Quantum computational chemistry

Sam McArdle,[1,*] Suguru Endo,[1] Alán Aspuru-Guzik,[2,3,4] Simon Benjamin,[1] and Xiao Yuan[1,†]

[1] *Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom*
[2] *Department of Chemistry and Department of Computer Science, University of Toronto, Toronto, Ontario M5S 3H6, Canada*
[3] *Vector Institute for Artificial Intelligence, Toronto, Ontario M5S 1M1, Canada*
[4] *Canadian Institute for Advanced Research (CIFAR) Senior Fellow, Toronto, Ontario M5S 1M1, Canada*

(Dated: August 31, 2018)

One of the most promising applications of quantum computing is solving classically intractable chemistry problems. As a result, quantum computational chemistry is rapidly emerging as an interdisciplinary field requiring knowledge of both quantum information and computational chemistry. This work provides a comprehensive introduction to both fields, bridging the current knowledge gap. We review the key developments in this area, with a focus on near-term quantum computation. We illustrate the methods discussed by explicitly demonstrating how to map chemical problems onto a quantum computer, and solve them. We conclude with an outlook for this nascent field.

CONTENTS

* samuel.mcardle@materials.ox.ac.uk
† xiao.yuan.ph@gmail.com

I. INTRODUCTION

Quantum mechanics underpins all of modern chemistry. One might therefore imagine that we could use this theory to predict the behaviour of any chemical compound. This is not the case. As Dirac noted; "*The exact application of these laws leads to equations much too complicated to be soluble.*" [1]. The problem described by Dirac is that the wavefunction of a quantum system grows exponentially with the number of particles. This leaves classical computers unable to exactly simulate quantum systems in an efficient way. Feynman proposed a solution to this problem; using quantum hardware as the simulation platform, remarking that *"If you want to make a simulation of nature, you'd better make it quantum mechanical, and by golly it's a wonderful problem, because it doesn't look so easy."* [2]. Building a quantum computer has taken over 30 years, but Feynman's vision may soon be fulfilled, following recent developments in quantum hardware (such as ion traps [3–6], superconducting systems [7, 8], and photonic systems [9, 10]). It is believed that using quantum systems as our simulation platform will yield unprecedented developments in chemistry [11], biology [12] and materials science [13].

To date, several efficient quantum algorithms have been proposed to solve problems in chemistry [14–17]. The runtime and physical resources required by these algorithms scale polynomially with the size of the system simulated. Recent experimental developments have accompanied these theoretical milestones, with many groups demonstrating proof of principle chemistry calculations [18–31]. However, limited by hardware capabilities, these experiments focus only on small molecules that we are already able to simulate classically. Moreover, the gate counts required for transformative chemistry simulations may mandate the need for fault-tolerance, which requires considerably more qubits than are currently available. New developments are needed to solve classically intractable chemistry problems in the near-future.

These breakthroughs will be achieved by connecting researchers working in quantum information with those working in computational chemistry. We seek to aid this connection with this succinct, yet comprehensive, review of quantum computational chemistry and its foundational fields.

Although quantum algorithms can solve a range of problems in chemistry, we focus predominantly on the problem of finding the ground state energy of molecules. There are two reasons for this restriction of scope. Primarily, it is because the machinery developed to solve this problem on quantum computers is easily applied to other types of problems. Moreover, most of the prior work in quantum computational chemistry has focused on this problem. As such, it provides an ideal context in which to explain the most important details of quantum computational chemistry.

We first provide a brief overview of quantum computing and simulation in Sec. II. We then introduce the key methods and terminology used in classical computational chemistry in Sec. III. Sec. IV describes the methods developed to merge these two fields, including mapping chemistry problems onto a quantum computer. We continue our discussion of quantum computational chemistry in Sec. V by describing algorithms for finding the ground and excited states of chemical systems. Sec. VI highlights the techniques developed to mitigate the effects of noise in non-error corrected quantum computers, which will be crucial for achieving accurate simulations in the near-future.

In Sec. VII we provide several examples of how to map chemistry problems onto a quantum computer. We discuss techniques that can be used to reduce the simulation resources required, and the quantum circuits that can be used. This section seeks to illustrate the techniques described throughout the rest of the paper, providing worked examples for the reader. We conclude this work with a comparison between classical and quantum techniques, and resource estimations for the different quantum methods. This section aims to help the reader to understand when and how quantum computational chemistry may surpass its classical counterpart.

II. QUANTUM COMPUTING AND SIMULATION

In this section, we introduce the basic elements of quantum computing and quantum simulation. We refer the reader to Refs. [32, 33] for more detailed introductions.

A. Quantum computing

In this paper, we focus on the circuit model of quantum computation [32]. It is also possible to do: adiabatic quantum computing [34, 35] (including in the context of chemistry simulation [36]), one-way or measurement based quantum computing [37–39], and continuous-variable quantum computing [40, 41].

In the circuit model of quantum computation, the basic unit of information is called a qubit, which lives in a two-dimensional Hilbert space. The basis vectors of the space are denoted as $\{|0\rangle, |1\rangle\}$, which are referred to as the computational basis states,

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{1}$$

A general single qubit state is described by

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle = \begin{bmatrix}\alpha\\ \beta\end{bmatrix} \tag{2}$$
$$\alpha, \beta \in \mathbb{C},$$
$$|\alpha|^2 + |\beta|^2 = 1.$$

When quantum logic gates act on the qubits, they manipulate both basis state vectors at the same time, providing (measurement limited) parallelism. Although the qubit is in a quantum superposition during the algorithm, when it is measured, it will be found in state $|0\rangle$ or state $|1\rangle$, not in a superposition. These measurement outcomes occur with probability $|\alpha|^2$ and $|\beta|^2$, respectively.

If there are N qubits in the system, the state is described by a vector in the 2^N dimensional Hilbert space formed by taking the tensor product of the Hilbert spaces of the individual qubits. States can be classified as either 'product' or 'entangled'. Product states can be decomposed into tensor products of fewer qubit wavefunctions, such as

$$\frac{1}{\sqrt{2}}(|00\rangle + |01\rangle) = |0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle). \tag{3}$$

Entangled states cannot be decomposed into tensor products, such as the state

$$\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle). \tag{4}$$

In this work, we refer to the leftmost qubit in a vector as the $(N-1)^{\text{th}}$ qubit, and the rightmost qubit as the zeroth qubit. A quantum circuit consists of a number of single and two qubit gates acted on the qubits. The qubits are initialised in a well defined state, such as the $|\bar{0}\rangle$ state ($|\bar{0}\rangle = |0\rangle^{\otimes n} = |0\rangle \otimes |0\rangle \otimes \cdots \otimes |0\rangle$). The circuit concludes with measurements to extract information. From a mathematical perspective, the qubit gates are unitary matrices. Typical gates include the Pauli gates

$$X = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \quad Y = \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}, \quad Z = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}, \tag{5}$$

the single qubit rotation gates

$$R(\theta)_{O=X,Y,Z} = \exp\left(\frac{-i\theta}{2}O\right), \tag{6}$$

the Hadamard and T gates

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}, \quad T = \begin{bmatrix}1 & 0\\ 0 & e^{i\pi/8}\end{bmatrix}, \tag{7}$$

and multi-qubit entangling gates, such as the two qubit controlled-NOT (CNOT) gate

where '•' denotes the control qubit and '⊕' denotes the target qubit, which can be written mathematically as

$$I_T \otimes |0\rangle\langle 0|_C + X_T \otimes |1\rangle\langle 1|_C\,, \tag{8}$$

where T denotes the target qubit, and C denotes the control qubit.

These gates are used to create an example quantum circuit in Fig. 1. This circuit generates an entangled state of 2 qubits, meaning the state cannot be written as a tensor product of individual qubit states.

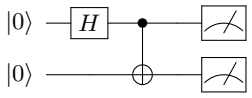

FIG. 1. A quantum circuit that generates the entangled state $(|00\rangle+|11\rangle)/\sqrt{2}$ and measures each qubit in the computational basis. Here, H is the Hadamard gate, defined in the main text. Half of the time both of the qubits will be measured to be 0, and the other half, both 1.

With only single qubit operations and the CNOT gate, it is possible to approximate an arbitrary multi-qubit gate to any desired accuracy [42]. As a result, the circuit model of quantum computing typically decomposes all algorithms into single and two qubit gates. We denote each gate by a unitary operator $U^{i,j}(\vec{\theta})$, where i, j are the indices of the qubits the gates act on ($i = j$ for single qubit operations), and $\vec{\theta}$ are gate parameters (although the gates do not have to be parametrised, such as the Pauli gates). We can then mathematically describe a quantum circuit by

$$|\psi\rangle = \prod_k U_k^{i_k,j_k}(\vec{\theta}_k)\,|\bar{0}\rangle\,, \tag{9}$$

where k denotes the k^{th} gate in the circuit. The gates are ordered right to left. For example, the circuit in Fig. 1 would be written as

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \text{CNOT}^{0,1}H^0\,|00\rangle\,. \tag{10}$$

We extract information from the circuits by performing measurements of observables, O, which are Hermitian matrices. We seek the average value over many measurements, $\bar{O}$, given by

$$\bar{O} = \langle\psi|\,O\,|\psi\rangle\,. \tag{11}$$

Qubits are measured in the computational basis. Measuring qubit i in the computational basis corresponds to $\langle\psi|\,Z_i\,|\psi\rangle$. In order to measure qubits in the X or Y basis, single qubit rotations are first applied to change the basis of the relevant qubits, which are then measured in the Z basis. The Pauli operators form a complete basis for any Hermitian operator. Therefore any observable

can be expanded into strings of Pauli operators which we can measure efficiently with a quantum computer.

It is important to distinguish between the number of *physical* and *logical* qubits in a quantum computer. In order to protect our quantum state from decoherence caused by coupling to the environment [43, 44], we use quantum error correcting codes. This means that we encode a single logical qubit in several physical qubits. We check for and correct errors in the physical qubits, thus protecting the encoded qubit [45, 46]. We examine groups of qubits together, as measuring a single qubit to see if an error had occured would destroy the quantum superposition. By examining several groups and allocating the qubits to the groups in an intelligent manner, it can be determined if any single qubit has undergone an error. If the error rate is below a certain (code dependant) threshold, the error correction procedure can be continuously applied, enabling an arbitrarily long computation - even if the error checking measurements cause errors. [47]. This is known as fault-tolerant quantum computation [48, 49].

The theory of error correcting codes was extended by Kitaev, who developed topological error correcting codes [50, 51]. These codes have the highest threshold values [52–54], and are particularly practical for superconducting qubits, for which a 2D grid is an achievable architecture [55]. Unfortunately, these topological codes restrict the quantum gates that can be implemented to those that are classically simulable - seemingly voiding the promised advantage of quantum computers [56, 57]. This can be overcome by creating 'magic states', which can be used to implement a T-gate [58], restoring the proposed power of a quantum computer. These magic states are created by a process known as distillation, which is the most costly aspect of error correction [59]. This can increase the qubit requirement by a factor of 1000 or more, placing the number of physical qubits required for useful applications in the millions [12, 59].

In contrast, current quantum computers are described in terms of tens of physical qubits, realised in a range of different architectures. Trapped ion and superconducting qubits have recently been demonstrated with sub-threshold error rates of 0.1 % [5, 6] and 0.5 % [7], respectively. There is free cloud access to quantum computers with up to 20 qubits [60]. Various groups have announced quantum computers with over 50 qubits (although these have yet to be benchmarked). A quantum computer of this size is too large to simulate classically, and may thus be capable of solving problems which are intractable on even the largest classical supercomputers. However, these problems are typically contrived examples, rather than real-world problems [61].

The dichotomy between the resources needed for tackling real problems, and the 'superiority' of a machine with more than 50 qubits poses the question; *'What, if anything, will near-term quantum computers be useful for?'*. The answer may lie with Feynman's original proposal; using quantum systems to simulate quantum systems.

B. Quantum simulation

In this work, we focus on the digital quantum simulation of many-body quantum systems - specifically molecules. Digital quantum simulation maps the target problem onto a set of gates which can be implemented by a quantum computer. A universal quantum computer can be programmed to perform many different simulations. This can be contrasted with analog quantum simulation, where the simulator emulates a specific real system of interest. Analog simulators are less controllable, but are more robust to noise, and therefore easier to construct. To date, there have been several proposals for the simulation of chemistry using analog simulators [17, 62, 63], some of which have been experimentally realised [30, 31]. Digital quantum simulation is more vulnerable to noise and device imperfections. Although such imperfections can be addressed via error correction, this requires additional qubits and places stringent requirements on gate fidelities [59]. In this work we focus solely on digital quantum simulation of chemistry problems. We refer the reader to Ref. [33] for information about quantum simulation of other physical systems, and analog quantum simulation.

The numerous problems in chemistry that can be simulated on a quantum computer can be divided into static and dynamics problems. Here, we use 'dynamics' to mean evolving wavefunctions in time and seeing how certain observables vary, as opposed to chemical reaction dynamics, which are discussed separately below.

Methods for solving dynamics problems were formalised by Lloyd [64] and further developed by Abrams and Lloyd [65]. As illustrated in Fig. 2, we can map the system Hamiltonian, H_s, to a qubit Hamiltonian, H_q. We similarly map the initial system wavefunction $|\psi_s^i\rangle$ to a qubit representation $|\psi_q^i\rangle$. We can then evolve the qubit wavefunction in time by mapping the system evolution operator, e^{-itH_s}, to a series of gates. This can be achieved using a Lie-Trotter-Suzuki decomposition (Trotterization) [66], which means that if the Hamiltonian of the system, H_s is local, and can be written as

$$H_s = \sum_i h_i \tag{12}$$

where h_i are local terms which act on a small subset of the particles in the system, then a small time evolution under the Hamiltonian can be decomposed as

$$e^{-iH_s\delta t} = e^{-i\sum_j h_j \delta t} \approx \prod_j e^{-ih_j\delta t} + O(\delta t^2). \tag{13}$$

The number of terms in the Hamiltonian scales polynomially with the number of particles for systems of interest, such as molecules or the Fermi-Hubbard model.

