}_{\bar{1}},\ \hat{\kappa}^{4\bar{4}}_{1\bar{1}},\ \hat{\kappa}^{4}_{1}+\hat{\kappa}^{\bar{4}}_{\bar{1}},\ \hat{\kappa}^{4\bar{4}}_{3\bar{3}},\ \hat{\kappa}^{4\bar{4}}_{1\bar{1}},\ \hat{\kappa}^{3\bar{3}}_{2\bar{2}},\ \hat{\kappa}^{3}_{2}+\hat{\kappa}^{\bar{3}}_{\bar{2}}\} \tag{C.13}$$

where the ordering relative to the initial state is

$$|\Psi(\boldsymbol{t})\rangle = e^{t_1\hat{\kappa}^{2\bar{2}}_{1\bar{1}}}\cdots e^{t_{13}(\hat{\kappa}^{3}_{2}+\hat{\kappa}^{\bar{3}}_{\bar{2}})}\,|\Phi_0\rangle\,. \tag{C.14}$$

BHPT calculations were performed to identify the optimal continuous coordinates at each bond length. This fixed ansatz gives exact energies at the majority of bond lengths considered and has a maximum error of $0.379\mu\mathrm{E_h}$ (Fig. C.4). The accuracy of this ansatz demonstrates the transferability of optimal DISCO-VQE wave functions across different geometries.

These results demonstrate the potential of DISCO-VQE to discover compact operator orderings for defining fixed ansätze for practical VQE implementations. In this case, our ansatz provides an essentially exact representation of the ground state with a total CNOT count of 194, estimated using the circuit costs outlined in Section C.2. The optimised variational parameters at selected bond lengths are given

Table C.1: Variational parameters and energies corresponding to the optimised solutions obtained for the fixed linear H_4 (STO-3G) ansatz constructed from the operator sequence defined in Eq. (C.13) binding curve.

Operator	0.5 Å	1.0 Å	1.5 Å	2.0 Å	2.5 Å	3.0 Å	3.5 Å	4.0 Å
$\hat{\kappa}_{1\bar{1}}^{2\bar{2}}$	−0.42495	2.812 45	2.868 19	−1.38418	2.126 27	2.502 54	1.827 73	−2.38227
$\hat{\kappa}_{2}^{4} + \hat{\kappa}_{\bar{2}}^{\bar{4}}$	−2.09072	−0.45568	2.754 04	0.551 72	−0.56157	−0.46455	2.493 77	−2.70864
$\hat{\kappa}_{1\bar{1}}^{4\bar{4}}$	−3.06248	3.075 36	2.923 22	−2.20230	−1.83494	−2.07241	−1.58757	2.070 07
$\hat{\kappa}_{2\bar{2}}^{4\bar{4}}$	1.59293	−1.48211	1.732 29	−1.98399	−1.29499	−1.45016	−1.11968	−1.72557
$\hat{\kappa}_{2}^{4} + \hat{\kappa}_{\bar{2}}^{\bar{4}}$	−2.61363	1.152 68	−1.27434	0.774 37	0.664 47	−2.59082	0.795 37	−2.62754
$\hat{\kappa}_{1\bar{1}}^{4\bar{4}}$	1.54387	−2.24225	2.547 34	0.568 14	−0.81727	2.685 39	−0.81703	2.679 07
$\hat{\kappa}_{1}^{3} + \hat{\kappa}_{\bar{1}}^{\bar{3}}$	−2.84172	0.114 16	0.191 35	1.240 63	−1.17288	1.959 48	−1.19036	−1.93316
$\hat{\kappa}_{1\bar{1}}^{4\bar{4}}$	2.37287	3.017 00	−1.71874	−0.23077	−0.01791	−1.61373	0.037 10	−0.06678
$\hat{\kappa}_{1}^{4} + \hat{\kappa}_{\bar{1}}^{\bar{4}}$	−2.35620	−2.35619	−2.35619	−0.78540	−0.78540	−2.35619	−2.35619	0.785 40
$\hat{\kappa}_{3\bar{3}}^{4\bar{4}}$	3.14066	3.140 13	−3.14017	3.140 84	3.102 67	3.126 67	3.136 50	3.140 39
$\hat{\kappa}_{1\bar{1}}^{4\bar{4}}$	0.78540	−2.35619	0.785 40	0.785 40	0.785 70	−2.35615	−2.35619	−2.35619
$\hat{\kappa}_{2\bar{2}}^{3\bar{3}}$	1.47618	2.942 89	−2.73118	0.582 83	2.984 14	−2.88999	−2.11749	−2.98211
$\hat{\kappa}_{2}^{3} + \hat{\kappa}_{\bar{2}}^{\bar{3}}$	−1.55906	0.068 97	−0.17689	1.301 91	−0.29982	2.834 91	−1.87839	1.263 16
Energy /E_h	−1.653117	−2.1663387	−1.996150	−1.897781	−1.872216	−1.867291	−1.866463	−1.866344

in Table C.1, providing the information required to implement these wave functions as practical quantum circuits. The non-linearity of the s-UPS wave functions can lead to equivalent minima on the continuous energy landscape, and thus individual optimisations at sequential geometries can jump between equivalent representations of the ground state even though the wave function evolves continuously.

C.4.5 Comparison of DISCO-VQE and ADAPT-VQE binding curves

The energetic error for truncated DISCO-VQE approximations is compared to ADAPT-VQE across the full linear H_4 (STO-3G) binding curve in Fig. C.5. ADAPT-VQE calculations, implemented in `GMIN`, were performed with the same operator pool with no restriction on the number of operators and were considered converged when the energy no longer changed on each iteration. DISCO-VQE calculations with six operators give more accurate energies than ADAPT-VQE, giving a non-parallelity error (NPE) of $12.02\,\mathrm{m}E_\mathrm{h}$ compared to $33.05\,\mathrm{m}E_\mathrm{h}$ for ADAPT-VQE. Ten operators is sufficient for DISCO-VQE to give a chemically accurate binding curve with NPE $= 0.29\,\mathrm{m}E_\mathrm{h}$.

C.5 Triplet ground-state of O_2 (STO-3G)

C.5.1 Computational Details

The ground-state $\langle S_z \rangle = 1$ ROHF solution at $R(\mathrm{O}{-}\mathrm{O}) = 2.55\,\text{Å}$ was obtained using Q-Chem 5.4[436] and was verified to have the same energy as the reference state in Ref. 391. Subsequent Hamiltonian matrix elements were computed with two frozen core orbitals. DISCO-VQE calculations to investigate the energy convergence with respect to the number of operators employed the following parameters:

- 12 BHPT replicas at temperatures distributed exponentially from 0.0001 to $0.1\,E_\mathrm{h}$. Replica exchange was attempted after every 10 continuous basin-hopping steps. The same fictitious temperature was used for basin-hopping steps and discrete uphill steps ($T_\mathrm{BH} = T_\mathrm{d}$);
- Maximum continuous step size of 0.5, updated dynamically to give an average 50 % acceptance;
- $N_\mathrm{BH} = 100$ basin-hopping steps for each DISCO-VQE macrocycle;
- Taboo delay of $N_\mathrm{t} = 20$.
- Total of 20 000 basin-hopping steps;

We employed more parallel BHPT replicas for these calculations compared to H_4 to offset the smaller total number of basin-hopping steps that can achieved for this larger system, while still ensuring sufficient exploration of the space to provide reliable global optimisation results. Corresponding ADAPT-VQE calculations were performed for up to 1000 operators to remove the effect of choosing a gradient-based convergence threshold.

C.6 Linear H_6 binding curve (STO-3G)

C.6.1 Computational Details

DISCO-VQE calculations for the linear H_6 binding curve employed the following parameters:

- 8 BHPT replicas at temperatures distributed exponentially from 0.0001 to 0.01 E_h. Replica exchange considered after every 10 continuous basin-hopping steps. Same fictitious temperature for basin-hopping steps and discrete uphill steps ($T_\mathrm{BH} = T_\mathrm{d}$);
- Maximum continuous step size of 0.5, updated dynamically to give an average 50 % acceptance;
- $N_\mathrm{BH} = 100$ basin-hopping steps for each DISCO-VQE macrocycle;
- Taboo delay of $N_\mathrm{t} = 20$.

To improve the probability of finding the global minimum in this benchmarking exercise, calculations for $R(\mathrm{H{-}H) \leq 1.2\,\text{Å}}$ with 10, 20, or 30 operators were performed with 40 000 basin-hopping steps. Calculations for $R(\mathrm{H{-}H) > 1.2\,\text{Å}}$ with 10, 20, or 30 operators were performed with 20 000 basin-hopping steps, which reliably obtained the global minimum while retaining a suitable computational cost. 20, 000 basin-hopping steps were used for all calculations with 40 operators, regardless of the bond length. ADAPT-VQE calculations are performed up to 1000 operators to remove the effect of choosing a gradient-based convergence threshold.

C.7 Symmetric H_2O binding curve (STO-3G)

C.7.1 Computational Details

The number of operators in the s-UPS wave function was fixed to the size of the operator pool, 42. DISCO-VQE calculations for the H_2O symmetric stretch employed the following parameters:

- 8 BHPT replicas at temperatures distributed exponentially from 0.0001 to 0.01 E_h. Replica exchange considered after every 10 continuous basin-hopping steps. Same fictitious temperature for basin-hopping steps and discrete uphill steps ($T_\mathrm{BH} = T_\mathrm{d}$);
- Maximum continuous step size of 0.5, updated dynamically to give an average 50 % acceptance;

- Total of 20 000 basin-hopping steps;
- $N_{\mathrm{BH}} = 100$ basin-hopping steps for each DISCO-VQE macrocycle;
- Taboo delay of $N_{\mathrm{t}} = 20$.

The use of 20 000 basin-hopping steps was the largest feasible value within the computational resources available. ADAPT-VQE calculations are performed up to 1000 operators to remove the effect of choosing a gradient-based convergence threshold.

C.8 N_2 binding curve (STO-3G)

C.8.1 Computational Details

The number of operators in the s-UPS wave function was fixed to the size of the operator pool (30) and the four lowest molecular orbitals in the STO-3G were frozen. DISCO-VQE calculations for the N_2 binding curve employed the following parameters:

- 8 BHPT replicas at temperatures distributed exponentially from 0.0001 to 0.01 E_{h}. Replica exchange considered after every 10 continuous basin-hopping steps. Same fictitious temperature for basin-hopping steps and discrete uphill steps ($T_{\mathrm{BH}} = T_{\mathrm{d}}$);
- Maximum continuous step size of 0.5, updated dynamically to give an average 50 % acceptance;
- Total of 20 000 basin-hopping steps;
- $N_{\mathrm{BH}} = 100$ basin-hopping steps for each DISCO-VQE macrocycle;
- Taboo delay of $N_{\mathrm{t}} = 20$.

The use of 20 000 basin-hopping steps was the largest feasible value within the computational resources available.

C.8.2 Comparison of binding curve accuracy

Table C.2 compares the mean-average error for related fermionic ansatz across the N_2 binding curve with the number of variational parameters in the wave function. The un-Trotterised k-UpCCGSD ansatz[393] requires 180 variational parameters ($k = 4$) to obtain a better mean-average error (MAE) than DISCO-VQE and 135 parameters ($k = 3$) to get a more accurate NPE. ADAPT-VQE calculations using the same operator pool do not give a qualitatively correct binding curve, giving a large MAE and NPE. Depending on the bond length, ADAPT-VQE calculations converge with between 10 to 100 operators, after which adding more operators leaves the energy unchanged.