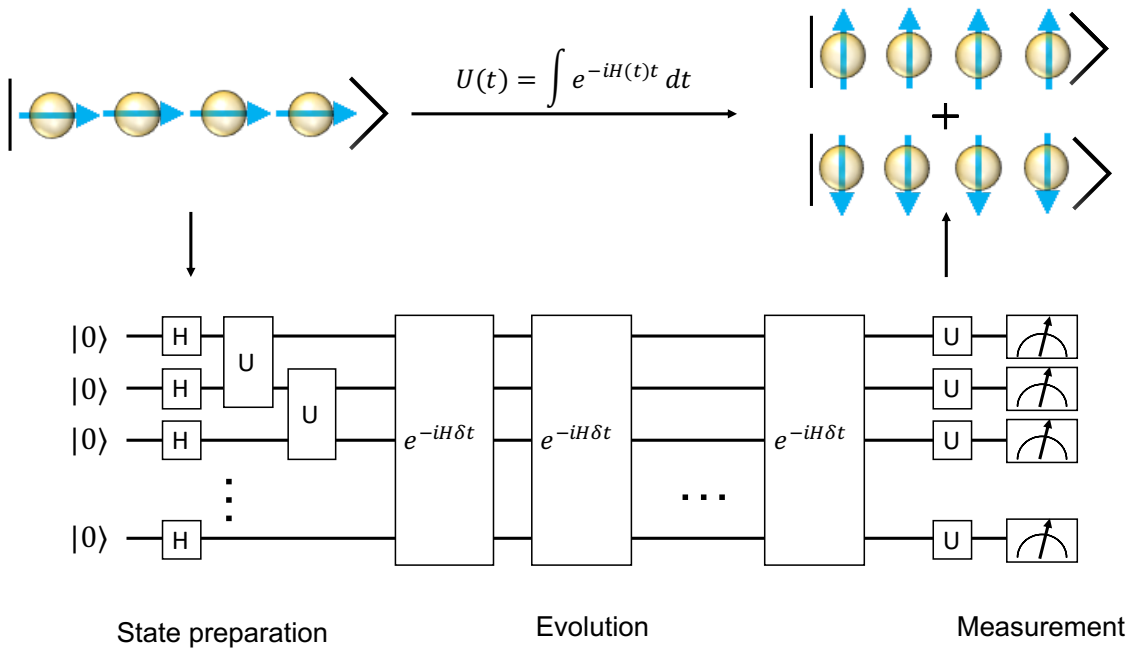

FIG. 2. Digital quantum simulation of time evolution of a spin chain, using a canonical Trotter-type method. We first map the system Hamiltonian, H_s, to a qubit Hamiltonian, H_q. Then the initial system wavefunction $|\psi_s^i\rangle$ is mapped to a qubit wavefunction $|\psi_q^i\rangle$. The time evolution of the system can be mapped to a Trotterized circuit that acts on the initial qubit wavefunction. Finally, well chosen measurements are applied to extract the desired information, such as particle correlation functions.

Each of the exponential terms in Eq. (13) can be realised efficiently using a quantum computer. It was recently shown that time evolution can also be simulated using variational approaches [67]. We can extract useful dynamical quantities from these simulations, such as the electronic charge density distribution, or two-particle correlation functions [65].

When studying chemical reaction dynamics, the greatest efficiency is obtained by treating the electrons and nuclei on an equal footing. This is not possible using the method shown in Fig. 2. Kassal *et al.* developed a quantum algorithm for studying chemical reaction dynamics in 2008 [15]. They showed that grid based methods enable the explicit simulation of chemical reactions, using resources which scale polynomially with the system size. We discuss this further in Sec. IV A.

We can obtain static properties by mapping the equilibrium wavefunction of the system onto a qubit wavefunction. We can then use the quantum computer to calculate the desired observable, $\langle\psi_q| O_q |\psi_q\rangle$. In particular, Abrams and Lloyd showed that Kitaev's phase estimation algorithm [68] can be used to find the energies of quantum systems [69]. As stated in the introduction, we focus on finding the ground state energy of molecules. This problem has received significant attention since it was first introduced in the context of quantum computational chemistry by Aspuru-Guzik *et al.* in 2005 [14]. It is a difficult problem for classical computers, and is widely considered to be one of the first applications of quantum computing. Finding the ground state energy enables the calculation of reaction rates via the Eyring equation [22], and the determination of molecular geometries [70, 71].

Before discussing how the ground state problem can be solved using a quantum computer, we first summarise the classical methods used to solve this problem. Many of these methods have formed the basis of the work done thus far in quantum computational chemistry.

III. CLASSICAL COMPUTATIONAL CHEMISTRY

In this section, we introduce the techniques used in classical computational chemistry. As discussed in the introduction, we focus on tools developed to find the ground state energy of molecules. This is known as the electronic structure problem. The problem is formulated in Sec. III A, and translated into the language of first and second quantisation in Sec. III B. In Sec. III C we describe the different approximations that can be used

to make this problem tractable for classical computers. In Sec. III D we review some of the common spin-orbital basis functions used in the second quantised approach. We discuss orbital basis changes, and their use in reducing the simulation resources in Sec. III E. We have sought to produce a self-contained summary of the essential knowledge required for quantum computational chemistry, and we refer the reader to Refs. [72, 73] for further information.

A. The electronic structure problem

The Hamiltonian of a molecule consisting of M nuclei and N electrons is

$$H = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_{i,I}\frac{e^2}{4\pi\epsilon_0}\frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J}\frac{e^2}{4\pi\epsilon_0}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \tag{14}$$

where M_I, $\mathbf{R}_I$, and Z_I denote the mass, position, and charge of the I^{th} nucleus, and $\mathbf{r}_i$ is the position of the i^{th} electron. The first two sums in H are the kinetic terms of the electrons and nuclei, respectively. The final three sums represent the Coulomb repulsion between: the electrons and nuclei, the electrons themselves, and the nuclei themselves, respectively. For conciseness, we work in atomic units, where the unit of length is $a_0 = 1$ Bohr (0.529167×10^{-10} m), the unit of mass is the electron mass m_e, and the unit of energy is 1 Hartree (1 Hartree $= e^2/4\pi\epsilon_0 a_0 = 27.2113$ eV). Denoting $M_I' = M_I/m_e$, the molecular Hamiltonian in atomic units becomes

$$H = -\sum_i \frac{\nabla_i^2}{2} - \sum_I \frac{\nabla_I^2}{2M_I'} - \sum_{i,I}\frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \tag{15}$$

We are predominantly interested in the electronic structure of the molecule. As a nucleon is over one thousand times heavier than an electron, we apply the Born-Oppenheimer approximation, treating the nuclei as classical point charges. As a result, for a given nuclear configuration one only needs to solve the electronic Hamiltonian

$$H_e = -\sum_i \frac{\nabla_i^2}{2} - \sum_{i,I}\frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{16}$$

Our aim is to find energy eigenstate $|E_i\rangle$ and the corresponding energy eigenvalue E_i of the Hamiltonian H_e. In the rest of this work, we drop the subscript e. In particular, we are interested in the ground state energy and the lowest excited state energies. We can solve this equation for a range of nuclear configurations to map out the potential energy surfaces of the molecule. We note that mapping out these potential energy curves explicitly is exponentially costly in the degrees of freedom of the molecule, and that there are a variety of methods being developed to solve this difficult problem more efficiently [74].

We wish to measure the energy to an accuracy of at least 1.6×10^{-3} Hartree, known as 'chemical accuracy'. If the energy is known to chemical accuracy, then the chemical reaction rate at room temperature can be predicted to within an order of magnitude using the Eyring equation [71]

$$\text{Rate} \propto e^{-\Delta E/k_B T}. \tag{17}$$

In computational chemistry, we are often more interested in the relative energies of two points on the potential energy surface than the absolute energy of a single point. Even if the individual energy values cannot be measured to within chemical accuracy, there is often a fortuitous cancellation of errors, which leads to the energy difference being found to chemical accuracy. However, in this work we consider chemical accuracy to mean an error of less than 1.6×10^{-3} Hartree in the energy value at a single point on the potential energy surface.

B. First and Second quantisation

1. *First quantisation*

The electronic structure problem as shown in Eq. (16) is said to be in the first quantised representation. We focus below on classical first quantised simulation methods. Discussion of first quantised chemistry simulation on quantum computers is postponed until Sec. IV A.

Here we consider the wavefunction in the position representation, $\{|\mathbf{r}\rangle\}$, which must be explicitly antisymmetrised to enforce exchange symmetry [65]. Mathematically, we describe the N electron wavefunction as

$$|\psi\rangle = \int_{\mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N}} \psi(\mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N})\,|\mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N}\rangle\,, \tag{18}$$

where $\mathbf{r}_i = (x_i, y_i, z_i)$ gives the position of the i^{th} electron and $\psi(\mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N}) = \langle \mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N}|\psi\rangle$. We can simulate this system on a classical computer by evaluation of the wavefunction on a discretised spatial grid. However, the cost of storing the wavefunction scales exponentially with the number of electrons, N. Suppose each axis of space is discretised into P equidistant points. The wavefunction is given by

$$|\psi\rangle = \sum_{\mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N}} \psi(\mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N})\,|\mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N}\rangle\,, \tag{19}$$

where $\mathbf{r}_i = (x_i, y_i, z_i), \forall i \in \{1, 2, \ldots, N\}$ and $x_i, y_i, z_i \in \{0, 1, \ldots, P-1\}$. In total, there are P^{3N} complex am-

plitudes, showing that the memory required scales exponentially with the size of the simulated system. This makes it classically intractable to simulate more than a few particles in the first quantisation using a classical computer.

Grid based methods have several advantages over their 2nd quantised counterparts. First quantised simulations can be made arbitrarily precise, and do not introduce basis set discretisation errors, so can be made more accurate than second quantised approaches. Moreover, they are very useful when considering chemical dynamics, or when simulating systems for which the Born-Oppenheimer approximation is not appropriate. In these scenarios, we must include the motion of the nuclei. As described above, if we consider the nuclear motion separately, we need to obtain the potential energy surfaces from electronic structure calculations. As mentioned in the previous section, mapping out these potential energy surfaces is exponentially costly. As such, it is often better to treat the nuclei and electrons on an equal footing, which is only possible with first quantised methods. This is discussed further in Ref. [15].

However, the first quantised method directly stores the wavefunction without exploiting any prior knowledge of the molecule. In contrast, the second quantised method exploits our knowledge that the electrons predominantly occupy the lowest energy levels, as well as our knowledge of the general spatial form of the orbitals. This reduces the resources needed to simulate molecules. Because of the low number of qubits available in current quantum computers, the majority of the work in quantum computational chemistry to date has considered second quantised methods, which are described in the following sections.

2. *Second quantisation*

In the second quantised formalism, we project the Hamiltonian onto M basis wavefunctions, $\{\phi_p(\mathbf{x_i})\}$ (where $\mathbf{x_i}$ is the spatial and spin coordinate of the i^{th} electron, $\mathbf{x_i} = (\mathbf{r_i}, s_i)$), which approximate electron spin-orbitals. We write the many electron wavefunction as a Slater determinant, which is an antisymmetrised product of the single electron basis functions. The wavefunction is given by

$$\psi(\mathbf{x_0}\dots\mathbf{x_{N-1}}) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \phi_0(\mathbf{x_0}) & \phi_1(\mathbf{x_0}) & \dots & \phi_{M-1}(\mathbf{x_0}) \\ \phi_0(\mathbf{x_1}) & \phi_1(\mathbf{x_1}) & \dots & \phi_{M-1}(\mathbf{x_1}) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ \phi_0(\mathbf{x_{N-1}}) & \phi_1(\mathbf{x_{N-1}}) & \dots & \phi_{M-1}(\mathbf{x_{N-1}}) \end{vmatrix}. \tag{20}$$

Swapping the positions of any two electrons is equivalent to interchanging two rows of the Slater determinant, which changes the sign of the wavefunction. This provides the correct exchange symmetry for the fermionic wavefunction. While the number of orbitals considered, M, is typically larger than the number of electrons in the molecule, N, the electrons can only occupy N of the orbitals in a given Slater determinant. As a result, the Slater determinant only contains the N occupied orbitals. This means that to write down a Slater determinant, we only need to indicate which orbitals are occupied by electrons. This enables the introduction of a convenient shorthand for Slater determinants [73]

$$\begin{aligned}\psi(\mathbf{x_0}\dots\mathbf{x_{N-1}}) = \\ |f_{M-1},\dots,f_p,\dots,f_0\rangle = |f\rangle\end{aligned} \tag{21}$$

where $f_p = 1$ when ϕ_p is occupied (and therefore present in the Slater determinant), and $f_p = 0$ when ϕ_p is empty (and therefore not present in the determinant). The vector $|f\rangle$ is known as an occupation number vector. The second quantised formalism is concerned with manipulating these occupation number vectors. As these occupation number vectors are a convenient short-hand for Slater determinants, we will refer to them throughout this paper as Slater determinants. This is common practice in computational chemistry [73].

Electrons are excited into the single electron orbitals by fermionic creation operators, $a_p^\dagger$. They are de-excited by annihilation operators, a_p. These operators obey fermionic anti-commutation relations

$$\begin{aligned}\{a_p, a_q^\dagger\} = a_p a_q^\dagger + a_q^\dagger a_p = \delta_{pq}, \\ \{a_p, a_q\} = \{a_p^\dagger, a_q^\dagger\} = 0.\end{aligned} \tag{22}$$

The determinants $|f\rangle$ form an orthonormal basis in the Fock space of the system. The actions of the fermionic operators on the determinants are given by

$$\begin{aligned} &a_p\,|f_{M-1}, f_{M-2},\dots,f_0\rangle \\ =&\delta_{f_p,1}(-1)^{\sum_{i=0}^{p-1} f_i}\,|f_{M-1}, f_{M-2},\dots,f_p\oplus 1,\dots,f_0\rangle\,, \\ &a_p^\dagger\,|f_{M-1}, f_{M-2},\dots,f_0\rangle \\ =&\delta_{f_p,0}(-1)^{\sum_{i=0}^{p-1} f_i}\,|f_{M-1}, f_{M-2},\dots,f_p\oplus 1,\dots,f_0\rangle\,, \end{aligned} \tag{23}$$

where $\oplus$ denotes addition modulo 2. The phase term $(-1)^{\sum_{i=0}^{p-1} f_i}$ enforces the exchange anti-symmetry of fermions. The orbital occupation operator is given by

$$\begin{aligned}\hat{n}_i = a_i^\dagger a_i \\ \hat{n}_i\,|f_{M-1},\dots,f_i,\dots,f_0\rangle = f_i\,|f_{M-1},\dots,f_i,\dots,f_0\rangle\,,\end{aligned} \tag{24}$$

and counts the number of electrons in a given orbital.

Observables must be independent of the representation used. Therefore, the expectation values of second

quantised operators must be equivalent to the expectation values of the corresponding first quantised operators. As first quantised operators conserve the number of electrons, the second quantised operators must contain an equal number of creation and annihilation operators. We can use these requirements to obtain the second quantised form of the electronic Hamiltonian [72, 73].

$$H = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \tag{25}$$

with

$$\begin{aligned} h_{pq} &= \int d\mathbf{x} \phi_p^*(\mathbf{x}) \left(-\frac{\nabla_i^2}{2} - \sum_I \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \right) \phi_q(\mathbf{x}), \\ h_{pqrs} &= \int d\mathbf{x}_1 d\mathbf{x}_2 \frac{\phi_p^*(\mathbf{x}_1)\phi_q^*(\mathbf{x}_2)\phi_s(\mathbf{x}_1)\phi_r(\mathbf{x}_2)}{|\mathbf{x}_1 - \mathbf{x}_2|}. \end{aligned} \tag{26}$$

This Hamiltonian contains up to M^4 terms, and becomes extremely difficult to solve as the number of basis functions increases. Before examining the form of these basis functions and how to select them in Sec. III D, we first consider general and approximate solutions of the electronic Hamiltonian.

If the electron-electron Coulomb interaction term in Eq. (16) is neglected, we obtain a new Hamiltonian which describes the behaviour of N independent electrons. The solutions of this new Hamiltonian are anti-symmetrised tensor products of the single-particle energy eigenfunctions. This is done by taking Slater determinants of the single particle energy eigenfunctions. These single particle energy eigenfunctions are typically chosen to be orthonormal molecular orbitals (MOs) (whose exact form we will discuss in Sec. III D). These MOs are constructed from linear combinations of atomic orbitals (LCAO method) [72, 73].

As they are the eigenstates of a Hermitian operator, these Slater determinants form a complete basis of the problem Hilbert space. As a result, the eigenstates of the true Hamiltonian can be expressed as linear combinations of these Slater determinants, written as

$$|\psi\rangle = \sum_f \alpha_f |f\rangle, \tag{27}$$

where α_f are complex coefficients which we refer to herein as the determinant amplitudes. These solutions are exact, provided that the molecular orbitals form a complete basis for the single particle states, and the N-electron wavefunction contains all the determinants that these MOs can generate [72, 73]. If all $\binom{M}{N}$ determinants are included, the wavefunction is known as the full configuration interaction (FCI) wavefunction. However, this has a number of determinants which scales exponentially with the number of orbitals, making the calculations classically intractable. One way to make the calculation classically tractable is to approximate the exact ground state wavefunction by considering a restricted number of Slater determinants, as will be discussed in Sec. III C. Alternatively, one may consider only the most important MOs, which we discuss in Sec. III D.

C. Classical computation methods

One route to tractable classical chemistry simulation is to approximate the ground state wavefunction by considering a reduced number of determinants. In this section, we review four such methods for this approximation; the Hartree-Fock (HF), multiconfigurational self-consistent field (MCSCF), configuration interaction (CI), and coupled cluster (CC) methods. These methods create parametrised trial states, which can then be optimised to approach the ground state (to an accuracy determined by the approximations made). In this section we assume that we are working in the full MO basis for our molecule, although in practice this is intractable. The errors resulting from truncation of the basis will be discussed in the next section.

1. *Hartree–Fock*

The Hartree–Fock (HF) method aims to find the dominant Slater determinant in the system wavefunction. This is achieved by optimising the spatial form of the spin-orbitals to minimise the energy of the wavefunction. We generally consider a set of orbitals, M, that is larger than the number of electrons in the molecule, N. As we only consider a single Slater determinant, we are essentially assuming that N of the orbitals are occupied, and $M-N$ are left unoccupied, or virtual. In the HF method, we first neglect the Coulomb repulsion term in Eq. (16), reducing the problem to one of N independent electrons. We then assume that each electron moves in the average charge distribution of all of the other electrons, which introduces an effective potential. We can solve the N coupled equations iteratively; first calculating the position of each electron, then updating the potential, and repeating this process until the orbitals converge. In the second quantised formalism, this procedure is carried out by repeatedly updating the orbitals to construct the Fock operator, and diagonalising the Fock operator to obtain new orbitals. This process is repeated until the orbitals converge, and so HF is also referred to as the self-consistent field method (SCF). The Fock operator, $\hat{f}$, is given by [72]

$$\begin{aligned} \hat{f} &= \sum_{i,j} (h_{ij} + V_{ij}) a_i^\dagger a_j, \\ V_{ij} &= \sum_{k \in occ} (h_{ijkk} - h_{ikkj}), \end{aligned} \tag{28}$$

where V_{ij} describes the effective potential, and *occ* is the set of occupied orbitals. We see that the Fock opera-

tor depends on the spatial form of the orbitals through h_{ij}, h_{ijkk}, and h_{ikkj} which are obtained by calculating the integrals in Eq. (26). When performing a HF calculation, we input a set of atomic orbitals, which are localised around each atom. These orbitals are used to calculate the Fock operator, which is then diagonalised to obtain new orbitals (which are linear combinations of the old orbitals). This process is repeated until the orbitals converge [73]. The new orbitals obtained are referred to as the canonical orbitals. This procedure generates single particle molecular orbitals from combinations of the single particle atomic orbitals.

The HF method accounts for exchange effects (also called Fermi correlation) through the term h_{ikkj}, while the term h_{ijkk} describes the interaction of an electron with the charge distribution of the other electrons. However, as a mean-field solution, the HF method neglects the effects of dynamical correlation (correlation between the electrons due to their Coulomb repulsion) and static correlation (correlation arising from near degeneracies of electronic configurations). As a result, this method is inaccurate when applied to strongly correlated molecules [72].

The Slater determinant generated from a HF calculation is typically taken as the reference state for post-HF methods, such as configuration interaction and coupled cluster, which seek to capture some of the electron correlation energy by including additional determinants, describing excitations above the HF state. In these excitations, the electrons are excited into the virtual orbitals described above. However, in practice the form of the virtual orbitals is found by performing correlated calculations on atomic systems, rather than uncorrelated HF calculations on molecular systems [72].

2. *Multiconfigurational self-consistent field*

As discussed above, the HF method performs poorly for strongly correlated systems. Moreover, the configuration interaction and coupled cluster methods are only effective at recovering dynamic correlation. For states where multiple Slater determinants are equally important, static correlation dominates. These include excited states, systems at the dissociation limit, transition metals, large systems and reaction pathways [75]. One method to treat these systems is to use a multiconfigurational self-consistent field (MCSCF) approach. The MCSCF approach considers a wavefunction with several Slater determinants, and variationally optimises both the molecular orbitals, and the determinant amplitudes simultaneously [76]. MCSCF can be considered the best approximation to the exact wavefunction for a given number of determinants [75]. It is not possible to perform an MCSCF calculation on systems with more than a few orbitals, as the number of determinants scales exponentially with the number of orbitals. For large systems, one can instead use chemical intuition to select the most important Slater determinants, and perform an MCSCF calculation on this restricted number of determinants. Alternatively, we can use the complete active space self-consistent field (CASSCF) method [76]. This considers only the most important orbitals (an active space, see Sec. III E) and performs an MCSCF calculation on all of the determinants that could be generated from distributing a certain number of electrons in these orbitals. Both MCSCF and CASSCF calculations are computationally expensive, with the cost dominated by basis transformation of the two electron integrals [75]. However, they are the most effective methods at treating systems with strong static correlation [72].

3. *Configuration interaction*

The configuration interaction (CI) method generates a correlated wavefunction by considering excitations above a reference state, typically the HF state. If all determinants are included, we recover the full configuration interaction (FCI) wavefunction. The FCI wavefunction can be generated by considering all excitations above the HF wavefunction

$$|\psi_{\mathrm{FCI}}\rangle = \left(I + \sum_{i,\alpha} C_{i\alpha} a_i^\dagger a_\alpha + \sum_{i,j,\alpha,\beta} C_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta + ...\right) |\psi_{\mathrm{HF}}\rangle, \tag{29}$$

where C are parameters to be variationally optimised. As considering all determinants is classically intractable, the CI method is typically limited to including a small number of excitations above the reference state; single excitations (CIS), double excitations (CISD), and occasionally triple excitations (CISDT). However, as low energy excitations dominate the ground state wavefunction, these truncations produce good approximations to the ground state energy [72, 73]. If the reference state is a MCSCF state, the method is known as multireference configuration interaction (MRCI). The CI method suffers from two major limitations. The method converges slowly to the FCI wavefunction, as a result of its linear parametrisation. Furthermore, the energy obtained from a CI calculation is not proportional to the size of the system (not 'size extensive') [71].

4. *Coupled cluster*

The coupled cluster (CC) method also includes additional determinants to recover the correlation energy, but uses a product parametrisation. This overcomes the size-extensivity problem of the CI method. The CC wavefunc-

tion is given by

$$|\psi_{\mathrm{CC}}\rangle = \prod_{i,\alpha}\left(I + \sum_{i,\alpha} C_{i\alpha} a_i^\dagger a_\alpha\right) \times \prod_{i,j,\alpha,\beta}\left(I + \sum_{i,j,\alpha,\beta} C_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta\right) ...|\psi_{\mathrm{HF}}\rangle. \tag{30}$$

This formula can be recast in an exponential form, written as

$$|\psi_{\mathrm{CC}}\rangle = e^T |\psi_{\mathrm{HF}}\rangle, \tag{31}$$

where $T = \sum_i T_i$,

$$\begin{aligned} T_1 &= \sum_{i\in virt,\alpha\in occ} t_{i\alpha} a_i^\dagger a_\alpha, \\ T_2 &= \sum_{i,j\in virt,\alpha,\beta\in occ} t_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta, \\ &..., \end{aligned} \tag{32}$$

where *occ* denotes orbitals that are occupied in the Hartree-Fock state, *virt* denotes orbitals that are unoccupied (virtual) in the Hartree-Fock state, and t are excitation amplitudes which are variationally optimised. When all of the excitation operators T_i are included, the CC method recovers the FCI wavefunction, but this is classically intractable. As a result, the method is normally truncated at a lower excitation level; typically single and double excitations (CCSD). Because of its product parametrisation, the CCSD method generates a trial wavefunction which includes all possible determinants, albeit with an incorrect parametrisation. It therefore provides faster convergence than the CI method. However, the CC method is not without its own shortcomings. Most notably, the CC operator is not unitary, and therefore the wavefunction generated doesn't obey the Rayleigh-Ritz variational principle [72]. Furthermore, the CC method can't be used with MCSCF states, and so struggles with systems displaying strong static correlation [71]. In Sec. V B we describe a modified form of the CC method, known as Unitary Coupled Cluster (UCC). This method is both variational and suitable for multireference states. While it is classically intractable, this method is efficient to implement using a quantum computer [71].

This section has treated the inaccuracies which result from considering a reduced number of determinants, while including all MOs. The following section will discuss the converse case; we consider only a limited number of MOs, but assume that we include all possible determinants that they can generate, unless explicitly stated.

D. Chemical basis sets

In the second quantised formalism the Schrödinger equation is projected onto M basis functions, as discussed in Sec. III B. In this section, we describe some of the conventional basis sets used in classical computational chemistry. Throughout this section, we refer to the 'true' orbitals of the system. These can be obtained by numerically solving the Schrödinger equation using first quantised methods, which is only possible for small atoms or simple molecules. The orbital functions introduced in this section are approximations of these true orbitals. This immediately introduces an error in the second quantised approach which is not present in first quantised simulations. To understand the form of the single-particle atomic orbitals, it is helpful to first revisit the most simple atomic system, the non-relativistic hydrogen atom. The Hamiltonian of this system is that of a single particle in a central Coulomb potential, and has solutions of the form

$$\psi_{nlm} = R_{nl}(r) Y_{lm}(\theta, \phi), \tag{33}$$

where n denotes the energy level of the orbital, l and m describe the angular momentum, $R_{nl}(r)$ are products of Laguerre polynomials and a term decaying exponentially with distance r, and $Y_{lm}(\theta, \phi)$ are spherical harmonics [77]. While these solutions are exact for one electron atoms, they perform poorly for many-electron atoms. They rapidly become diffuse, so can't describe the behaviour of the core electrons well. As a result, functions of a different form are used as the basis states in computational chemistry.

1. *Slater and Gaussian type orbitals*

A better basis is obtained by retaining only the term in $R_{nl}(r)$ with the highest power of r (thus we discard the l index), and including an additional parameter ζ. These functions are known as Slater-type orbitals (STO)

$$R_n^{\mathrm{STO}}(r) \propto (\zeta r)^{n-1} e^{-\zeta r}, \tag{34}$$

where n is the energy level and ζ is a fitting parameter. By using different values of ζ for each orbital, we can generate a good basis [72]. Unlike the true atomic orbitals, these functions do not display oscillatory behaviour. Consequently, linear combinations of STO's are required to approximate the true orbitals. It is possible to only introduce a single basis function for each considered orbital in the molecule, and give each basis function a different ζ value. This is known as a single-zeta representation. Alternatively, we can introduce n basis functions (where n is not the energy level of the orbital, but a number defining the number of basis functions we wish to include), each with a different ζ value, for each orbital. This is known as an n-zeta representation. Introducing additional basis functions in this way increases the radial

flexibility of the wavefunction. While the STO functions exhibit many desirable features, they make evaluating the two-electron integrals in Eq. (26) computationally difficult. As a result, they are not used as basis functions in practice.