Table C.2: Comparison of mean-average error (MAE) and non-parallelity error (NPE) of related fermionic ansatz for the N_2 (STO-3G) binding curve. The un-Trotterised k-UpCCGSD results are taken from Ref. 393.

Method	# Parameters	MAE /mE_h	NPE /mE_h
ADAPT-VQE (paired)	10–100	87.88	242.55
DISCO-VQE	30	0.77	5.08
1-UpCCGSD	45	43.16	41.09
2-UpCCGSD	90	13.16	8.85
3-UpCCGSD	135	2.81	2.45
4-UpCCGSD	180	0.60	1.51
5-UpCCGSD	225	0.31	0.77
6-UpCCGSD	270	0.15	0.36

C.9 Two-dimensional Hubbard lattice

DISCO-VQE calculations on the 4×2 Hubbard lattice employed periodic boundary conditions along the long axis and open boundary conditions along the short axis. Since the 4900-dimensional Hilbert space is computationally expensive for BHPT, we perform DISCO-VQE using a dynamic temperature update and divide the discrete step search over parallel processes. In particular, after each set of 50 uphill discrete steps, the corresponding temperature T_d is either increased or decreased to keep the average acceptance ratio close to 50 %. This temperature for discrete uphill steps is not necessarily the same as the temperature T_{BH} for

the continuous Metropolis criteria used in the basin-hopping steps. The input parameters for these calculations are:

- Maximum continuous step size of 0.5, updated dynamically to give an average 50 % acceptance;
- Temperature for continuous basin-hopping steps $T_{\mathrm{BH}} = 0.001\,t$;
- Initial temperature for uphill discrete steps $T_{\mathrm{d}} = 0.01\,t$, updated dynamically to give 50 % acceptance;
- $N_{\mathrm{BH}} = 100$ basin-hopping steps for each DISCO-VQE macrocycle;
- Total of 5 000 basin-hopping steps;
- Taboo delay of $N_{\mathrm{t}} = 20$.

The use of 5 000 basin-hopping steps was the largest feasible value within the computational resources available. Corresponding ADAPT-VQE simulations used a convergence threshold corresponding to a discrete gradient root-mean-square threshold of $10^{-4}\,t$.

C.9.1 Investigating the 1/8-filling fraction

Doping fractions away from half-filling are also important for understanding the phase diagram of the two-dimensional Hubbard lattice. However, the sign problem in quantum Monte–Carlo, and the open-shell character, makes these systems difficult to predict using standard computational techniques.[410] To investigate the dependence on the filling fraction, DISCO-VQE calculations were performed on the 4×2 Hubbard lattice at the 1/8-filling fraction with periodic boundary conditions along the long axis and open boundary conditions along the short axis. The input parameters for these calculations are:

- 8 BHPT replicas at temperatures distributed exponentially from 0.0001 to $0.01\,t$. Replica exchange considered after every 10 continuous basin-hopping steps. Same fictitious temperature for basin-hopping steps and discrete uphill steps ($T_{\mathrm{BH}} = T_{\mathrm{d}}$);

- Maximum continuous step size of 0.5, updated dynamically to give an average 50 % acceptance;
- Total of 20 000 basin-hopping steps;
- $N_{\mathrm{BH}} = 100$ basin-hopping steps for each DISCO-VQE macrocycle;
- Taboo delay of $N_{\mathrm{t}} = 20$.

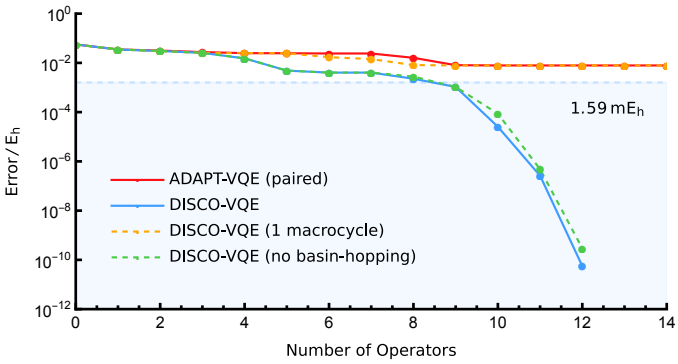


Figure C.2: A single macrocycle of DISCO-VQE is sufficient to match the accuracy of ADAPT-VQE using the same operator pool, while the discrete search can still obtain an exact representation without any continuous basin-hopping steps. These properties are illustrated for the convergence with respect to the number of operators for linear H_4 (STO-3G) at $R(\mathrm{H{-}H}) = 0.90\,\text{Å}$.

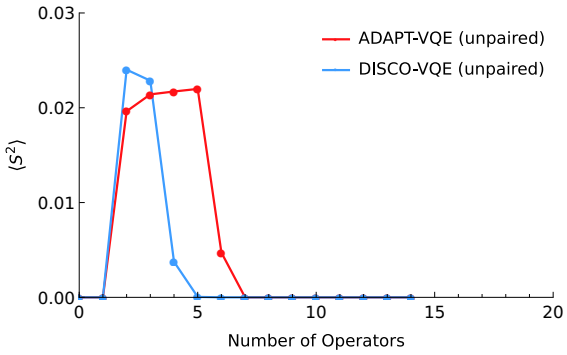


Figure C.3: Using a pool containing spin-complemented unpaired two-body operators for linear H_4 (STO-3G) at $R(\mathrm{H{-}H}) = 0.90\,\text{Å}$ results in spin-contaminated truncated wave functions, with a non-zero expectation value $\langle \hat{S}^2 \rangle$. The spin-contamination decays more rapidly using DISCO-VQE compared to ADAPT-VQE optimisation.

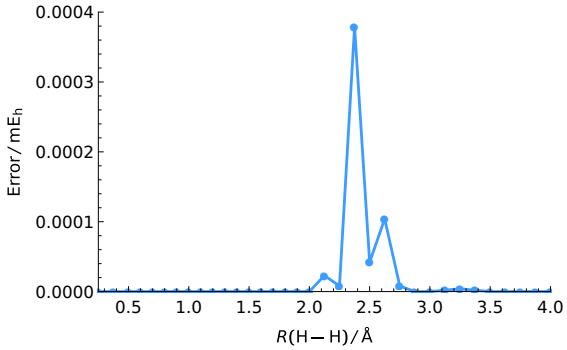


Figure C.4: Error in the total energy for linear H_4 (STO-3G) using the optimal DISCO-VQE(13) ansatz identified at $R(\mathrm{H{-}H}) = 0.90\,\text{Å}$.

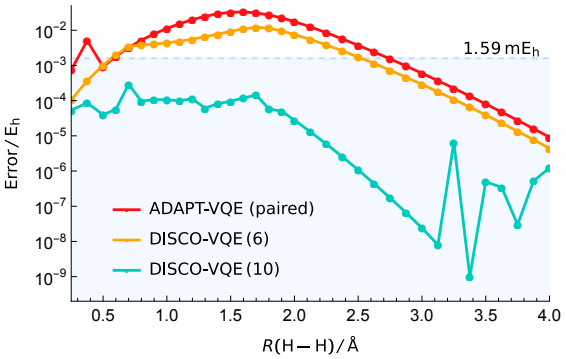


Figure C.5: Comparison of the energy error for linear H_4 (STO-3G) using truncated DISCO-VQE approximations and ADAPT-VQE. The NPE for ADAPT-VQE, DISCO-VQE(6) and DISCO-VQE(10) are 33.05, 12.02, and 0.29 mE_h respectively.

References

[1] F. Arute et al., *Quantum supremacy using a programmable superconducting processor*, Nature **574**, 505–510 (2019).

[2] H. S. Zhong et al., *Quantum computational advantage using photons*, Science **370**, 1460–1463 (2020).

[3] A. Morvan et al., *Phase transitions in random circuit sampling*, Nature **634**, 328–333 (2024).

[4] S. Boixo et al., *Characterizing quantum supremacy in near-term devices*, Nature Physics **14**, 595–600 (2018).

[5] S. Aaronson and L. Chen, *Complexity-theoretic foundations of quantum supremacy experiments*, Leibniz International Proceedings in Informatics, LIPIcs **79** (2016).

[6] S. Aaronson and S. H. Hung, *Certified randomness from quantum supremacy*, Proceedings of the Annual ACM Symposium on Theory of Computing, 933–944 (2023).

[7] S. Aaronson and Y. Zhang, *On verifiable quantum advantage with peaked circuit sampling*, arXiv:2404.14493 (2024).

[8] Y. Liu et al., *Verifying quantum advantage experiments with multiple amplitude tensor network contraction*, Physical Review Letters **132**, 030601 (2024).

[9] H. Luo and A. Alavi, *Combining the Transcorrelated Method with Full Configuration Interaction Quantum Monte Carlo: Application to the Homogeneous Electron Gas*, J. Chem. Theory Comput. **14**, 1403 (2018).

[10] K. T. Schütt et al., *Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions*, Nature Communications **10**, 5024 (2019).

[11] L. Yang et al., *Deep learning-enhanced variational Monte Carlo method for quantum many-body physics*, Physical Review Research **2**, 012039 (2020).

[12] J. Hermann, Z. Schätzle, and F. Noé, *Deep-neural-network solution of the electronic Schrödinger equation*, Nature Chemistry **12**, 891–897 (2020).

[13] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, *Ab initio solution of the many-electron Schrödinger equation with deep neural networks*, Physical Review Research **2**, 033429 (2020).

[14] J. Hermann et al., *Ab initio quantum chemistry with neural-network wavefunctions*, Nature Reviews Chemistry **7**, 692–709 (2023).

[15] G. K.-L. Chan and S. Sharma, *The density matrix renormalization group in quantum chemistry*, Annual Review of Physical Chemistry **62**, 465–481 (2011).

[16] S. Sharma and G. K. L. Chan, *Spin-adapted density matrix renormalization group algorithms for quantum chemistry*, The Journal of Chemical Physics **136**, 124121 (2012).

[17] S. Sharma, K. Sivalingam, F. Neese, and G. K. L. Chan, *Low-energy spectrum of iron-sulfur clusters directly from many-particle quantum mechanics*, Nature Chemistry **6**, 927–933 (2014).

[18] Z. Li, S. Guo, Q. Sun, and G. K. L. Chan, *Electronic landscape of the P-cluster of nitrogenase as revealed through many-electron quantum wavefunction simulations*, Nature Chemistry **11**, 1026–1033 (2019).

[19] D. B. Williams-Young, N. M. Tubman, C. Mejuto-Zaera, and W. A. de Jong, *A parallel, distributed memory implementation of the adaptive sampling configuration interaction method*, Journal of Chemical Physics **158**, 214109 (2023).

[20] H. Gao, S. Imamura, A. Kasagi, and E. Yoshida, *Distributed implementation of full configuration interaction for one trillion determinants*, Journal of Chemical Theory and Computation **20**, 1185–1192 (2024).

[21] A. Menczer et al., *Parallel implementation of the density matrix renormalization group method achieving a quarter petaFLOPS performance on a single DGX-H100 GPU node*, Journal of Chemical Theory and Computation **20**, 8397–8404 (2024).

[22] A. E. Deprince and J. R. Hammond, *Coupled cluster theory on graphics processing units I. The coupled cluster doubles method*, Journal of Chemical Theory and Computation **7**, 1287–1295 (2011).