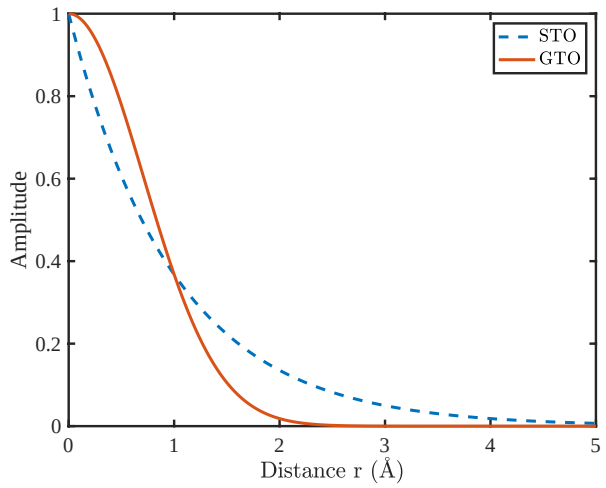

FIG. 3. Comparing the shapes of Slater (STO) and Gaussian (GTO) type orbitals. Normalisation factors are neglected.

To simplify the two-electron integrals, Gaussian basis functions are used. The Gaussian basis functions are obtained by considering the Schrödinger equation with a three dimensional Harmonic oscillator potential. The form of the Gaussian-type orbitals (GTOs) is given by

$$R_{nl}^{\mathrm{GTO}}(r) \propto (\sqrt{\alpha_{nl}}r)^l e^{-\alpha_{nl}r^2}, \tag{35}$$

where α_{nl} is a fitting parameter. As illustrated in Fig. 3, because of the dependence on r^2 in the exponent, GTOs are more localised than STOs. As a result, GTOs do not approximate the atomic charge distribution as well, so more are required to describe a given orbital. However, this limitation is compensated by the ease of integral evaluation. Furthermore, the disadvantages of GTOs are less prominent in molecular calculations [72].

The most common basis sets construct approximate STOs from linear combinations of GTOs. These approximate STOs are used as the basis functions for our atomic orbitals. The number and type of orbitals defines the basis set. There is a compromise between the accuracy obtained and the number of basis functions used. The number of orbitals considered determines the runtime and memory requirements of classical chemistry algorithms. In the case of quantum computational chemistry, the number of basis functions determines the number of qubits and gate operations required to solve the problem.

2. *STO-nG basis sets*

The most simple bases are the STO-nG basis sets (Slater Type Orbital-n Gaussians). In an STO-nG basis, each atomic orbital is considered to be an STO. These STOs are approximated using n GTOs. These basis sets are often called minimal basis sets, as they contain only the orbitals required to write the Hartree–Fock (HF) state (and those orbitals of similar energy). Calculations using minimal basis sets are of limited accuracy, giving only a qualitative description of the system. It is important to note that when carrying out a HF calculation in an STO-nG basis, the true HF energy will not be obtained, as these basis sets only approximate the true HF orbitals. As an example of an STO-nG basis set we consider lithium, which has 3 electrons, of which 2 can reside in the $1s$ orbital, leaving 1 in the second energy level. We include in the minimal basis set $\{1s, 2s, 2p_x, 2p_y, 2p_z\}$ orbitals. We include both the $2s$ and $2p$ orbitals because they are of the same energy level. On a quantum computer, we must use two qubits for each orbital, due to the electron spin. As a result, without any reduction we would require 10 qubits to simulate lithium on a quantum computer.

3. *Split-valence basis sets*

Split-valence (or Pople [78]) basis sets, such as the 6-31G basis, can be used to obtain more accurate results. These basis sets again include only the minimal orbitals, but better approximate the true orbitals than the STO-nG bases do, as they introduce increased radial flexibility for the valence electrons. In the case of the 6-31G basis, the core orbitals are described by one approximate STO, constructed from a linear combination of six GTOs. However, each valence shell orbital has a double-zeta representation; we introduce two approximate STOs for each valence orbital. The more localised STO is composed of three GTOs, while the more diffuse is represented by a single GTO. For example, lithium in the 6-31G basis has a single-zeta representation of the core $1s$ orbital, and a double-zeta representation of the valence orbitals. As a result, the 6-31G basis for lithium consists of $\{1s, 2s, 2s', 2p_x, 2p_y, 2p_z, 2p'_x, 2p'_y, 2p'_z\}$, where the prime denotes that the orbital has a different exponent. This would require 18 qubits to simulate on a quantum computer.

The 6-31G basis is inaccurate when describing molecular systems, as it does not take into account the polarisation of atomic charge caused by bonding. This is somewhat rectified by considering a polarised Pople basis set, such as the 6-31G* basis (which is not a minimal basis set). This basis set includes orbitals with higher angular momenta, which make the angular part of the wavefunction more flexible. These additional orbitals are referred to as 'polarisation functions', as they describe the polarisation of electronic charge. Nevertheless, as Pople basis sets were designed for HF calculations, even large, polarised Pople basis sets aren't well suited for correlated post-HF calculations [72].

4. Correlation-consistent basis sets

Additional accuracy can be obtained by using cc-PVnZ basis sets (correlation consistent polarised valence n zeta), introduced by Dunning [79]. These include additional unoccupied ('virtual') orbitals to recover the correlation energy. The virtual orbitals are generated from correlated calculations on atoms. The core orbitals have a single-zeta representation, while the valence orbitals have an n-zeta representation. The virtual orbitals considered are polarisation functions, with higher angular momenta than the valence orbitals. The polarisation functions are selected by the size of their contribution to the correlation energy. Higher accuracy can be obtained by correlating both the core and valence electrons (cc-PCVnZ bases), but the cost typically outweighs the benefits. The contribution of the core orbitals to the correlation energy is approximately constant over the potential energy surface, so it can be removed by taking relative energies [72]. We illustrate the number of basis functions included in the cc-PVnZ basis by considering several examples.

For atomic hydrogen in the cc-PVDZ ($n = 2$, D = double) the highest occupied electron level (the valence level) is the 1$^{\text{st}}$ level, and so we take a double-zeta representation of the $1s$ state, considering $\{1s,\ 1s'\}$ orbitals. The $1s'$ orbital is often referred to as a $2s$ orbital. This is because the additional function chosen to describe the valence orbital has the same angular momentum as the ordinary $1s$ orbital, but is more diffuse – so it resembles a $2s$ orbital. We then include polarisation functions, which have a higher angular momentum value than the valence functions. In total, there are five basis functions for cc-PVDZ hydrogen: $\{1s,\ 1s',\ 2p_x,\ 2p_y,\ 2p_z\}$, requiring 10 qubits to simulate. These are shown in Fig. 4.

For lithium in the cc-PVDZ basis, the core orbital is $\{1s\}$. The valence orbitals (which have a double-zeta representation) are $\{2s,\ 2p_x,\ 2p_y,\ 2p_z,\ 2s',\ 2p'_x,\ 2p'_y,\ 2p'_z\}$, and the polarisation functions are $\{3d_{zz},\ 3d_{xz},\ 3d_{yz},\ 3d_{xy},\ 3d_{x^2-y^2}\ \}$, which we write as $\{5 \times 3d\}$. This yields 14 basis functions, requiring 28 qubits.

Argon has 18 electrons, which completely fill the $1s$, $2s$, $2p$, $3s$ and $3p$ states. The core orbitals are $\{1s,\ 2s,\ 2p_x,\ 2p_y,\ 2p_z\}$. The valence orbitals are $\{3s,\ 3p_x,\ 3p_y,\ 3p_z,\ 3s',\ 3p'_x,\ 3p'_y,\ 3p'_z\}$. We include $\{5 \times 3d\}$ polarisation functions for angular flexibility. This gives a total of 18 basis functions for argon, requiring 36 qubits to simulate.

For lithium in the cc-PVTZ basis ($n = 3$, T = triple), we first include the 14 orbitals above. As we consider a triple-zeta representation of the valence orbitals, we need additional $\{2s'', 2p''_x, 2p''_y, 2p''_z\}$ orbitals. We then include additional polarisation functions; $\{5 \times 3d', 7 \times 4f\}$. This leads to a total of 30 orbitals, requiring 60 qubits to simulate.

The number of orbitals included in a cc-PVnZ basis scales approximately as $n^3/3$ [72]. It is important to highlight that cc-PVnZ basis sets with higher values of n contain orbitals that better approximate the true atomic orbitals than those with lower n values. However, even large ($n = 5$) basis sets struggle to exactly represent the true HF orbitals of simple molecules such as N_2 [72]. This limitation can be overcome by measuring the ground state energy in several different cc-PVnZ bases, and then extrapolating to the basis set limit.

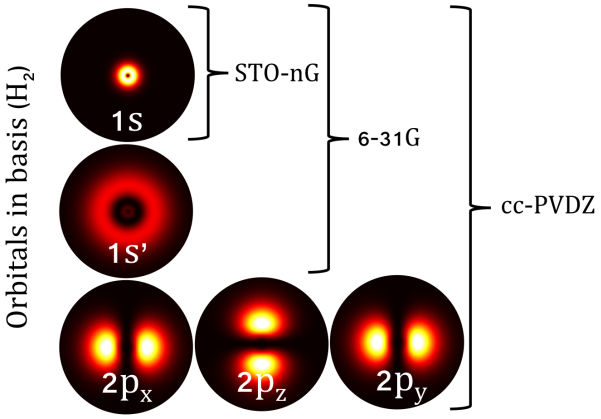

FIG. 4. The orbitals included in different basis sets for the Hydrogen atom. The $1s'$ orbital is often written as $2s$. The plots show the radial probability distributions for the true Hydrogenic orbitals, which the basis orbitals approximate.

5. Plane wave basis sets

While the afforementioned basis sets have a long history of use in classical computational chemistry (and as a result, early work in quantum computational chemistry), they are not necessarily optimal basis sets for calculations performed on quantum computers. These basis sets were designed for ease of performing the two-body integrals, which is no longer a major bottleneck for modern supercomputers. As a result, there is some freedom to develop basis sets which may be more useful for quantum computational chemistry. Two examples of such bases are the plane wave and plane wave dual basis sets introduced for quantum computing in Ref. [13]. The plane wave basis functions, $\phi_\nu(r)$, are given by

$$\phi_\nu = \sqrt{\frac{1}{V}} \exp\left(\frac{2\pi i \nu r}{L}\right) \tag{36}$$

for a wave with wavevector corresponding to the ν^{th} harmonic of the computational cell with length L and volume V. The plane wave dual basis is obtained by taking the discrete Fourier transform of the plane wave basis states. These basis sets diagonalise the kinetic and potential operators, respectively. Multiple benefits are obtained from diagonalising these terms – most notably a

reduction in the asymptotic scaling of quantum chemistry algorithms, which will be discussed in more detail in Sec. V. Plane wave basis sets are well suited to periodic systems, and have a long history of use in classical density functional theory calculations. However, to describe molecular systems, approximately 10 times as many plane wave basis functions are required as GTOs – thus the number of qubits required may be much greater when using the plane wave basis sets. A similar improvement in algorithmic scaling might be obtained using the recently proposed gausslet basis sets [80], which have a smaller multiplicative overhead. Creating efficient basis sets for quantum computational chemistry remains an open and fundamental problem.

E. Reduction of orbitals

It is often the case that certain orbitals are very likely to be either occupied or virtual in all Slater determinants in the wavefunction. As calculating the ground state energy is essentially a question of distributing electrons among orbitals, we can simplify our calculation by using this information. Specifically, we are able to remove orbitals from the calculation if their expected occupation number is close to 0 or 1. Our calculation is reduced to including only the most important (ambiguously occupied) orbitals. This is known as performing the calculation in a reduced active space.

In order to determine the occupation of orbitals, we use the reduced density matrices (RDMs) of the system. The expectation value of any 1- or 2-electron Hermitian operator, O, with a state $|\psi\rangle = \sum_f \alpha_f |f\rangle$, is given by [72]

$$\langle\psi| O |\psi\rangle = \sum_{i,j} O_{ij}\rho^1_{ij} + \sum_{i,j,k,l} V_{ijkl}\rho^2_{ijkl}, \tag{37}$$

$$\rho^1_{ij} = \langle\psi| a_i^\dagger a_j |\psi\rangle, \quad \rho^2_{ijkl} = \langle\psi| a_i^\dagger a_k^\dagger a_l a_j |\psi\rangle,$$

where ρ^1 is the single-particle reduced density matrix (1-RDM), ρ^2 is the two-particle reduced density matrix (2-RDM), and O_{ij} and V_{ijkl} are defined in a similar way to the coefficients in Eq. (26). These RDMs contain all of the information required to evalulate $\langle O\rangle_\psi$. From the definition above, we can see that the diagonal elements of ρ^1 are the expectation of the number operator for the corresponding orbitals. As ρ^1 is a Hermitian operator, we can diagonalise it with a unitary transform. This is a basis change from the canonical orbitals to the 'natural molecular orbitals'. The diagonal elements of the basis transformed ρ^1 are called the natural orbital occupation numbers (NOONs).

Orbitals with a NOON close to 0 or 1 (compared to the other NOONs) can be assumed to be empty or occupied, respectively. As a result, we can reduce our problem by considering only the ambiguously occupied orbitals. This was used in Ref. [28] to reduce the number of qubits required for simulation. In Sec. VII we provide an explicit example of how this method can be used to reduce the number of orbitals required to simulate lithium hydride in an STO-3G basis, and H_2 in a cc-PVDZ basis set.

This section has introduced the concepts in classical computational chemistry necessary to understand the early work in quantum computational chemistry. The following sections introduce methods developed to solve chemistry problems using quantum computers. We return to classical computational chemistry methods in Sec. VIII, where we assess the strengths, weaknesses and limits of the methods introduced here.

IV. QUANTUM COMPUTATIONAL CHEMISTRY MAPPINGS

In this section, we describe the techniques developed to enable quantum computers to solve problems in chemistry. In Sec. IV A and Sec. IV B we introduce the methods of encoding fermions into qubits, which maps the chemistry problem onto a quantum computer. We then describe methods which take advantage of symmetries to reduce the number of qubits required, in Sec. IV C.

A. 1st quantised fermion encoding methods

Here we give a brief overview of first quantised quantum simulation. As discussed in Sec. III B 1, the wavefunction of an N-particle system can be represented in real space on a discretised grid of P points per axis, and is given by

$$|\psi\rangle = \sum_{\mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N}} \psi(\mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N}) |\mathbf{r_1},\mathbf{r_2},\ldots,\mathbf{r_N}\rangle, \tag{38}$$

where $\mathbf{r}_i = (x_i, y_i, z_i), \forall i \in \{1, 2, \ldots, N\}$ and $x_i, y_i, z_i \in \{0, 1, \ldots, P-1\}$. We consider the case where $P = 2^m$, where m is an arbitrary number which determines the precision of our simulation. While it is classically intractable to store the required 2^{3mN} complex amplitudes for large quantum systems, it is possible using a quantum computer. If we write the basis vector $|x = 2^m - 1\rangle$ in binary as $|11....11\rangle$, we note that it only requires m bits. Furthermore, an m qubit register can be in a superposition of 2^m possible states. As a result, it only requires $3mN$ qubits to store the N electron wavefunction described in Eq. (38). This makes it efficient to represent quantum systems on quantum computers in first quantisation. Moreover, the simulation accuracy increases exponentially with the number of qubits used.

This approach was first introduced for general quantum systems by Wiesner [81] and Zalka [82]. It was then adapted for simulating problems in chemistry by Lidar and Wang [83], and Kassal *et al.* [15]. Once qubits have

been used to create a discretised grid, physically relevant states can be efficiently prepared [15, 84], and the relevant observables measured [15, 85].

As discussed in the previous section, simulation of reaction dynamics is more efficient in the first quantised representation, and without making the Born-Oppenheimer approximation, which necessitates costly classical precomputation [15]. However, the simulation of molecules in the first quantised approach requires considerably more qubits than in the second quantised approach [15, 84, 86]. For example, it would require around 100 logical qubits to simulate the lithium atom in the first quantisation (a task beyond classical computers using numerical grid based methods) [15]. However, it is important to note that this simulation would be exact, while a second quantised simulation would include both basis set errors, and the Born-Oppenheimer approximation.

Consequently, the first quantised approach is typically considered less suitable for near-term quantum computers, which will have small numbers of qubits. We refer the reader to Refs. [15, 81, 82, 84–87] for more details on the first quantised approach.

B. 2nd quantised fermion encoding methods

To simulate chemical systems in the 2nd quantised representation on a quantum computer, we need to map from operators which act on indistinguishable fermions to distinguishable qubits. An encoding method is a map from the fermionic Fock space to the qubits' Hilbert space, such that every fermionic state can be represented by a qubit state. Under this mapping, the fermionic creation and annihilation operators are represented by unitary operators on the qubits. There are three main methods of encoding, which we describe below: the Jordan–Wigner (JW), Parity, and Bravyi–Kitaev (BK) mappings. In this section we consider a molecule with M spin-orbitals and N electrons.

1. *Jordan-Wigner encoding*

In the Jordan–Wigner (JW) encoding, we store the occupation number of an orbital in the $|0\rangle$ or $|1\rangle$ state of a qubit (unoccupied and occupied, respectively). More rigourously,

$$|f_{M-1}, f_{M-2}, \ldots, f_0\rangle \rightarrow |q_{M-1}, q_{M-2}, \ldots, q_0\rangle, \quad q_p = f_p \in \{0, 1\}. \tag{39}$$

The fermionic creation and annhilation operators increase or decrease the occupation number of an orbital by 1, and also introduce a multiplicative phase factor (see Eq. (23)). The qubit mappings of the operators preserves these features, and are given by,

$$\begin{aligned} a_p &= Q_p \otimes Z_{p-1} \otimes \cdots \otimes Z_0, \\ a_p^\dagger &= Q_p^\dagger \otimes Z_{p-1} \otimes \cdots \otimes Z_0, \end{aligned} \tag{40}$$

where $Q = |0\rangle\langle 1| = \frac{1}{2}(X + iY)$ and $Q^\dagger = |1\rangle\langle 0| = \frac{1}{2}(X - iY)$. The $Q/Q^\dagger$ operator changes the occupation number of the target orbital, while the string of Z operators recovers the exchange phase factor $(-1)^{\sum_{i=0}^{p-1} f_i}$. An example JW mapping is shown in Table. I.

Fermion	Qubit
$a\|001\rangle + b\|010\rangle + c\|100\rangle$	$a\|001\rangle + b\|010\rangle + c\|100\rangle$
$a_0, \quad a_1, \quad a_2$	$Q_0, \quad Q_1Z_0, \quad Q_2Z_1Z_0$
$a_0^\dagger, \quad a_1^\dagger, \quad a_2^\dagger$	$Q_0^\dagger, \quad Q_1^\dagger Z_0, \quad Q_2^\dagger Z_1 Z_0$
$n_i = a_i^\dagger a_i$	$Q_i^\dagger Q_i = \|1\rangle\langle 1\|_i$

TABLE I. An example Jordan–Wigner mapping of a fermionic Fock state and its fermionic operators onto the corresponding qubit state, and qubit operators. n_i is the fermionic number operator.

The primary advantage of the JW encoding is its simplicity. However, while the occupation of an orbital is stored locally, the parity is stored non-locally. The string of Z operators means that it takes $O(M)$ qubit operations to apply a fermionic operator.

Working in the JW basis, it is easy to see the advantage that quantum computers have over their classical counterparts for chemistry problems. As discussed in Sec. III B 2, the full configuration interaction wavefunction contains a number of determinants which scales exponentially with the number of orbitals, M. As such, it requires an amount of memory that scales exponentially with the system size. However, using a quantum computer, we can instead store the FCI wavefunction using only M qubits [14]. A register of M qubits can be in a superposition of 2^M computational basis states. In the JW basis, every Slater determinant required for the FCI wavefunction can be written as one of these basis states. As such, quantum computers can efficiently store the FCI wavefunction. This is also true for the other 2nd encodings.

2. *Parity encoding*

In the parity encoding, instead of directly storing the occupation number, one can instead use the p^{th} qubit to store the parity of the first p modes [88],

$$|f_{M-1}, f_{M-2}, \ldots, f_0\rangle \rightarrow |q_{M-1}, q_{M-2}, \ldots, q_0\rangle \quad q_p = \left[\sum_i^p f_i\right] \pmod 2. \tag{41}$$

The creation and annihilation operators are

$$\begin{aligned} a_p =& X_{M-1} \otimes \cdots \otimes X_{p+1} \\ & \otimes (Q_p \otimes |0\rangle \langle 0|_{p-1} - Q_p^\dagger \otimes |1\rangle \langle 1|_{p-1}), \\ a_p^\dagger =& X_{M-1} \otimes \cdots \otimes X_{p+1} \\ & \otimes (Q_p^\dagger \otimes |0\rangle \langle 0|_{p-1} - Q_p \otimes |1\rangle \langle 1|_{p-1}). \end{aligned} \tag{42}$$

These operators check the parity of the first $(p-1)^{\text{th}}$ modes, and update q_p accordingly using Q_p or $Q_p^\dagger$. The string of X gates then updates all of the qubits which store the parity of qubit p. The minus sign is introduced to recover the exchange phase factor $(-1)^{\sum_{i=0}^{p-1} f_i}$ in Eq. (23).

Fermion	Qubit
$a\,\|001\rangle + b\,\|010\rangle + c\,\|100\rangle$	$a\,\|111\rangle + b\,\|110\rangle + c\,\|100\rangle$
a_0	$X_2X_1Q_0$
a_1	$X_2(Q_1\,\|0\rangle\langle 0|_0 - Q_1^\dagger\,\|1\rangle\langle 1|_0)$
a_2	$Q_2\,\|0\rangle\langle 0|_1 - Q_2^\dagger\,\|1\rangle\langle 1|_1$
$a_0^\dagger$	$X_2X_1Q_0^\dagger$
$a_1^\dagger$	$X_2(Q_1^\dagger\,\|0\rangle\langle 0|_0 - Q_1\,\|1\rangle\langle 1|_0)$
$a_2^\dagger$	$Q_2^\dagger\,\|0\rangle\langle 0|_1 - Q_2\,\|1\rangle\langle 1|_1$
$n_i = a_i^\dagger a_i$	$\frac{1}{2}(I - Z_jZ_{j-1})$

TABLE II. An example parity mapping of a fermionic Fock state and its fermionic operators onto the corresponding qubit state, and qubit operators. n_i is the fermionic number operator.

It is simple to transform between the JW basis and the parity basis. This can be done using a lower triangular matrix, as shown below for the fermionic state described in Table. I and Table. II.

$$\begin{aligned} & P \cdot (a\,|001\rangle + b\,|010\rangle + c\,|100\rangle) \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \cdot \left(a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) \\ &= \left(a \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) \\ &= a\,|111\rangle + b\,|110\rangle + c\,|100\rangle \end{aligned} \tag{43}$$

In contrast to the JW transform, the parity mapping stores the parity information locally, and the occupation number non-locally. Like the JW mapping, it requires $O(M)$ qubit operations to carry out a single fermionic operation.

3. *Bravyi–Kitaev encoding*

The Bravyi–Kitaev (BK) encoding [89] is a midway point between the JW and parity encoding methods, in that it compromises on the locality of occupation number and parity information. The orbitals store partial sums of occupation numbers. The occupation numbers included in each partial sum are defined by the BK matrix, β_{pq}.

$$\begin{aligned} |f_{M-1}, f_{M-2}, \ldots, f_0\rangle \rightarrow |q_{M-1}, q_{M-2}, \ldots, q_0\rangle\,, \\ q_p = \left[\sum_{q=0}^{p} \beta_{pq} f_q \right] \pmod 2. \end{aligned} \tag{44}$$

It is defined recursively [88, 89] via

$$\begin{aligned} \beta_1 &= [1] \\ \beta_{2^{x+1}} &= \begin{bmatrix} \beta_{2^x} & \mathbf{0} \\ \mathbf{A} & \beta_{2^x} \end{bmatrix} \end{aligned} \tag{45}$$

where $\mathbf{A}$ is an $(2^x \times 2^x)$ matrix of zeros, with the bottom row filled with ones, and $\mathbf{0}$ is a $(2^x \times 2^x)$ matrix of zeros. As an example, when $M = 8$ $(x = 2)$, the matrix β_{pq} is

$$\beta_8 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{46}$$

When the number of qubits is not a power of two, the BK encoding is carried out by creating the BK matrix for the next largest power of two, and only using the first M rows. The qubit operators for the BK encoding are considerably more complicated than those in the JW or parity encodings, and we refer to Refs. [88, 90] for a complete derivation of them. The advantage of the BK encoding is a reduction in the number of qubit operations needed to carry out a fermionic operation. By balancing the locality of occupation and parity information, the number of terms needed to realise a fermionic operator scales as $O(\log_2 M)$.

We remark that another version of the BK encoding also exists in the literature. This is referred to as the BK-tree method, as it takes its inspiration from a classical data structure known as a Fenwick tree [91]. We explicitly show how to use this mapping with molecules in Sec. VII.

As with the standard BK mapping, the BK-tree encoding balances how it stores occupation and parity information. As a result, it too only requires $O(\log_2 M)$ qubit operations to realise a fermionic operator. However, there

are subtle differences between the two mappings. It has been noted that the BK-tree mapping produces qubit operators with a greater weight than the standard BK mapping [92]. This would suggest that it is less suitable for near-term quantum computation. However, the BK-tree mapping also possesses advantages over the standard BK encoding. The BK-tree mapping is uniquely defined even when the number of orbitals, M, is not a power of 2. As a result, when using the BK-tree mapping we are always able to use the qubit reduction by symmetry technique, which is discussed in Sec. IV C. We have observed that it is only possible to use this technique with the standard BK mapping when the number of orbitals is a power of two. As a result, it is important to consider the benefits of both mappings, before choosing which one to use.

4. *Other encoding methods*

There are also other possible encodings, although these are less widely discussed in the literature, and have not yet been experimentally implemented. There are non-linear encoding methods which can be used to encode M orbitals into $M' < M$ qubits [93]. In some cases, the number of qubits saved can be exponential in M, although this introduces the need for $O(M)$-controlled gates [93].

In contrast to these non-linear encodings, other mappings have been developed which exchange an increased number of qubits for a lower gate count. Verstraete and Cirac developed a scheme to eliminate the strings of Z operators introduced by the JW transform, resulting in qubit operators with the same locality as the fermionic operators [94]. This is achieved by doubling the number of qubits. This was then generalised by Whitfield *et al.* , who refer to these mappings as 'auxiliary fermion transforms' [95]. They too introduce additional qubits (at most doubling the number of qubits) to map local fermionic operators to local qubit operators.

There is also another variant of the BK transform, known as the Bravyi-Kitaev superfast transform [96]. This mapping first represents each orbital by a vertex on a graph, and each interaction term in the Hamiltonian as an edge on the graph. Qubits are then associated to the edges. In general, a graph will have more edges than vertices, so this increases the number of qubits required. However, the number of gates required to implement a fermionic operator will scale as $O(d)$ where d is the degree of the graph. Assuming fairly local interactions for a molecule, the degree of the graph will be less than the number of vertices. As a result, the BKSF transform will require fewer gates than the JW or parity mapping.

C. Hamiltonian reduction

In this section, we focus on methods used to reduce the number of qubits required for the 2nd quantised approach, using $\mathbb{Z}_2$ symmetries. Methods have also been developed to reduce qubits in the first quantised approach, and in the 2nd quantised approach using error correcting codes or stabilisers [97]. Alternatively, one may make use of quantum autoencoders, which can compress the wavefunction into a subspace of the full Hilbert space [98].