[23] M. Alkan et al., *LibERI - a portable and performant multi-GPU accelerated library for electron repulsion integrals via OpenMP offloading and standard language parallelism*, Journal of Chemical Physics **161**, 082501 (2024).

[24] C. Xiang, W. Jia, W. H. Fang, and Z. Li, *Distributed Multi-GPU ab initio density matrix renormalization group algorithm with applications to the P-Cluster of nitrogenase*, Journal of Chemical Theory and Computation **20**, 775–786 (2024).

[25] R. Li, Q. Sun, X. Zhang, and G. K.-L. Chan, *Introducing GPU-acceleration into the Python-based simulations of chemistry framework*, The Journal of Physical Chemistry A **129** (2024).

[26] M. Ganahl et al., *Density matrix renormalization group with tensor processing units*, PRX Quantum **4**, 010317 (2023).

[27] Q. Sun et al., *Recent developments in the PySCF program package*, The Journal of Chemical Physics **153**, 024109 (2020).

[28] E. Bernstein and U. Vazirani, *Quantum complexity theory*, SIAM Journal on Computing **26**, 1411–1473 (1997).

[29] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM Journal on Computing **26**, 1484–1509 (1997).

[30] Y. Manin, *Computable and non-computable* (Sovetskoe Radio, Moscow, 1980).

[31] R. P. Feynman, *Simulating physics with computers*, International Journal of Theoretical Physics **21**, 467–488 (1982).

[32] S. McArdle et al., *Quantum computational chemistry*, Reviews of Modern Physics **92**, 15003 (2020).

[33] Y. Cao et al., *Quantum Chemistry in the Age of Quantum Computing*, Chem. Rev. **119**, 10856 (2019).

[34] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, *Quantum Algorithms for Quantum Chemistry and Quantum Materials Science*, Chem. Rev. **120**, 12685 (2020).

[35] R. J. Bartlett and M. Musia, *Coupled-cluster theory in quantum chemistry*, Rev. Mod. Phys. **79**, 291 (2007).

[36] K. Burke, *Perspective on density functional theory*, Journal of Chemical Physics **136**, 150901 (2012).

[37] A. Szabo and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory* (Dover Publications, 2000).

[38] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular electronic-structure theory* (John Wiley and Sons, 2014).

[39] J. Kempe, A. Kitaev, and O. Regev, *The complexity of the local hamiltonian problem*, SIAM Journal on Computing **35**, 1070–1097 (2006).

[40] B. OGorman, S. Irani, J. Whitfield, and B. Fefferman, *Intractability of Electronic Structure in a Fixed Basis*, PRX Quantum **3**, 020322 (2022).

[41] S. Lee et al., *Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry*, Nature Communications **14**, 1–7 (2023).

[42] J. F. Gonthier et al., *Measurements as a roadblock to near-term practical quantum advantage in chemistry: resource analysis*, Physical Review Research **4**, 033154 (2022).

[43] R. N. Tazhigulov et al., *Simulating models of challenging correlated molecules and materials on the sycamore quantum processor*, PRX Quantum **3**, 040318 (2022).

[44] K. Dalton et al., *Quantifying the effect of gate errors on variational quantum eigensolvers for quantum chemistry*, npj Quantum Information 2024 10:1 **10**, 1–11 (2024).

[45] K. Bharti et al., *Noisy intermediate-scale quantum algorithms*, Reviews of Modern Physics **94**, 015004 (2022).

[46] A. M. Dalzell et al., *Quantum algorithms: a survey of applications and end-to-end complexities*, arXiv:2310.03011 (2023).

[47] Y. Zhang et al., *Fault-tolerant quantum algorithms for quantum molecular systems: A survey*, arXiv:2502.02139 (2025).

[48] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, 2010).

[49] I. Shavitt and R. Bartlett, *Many-Body Methods in Chemistry and Physics* (Cambridge University Press, 2009).

[50] D. Marti-Dafcik, N. Lee, H. G. A. Burton, and D. P. Tew, *Spin-coupled molecular orbitals: chemical intuition meets quantum chemistry*, arXiv:2402.08858 (2024).

[51] D. Marti-Dafcik, H. G. A. Burton, and D. P. Tew, *Spin coupling is all you need: Encoding strong electron correlation in molecules on quantum computers*, Physical Review Research **7**, 013191 (2025).

[52] H. G. Burton, D. Marti-Dafcik, D. P. Tew, and D. J. Wales, *Exact electronic states with shallow quantum circuits from global optimisation*, npj Quantum Information 2023 9:1 **9**, 1–8 (2023).

[53] Y. Chi et al., *A programmable qudit-based quantum processor*, Nature Communications **13**, 1–10 (2022).

[54] A. Gilyen, Y. Su, G. H. Low, and N. Wiebe, *Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics*, Proceedings of the Annual ACM Symposium on Theory of Computing, 193–204 (2019).

[55] L. K. Grover, *A fast quantum mechanical algorithm for database search*, Proceedings of the Annual ACM Symposium on Theory of Computing, 212–219 (1996).

[56] A. M. Childs and N. Wiebe, *Hamiltonian Simulation using linear combinations of unitary operations*, Quantum Information and Computation **12**, 901–924 (2012).

[57] G. H. Low and I. L. Chuang, *Optimal Hamiltonian simulation by quantum signal processing*, Physical Review Letters **118**, 010501 (2017).

[58] G. H. Low and I. L. Chuang, *Hamiltonian simulation by qubitization*, Quantum **3**, 163 (2019).

[59] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, *Quantum amplitude amplification and estimation*, arXiv:quant-ph/0005055, 53–74 (2000).

[60] T. Scholten et al., *Assessing the benefits and risks of quantum computers*, IEEE Security & Privacy (2024).

[61] E. Campbell, A. Khurana, and A. Montanaro, *Applying quantum algorithms to constraint satisfaction problems*, Quantum **3**, 167 (2019).

[62] R. Babbush et al., *Focus beyond quadratic speedups for error-corrected quantum advantage*, PRX Quantum **2**, 010103 (2021).

[63] T. Hoefler, T. Häner, and M. Troyer, *Disentangling hype from practicality: on realistically achieving quantum advantage*, Communications of the ACM **66**, 82–87 (2023).

[64] M. Reiher et al., *Elucidating reaction mechanisms on quantum computers*, Proceedings of the National Academy of Sciences of the United States of America **114**, 7555–7560 (2017).

[65] J. Lee et al., *Even more efficient quantum computations of chemistry through tensor hypercontraction*, PRX Quantum **2**, 030305 (2021).

[66] S. Lloyd, *Universal quantum simulators*, Science **273**, 1073–1078 (1996).

[67] A. Y. Kitaev, *Quantum measurements and the abelian stabilizer problem*, arXiv:quant-ph/9511026 (1995).

[68] D. S. Abrams and S. Lloyd, *Simulation of many-body Fermi systems on a universal quantum computer*, Physical Review Letters **79**, 2586–2589 (1997).

[69] D. W. Berry et al., *Improved techniques for preparing eigenstates of fermionic Hamiltonians*, npj Quantum Information **4**, 22 (2018).

[70] P. Jordan and E. Wigner, *Über das Paulische Äquivalenzverbot*, Z. Phys. **47**, 631 (1928).

[71] R. Somma et al., *Simulating physical phenomena by quantum networks*, Physical Review A **65**, 17 (2002).

[72] D. S. Abrams and S. Lloyd, *Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors*, Physical Review Letters **83**, 5162–5165 (1999).

[73] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, *Quantum computation by adiabatic evolution*, arXiv:quant-ph/0001106 (2000).

[74] E. Farhi et al., *A quantum adiabatic evolution algorithm applied to random instances of an np-complete problem*, Science **292**, 472–476 (2001).

[75] D. Aharonov et al., *Adiabatic quantum computation is equivalent to standard quantum computation*, SIAM Rev **50**, 755–787 (2008).

[76] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, *Simulated quantum computation of molecular energies*, Science **309**, 1704–1707 (2005).

[77] D. Wecker et al., *Gate-count estimates for performing quantum chemistry on small quantum computers*, Physical Review A - Atomic, Molecular, and Optical Physics **90**, 022305 (2014).

[78] D. Poulin et al., *The Trotter step size required for accurate quantum simulation of quantum chemistry*, Quantum Information and Computation **15**, 361–384 (2014).

[79] J. R. McClean, R. Babbush, P. J. Love, and A. Aspuru-Guzik, *Exploiting locality in quantum computation for quantum chemistry*, Journal of Physical Chemistry Letters **5**, 4368–4380 (2014).

[80] R. Babbush et al., *Chemical basis of Trotter-Suzuki errors in quantum chemistry simulation*, Physical Review A **91**, 022311 (2015).

[81] R. Babbush et al., *Exponentially more precise quantum simulation of fermions in second quantization*, New Journal of Physics **18**, 033032 (2016).

[82] R. Babbush et al., *Exponentially more precise quantum simulation of fermions in the configuration interaction representation*, Quantum Science and Technology **3**, 15006 (2018).

[83] R. Babbush et al., *Encoding electronic spectra in quantum circuits with linear t complexity*, Physical Review X **8**, 041015 (2018).

[84] M. Motta et al., *Low rank representations for quantum simulation of electronic structure*, npj Quantum Information 2021 7:1 **7**, 1–7 (2018).

[85] D. W. Berry et al., *Qubitization of arbitrary basis quantum chemistry leveraging sparsity and low rank factorization*, Quantum **3**, 208 (2019).

[86] I. D. Kivlichan, C. E. Granade, and N. Wiebe, *Phase estimation with randomized Hamiltonians*, arXiv:1907.10070 (2019).

[87] K. Wan, M. Berta, and E. T. Campbell, *Randomized quantum algorithm for statistical phase estimation*, Physical Review Letters **129**, 030503 (2022).

[88] V. von Burg et al., *Quantum computing enhanced computational catalysis*, Physical Review Research **3**, 033055 (2021).

[89] C. L. Cortes et al., *Assessing the query complexity limits of quantum phase estimation using symmetry-aware spectral bounds*, Physical Review A **110**, 022420 (2024).

[90] A. Caesura et al., *Faster quantum chemistry simulations on a quantum computer with improved tensor factorization and active volume compilation*, arXiv:2501.06165 (2025).

[91] A. Peruzzo et al., *A variational eigenvalues solver on a photonic quantum processor*, Nat. Comm. **5**, 4213 (2014).

[92] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, *The theory of variational hybrid quantum-classical algorithms*, New J. Phys. **18**, 023023 (2016).

[93] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, *Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states*, Physical Review A **95**, 042308 (2017).

[94] J. I. Colless et al., *Computation of molecular spectra on a quantum processor with an error-resilient algorithm*, Physical Review X **8**, 011021 (2018).

[95] M. Motta et al., *Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution*, Nature Physics **16**, 205–210 (2019).

[96] R. M. Parrish and P. L. McMahon, *Quantum filter diagonalization: quantum eigendecomposition without full quantum phase estimation*, arXiv:1909.08925 (2019).

[97] W. J. Huggins et al., *A non-orthogonal variational quantum eigensolver*, New Journal of Physics **22**, 073009 (2020).

[98] N. H. Stair, R. Huang, and F. A. Evangelista, *A multireference quantum Krylov algorithm for strongly correlated electrons*, Journal of Chemical Theory and Computation **16**, 2236–2245 (2020).

[99] K. Seki and S. Yunoki, *Quantum power method by a superposition of time-evolved states*, PRX Quantum **2**, 010333 (2021).

[100] K. Klymko et al., *Real-time evolution for ultracompact hamiltonian eigenstates on quantum hardware*, PRX Quantum **3**, 020323 (2022).

[101] E. N. Epperly, L. Lin, and Y. Nakatsukasa, *A theory of quantum subspace diagonalization*, SIAM Journal on Matrix Analysis and Applications **43**, 1263–1290 (2022).

[102] C. L. Cortes and S. K. Gray, *Quantum Krylov subspace algorithms for ground- and excited-state energy estimation*, Physical Review A **105**, 022417 (2022).

[103] W. Kirby, M. Motta, and A. Mezzacapo, *Exact and efficient Lanczos method on a quantum computer*, Quantum **7**, 1018 (2023).

[104] L. Lin and Y. Tong, *Near-optimal ground state preparation*, Quantum **4**, 372 (2020).

[105] L. Lin and Y. Tong, *Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers*, PRX Quantum **3**, 010318 (2022).

[106] J. J. Goings et al., *Reliably assessing the electronic structure of cytochrome P450 on today's classical computers and tomorrow's quantum computers*, Proceedings of the National Academy of Sciences of the United States of America **119**, e2203533119 (2022).

[107] N. A. Bogdanov et al., *Enhancement of superexchange due to synergetic breathing and hopping in corner-sharing cuprates*, Nature Physics **18**, 190–195 (2021).

[108] H. R. Larsson, H. Zhai, C. J. Umrigar, and G. K. L. Chan, *The chromium dimer: closing a chapter of quantum chemistry*, Journal of the American Chemical Society **144**, 15932–15937 (2022).

[109] N. C. Rubin et al., *Fault-tolerant quantum simulation of materials using bloch orbitals*, PRX Quantum **4**, 040303 (2023).

[110] S. R. White, *Density matrix formulation for quantum renormalization groups*, Physical Review Letters **69**, 2863 (1992).

[111] S. R. White, *Density-matrix algorithms for quantum renormalization groups*, Physical Review B **48**, 10345 (1993).

[112] Z. Li et al., *The electronic complexity of the ground-state of the FeMo cofactor of nitrogenase as relevant to quantum simulations*, The Journal of Chemical Physics **150**, 024302 (2019).

[113] R. Santagati et al., *Drug design on quantum computers*, Nature Physics **20**, 549–557 (2024).

[114] Y. Li and S. C. Benjamin, *Efficient variational quantum simulator incorporating active error minimization*, Physical Review X **7**, 021050 (2017).

[115] S. Mcardle, X. Yuan, and S. Benjamin, *Error-mitigated digital quantum simulation*, Physical Review Letters **122**, 180501 (2019).

[116] Z. Cai et al., *Quantum error mitigation*, Reviews of Modern Physics **95**, 045005 (2023).

[117] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, *Surface codes: towards practical large-scale quantum computation*, Physical Review A - Atomic, Molecular, and Optical Physics **86**, 032324 (2012).

[118] C. Gidney and A. G. Fowler, *Efficient magic state factories with a catalyzed* $|CCZ\rangle$ *to* $2\,|T\rangle$ *transformation*, Quantum **3**, 135 (2019).

[119] P. A. M. Dirac, *Quantum mechanics of many-electron systems*, Proceedings of the Royal Society of London **123**, 714–733 (1929).

[120] E. Schrödinger, *Quantisierung als Eigenwertproblem*, Annalen der Physik **385**, 437–490 (1926).

[121] E. Schrödinger, *Der stetige Übergang von der Mikro- zur Makromechanik*, Die Naturwissenschaften **14**, 664–666 (1926).

[122] E. Schrödinger, *An undulatory theory of the mechanics of atoms and molecules*, Physical Review **28**, 1049 (1926).

[123] P. A. M. Dirac, *The quantum theory of the emission and absorption of radiation*, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **114**, 243–265 (1927).

[124] J. W. S. B. Rayleigh, *The theory of sound* (Macmillan, 1877).

[125] W. Ritz, *Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik.* Journal für die reine und angewandte Mathematik **135**, 1–61 (1909).

[126] W. Heisenberg, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen.* Zeitschrift für Physik **33**, 879–893 (1925).

[127] M. Born and P. Jordan, *Zur Quantenmechanik*, Zeitschrift für Physik **34**, 858–888 (1925).

[128] M. Born, W. Heisenberg, and P. Jordan, *Zur Quantenmechanik. II.* Zeitschrift für Physik **35**, 557–615 (1926).

[129] M. Born and R. Oppenheimer, *Zur Quantentheorie der Molekeln*, Ann. Phys. **389**, 457 (1927).

[130] D. R. Hartree, Hartree, and D. R., *The wave mechanics of an atom with a non-Coulomb central field. Part I. Theory and methods*, PCPS **24**, 89 (1928).

[131] V. Fock, *Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems*, Zeitschrift für Physik **61**, 126–148 (1930).

[132] I. Waller and D. R. Hartree, *On the intensity of total scattering of X-rays*, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **124**, 119–142 (1929).

[133] J. C. Slater, *The theory of complex spectra*, Physical Review **34**, 1293 (1929).

[134] A. D. Becke, *Density-functional thermochemistry. III. The role of exact exchange*, The Journal of Chemical Physics **98**, 5648–5652 (1993).

[135] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple*, Physical Review Letters **77**, 3865–3868 (1996).

[136] N. Mardirossian and M. Head-Gordon, *Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals*, Molecular Physics **115**, 2315–2372 (2017).

[137] G. H. Golub and C. F. V. Loan, *Matrix computations* (Johns Hopkins University Press, Nov. 2013).

[138] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J.Res.Natl.Bur.Stand.B **45**, 255–282 (1950).

[139] E. R. Davidson, *The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices*, J. Comput. Phys. **17**, 87 (1975).

[140] J. Olsen et al., *Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces*, The Journal of Chemical Physics **89**, 2185–2192 (1988).

[141] J. Olsen, P. Jørgensen, and J. Simons, *Passing the one-billion limit in full configuration-interaction (FCI) calculations*, Chemical Physics Letters **169**, 463–472 (1990).

[142] N. H. Beebe and J. Linderberg, *Simplifications in the generation and transformation of two-electron integrals in molecular calculations*, International Journal of Quantum Chemistry **12**, 683–705 (1977).

[143] F. Aquilante, T. B. Pedersen, and R. Lindh, *Low-cost evaluation of the exchange Fock matrix from Cholesky and density fitting representations of the electron repulsion integrals*, Journal of Chemical Physics **126**, 194106 (2007).

[144] N. J. Higham, *Accuracy and stability of numerical algorithms* (Society for Industrial and Applied Mathematics Publications, 2002), pp. 1–656.

[145] T. G. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM Review **51**, 455–500 (2009).

[146] J. Gray and S. Kourtis, *Hyper-optimized tensor network contraction*, Quantum **5**, 410 (2021).

[147] M. C. Bañuls, *Tensor network algorithms: a route map*, Annual Review of Condensed Matter Physics **14**, 173–191 (2023).

[148] N. Glaser and M. Reiher, *Solving intractable chemical problems by tensor decomposition*, Chimia **78**, 215–221 (2024).

[149] E. G. Hohenstein, R. M. Parrish, and T. J. Martínez, *Tensor hypercontraction density fitting. I. Quartic scaling second- and third-order Møller-Plesset perturbation theory*, Journal of Chemical Physics **137**, 1085 (2012).

[150] R. M. Parrish, E. G. Hohenstein, T. J. Martínez, and C. D. Sherrill, *Tensor hypercontraction. II. Least-squares renormalization*, Journal of Chemical Physics **137**, 224106 (2012).

[151] E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, and T. J. Martínez, *Communication: Tensor hypercontraction. III. Least-squares tensor hypercontraction for the determination of correlated wavefunctions*, Journal of Chemical Physics **137**, 221101 (2012).

[152] P. Pulay, *Convergence acceleration of iterative sequences. the case of scf iteration*, Chem. Phys. Lett. **73**, 393 (1980).

[153] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes. the art of scientific computing*, Vol. 53 (Cambridge University Press, 2007), pp. 1–1262.

[154] T. H. Dunning Jr., *Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen*, J. Chem. Phys. **90**, 1007 (1989).

[155] R. A. Kendall, T. H. Dunning, and R. J. Harrison, *Electron affinities of the firstrow atoms revisited. Systematic basis sets and wave functions*, The Journal of Chemical Physics **96**, 6796–6806 (1992).

[156] F. Weigend and R. Ahlrichs, *Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy*, Physical Chemistry Chemical Physics **7**, 3297–3305 (2005).

[157] S. Obara and A. Saika, *Efficient recursive computation of molecular integrals over Cartesian Gaussian functions*, The Journal of Chemical Physics **84**, 3963–3974 (1986).

[158] M. Head-Gordon and J. A. Pople, *A method for two-electron Gaussian integral and integral derivative evaluation using recurrence relations*, The Journal of Chemical Physics **89**, 5777–5786 (1988).

[159] P. M. Gill, *Molecular integrals over Gaussian basis functions*, Advances in Quantum Chemistry **25**, 141–205 (1994).

[160] F. A. Evangelista, *Perspective: Multireference coupled cluster theories of dynamical electron correlation*, J. Chem. Phys. **149**, 030901 (2018).

[161] A. I. Krylov, *Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker's Guide to Fock Space*, Annu. Rev. Phys. Chem. **59**, 433 (2008).

[162] B. O. Roos, P. R. Taylor, and P. E. M. Sigbahn, *A complete active space SCF method (CASSCF) using a density matrix formulated super-CI*, Chem. Phys. **48**, 157 (1980).

[163] G. H. Booth, A. J. W. Thom, and A. Alavi, *Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space*, J. Chem. Phys. **131**, 054106 (2009).

[164] L. Lin, *Lecture notes on quantum algorithms for scientific computation*, arXiv:2201.08309 (2022).

[165] J. M. Martyn, Z. M. Rossi, A. K. Tan, and I. L. Chuang, *Grand unification of quantum algorithms*, PRX Quantum **2**, 040203 (2021).

[166] D. Motlagh and N. Wiebe, *Generalized quantum signal processing*, PRX Quantum **5**, 020368 (2024).

[167] G. H. Low and Y. Su, *Quantum eigenvalue processing*, 65th Annual Symposium on Foundations of Computer Science (FOCS) (2024).

[168] J. Preskill, *Quantum computing and the entanglement frontier*, arXiv:1203.5813 (2012).

[169] B. O'Gorman, W. J. Huggins, E. G. Rieffel, and K. B. Whaley, *Generalized swap networks for near-term quantum computing*, arxiv:1905.05118v1 (2019).

[170] A. M. Childs, *Universal computation by quantum walk*, Physical Review Letters **102**, 180501 (2009).

[171] D. W. Berry and A. M. Childs, *Black-box Hamiltonian simulation and unitary implementation*, Quantum Information and Computation **12**, 29–62 (2009).

[172] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, *Efficient quantum algorithms for simulating sparse Hamiltonians*, Communications in Mathematical Physics **270**, 359–371 (2007).

[173] D. W. Berry et al., *Exponential improvement in precision for simulating sparse Hamiltonians*, Proceedings of the Annual ACM Symposium on Theory of Computing, 283–292 (2014).