In the JW, parity and BK encoding methods, the number of qubits is equal to the number of spin-orbitals considered, M. However, as the molecular Hamiltonian possesses symmetries, the wavefunction can be stored in a smaller Hilbert space. In this paper, we describe the method of Ref. [97], which utilises two such symmetries: conservation of electron number and spin. This method enables the systematic reduction of two qubits when using the parity, BK, or BK-tree encoding. For a molecule with M spin-orbitals, we can arrange the orbitals such that the first $M/2$ orbitals describe spin up states ($|\uparrow\rangle$), and the last $M/2$ orbitals describe the spin down states ($|\downarrow\rangle$). For non-relativistic molecules, the total number electrons and the total spin are conserved. The parity encoding matrix for $M = 4$ is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}. \tag{47}$$

Because every element in the final row is one, the final element of a parity encoded vector, q_{M-1}, is equal to the number of electrons (mod 2). Similarly, half of the elements in the $M/2^{\text{th}}$ row are one. As a result, the $M/2^{\text{th}}$ element in the parity encoded vector, $q_{\frac{M}{2}-1}$, is equal to the number of spin up electrons (mod 2). As the electron number and total spin are conserved by the molecular Hamiltonian, these qubits are only acted on by the identity or Pauli Z operators. We can replace these operators by their corresponding eigenvalues (+1 for the identity, +1 for Z_{M-1} if the total number of orbitals is even, −1 for Z_{M-1} if the total number of orbitals is odd, +1 for $Z_{\frac{M}{2}-1}$ if the number of spin up orbitals is even, and −1 for $Z_{\frac{M}{2}-1}$ if the number of spin up orbitals is odd). The Hamiltonian then only acts on $(M-2)$ qubits, so two qubits can be removed from the simulation. Exactly the same method can be used for the BK and BK-tree encodings. We will explicity show how this method can be used to remove two qubits from molecular Hamiltonians in Sec. VII. We remark that while this transformation leaves the ground state of the system unchanged, it does alter the excited states of the Hamiltonian. In particular, we are restricted to finding those states with an electron number equal to the atomic number of the molecule.

V. QUANTUM COMPUTATIONAL CHEMISTRY ALGORITHMS

In this section, we focus on methods used to solve the electronic structure problem with a quantum computer. We describe the quantum phase estimation algorithm (QPE) and related methods in Sec. V A. We then give a comprehensive description of the variational quantum eigensolver (VQE) in Sec. V B. Both of these sections are concerned with finding the ground state energies of molecules. We conclude this section with a discussion of methods that can be used to find the excited states of molecules in Sec. V C.

It can be argued that the VQE and QPE, as presented herein, represent near-term and long-term methods (respectively) for solving chemistry problems with a quantum computer. However, in reality, aspects of each algorithm can be incorporated into the other, creating new methods [99, 100] which occupy the intermediate region in the quantum computational chemistry timeline.

A. Quantum phase estimation

Phase estimation [68] is used to find the lowest energy eigenstate, $|E_0\rangle$, of a qubit Hamiltonian acting on M qubits [69]. In the case of quantum computational chemistry, this qubit Hamiltonian encodes a molecular fermionic Hamiltonian, and may have been reduced using the methods described in the previous sections.

The phase estimation algorithm is described as follows [32], and shown in Fig. 5.

1. We initialise the qubit register in state $|\psi\rangle$, which has nonzero overlap with the true FCI ground state of the molecule. We require an additional register of ancilla qubits. The number of ancilla qubits required is equal to the number of bits of precision desired in the energy, P. We can expand the state $|\psi\rangle$ in terms of energy eigenstates of the molecular Hamiltonian, writing that $|\psi\rangle = \sum_i c_i |E_i\rangle$.

2. We apply a Hadamard gate to each ancilla qubit, placing the ancilla register in the superposition $\frac{1}{\sqrt{N}}\sum_x |x\rangle$, where x are all possible bit-strings that can be constructed from P bits. We then apply the controlled gates shown in Fig. 5:

$$\frac{1}{\sqrt{N}}\sum_i\sum_x |x\rangle c_i |E_i\rangle \rightarrow \frac{1}{\sqrt{N}}\sum_i\sum_x e^{-2\pi i E_i x} c_i |x\rangle |E_i\rangle . \tag{48}$$

3. Apply the quantum Fourier transform to the ancilla qubits in order to learn the phase, which encodes the information about the energy eigenvalue:

$$\frac{1}{\sqrt{N}}\sum_i\sum_x e^{-2\pi i E_i x} c_i |x\rangle |E_i\rangle \xrightarrow{QFT} \sum_i c_i |e_i\rangle |E_i\rangle . \tag{49}$$

4. Measure the ancilla qubits in the Z basis, which gives the ground state energy eigenvalue, e_0, as a binary bit-string, with probability $|c_0|^2$.

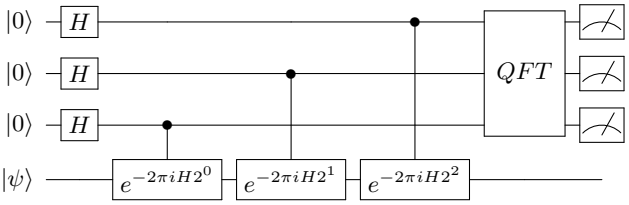

FIG. 5. The canonical QPE circuit with three ancilla qubits. When the ancilla qubits are in state $|x\rangle$, a control rotation $e^{-2\pi i H x}$ is applied to the target state $|\psi\rangle$. QFT denotes the quantum Fourier transform circuit. By measuring the ancilla qubits in the computational basis, they collapse to an eigenvalue of H and the register qubits collapse to the corresponding energy eigenstate.

The phase estimation algorithm described above can also be modified to use only a single ancilla qubit, which uses the single ancilla to measure each bit in the energy eigenvalue [14].

QPE requires the implementation of the time evolution operator, e^{-iHt}. If H can be decomposed as $H = \sum_i h_i$, then a Lie-Trotter-Suzuki approximation [66] of the time evolution is

$$e^{-iHt} = \left(\prod_i e^{-ih_i t/S}\right)^S + O(1/S). \tag{50}$$

In practice, S should be large in order to suppress the errors in the approximation.

Initialising the qubit register in a state which has a sufficiently large overlap with the ground state is a non-trivial problem. This is important, because otherwise we will be more likely to collapse the wavefunction to a higher energy excited state than the ground state. It has been shown that if our state preparation scheme is imperfect, unbiased ground state preparation becomes exponentially more difficult as the system increases in size [101]. Several techniques have been proposed for state preparation, including: using the variational methods discussed in the next section [99], quantum cooling [102], and most commonly, adiabatic state preparation [14]. We focus here on adiabatic state preparation.

For any Hamiltonian H_s, we can prepare a state $|\psi\rangle$ that is close to its ground state via adiabatic state preparation. To do so, we first start with a simple Hamiltonian

H_0 and prepare its ground state. We then slowly change the Hamiltonian from H_0 to H_s, thus preparing a state that is close to the ground state of H_s. The efficiency of adiabatic state preparation depends on the gap between the ground state and the first excited state along the path between H_0 and H_s. For molecules, this is commonly achieved by initialising the system in the ground state of the Hartree-Fock Hamiltonian (H_0), and interpolating between the initial and final Hamiltonians using an annealing schedule such as $H(t) = (1-t)H_0 + tH_s$, where t is time [14]. An evolution under this time-dependent Hamiltonian requires Trotterization.

Recently, new methods have been developed to realise the time evolution operator more efficiently, including: higher-order product formulas [103], quantum walk based methods [104, 105], Taylor series expansions [106–108], and quantum signal processing [109–111]. More information about these methods can be found in Ref. [112]. These methods have been considered for the problem of finding the ground states of electronic structure Hamiltonians, where they lead to a dramatic reduction in the resources required [113].

Despite recent progress, all of the methods discussed above require circuits with a large number of gates, resulting in an accumulation of errors. As a result, these methods are typically assumed to require fault-tolerance. As near-term quantum computers will not have enough physical qubits for error correction [59], the long gate sequences required by these algorithms make them unsuitable for near-term hardware. As such, alternative methods are required for near-future chemistry simulation.

B. Variational quantum eigensolver

The most promising chemistry algorithm for the noisy intermediate-scale quantum (NISQ) era [114] is the variational quantum eigensolver (VQE) [16, 20]. The VQE aims to find the lowest eigenvalue of a Hamiltonian, such as that of a molecule. The VQE is a hybrid quantum-classical algorithm. This means that it only uses the quantum computer for a classically intractable subroutine. This exchanges the long coherence times needed for phase estimation, for a polynomial overhead due to measurement repetitions and classical optimisation.

The VQE has been experimentally demonstrated on most leading quantum architectures [20, 22, 23, 25–28], and shows many desirable features. It appears to be robust against errors [16, 25], and capable of finding the ground state energies of small molecules using low depth circuits [25]. Despite the successes of the VQE, several challenges remain – most notably the difficulty of classical optimisation, the number of measurements required, the construction of suitable quantum circuits, and mitigating the effects of noise.

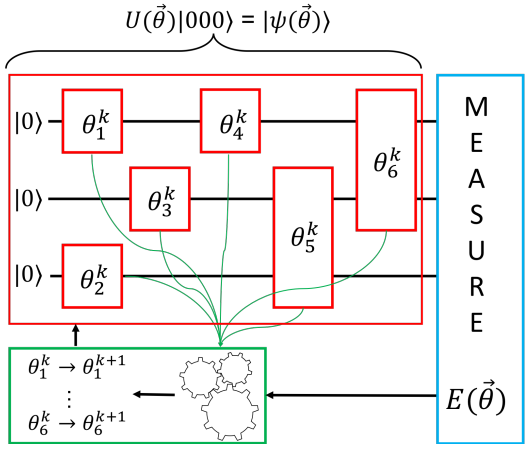

FIG. 6. A schematic of the variational quantum eigensolver (VQE). The classically intractable state preparation and measurement subroutines (red and blue) are performed on the small quantum computer. The current energy and parameter values are fed into a classical optimisation routine (green), which outputs new values of the parameters. The new parameters are then fed back into the quantum circuit. The gates acting on the qubits can be any parametrised gates, e.g. single qubit rotations or controlled rotations. Non-parametrised gates (e.g. X, Y, Z, CNOT) are also allowed. The circuit $U(\vec{\theta})$ and trial wavefunction it produces $|\psi(\vec{\theta})\rangle$ are both known as the VQE ansatz. The process is repeated until the energy converges.

1. *Implementation*

The VQE relies upon the Rayleigh-Ritz variational principle [115]. This states that for a parametrised trial wavefunction $|\psi(\vec{\theta})\rangle$

$$\langle\psi(\vec{\theta})|\, H\, |\psi(\vec{\theta})\rangle \geq E_0, \tag{51}$$

where E_0 is the lowest energy eigenvalue of the Hamiltonian H, and $\vec{\theta}$ is a vector of independent parameters, $\vec{\theta} = (\theta_1, ..., \theta_n)^T$. This implies that we can find the ground state wavefunction and energy by finding the values of the parameters which minimise the energy expectation value. As classical computers are unable to efficiently prepare, store and measure the wavefunction, we use the quantum computer for this subroutine. We then use the classical computer to update the parameters using an optimisation algorithm. This sequence is shown in Fig. 6. The qubit register is initialised in the zero state. We can optionally apply a non-parametrised set of gates to generate a mean-field or multi-reference state [75, 116] describing the chemical system of interest

$$|\psi_{\text{ref}}\rangle = U_{\text{prep}}|\bar{0}\rangle. \tag{52}$$

A series of parametrised gates $U(\vec{\theta}) = U_N(\theta_N)\dots U_k(\theta_k)\dots U_1(\theta_1)$ are then applied to the qubits. Here, $U_k(\theta_k)$ is the k^{th} single or two qubit

unitary gate, controlled by parameter θ_k. This circuit generates the trial wavefunction

$$|\psi(\vec{\theta})\rangle = U(\vec{\theta})|\psi_{\text{ref}}\rangle. \tag{53}$$

We refer to both $|\psi(\vec{\theta})\rangle$ and $U(\vec{\theta})$ as the ansatz. The set of all possible states that could be created by the circuit U is known as the 'ansatz space'.

Once the wavefunction has been generated, we need to measure the expectation value of the Hamiltonian. Molecular Hamiltonians can be written as a linear combination of products of local Pauli operators,

$$H = \sum_j h_j \prod_i \sigma_i^j, \tag{54}$$

where h_j is a scalar coefficient, σ_i^j represents one of I , X , Y or Z, i denotes which qubit the operator acts on, and j denotes the term in the Hamiltonian. We can then use the linearity of expectation values to write that

$$E(\vec{\theta}_k) = \sum_j^N h_j \langle\psi(\vec{\theta}_k)| \prod_i \sigma_i^j |\psi(\vec{\theta}_k)\rangle. \tag{55}$$

These state preparation and measurement steps should be repeated many times in order to measure the expectation value of every term in the Hamiltonian to the required precision. This is known as the Hamiltonian averaging method of calculating the energy [101], and requires $O(\frac{1}{\epsilon^2})$ measurements to determine the energy to a precision ϵ [71]. As the quantum computer is reinitialised for each repetition, the required coherence time is dramatically reduced compared to QPE.

Once the energy has been measured, it is used as the input for a classical optimisation routine, together with the current values of $\vec{\theta}_k$. The optimisation routine outputs new values of the circuit parameters, $\vec{\theta}_{k+1}$. These are used to prepare a new trial state, $|\psi(\vec{\theta}_{k+1})\rangle$, which is ideally lower in energy. These steps are repeated until the energy converges to a minimum. While the algorithmic description of the VQE is simple, effective implementation can be challenging – even for small chemical systems. One must select an ansatz appropriate for the capabilities of the available hardware, as well as a suitable classical optimisation routine. The merits, drawbacks and implementation of common ansatze are discussed in the following section. We then summarise previous investigations into classical optimisation routines for use in the VQE, as well as related methods which aid optimisation.

2. Ansatze

The parametrised circuits, or 'ansatze', for the VQE lie between two extremes; hardware-efficient and chemically inspired.

a. Hardware efficient ansatze

Hardware efficient ansatze were introduced by Farhi *et al.* for the Quantum Approximate Optimisation Algorithm (QAOA) [117], which is very similar to the VQE. These ansatze rely on short circuits, and utilise a limited selection of gates that are easy to implement on the available hardware. As such, they are well suited to the quantum computers currently available, which have short coherence times and constrained gate topologies. However hardware-efficient ansatze are unlikely to be suitable for large molecules, as they take into account no details of the chemical system being simulated. One such hardware-efficient ansatz was used to find the ground state energies of several small molecules by Kandala *et al.* [25].