[174] D. W. Berry et al., *Simulating Hamiltonian dynamics with a truncated Taylor series*, Physical Review Letters **114**, 090502 (2015).

[175] M. Motta, J. Shee, S. Zhang, and G. K. L. Chan, *Efficient ab initio auxiliary-field quantum Monte Carlo calculations in Gaussian bases via low-rank tensor decomposition*, Journal of Chemical Theory and Computation **15**, 3510–3521 (2019).

[176] B. Peng and K. Kowalski, *Highly efficient and scalable compound decomposition of two-electron integral tensor and its application in coupled cluster calculations*, Journal of Chemical Theory and Computation **13**, 4179–4192 (2017).

[177] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, *Simulation of electronic structure Hamiltonians using quantum computers*, Molecular Physics **109**, 735–750 (2011).

[178] R. Babbush, D. W. Berry, J. R. McClean, and H. Neven, *Quantum simulation of chemistry with sublinear scaling in basis size*, npj Quantum Information **5**, 1–7 (2019).

[179] G. H. Low and N. Wiebe, *Hamiltonian Simulation in the interaction picture*, arXiv:1805.00675 (2018).

[180] Y. Su et al., *Fault-tolerant quantum simulations of chemistry in first quantization*, PRX Quantum **2** (2021).

[181] N. M. Tubman et al., *Postponing the orthogonality catastrophe: efficient state preparation for electronic structure simulations on quantum devices*, arXiv:1809.05523 (2018).

[182] G. H. Low, V. Kliuchnikov, and L. Schaeffer, *Trading T gates for dirty qubits in state preparation and unitary synthesis*, Quantum **8**, 1375 (2024).

[183] S. Sim, J. Romero, J. F. Gonthier, and A. A. Kunitsa, *Adaptive pruning-based optimization of parametrized quantum circuits*, Quantum Sci. Technol. **6**, 025019 (2021).

[184] J. Hu et al., *Benchmarking variational quantum eigensolvers for quantum chemistry*, arXiv:2211.12775 (2022).

[185] R. DCunha, T. D. Crawford, M. Motta, and J. E. Rice, *Challenges in the use of quantum computing hardware-efficient ansätze in electronic structure theory*, Journal of Physical Chemistry A **127**, 3437–3448 (2023).

[186] L. Veis and J. Pittner, *Adiabatic state preparation study of methylene*, Journal of Chemical Physics **140**, 224109 (2014).

[187] V. Kremenetski, C. Mejuto-Zaera, S. J. Cotton, and N. M. Tubman, *Simulation of adiabatic quantum computing for molecular ground states*, The Journal of Chemical Physics **155**, 234106 (2021).

[188] K. Sugisaki et al., *Adiabatic state preparation of correlated wave functions with nonlinear scheduling functions and broken-symmetry wave functions*, Communications Chemistry 2022 5:1 **5**, 1–13 (2022).

[189] R. Pauncz, *Spin eigenfunctions: construction and use* (Springer US (New York), 1979).

[190] W. Dobrautz, S. D. Smart, and A. Alavi, *Efficient formulation of full configuration interaction quantum Monte Carlo in a spin eigenbasis via the graphical unitary group approach*, Journal of Chemical Physics **151** (2019).

[191] D. Poulin, A. Qarry, R. Somma, and F. Verstraete, *Quantum simulation of time-dependent Hamiltonians and the convenient illusion of Hilbert space*, Physical Review Letters **106**, 170501 (2011).

[192] G. H. Low, V. Kliuchnikov, and L. Schaeffer, *Trading t-gates for dirty qubits in state preparation and unitary synthesis*, arXiv:1812.00954 (2018).

[193] P. Shor, *Fault-tolerant quantum computation*, Proceedings of 37th Conference on Foundations of Computer Science, 56–65 (1996).

[194] E. Knill, R. Laflamme, and W. H. Zurek, *Resilient quantum computation*, Science **279**, 342–345 (1998).

[195] A. Kitaev, *Fault-tolerant quantum computation by anyons*, Annals of Physics **303**, 2–30 (2003).

[196] D. Aharonov and M. Ben-Or, *Fault-tolerant quantum computation with constant error rate*, SIAM Journal on Computing **38**, 1207–1282 (2008).

[197] Y. Cao, J. Romero, and A. Aspuru-Guzik, *Potential of quantum computing for drug discovery*, IBM Journal of Research and Development **62** (2018).

[198] N. S. Blunt et al., *Perspective on the current state-of-the-art of quantum computing for drug discovery applications*, Journal of Chemical Theory and Computation **18**, 7001–7023 (2022).

[199] Google Quantum AI, *Willow Spec Sheet*, (Dec. 2024)

[200] R. Acharya et al., *Quantum error correction below the surface code threshold*, Nature 2024, 1–7 (2024).

[201] J. Preskill, *Quantum computing in the NISQ era and beyond*, Quantum **2**, 79 (2018).

[202] J. Romero et al., *Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz*, Quantum Science and Technology **4** (2019).

[203] J. C. Slater, *Quantum theory of atomic structure*, Vol. 1 (McGraw-Hill, 1960).

[204] M. Reiher and A. Wolf, *Relativistic quantum chemistry: the fundamental theory of molecular science: second edition*, Relativistic Quantum Chemistry: The Fundamental Theory of Molecular Science: Second Edition, 1–737 (2015).

[205] B. M. Austin, D. Y. Zubarev, and W. A. Lester, *Quantum Monte Carlo and related approaches*, Chemical Reviews **112**, 263–288 (2012).

[206] W. M. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, *Quantum Monte Carlo simulations of solids*, Reviews of Modern Physics **73**, 33 (2001).

[207] I. Kassal et al., *Polynomial-time quantum algorithm for the simulation of chemical dynamics*, Proceedings of the National Academy of Sciences of the United States of America **105**, 18681–18686 (2008).

[208] I. D. Kivlichan, N. Wiebe, R. Babbush, and A. Aspuru-Guzik, *Bounding the costs of quantum simulation of many-body physics in real space*, Journal of Physics A: Mathematical and Theoretical **50**, 305301 (2017).

[209] R. Babbush et al., *Low-depth quantum simulation of materials*, Physical Review X **8**, 011044 (2018).

[210] W. J. Hehre, R. F. Stewart, and J. A. Pople, *Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals*, J. Chem. Phys. **51**, 2657 (1969).

[211] T. Helgaker, W. Klopper, and D. P. Tew, *Quantitative quantum chemistry*, Molecular Physics **106**, 2107–2143 (2010).

[212] L. Pauling, *The nature of the chemical bond. Application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules*, Journal of the American Chemical Society **44**, 13 (1931).

[213] H. G. Hiscock and A. J. W. Thom, *Holomorphic Hartree–Fock Theory and Configuration Interaction*, J. Chem. Theory Comput. **10**, 4795 (2014).

[214] Q. Sun et al., *Pyscf: the Pythonbased simulations of chemistry framework*, WIREs Computational Molecular Science **8** (2018).

[215] D. J. Thouless, *Stability Conditions and Nuclear Rotations in the Hartree–Fock Theory*, Nucl. Phys. **21**, 225 (1960).

[216] A. D. Gunasekera, N. Lee, and D. P. Tew, *Multi-reference coupled cluster theory using the normal ordered exponential ansatz*, Faraday Discussions **254**, 170–190 (2024).

[217] L. Ding, S. Knecht, Z. Zimborás, and C. Schilling, *Quantum correlations in molecules: from quantum resourcing to chemical bonding*, Quantum Science and Technology **8**, 015015 (2022).

[218] D. P. Tew, W. Klopper, and T. Helgaker, *Electron correlation: the many-body problem at the heart of chemistry*, Journal of Computational Chemistry **28**, 1307–1320 (2007).

[219] J. Lee and M. Head-Gordon, *Distinguishing artifical and essential symmetry breaking in a single determinant: approach and application to the C_{60}, C_{36}, and C_{20} fullerenes*, Phys. Chem. Chem. Phys. **21**, 4763 (2019).

[220] L. Ding et al., *Concept of orbital entanglement and correlation in quantum chemistry*, Journal of Chemical Theory and Computation **17**, 79–95 (2021).

[221] L. Cheng et al., *Bond dissociation energies for diatomic molecules containing 3d transition metals: benchmark scalar-relativistic coupled-cluster calculations for 20 molecules*, Journal of Chemical Theory and Computation **13**, 1044–1056 (2017).

[222] K. Andersson et al., *Second-order perturbation theory with a CASSCF reference function*, J. Phys. Chem. **93**, 5483 (1990).

[223] K. Andersson, P.-Å. Malmqvist, and B. O. Roos, *Second-order perturbation theory with a complete active space self-consistent field reference function*, J. Chem. Phys. **96**, 1218 (1992).

[224] R. Izsák et al., *Measuring electron correlation: the impact of symmetry and orbital transformations*, Journal of Chemical Theory and Computation **19**, 2703–2720 (2023).

[225] A. I. Krylov, *The Quantum Chemistry of Open-Shell Species*, in Reviews in Computational Chemistry (John Wiley & Sons, Ltd, 2017) Chap. 4, p. 151.

[226] N. M. Tubman et al., *A deterministic alternative to the full configuration interaction quantum Monte Carlo method*, Journal of Chemical Physics **145**, 44112 (2016).

[227] A. A. Holmes, N. M. Tubman, and C. J. Umrigar, *Heat-Bath Configuration Interaction: An Efficient Selected Configuration Interaction Algorithm Inspired by Heat-Bath Sampling*, J. Chem. Theory Comput. **12**, 3674 (2016).

[228] V. Abraham and N. J. Mayhall, *Selected Configuration Interaction in a Basis of Cluster State Tensor Products*, J. Chem. Theory Comput. **16**, 6098 (2020).

[229] V. Abraham and N. J. Mayhall, *Cluster many-body expansion: A many-body expansion of the electron correlation energy about a cluster mean field reference*, J. Chem. Phys. **155**, 054101 (2021).

[230] V. Abraham and N. J. Mayhall, *Coupled electron pair-type approximations for tensor product state wave functions*, Journal of Chemical Theory and Computation **18**, 4856–4864 (2022).

[231] J. Robledo-Moreno et al., *Chemistry beyond the scale of exact diagonalization on a quantum-centric supercomputer*, Science Advances **11** (2025).

[232] D. S. Levine et al., *CASSCF with extremely large active spaces using the adaptive sampling configuration interaction method*, Journal of Chemical Theory and Computation **16**, 2340–2354 (2020).

[233] G. H. Booth, D. Cleland, A. J. Thom, and A. Alavi, *Breaking the carbon dimer: The challenges of multiple bond dissociation with full configuration interaction quantum Monte Carlo methods*, The Journal of Chemical Physics **135**, 084104 (2011).

[234] G. K. L. Chan, *Low entanglement wavefunctions*, Wiley Interdisciplinary Reviews: Computational Molecular Science **2**, 907–920 (2012).

[235] M. B. Hastings, *An area law for one-dimensional quantum systems*, Journal of Statistical Mechanics: Theory and Experiment **2007**, P08024 (2007).

[236] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, *Matrix product states and projected entangled pair states: Concepts, symmetries, theorems*, Reviews of Modern Physics **93**, 045003 (2021).

[237] A. Anand et al., *A quantum computing view on unitary coupled cluster theory*, Chem. Soc. Rev. **51**, 1659 (2022).

[238] I. D. Kivlichan et al., *Quantum simulation of electronic structure with linear depth and connectivity*, Physical Review Letters **120**, 110501 (2018).

[239] S. B. Bravyi and A. Y. Kitaev, *Fermionic Quantum Computation*, Ann. Phys. **298**, 210 (2002).

[240] J. T. Seeley, M. J. Richard, and P. J. Love, *The Bravyi-Kitaev transformation for quantum computation of electronic structure*, Journal of Chemical Physics **137**, 224109 (2012).

[241] A. Miller et al., *Bonsai algorithm: grow your own fermion-to-qubit mappings*, PRX Quantum **4**, 030314 (2023).

[242] C. Zalka, *Simulating quantum systems on a quantum computer*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **454**, 313–322 (1998).

[243] N. J. Ward, I. Kassal, and A. Aspuru-Guzik, *Preparation of many-body states for quantum simulation*, The Journal of Chemical Physics **130**, 194105 (2009).

[244] H. H. S. Chan et al., *Grid-based methods for chemistry simulations on a quantum computer*, Science Advances **9** (2023).

[245] F. Bloch, *Über die Quantenmechanik der Elektronen in Kristallgittern*, Zeitschrift für Physik **52**, 555–600 (1929).

[246] W. Kohn, *Analytic properties of Bloch waves and Wannier functions*, Physical Review **115**, 809 (1959).

[247] N. Marzari et al., *Maximally localized Wannier functions: Theory and applications*, Reviews of Modern Physics **84**, 1419–1475 (2012).

[248] B. L. Higgins et al., *Entanglement-free Heisenberg-limited phase estimation*, Nature **450**, 393–396 (2007).

[249] E. Knill, G. Ortiz, and R. D. Somma, *Optimal quantum measurements of expectation values of observables*, Physical Review A - Atomic, Molecular, and Optical Physics **75**, 012328 (2007).

[250] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, *Quantum algorithms revisited*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **454**, 339–354 (1998).

[251] T. Albash and D. A. Lidar, *Adiabatic quantum computation*, Reviews of Modern Physics **90**, 015002 (2018).

[252] M. Motta et al., *Subspace methods for electronic structure simulations on quantum computers*, Electronic Structure **6**, 013001 (2024).

[253] S. Fomichev et al., *Initial state preparation for quantum chemistry on quantum computers*, arXiv:2310.18410 (2024).

[254] A. M. Childs et al., *Theory of Trotter error with commutator scaling*, Physical Review X **11**, 011020 (2021).

[255] A. M. Childs et al., *Toward the first quantum simulation with quantum speedup*, Proceedings of the National Academy of Sciences of the United States of America **115**, 9456–9461 (2018).

[256] M. Cerezo et al., *Cost function dependent barren plateaus in shallow parametrized quantum circuits*, Nat. Comm. **12**, 1791 (2021).

[257] S. Patel, P. Jayakumar, T.-C. Yen, and A. F. Izmaylov, *Quantum measurement for quantum chemistry on a quantum computer*, arXiv:2501.14968 (2025).

[258] E. R. Anschuetz and B. T. Kiani, *Quantum variational algorithms are swamped with traps*, Nature Communications **13**, 1–10 (2022).

[259] J. R. McClean et al., *Barren plateaus in quantum neural network training landscapes*, Nat. Comm. **9**, 4812 (2018).

[260] A. Kandala et al., *Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets*, Nature **549**, 242 (2017).

[261] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, *An adaptive variational algorithm for exact molecular simulations on a quantum computer*, Nat. Comm. **10**, 3007 (2019).

[262] Y. S. Yordanov, V. Armaos, C. H. W. Barnes, and D. R. M. Arvidsson-Shukur, *Qubit-excitation-based adaptive variational quantum eigensolver*, Commun. Phys. **4**, 228 (2021).

[263] P. L. Dallaire-Demers et al., *Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer*, Quantum Science and Technology **4**, 045005 (2019).

[264] A. Anand, M. Degroote, and A. Aspuru-Guzik, *Natural evolutionary strategies for variational quantum computation*, Mach. Learn.: Sci. Technol. **2**, 045012 (2021).

[265] S. Sim, P. D. Johnson, and A. AspuruGuzik, *Expressibility and entangling capability of parameterized quantum circuits for hybrid quantumclassical algorithms*, Advanced Quantum Technologies **2**, 1900070 (2019).

[266] Y. R. Sanders et al., *Compilation of fault-tolerant quantum heuristics for combinatorial optimization*, PRX Quantum **1**, 020312 (2020).

[267] W. K. Wootters and W. H. Zurek, *A single quantum cannot be cloned*, Nature **299**, 802–803 (1982).

[268] J. J. Sakurai, S. F. Tuan, and E. D. Commins, *Modern quantum mechanics (revised edition)*, American Journal of Physics **63**, 93–95 (1995).

[269] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Dover Publications Inc., 1989).

[270] R. Pauncz, *The construction of spin eigenfunctions: an exercise book* (2000).

[271] A. Alex, M. Kalus, A. Huckleberry, and J. von Delft, *A numerical algorithm for the explicit calculation of SU(N) and SL(N,C) Clebsch-Gordan coefficients*, Journal of Mathematical Physics **52** (2011).

[272] B. C. Hall, *Lie groups, Lie algebras, and representations* (Springer Chem, 2015).

[273] M. Kotani, A. Amemita, E. Ishiguro, and T. Kimura, *Tables of molecular integrals*, 2nd edition (Maruzen,Tokyo, 1963).

[274] R. Serber, *Extension of the Dirac vector model to include several configurations*, Physical Review **45**, 461 (1934).

[275] G. Rumer, *Zur Theorie der Spinvalenz*, Ges. d. Wiss. Nachrichten. Math.-Phys. Klasse **377**, 337–341 (1932).

[276] J. Paldus, *Group theoretical approach to the configuration interaction and perturbation theory calculations for atomic and molecular systems*, The Journal of Chemical Physics **61**, 5321–5330 (1974).

[277] J. Paldus, *Unitary-group approach to the many-electron correlation problem: Relation of Gelfand and Weyl tableau formulations*, Physical Review A **14**, 1620 (1976).

[278] I. Shavitt, *Graph theoretical concepts for the unitary group approach to the many-electron correlation problem*, International Journal of Quantum Chemistry **12**, 131–148 (1977).

[279] I. Shavitt, *Matrix element evaluation in the unitary group approach to the electron correlation problem*, International Journal of Quantum Chemistry **14**, 5–32 (1978).

[280] R. S. Mulliken, *The Assignment of Quantum Numbers for Electrons in Molecules. I*, Phys. Rev. **32**, 186 (1928).

[281] R. S. Mulliken, *The Assignment of Quantum Numbers for Electrons in Molecules. II*, Phys. Rev. **32**, 761 (1928).

[282] J. E. Lennard-Jones, *The determination of molecular orbitals*, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **198**, 1–13 (1949).

[283] G. G. Hall and J. Lennard-Jones, *The molecular orbital theory of chemical valency. III. Properties of molecular orbitals*, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **202**, 155–165 (1950).

[284] W. Kohn and L. J. Sham, *Self-consistent equations including exchange and correlation effects*, Physical Review **140**, A1133 (1965).

[285] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, 2000).

[286] C. C. J. Roothaan, *New Developments in Molecular Orbital Theory*, Rev. Mod. Phys. **23**, 69 (1951).

[287] J. C. Slater, *Ferromagnetism and the band theory*, Rev. Mod. Phys. **25**, 199–210 (1953).

[288] P. Lykos and G. W. Pratt, *Discussion on The Hartree–Fock Approximation*, Rev. Mod. Phys. **35**, 496 (1963).

[289] J. C. Slater, *Note on Molecular Structure*, Phys. Rev. **41**, 255 (1932).

[290] P. C. Varras and P. S. Gritzapis, *The Quantum Theory of Valence*, Rev. Mod. Phys. **7**, 167 (1935).

[291] B. O. Roos, P. R. Taylor, and P. E. Sigbahn, *A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach*, Chemical Physics **48**, 157–173 (1980).

[292] B. O. Roos, *The complete active space SCF method in a fock-matrix-based super-CI formulation*, International Journal of Quantum Chemistry **18**, 175–189 (1980).

[293] S. R. White and R. L. Martin, *Ab initio quantum chemistry using the density matrix renormalization group*, The Journal of Chemical Physics **110**, 4127–4130 (1999).

[294] B. Huron, J. P. Malrieu, and P. Rancurel, *Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wave functions*, J. Chem. Phys. **58**, 5745 (1973).

[295] Google AI Quantum, *Hartree-Fock on a superconducting qubit quantum computer*, Science **369**, 1084–1089 (2020).

[296] W. J. Huggins et al., *Unbiasing fermionic quantum Monte Carlo with a quantum computer*, Nature **603**, 416–420 (2022).

[297] D. Bluvstein et al., *Logical quantum processor based on reconfigurable atom arrays*, Nature 2023 **626**, 58–65 (2023).

[298] Y. Dong, L. Lin, and Y. Tong, *Ground-state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices*, PRX Quantum **3**, 040305 (2022).

[299] G. Wang et al., *Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision*, Quantum **7**, 1167 (2023).

[300] G. Wang, D. S. França, G. Rendon, and P. D. Johnson, *Faster ground state energy estimation on early fault-tolerant quantum computers via rejection sampling*, arXiv:2304.09827 (2023).

[301] W. Heitler and F. London, *Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik*, Z. Phys. **44**, 455 (1927).

[302] J. Gerratt, *General Theory of Spin-Coupled Wave Functions for Atoms and Molecules*, Adv. At. Mol. Phys. **7**, 141 (1971).

[303] I. William A. Goddard, T. H. D. Jr., W. J. Hunt, and P. J. Hay, *Generalized valence bond description of bonding in low-lying states of molecules*, **6**, 368 (1973).

[304] J. Li and W. Wu, *New algorithm for nonorthogonal ab initio valence-bond calculations I. New strategy and basic expressions*, Theor. Chim. Acta **89**, 105 (1994).

[305] J.-P. Malrieu, C. Angeli, and R. Cimiraglia, *On the Relative Merits of Non-Orthogonal and Orthogonal Valence Bond Methods Illustrated on the Hydrogen Molecule*, J. Chem. Educ. **85**, 150 (2008).

[306] W. K. Trygve Helgaker and D. P. Tew, *Quantitative quantum chemistry*, Molecular Physics **106**, 2107–2143 (2008).

[307] G. L. Manni, W. Dobrautz, and A. Alavi, *Compression of spin-adapted multiconfigurational wave functions in exchange-coupled polynuclear spin systems*, Journal of Chemical Theory and Computation **16**, 2202–2215 (2020).

[308] W. Dobrautz et al., *Combined unitary and symmetric group approach applied to low-dimensional Heisenberg spin systems*, Physical Review B **105**, 195123 (2022).

[309] G. L. Manni, D. Kats, and N. Liebermann, *Resolution of electronic states in Heisenberg cluster models within the unitary group approach*, Journal of Chemical Theory and Computation **19**, 1218–1230 (2023).

[310] W. D. Laidig, P. Saxe, and R. J. Bartlett, *The description of N_2 and F_2 potential energy surfaces using multireference coupled cluster theory*, J. Chem. Phys. **86**, 887 (1987).

[311] K. Kowalski and P. Piecuch, *Renormalized CCSD(T) and CCSD(TQ) approaches: Dissociation of the N_2 triple bond*, The Journal of Chemical Physics **113**, 5644–5652 (2000).