Recent work [118] has shown that using hardware efficient ansatze with random initial parameters leads to several problems. The trial states produced will cluster on a 'barren plateau' in Hilbert space, with energy value close to the average of a totally mixed state. The gradient is also zero among most directions of this space, making classical optimisation extremely difficult. These effects become exponentially more prominent as the number of qubits and circuit depth increases. This suggests that randomly initialised hardware efficient ansatze are not a scalable solution for problems in quantum computational chemistry.

However, there has been recent work seeking to improve hardware efficient ansatze [119], by introducing modifications to the circuit used in Ref. [25]. The two qubit entangling gates were replaced with particle number conserving entangling gates, which are natural to implement on certain hardware [119]. This enables the hardware efficient ansatz to be used with an initial Hartree–Fock state, avoiding the barren plateau problem described above.

b. Chemically inspired ansatze

Chemically inspired ansatze result from adapting classical computational chemistry algorithms to run efficiently on quantum computers. These ansatze prepare a trial state by considering the details of the chemical system under investigation. Most notably, the coupled cluster (CC) method discussed in Sec. III C has been extended to produce the unitary coupled cluster (UCC) ansatz. The UCC method creates a parametrised trial state by considering excitations above the initial reference state, and can be written as

$$U(\vec{\theta}) = e^{T - T^\dagger}, \tag{56}$$

where $T = \sum_i T_i$, and

$$\begin{aligned} T_1 &= \sum_{i \in virt, \alpha \in occ} t_{i\alpha} a_i^\dagger a_\alpha, \\ T_2 &= \sum_{i,j \in virt, \alpha,\beta \in occ} t_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta, \\ &\dots \end{aligned} \tag{57}$$

and *occ* are occupied orbitals in the Hartree–Fock state, and *virt* are orbitals that are initially unoccupied in

the Hartree–Fock state. The UCC method is intractable on classical computers, as the operator product $\left(e^{T^\dagger - T} H e^{T - T^\dagger}\right)$ requires an infinite expansion of the exponential terms to converge [71]. However, the UCC operator is unitary, as $T - T^\dagger$ is anti-Hermitian. Therefore it can be decomposed into a gate sequence and implemented on a quantum computer.

The UCC method retains all of the advantages of the CC method, with the added benefits of being fully variational, and able to converge when used with multi-reference ground states [71]. Multi-reference states such as MCSCF states can be prepared on a quantum computer, and provide better trial states for systems displaying strong static correlation, or molecular excited states [75]. As with the CC method, the UCC ansatz is typically truncated at a given excitation level – usually single and double excitations (known as UCCSD).

In order to implement the UCC operator on a quantum computer, it must be converted into single and two qubit unitary gates. This is done by first applying Trotterization to Eq. (56), and then mapping the exponentiated creation and annihilation operators into Pauli operators using a suitable encoding (as discussed in Sec. IV B). As a result, the UCC ansatz requires longer gate sequences than hardware-efficient ansatze. This places greater demands on coherence times and gate fidelities. The formal gate scaling of the UCC ansatz is approximately $O(M^3 N^2)$ (where M is the number of spin-orbitals, and N is the number of electrons) when using the Jordan-Wigner mapping [71]. This assumes that a single Trotter step can be used, which appears to yield accurate results [71, 119]. The number of parameters required scales formally as $O(M^2 N^2)$. Even for simple molecules, this can quickly approach thousands of free parameters, making classical optimisation difficult.

However, it is important to note that in reality the gate scaling is typically better than the upper bound given above, as many excitations are forbidden by the symmetry point groups of molecular orbitals [28]. Moreover, we can use classically tractable methods to get initial approximations for the remaining non-zero parameters [22, 71], which makes classical optimisation easier. We will discuss the practicality of UCC in more detail in Sec. VIII.

There has been recent work seeking to improve the UCC ansatz. Ref. [119] transforms the Hamiltonian such that it only measures the energy of excitations above the Hartree-Fock state (the correlation energy). Because only the correlation energy is calculated, fewer measurements are required and classical optimisation becomes easier. Overall, simulated VQE calculations on small molecules were sped up by a factor of 2-4 [119]. Ref. [120] introduces a nested decomposition of the UCC operator, which reduces the gate complexity to $O(M^4)$ with increasing molecular size, and $O(M^3)$ for a fixed molecular size and increasing basis set size.

c. Hamiltonian variational

There are also variational ansatze that lie between the two extremes described above. One important example is the Hamiltonian variational ansatz, proposed by Wecker *et al.* [121]. This ansatz was inspired by adiabatic state preparation and the QAOA. Specifically, the system is evolved under exponentiated groups of terms in the Hamiltonian. For a general fermionic Hamiltonian, there are 3 groups, composed of diagonal terms

$$H_d = \sum_p h_p a_p^\dagger a_p + \sum_{p,q} h_{pqqp} a_p^\dagger a_p a_q^\dagger a_q, \tag{58}$$

hopping terms

$$H_h = \sum_{p,q} h_{pq} a_p^\dagger a_q + \sum_{p,q,r} h_{prrq} a_p^\dagger a_q a_r^\dagger a_r, \tag{59}$$

and all other exchange terms

$$H_{ex} = \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s. \tag{60}$$

The ansatz approximates adiabatic evolution under the Hamiltonian, and is given by

$$\begin{aligned} |\psi(\vec{\theta})\rangle = \prod_{b=1}^{S} & U_{ex}\left(\frac{\theta_{ex}^b}{2}\right) U_h\left(\frac{\theta_h^b}{2}\right) \cdot \\ & U_d\left(\theta_d^b\right) U_h\left(\frac{\theta_h^b}{2}\right) U_{ex}\left(\frac{\theta_{ex}^b}{2}\right) |\psi_0\rangle, \end{aligned} \tag{61}$$

where

$$U_i = e^{i\theta_i^b H_i}, \tag{62}$$

and $|\psi_0\rangle$ is the ground state of H_d. The number of parameters in this ansatz scales linearly with the number of steps, S. Wecker *et al.* compared the performance of their ansatz to that of the UCC, and found that the UCC ansatz was able to more accurately find the ground state energies of small molecules. They attributed this to the large number of parameters in the UCC ansatz, which was comparable with the size of the Hilbert space. However, when both ansatze were applied to larger systems with stronger interactions, the Hamiltonian variational ansatz was able to obtain a higher overlap with the true ground state wavefunction, using fewer energy evaluations.

The introduction of the plane wave basis [13] has made the Hamiltonian variational algorithm more efficient to implement. This basis reduces the number of terms in the molecular Hamiltonian to $O(M^2)$. The recent work of Kivlichan *et al.* showed that for Hamiltonians of this form, we can implement Trotter steps in depth (M), using $O(M^2)$ two qubit gates [122]. This is a significant improvement over the scaling found by Wecker *et al.*, which was approximately $O(M^4)$. Similar improvements in asymptotic scaling were obtained using the nested decomposition method introduced in Ref. [120]. The

number of gates required to implement a Trotter step of the Hamiltonian is $O(M^2\log_2(M))$ with increasing molecular size, and $O(M^3)$ for fixed molecular size and increasing basis size. Notably, the nested decomposition method is applicable using the gaussian basis sets described in Sec. III D. Fewer basis functions (and therefore qubits) are needed to describe molecules using a gaussian basis set than using a plane wave basis set.

d. Low depth circuit ansatz

The recently proposed low depth circuit ansatz (LDCA) [116] is motivated by Bogoliubov unitary coupled cluster (BUCC) theory, which is a variational method employing a transformation to quasiparticle operators. BUCC is suitable for a more general Hamiltonian than Eq. (25), potentially including pairing terms (superconductivity) or three body terms (nuclear physics). It can be truncated at single and double excitations (BUCCSD). It is important to distinguish between the Bogoliubov transform and the BUCCSD ansatz. The Bogoliubov transform implements a change of basis of the fermionic modes, and can be implemented in depth $O(M)$. In contrast, the BUCCSD ansatz creates a variational trial state, with approximately $O(M^4)$ free parameters to optimise, requiring a gate sequence of approximately $O(M^5)$. The LDCA was created by augmenting a BUCCS (i.e. only single excitations) ansatz with additional nearest-neighbor phase couplings to mimic the effects of the quartic terms in a BUCCSD ansatz. The LDCA consists of a set of single qubit Z rotations, followed by a series of paired rotation gates

$$K_{\alpha\beta} = R_{\alpha\beta}^{-YX} R_{\alpha\beta}^{XY} R_{\alpha\beta}^{ZZ} R_{\alpha\beta}^{-YY} R_{\alpha\beta}^{XX}, \tag{63}$$

where

$$R_{\alpha\beta}^{\pm AB} = \exp(\pm i\theta_{\alpha\beta}^{AB} \sigma_A^\alpha \otimes \sigma_B^\beta), \tag{64}$$

in which α, β indicate which qubits the operators act on, and A, B denote the type of Pauli operator applied (X, Y or Z). The operators $K_{\alpha\beta}$ are applied first between the even neighbouring pairs of qubits, then between the odd neighbouring pairs. This sequence is repeated to generate a suitably flexible trial state, with different parameters for every $R_{\alpha\beta}^{\pm AB}$. The inverse Bogliubov transform is then applied in order to reduce the number of measurements required. A circuit diagram for the LDCA can be found in Ref. [116]. This circuit has a depth $O(M)$, and the number of cycles can be increased until the best estimate for the ground state energy is obtained; which occurs when cumulative gate errors dominate the precision on the result. In this sense, the ansatz can be considered both hardware-efficient and chemically motivated.

The LDCA was applied to the small molecule cyclobutadiene (C_4H_4), and was found to recover the exact ground state after only 2 cycles, outperforming UCCSD. It is noted that 2-cycle LDCA has a parameter space larger than the problem Hilbert space, so it is perhaps unsurprising that it provides a good estimate for the ground state energy. However, this is also true for UCCSD, so LDCA's superior performance suggests it may be able to create flexible trial states for chemical problems.

3. Classical optimisation

As discussed above, classical optimisation is a crucial aspect of the VQE. However, finding the global minimum of a complicated function, in a high dimensional parameter space, is in general very difficult. Classical optimisation routines must be both fast and accurate. They also need to be robust to stochastic noise, which will be significant in near-term quantum computers.

Classical optimisation algorithms can be broadly divided into two classes; direct search and gradient based methods. Direct search algorithms do not make use of the gradient of the objective function, and include: particle swarm optimisation, Nelder-Mead simplex, and Powell's conjugate direction algorithm. Gradient based methods use the gradient of the objective function in order to determine how to update the parameters. They include: gradient descent, the simultaneous perturbation stochastic approximation (SPSA) algorithm, and L-BFGS-B. Direct search algorithms are considered more robust to noise than gradient based methods, but may require more function evaluations [71, 116, 123].

In this section we summarise the results of previous investigations into classical optimisation algorithms used in quantum computational chemistry. We also discuss methods to assist the classical optimisation procedure.

a. Previous optimisation studies

Experimental VQE implementations are limited to small molecules by the number of qubits available. As a result, the parameter space to optimise over is relatively small, so previous results may not be indicative of how these optimisation algorithms will perform for large problems. However, these studies are able to demonstrate which methods cope best with the high noise rates of current hardware.

The original implementation of the VQE, by Peruzzo *et al.*, used the Nelder-Mead simplex method. They found this derivative-free algorithm to be superior to gradient descent, which was unable to converge to the ground state energy due to high noise rates [20]. Nelder-Mead was also used successfully [23] and partially successfully [28] in trapped ion implementations of the VQE. In the latter case, while Nelder-Mead was able to successfully find the ground state of the Hydrogen molecule, it became trapped in local minima when applied to Lithium Hydride. In order to overcome this, a hybrid algorithm combining Nelder-Mead with simulated annealing was used, which gave a better estimate of the ground state energy. Derivative-free methods were also used in Refs [26, 27], which used

VQE based methods to calculate both the ground and excited states of small molecules. Both works used particle swarm optimisation, due to its resilience to noise and avoidance of local minima. The first successful experimental use of a derivative based method in the VQE was by Kandala *et al.* [25]. They used the SPSA algorithm, because of its purported resilience to noise and low number of required function evaluations [124].

To date, there have been several numerical studies comparing different optimisation algorithms for the VQE. McClean *et al.* compared four direct search algorithms: Nelder-Mead simplex, and TOMLAB's glcCluster, LGO, and multiMin algorithms [16]. Their numerical simulations included measurement noise to determine the suitability of the algorithms for stochastic objective functions. They found LGO to have the fastest convergence (up to 1000 times faster than simplex), and glcCluster to converge to the most accurate result, especially as the measurement precision was increased.

Romero *et al.* also compared four optimisation algorithms: Nelder-Mead simplex, Powell, COBYLA, and L-BFGS-B (L-BFGS-B is the only gradient based algorithm of the four) [71]. They found that L-BFGS-B outperformed the other algorithms, converging most rapidly, and often to the lowest energy value. Simplex performed very poorly, often not converging, and typically achieving the worst energy estimate of the four algorithms when it did converge. The authors found that they were able to markedly improve the performance of the algorithms by using a Moller-Plesset guess for the UCC excitation amplitudes. These simulations neglected both shot and gate noise, supporting the claim that gradient based methods converge more rapidly than direct search algorithms in low noise environments.

Heuristic algorithms have also been used in numerical studies of the VQE. Wecker *et al.* introduced the 'global variational' method for optimisation, alongside their Hamiltonian variational ansatz [121]. They found that their algorithm converged quickly, but was susceptible to becoming trapped in local minima. Moreover, they noted that simulations with a greater number of parameters became trapped more often. They attributed this to motion in the 'wrong' direction being more likely in larger parameter spaces.

b. *Related methods of optimisation*

Methods which aid classical optimisation, but that are not optimisation algorithms in their own right, have also been proposed. We discuss these methods in the following section.

Quantum gradient finding

Quantum circuits have been proposed which calculate the analytic gradient of the energy with respect to one of the parameters [71, 116]. This avoids the use of finite difference formulae, which restrict the accuracy of gradient evaluation, as the finite difference considered is limited by the noise in the energy evaluation. The quantum gradient method makes use of the differentiability of parametrised unitary operators. Parametrised unitaries can be written as exponentials of the parameters and an anti-Hermitian operator, which are simple to differentiate. A circuit to obtain the gradient of a toy VQE simulation is shown in Fig. 7.