[312] P. Piecuch, S. A. Kucharski, and K. Kowalski, *Can ordinary single-reference coupled-cluster methods describe the potential energy curve of N_2? The renormalized CCSDT(Q) study*, Chemical Physics Letters **344**, 176–184 (2001).

[313] H. Larsen, J. Olsen, P. Jørgensen, and O. Christiansen, *Full configuration interaction benchmarking of coupled-cluster models for the lowest singlet energy surfaces of N_2*, The Journal of Chemical Physics **113**, 6677–6686 (2000).

[314] G. K. L. Chan, M. Kállay, and J. Gauss, *State-of-the-art density matrix renormalization group and coupled cluster theory studies of the nitrogen binding curve*, The Journal of Chemical Physics **121**, 6110–6116 (2004).

[315] L. Bytautas et al., *Seniority and orbital symmetry as tools for establishing a full configuration interaction hierarchy*, Journal of Chemical Physics **135** (2011).

[316] J. A. Parkhill, K. Lawler, and M. Head-Gordon, *The perfect quadruples model for electron correlation in a valence active space*, Journal of Chemical Physics **130**, 84101 (2009).

[317] G. Gidofalvi and D. A. Mazziotti, *Active-space two-electron reduced-density-matrix method: Complete active-space calculations without diagonalization of the N-electron Hamiltonian*, Journal of Chemical Physics **129**, 134108 (2008).

[318] D. W. Small and M. Head-Gordon, *Tractable spin-pure methods for bond breaking: local many-electron spin-vector sets and an approximate valence bond model*, Journal of Chemical Physics **130**, 84103 (2009).

[319] A. J. W. Thom, *Stochastic coupled cluster theory*, Physical Review Letters **105**, 263004 (2010).

[320] T. V. Voorhis and M. Head-Gordon, *Two-body coupled cluster expansions*, The Journal of Chemical Physics **115**, 5033–5040 (2001).

[321] P. Å. Malmqvist, A. Rendell, and B. O. Roos, *The restricted active space self-consistent-field method, implemented with a split graph unitary group approach*, Journal of Physical Chemistry **94**, 5477–5482 (1990).

[322] H. Fukutome, *Theory of Resonating Quantum Fluctuations in a Fermion System: Resonating Hartree-Fock Approximation*, Prog. Theor. Phys. **80**, 417 (1988).

[323] P. Y. Ayala and H. B. Schlegel, *A nonorthogonal CI treatment of symmetry breaking in sigma formyloxyl radical*, J. Chem. Phys. **108**, 7560 (1998).

[324] E. J. Sundstrom and M. Head-Gordon, *Non-orthogonal configuration interaction for the calculation of multielectron excited states*, J. Chem. Phys. **140**, 114103 (2014).

[325] A. J. W. Thom and M. Head-Gordon, *Hartree–Fock solutions as a quasidiabatic basis for nonorthogonal configuration interaction*, J. Chem. Phys. **131**, 124113 (2009).

[326] H. G. A. Burton and A. J. W. Thom, *General Approach for Multireference Ground and Excited States using Nonorthogonal Configuration Interaction*, J. Chem. Theory Comput. **15**, 4851 (2019).

[327] W. D. Laidig, P. Saxe, and H. F. Schaefer, III, *Multiconfiguration selfconsistentfield study of the importance of triply and quadruply excited electronic configurations in the water molecule*, J. Chem. Phys. **73**, 1765 (1980).

[328] P. Saxe, H. F. Schaefer III, and N. C. Handy, *Exact solution (within a double-zeta basis set) of the schrodinger electronic equation for water*, Chem. Phys. Lett. **79**, 202 (1981).

[329] M. Motta et al., *Towards the solution of the many-electron problem in real materials: equation of state of the hydrogen chain with state-of-the-art many-body methods*, Physical Review X **7**, 031059 (2017).

[330] P. W. Anderson, *Theory of magnetic exchange interactions: exchange in insulators and semiconductors*, in , Vol. 14, edited by F. Seitz and D. Turnbull, Solid State Physics (Academic Press, 1963), pp. 99–214.

[331] C. L. Cleveland and R. Medina A., *Obtaining a Heisenberg Hamiltonian from the Hubbard model*, American Journal of Physics **44**, 44–46 (1976).

[332] V. von Burg et al., *Quantum computing enhanced computational catalysis*, Physical Review Research **3**, 033055 (2021).

[333] K. Sugisaki et al., *Quantum chemistry on quantum computers: a method for preparation of multiconfigurational wave functions on quantum computers without performing post-hartreefock calculations*, ACS Central Science **5**, 167–175 (2019).

[334] N. H. Stair et al., *Stochastic quantum krylov protocol with double-factorized Hamiltonians*, Physical Review A **107**, 032414 (2023).

[335] P. G. Szalay et al., *Multiconfiguration Self-Consistent Field and Multireference Configuration Interaction Methods and Applications*, Chem. Rev. **112**, 108 (2012).

[336] I. Mayer, *Simple Theorems, Proofs, and Derivations in Quantum Chemistry* (Springer, 2003).

[337] H. G. A. Burton, *Generalized nonorthogonal matrix elements: Unifying Wick's theorem and the Slater–Condon rules*, J. Chem. Phys. **154**, 144109 (2021).

[338] H. G. A. Burton, *Generalized nonorthogonal matrix elements. II: Extension to arbitrary excitations*, J. Chem. Phys. **157**, 204109 (2022).

[339] W. Kohn, *Nobel Lecture: Electronic structure of matter-wave functions and density functionals*, Reviews of Modern Physics **71**, 1253 (1999).

[340] J. H. V. Vleck, *Nonorthogonality and ferromagnetism*, Physical Review **49**, 232–240 (1936).

[341] O. Higgott, D. Wang, and S. Brierley, *Variational quantum computation of excited states*, Quantum **3**, 156 (2019).

[342] Z. Cai, *Resource estimation for quantum variational simulations of the hubbard model*, Physical Review Applied **14**, 14059 (2020).

[343] B. V. Straaten and B. Koczor, *Measurement cost of metric-aware variational quantum algorithms*, PRX Quantum **2**, 030324 (2021).

[344] G.-L. R. Anselmetti, D. Wierichs, C. Gogolin, and R. M. Parrish, *Local, expressive, quantum-number-preserving VQE ansätze for fermionic systems*, New J. Phys. **23**, 113010 (2021).

[345] J. M. Arrazola et al., *Universal quantum circuits for quantum chemistry*, Quantum **6**, 742 (2022).

[346] U. Baek et al., *Say no to optimization: a nonorthogonal quantum eigensolver*, PRX Quantum **4**, 030307 (2023).

[347] C. Gustiani, R. Meister, and S. C. Benjamin, *Exploiting subspace constraints and ab initio variational methods for quantum chemistry*, New Journal of Physics **25**, 073019 (2023).

[348] Y. Shen et al., *Real-time Krylov theory for quantum computing algorithms*, Quantum **7**, 1066 (2023).

[349] W. Kirby, *Analysis of quantum Krylov algorithms with errors*, arXiv:2401.01246 (2024).

[350] G. Wang, S. Sim, and P. D. Johnson, *State preparation boosters for early fault-tolerant quantum computation*, Quantum **6**, 829 (2022).

[351] K. Gratsea, C. Sun, and P. D. Johnson, *When to reject a ground state preparation algorithm*, arXiv:2212.09492 (2022).

[352] K. Gratsea, J. S. Kottmann, P. D. Johnson, and A. A. Kunitsa, *Comparing classical and quantum ground state preparation heuristics*, arXiv:2401.05306 (2024).

[353] A. Bärtschi and S. Eidenbenz, *Deterministic preparation of Dicke states*, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **11651 LNCS**, 126–139 (2019).

[354] J. Lee, D. W. Small, and M. Head-Gordon, *Excited states via coupled cluster theory without equation-of-motion methods: Seeking higher roots with application to doubly excited states and double core hole states*, J. Chem. Phys. **151**, 214103 (2019).

[355] D. Aliverti-Piuri et al., *What can quantum information theory offer to quantum chemistry?*, Faraday Discussions (2024).

[356] R. Izsák et al., *Measuring electron correlation: the impact of symmetry and orbital transformations*, Journal of Chemical Theory and Computation **19**, 2703–2720 (2023).

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

[358] K. Sugisaki et al., *Quantum chemistry on quantum computers: a polynomial-time quantum algorithm for constructing the wave functions of open-shell molecules*, Journal of Physical Chemistry A **120**, 6459–6466 (2016).

[359] A. Carbone, D. E. Galli, M. Motta, and B. Jones, *Quantum circuits for the preparation of spin eigenfunctions on quantum computers*, Symmetry 2022, Vol. 14, Page 624 **14**, 624 (2022).

[360] K. Sugisaki et al., *Quantum chemistry on quantum computers: a method for preparation of multiconfigurational wave functions on quantum computers without performing post-Hartree-Fock calculations*, ACS Central Science **5**, 167–175 (2019).

[361] A. Bärtschi and S. Eidenbenz, *Short-depth circuits for Dicke state preparation*, 2022 IEEE International Conference on Quantum Computing and Engineering (QCE) (2022).

[362] T. Bastin et al., *Operational determination of multiqubit entanglement classes via tuning of local operations*, Physical Review Letters **102**, 053601 (2009).

[363] D. B. Hume, C. W. Chou, T. Rosenband, and D. J. Wineland, *Preparation of Dicke states in an ion chain*, Physical Review A - Atomic, Molecular, and Optical Physics **80**, 052302 (2009).

[364] D. Bacon, I. L. Chuang, and A. W. Harrow, *Efficient quantum circuits for Schur and Clebsch-Gordan transforms*, Physical Review Letters **97**, 170502 (2006).

[365] S. Aktar, A. Bartschi, A.-H. A. Badawy, and S. Eidenbenz, *A divide-and-conquer approach to Dicke state preparation*, IEEE Transactions on Quantum Engineering **3**, 1–16 (2022).

[366] S. Aktar, A. H. A. Badawy, A. Bärtschi, and S. Eidenbenz, *Scalable experimental bounds for Dicke and GHZ states fidelities*, Proceedings of the 20th ACM International Conference on Computing Frontiers 2023, CF 2023, 176–184 (2023).

[367] A. Bartschi and S. Eidenbenz, *Grover mixers for QAOA: shifting complexity from mixer design to state preparation*, Proceedings - IEEE International Conference on Quantum Computing and Engineering, QCE 2020, 72–82 (2020).

[368] A. Khamoshi, R. Dutta, and G. E. Scuseria, *State preparation of antisymmetrized geminal power on a quantum computer without number projection*, Journal of Physical Chemistry A **127**, 4005–4014 (2023).

[369] H. G. A. Burton, *Accurate and gate-efficient quantum ansätze for electronic states without adaptive optimisation*, Physical Review Research **6**, 023300 (2024).

[370] M. Scheurer et al., *Tailored and externally corrected coupled cluster with quantum inputs*, Journal of Chemical Theory and Computation (2024).

[371] O. Leimkuhler and K. B. Whaley, *A quantum eigenvalue solver based on tensor networks*, arXiv:2404.10223 (2024).

[372] P. J. Ollitrault et al., *Enhancing initial state overlap through orbital optimization for faster molecular electronic ground-state energy estimation*, Physical Review Letters **133**, 250601 (2024).

[373] D. Cleland, G. H. Booth, and A. Alavi, *Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo*, Journal of Chemical Physics **132**, 41103 (2010).

[374] M. Möttönen, J. J. Vartiainen, V. Bergholm, and M. M. Salomaa, *Transformation of quantum states using uniformly controlled rotations*, Quantum Information and Computation **5**, 467–473 (2004).

[375] C. Schön et al., *Sequential generation of entangled multiqubit states*, Physical Review Letters **95**, 110503 (2005).

[376] M. Moerchen et al., *Classification of electronic structures and state preparation for quantum computation of reaction chemistry*, arXiv:2409.08910 (2024).

[377] W. A. Goddard, *Improved quantum theory of many-electron systems. I. Construction of eigenfunctions of $\hat{S}^2$ which satisfy Pauli's principle*, Physical Review **157**, 73 (1967).

[378] W. A. Goddard, *Improved Quantum Theory of Many-Electron Systems. II. The Basic Method*, Physical Review **157**, 81 (1967).

[379] W. A. Goddard, R. Paul, A. E. D. Vries, and W. A. G. Iii, *Improved Quantum Theory of ManyElectron Systems. III. The GF Method*, The Journal of Chemical Physics **48**, 450 (2003).

[380] W. A. Goddard, *Improved Quantum Theory of ManyElectron Systems. IV. Properties of GF Wavefunctions*, The Journal of Chemical Physics **48**, 5337 (2003).

[381] R. C. Ladner and W. A. Goddard, *Improved Quantum Theory of ManyElectron Systems. V. The Spin-Coupling Optimized GI Method*, The Journal of Chemical Physics **51**, 1073 (2003).

[382] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, *Simulated Quantum Computation of Molecular Energies*, Science **309**, 1704 (2005).

[383] F. A. Evangelista, G. K.-L. Chan, and G. E. Scuseria, *Exact parameterization of fermionic functions via unitary coupled cluster theory*, J. Chem. Phys. **151**, 244122 (2019).

[384] A. F. Izmaylov, M. Díaz-Tinoco, and R. A. Lang, *On the order problem in construction of unitary operators for the variational quantum eigensolver*, Phys. Chem. Chem. Phys. **22**, 12980 (2020).

[385] H. R. Grimsley et al., *Is the Trotterized UCCSD Ansatz Chemically Well-Defined?*, J. Chem. Theory Comput. **16**, 1 (2020).

[386] T. Tsuchimochi, Y. Mori, and S. L. Ten-no, *Spin-projection for quantum computation: A low-depth approach to strong correlation*, Phys. Rev. Research **2**, 043142 (2020).

[387] H. L. Tang et al., *Qubit-ADAPT-VQE: An Adaptive Algorithm for Constructing Hardware-Efficient Ansätze on a Quantum Processor*, PRX Quantum **2**, 020310 (2021).

[388] H. H. S. Chan et al., *Molecular excited state calculations with adaptive wavefunctions on a quantum eigensolver emulation: reducing circuit depth and separating spin states*, Phys. Chem. Chem. Phys. **23**, 26438–26450 (2021).

[389] T. Tsuchimochi, M. Taii, T. Nishimaki, and S. L. Ten-no, *Adaptive construction of shallower quantum circuits with quantum spin projection for fermionic systems*, Phys. Rev. Research **4**, 033100 (2022).

[390] V. O. Shkolnikov, N. J. Mayhall, S. E. Economou, and E. Barnes, *Avoiding symmetry roadblocks and minimizing the measurement overhead of adaptive variational quantum eigensolvers*, arxiv:2109.05340 (2021).

[391] N. C. Rubin, J. Lee, and R. Babbush, *Compressing Many-Body Fermion Operators under Unitary Constraints*, J. Chem. Theory Comput. **18**, 1480 (2022).

[392] R. Gilmore, *Lie Groups, Physics, and Geometry: An Introduction for Physicists, Engineers, and Chemists*, first (Dover Publications Inc., 2008).

[393] J. Lee, W. J. Huggins, M. Head-Gordon, and K. B. Whaley, *Generalized Unitary Coupled Cluster Wave functions for Quantum Computation*, J. Chem. Theory Comput. **15**, 311 (2019).

[394] H. R. Grimsley et al., *ADAPT-VQE is insensitive to rough parameter landscapes and barren plateaus*, arxiv:2204.07179 (2022).

[395] S.-X. Zhang, C.-Y. Hsieh, S. Zhang, and H. Yao, *Differentiable quantum architecture search*, Quantum Sci. Technol. **7**, 045023 (2022).

[396] D. Chivilikhin et al., *MoG-VQE: Multiobjective genetic variational quantum eigensolver*, arxiv:2007.04424 (2020).

[397] A. G. Rattew et al., *A Domain-agnostic, Noise-resistant, Hardware-efficient, Evolutional Variational Quantum Eigensolver*, arXiv:1910.09694 (2019).

[398] M. O. E. Grant and M. Benedetti, *Structure optimization for parameterized quantum circuits*, Quantum **5**, 391 (2021).

[399] D. Schebarchov and D. J. Wales, *Communication: A new paradigm for structure prediction in multicomponent systems*, J. Chem. Phys. **139**, 221101 (2013).

[400] D. Schebarchov and D. J. Wales, *Structure Prediction for Multicomponent Materials Using Biminima*, Phys. Rev. Lett. **113**, 156102 (2014).

[401] K. Röder and D. J. Wales, *Mutational Basin-Hopping: Combined Structure and Sequence Optimization for Biomolecules*, J. Phys. Chem. Lett. **9**, 6169 (2018).

[402] Z. Li and H. A. Scheraga, *Monte Carlo-minimization approach to the multiple-minima problem in protein folding*, Proceedings of the National Academy of Sciences of the United States of America **84**, 6611 (1987).

[403] D. J. Wales and J. P. K. Doye, *Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms*, J. Phys. Chem. A **101**, 5111 (1997).

[404] D. A. Mazziotti, *Anti-Hermitian Contracted Schrödinger Equation: Direct Determination of the Two-Electron Reduced Density Matrices of Many-Electron Molecules*, Phys. Rev. Lett. **97**, 143002 (2006).

[405] J. Liu, Z. Li, and J. Yang, *An efficient adaptive variational quantum solver of the Schrödinger equation based on reduced density matrices*, J. Chem. Phys. **154**, 244112 (2021).

[406] Y. Matsuzawa and Y. Kurashige, *Jastrow-type Decomposition in Quantum Chemistry for Low-Depth Quantum Circuits*, J. Chem. Theory Comput. **16**, 944 (2020).

[407] M. D. Sapova and A. K. Fedorov, *Variational quantum eigensolver techniques for simulating carbon monoxide oxidation*, Nat. Comm. **5**, 199 (2022).

[408] Y. S. Yordanov, D. R. M. Arvidsson-Shukur, and C. H. W. Barnes, *Efficient quantum circuits for quantum computational chemistry*, Phys. Rev. A **102**, 062612 (2020).

[409] J. P. F. LeBlanc et al., *Solutions of the Two-Dimensional Hubbard Model: Benchmarks and Results from a Wide Range of Numerical Algorithms*, Phys. Rev. X **5**, 041041 (2015).

[410] B.-X. Zheng et al., *Stripe order in the underdoped region of the two-dimensional Hubbard model*, Science **358**, 1155 (2017).

[411] Y. S. Yordanov, D. R. Arvidsson-Shukur, and C. H. Barnes, *Efficient quantum circuits for quantum computational chemistry*, Physical Review A **102**, 062612 (2020).

[412] P. Selinger, *Efficient Clifford+T approximation of single-qubit operators*, Quantum Information and Computation **15**, 159–180 (2012).

[413] N. J. Ross and P. Selinger, *Optimal ancilla-free Clifford+T approximation of Z-rotations*, Quantum Information and Computation **16**, 0901–0953 (2014).

[414] A. Bocharov, M. Roetteler, and K. M. Svore, *Efficient synthesis of universal repeat-until-success quantum circuits*, Physical Review Letters **114**, 080502 (2015).

[415] G. H. Low, *Halving the cost of quantum multiplexed rotations*, arXiv:2110.13439 (2021).

[416] C. Gidney, *Halving the cost of quantum addition*, Quantum **2**, 74 (2018).

[417] D. Malz, G. Styliaris, Z.-Y. Wei, and J. I. Cirac, *Preparation of matrix product states with log-depth quantum circuits*, Physical Review Letters **132**, 040404 (2024).

[418] S. Bravyi, M. B. Hastings, and F. Verstraete, *Lieb-Robinson bounds and the generation of correlations and topological quantum order*, Physical Review Letters **97**, 050401 (2006).

[419] M. Ben-Dov, D. Shnaiderov, A. Makmal, and E. G. D. Torre, *Approximate encoding of quantum states using shallow circuits*, npj Quantum Information 2024 **10**, 1–8 (2024).

[420] M. S. Rudolph et al., *Decomposition of matrix product states into shallow quantum circuits*, Quantum Science and Technology **9**, 015012 (2023).

[421] S. J. Ran, *Encoding of matrix product states into quantum circuits of one- a nd two-qubit gates*, Physical Review A **101**, 032310 (2020).

[422] J. Nocedal, *Updating quasi-Newton matrices with limited storage*, Math. Comp. **35**, 773 (1980).

[423] C. G. Broyden, *The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations*, IMA J. Appl. Math. **6**, 76 (1970).

[424] R. Fletcher, *A new approach to variable metric algorithms*, Comput. J. **13**, 317 (1970).

[425] D. Goldfarb, *A family of variable-metric methods derived by variational means*, Math. Comp. **24**, 23 (1970).

[426] D. F. Shanno, *Conditioning of quasi-Newton methods for function minimization*, Math. Comp. **24**, 647 (1970).

[427] `Gmin`*: a program for finding global minima and calculating thermodynamic properties*, http://www-wales.ch.cam.ac.uk/software.html, http://www-wales.ch.cam.ac.uk/software.html.

[428] D. J. Wales, *Energy landscapes: applications to clusters, biomolecules and glasses* (Cambridge University Press, Cambridge, 2004).

[429] N. Metropolis et al., *Equation of State Calculations by Fast Computing Machines*, J. Chem. Phys. **21**, 1087 (1953).

[430] W. K. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika **57**, 97 (1970).

[431] B. Strodel, J. W. L. Lee, C. S. Whittleston, and D. J. Wales, *Transmembrane structures for alzheimers aβ1-42 oligomers*, J. Am. Chem. Soc. **132**, 13300–13312 (2010).

[432] R. H. Swendsen and J.-S. Wang, *Replica Monte Carlo Simulation of Spin-Glasses*, Phys. Rev. Lett. **57**, 2607 (1986).

[433] Q. Sun et al., *Recent developments in the PySCF program package*, J. Chem. Phys. **153**, 024109 (2020).

[434] J. R. McClean et al., *OpenFermion: the electronic structure package for quantum computers*, Quantum Sci. Technol. **5**, 034014 (2020).

[435] S. E. Smart and D. A. Mazziotti, *Quantum Solver of Contracted Eigenvalue Equations for Scalable Molecular Simulations on Quantum Computing Devices*, Phys. Rev. Lett. **126**, 070504 (2021).

[436] Y. Shao et al., *Advances in molecular quantum chemistry contained in the Q-Chem 4 program package*, Mol. Phys. **113**, 184 (2015).