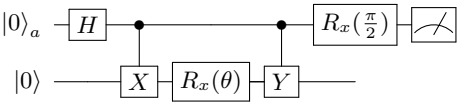

FIG. 7. A quantum circuit to calculate the gradient of a toy VQE simulation. In this toy problem, the ansatz used is $|\psi\rangle = R_x(\theta)\,|0\rangle$, and the Hamiltonian is $H = Y$. The energy is given by $E(\theta) = \langle\psi(\theta)|\,H\,|\psi(\theta)\rangle = \langle 0|\,R_x^\dagger(\theta)YR_x(\theta)\,|0\rangle$. The energy gradient, $\frac{\partial E}{\partial\theta} = \frac{i}{2}(\langle 0|\,XR_x^\dagger(\theta)YR_x(\theta)\,|0\rangle - \langle 0|\,R_x^\dagger(\theta)YR_x(\theta)X\,|0\rangle)$. This is the expectation value generated by the circuit above.

Annealed and morphed Hamiltonians

Several works have used concepts from adiabatic quantum computing to aid the classical optimisation procedure. Wecker *et al.* proposed an 'annealed variational' method alongside their Hamiltonian variational ansatz [121]. They first decompose the molecular Hamiltonian of interest into

$$H_s = H_0 + sH_1, \tag{65}$$

where H_0 is a Hamiltonian that is easy to solve, and H_1 is a difficult Hamiltonian to solve. When $s = 1$, the Hamiltonian is equivalent to the problem Hamiltonian. As described above, the Hamiltonian variational ansatz is comprised of a number of steps (blocks of gates, see Eq. (61)). The annealed variational method works by considering the S steps as separate, distinct problems. The input state to the first step is the ground state of $H_{s=0}$. The first step uses the Hamiltonian variational method to find the solution of $H_{s=1/S}$. This state is then the input into step 2. Step 2 targets the ground state of $H_{s=2/S}$. This process is repeated until the final step, which takes the ground state of $H_{s=(S-1)/S}$ as its input, and targets the ground state of $H_{s=1}$. All of these steps are then combined, and used as the starting point for the standard Hamiltonian variational approach, as described above. This procedure was useful for avoiding local minima. A similar technique has recently been proposed in Ref. [125].

Variational imaginary time evolution

Imaginary time evolution under a Hamiltonian, H, is defined by $|\psi(\tau)\rangle = e^{-H\tau}\,|\psi(0)\rangle$. If the initial state has a nonzero overlap with the ground state, the system deterministically propagates to the ground state as $\tau \to \infty$. As imaginary time evolution is a dissipative process, it

cannot be directly simulated with a unitary quantum circuit. However, a variational approach has recently been proposed which simulates imaginary time evolution on a quantum computer [126]. When the ansatz used is sufficiently powerful, imaginary time evolution is able to avoid local minima, and converge to the ground state of the system.

C. Evaluation of excited states

In this section, we discuss methods used to evaluate the excited states of molecular Hamiltonians.

1. *Witness-assisted variational eigenspectra solver*

The witness-assisted variational eigenspectra solver (WAVES) [27] combines the variational method with phase estimation to find the excited states of Hamiltonians. The method works as follows. We first use the circuit shown in Fig. 8 alongside the VQE, to minimise a cost function which depends on both the energy of the register qubits, E, and the Von-Neumann entropy, S, of the ancilla qubit.

$$F_{\text{cost}} = E(\vec{\theta}) + aS(\vec{\theta}), \tag{66}$$

where a is a constant which determines the trade-off between minimising energy and entropy. Minimising the entropy forces the system into an eigenstate of the Hamiltonian, H_s.

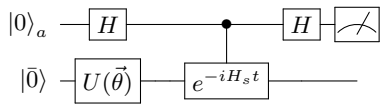

FIG. 8. The initial circuit used in the WAVES protocol, which forces the system into an eigenstate of the Hamiltonian, H_s.

We then apply the circuit shown in Fig. 9, where the state $U_e U(\vec{\theta}_{\text{min}})\,|0\rangle$ approximates an excited state of the Hamiltonian. By varying $\vec{\theta}$, we can minimise the cost function, and thus obtain an appproximation of an excited energy eigenstate. We can then use phase estimation to force the system to collapse to the true energy eigenstate, and discover its eigenvalue.

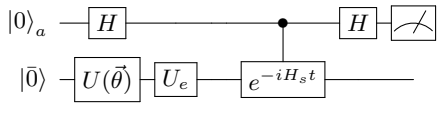

FIG. 9. The second circuit used in the WAVES protocol. The operator U_e excites the system from its ground state to being close to an excited state.

The WAVES method has been experimentally demonstrated using a silicon photonics device [27]. However, it is unlikely to be suitable for near-term hardware, due to the need to realise the time evolution operator, e^{-iHt} and phase estimation. Moreover, it is necessary to use an operator U_e which closely approximates an excitation from the ground state to the desired excited state, which may be difficult to determine *a priori*.

2. *The SWAP-test method*

In this section, we discuss a recently proposed method which makes use of variational algorithms, and the SWAP test to find excited states [127, 128].

Given the ground state $|E_0\rangle$ of a Hamiltonian H, we replace the Hamiltonian with

$$H' = H + \alpha\,|E_0\rangle\,\langle E_0|\,, \tag{67}$$

where α is chosen to be sufficiently large compared to the energy of the Hamiltonian. The ground state of the updated Hamiltonian H' is no longer $|E_0\rangle$, but the first excited state $|E_1\rangle$ of the original Hamiltonian H. This process can be repeated to obtain higher energy eigenstates. The energy of the updated Hamiltonian, $\langle\psi(\vec{\theta})|\,H'\,|\psi(\vec{\theta})\rangle = \langle\psi(\vec{\theta})|\,H\,|\psi(\vec{\theta})\rangle + \alpha\,\langle\psi(\vec{\theta})|E_0\rangle\,\langle E_0|\psi(\vec{\theta})\rangle$ can be obtained by measuring each term separately. We can measure the first term using the Hamiltonian averaging procedure described in the previous section. The second term can be calculated using the SWAP test, which has recently been modified to use a more shallow circuit [129, 130].

As this method uses only low depth circuits, has the potential for error mitigation [127, 128] and uses much of the machinery underlying the VQE, it is suitable for use with near-future hardware.

3. *The folded spectrum method*

The folded spectrum method can be used for finding the excited states of molecular Hamiltonians [16]. By replacing the Hamiltonian H with $(H-\alpha I)^2$, the ground state we obtain becomes the eigenstate with eigenvalue closest to α. Gradually changing α allows us to find the energy spectrum of the Hamiltonian H. When α is equal to an eigenvalue of H, the minimum expectation value of $(H-\alpha I)^2$ is 0, at which point the trial state $|\psi(\vec{\theta})\rangle$ is the eigenstate with eigenvalue α. We note that the folded spectrum method may require many iterations in order to locate an eigenstate. Moreover, as the energy gaps are not known *a priori*, it may be difficult to choose α in such a way that eigenstates are not missed. Furthermore, measuring the energy of the operator H^2 is notably more costly than measuring H, which has already been described as prohibitively expensive [121]. As such, while the folded spectrum method could be used on near-term hardware, it is likely far too expensive to be practical.

4. Quantum subspace expansion

The quantum subspace expansion method uses a polynomial number of additional measurements to find the excited states of a quantum system [131]. The motivation for this method is that the 3- and 4-particle reduced density matrices (RDM), which can be used to find the excited eigenstates, can be recovered by expanding the ground state in a subspace.

The authors consider the linear response subspace around the fermionic ground state. This subspace is spanned by the states $a_i^\dagger a_j \left|E_0\right\rangle$ for all possible i, j. This is designed to target the low-lying excited states, which are assumed to differ from the ground state by a small number of excitations.

The excited states can be found by solving a generalised eigenvalue problem in fermionic Fock space

$$H^{\mathrm{QSE}} C = S^{\mathrm{QSE}} C E, \tag{68}$$

with eigenvectors C, and a diagonal matrix of eigenvalues E. The Hamiltonian projected into the subspace is given by

$$H^{\mathrm{QSE}}_{ij,kl} = \left\langle E_0\right| a_i a_j^\dagger H a_k^\dagger a_l \left|E_0\right\rangle . \tag{69}$$

The overlap matrix, required because the subspace states are not orthogonal to each other, is given by

$$S^{\mathrm{QSE}}_{ij,kl} = \left\langle E_0\right| a_i a_j^\dagger a_k^\dagger a_l \left|E_0\right\rangle . \tag{70}$$

We provide more information on the QSE method in Sec. VI, where we discuss how it can be used to mitigate the effects of errors.

VI. ERROR MITIGATION FOR CHEMISTRY

All of the algorithms discussed thus far have ignored the occurrence of errors in our quantum hardware. If these errors are not dealt with, they will corrupt the results of our algorithms, rendering the calculations meaningless [132]. Circuits with a large number of gates can only be protected using error correction. The main techniques of error correction are well understood, and are discussed in detail in Refs. [48, 133, 134]. However, as discussed in Sec. II, error correction requires a large qubit overhead that is beyond the reach of current quantum computers.

New methods have recently been developed which seek to mitigate errors, rather than correct them [67, 131, 135–139]. These techniques are only applicable for low depth circuits. However, the additional resources required are much more modest than for full error correction. In general, these techniques only require a multiplicative overhead in the number of measurements required, if the error rate is sufficiently low. In this section, we review the extrapolation [67, 135, 136], probabilistic error cancellation [135, 136], quantum subspace expansion [131], QVECTOR [137] and stabiliser [138, 139] methods of error mitigation.

As we are dealing with errors, it becomes necessary to consider mixed states, rather than just pure states. As such, we now switch to the density matrix formalism of quantum mechanics [32].

We consider a quantum circuit that consists of N unitary gates applied to an initial reference state $\left|\bar{0}\right\rangle$. The output state if errors do not occur is given by

$$\rho_0 = \mathcal{U}_N \ldots \mathcal{U}_2 \mathcal{U}_1 (\left|\bar{0}\right\rangle \left\langle\bar{0}\right|). \tag{71}$$

We extract information from the circuit by measuring a Hermitian observable, O

$$\bar{O}_0 = \mathrm{Tr}[\rho_0 O] \tag{72}$$

If each gate is affected by a noise channel $\mathcal{N}_i$, the prepared state becomes

$$\rho = \prod_i \mathcal{N}_i (\mathcal{U}_i (\left|\bar{0}\right\rangle \left\langle\bar{0}\right|)), \tag{73}$$

and the measurement result becomes $\bar{O} = \mathrm{Tr}[\rho O]$. In practice, we cannot recover the noiseless state ρ_0 from the noisy state ρ without error correction. However, the error mitigation methods discussed below can approximate the noiseless measurement result $\bar{O}_0$ from the noisy measurement result, $\bar{O}$ when the error rate is sufficiently low.

A. Error suppression in the VQE

Even without additional error mitigation techniques, the VQE is inherently robust to coherent errors, such as qubit over-rotation [16].

Suppose the ansatz for the VQE is described by the operator $U(\vec{\theta})$. Due to the effect of noise, the actual operation is $\tilde{U}(\vec{\theta})$. If there exists a parameter set $(\vec{\theta} + \vec{\alpha})$ such that $||U(\vec{\theta}) - \tilde{U}(\vec{\theta} + \vec{\alpha})|| < \epsilon$ for sufficiently small $\epsilon > 0$, the classical optimiser will converge to $(\vec{\theta_{\min}} + \vec{\alpha})$, and we recover the ground state energy. The resilience of the VQE to coherent noise was observed in Ref. [22].

However, if a coherent error changes a conserved quantity (such as electron number), the argument discussed above cannot be applied. However, this problem can be resolved by optimising the modified cost function

$$L = H + \sum_j \beta_j (Q_j - q_j I)^2, \tag{74}$$

where $\{Q_j\}$ is the set of the operators for conserved quantities, q_j is the corresponding ideal expectation value of Q_j, and β_j is the penalty coefficient which should be sufficiently large.

B. Extrapolation

The extrapolation method [67, 135, 136] works by intentionally increasing the error rate, and inferring the error free result by extrapolation. The technique is based on Richardson extrapolation [140], which to first order corresponds to linear extrapolation using two points. Exponential extrapolation was introduced [136] as a more appropriate extrapolation technique for large quantum circuits. A comparison between the two extrapolation methods is shown in Fig. 10.

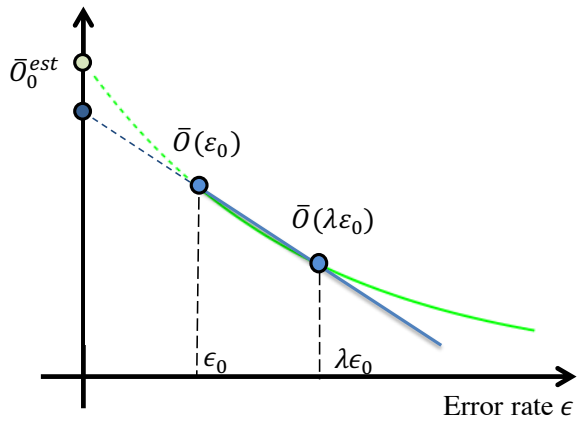

FIG. 10. A comparison of linear and exponential extrapolation. The horizontal axis is the error rate of each gate and the vertical axis is the expectation value of the measured observable, $\bar{O}$.

1. Richardson extrapolation

As an example of how Richardson extrapolation can be used in a quantum circuit, we consider a stochastic noise model

$$\mathcal{N}_i = (1-\epsilon)\mathcal{I} + \epsilon\mathcal{E} \tag{75}$$

where ϵ is the probability of the noise channel $\mathcal{E}$ occurring. In general, we can make use of the error twirling technique to convert an arbitrary noise channel to a stochastic noise channel [67]. Under our noise model, the state produced by the circuit is

$$\rho(\epsilon_0 \cdots \epsilon_n) = \prod_i ((1-\epsilon_i)\mathcal{I} + \epsilon_i\mathcal{E})\mathcal{U}_i(|\bar{0}\rangle\langle\bar{0}|). \tag{76}$$

If we can proportionally increase the noise rate ϵ_i to $\lambda\epsilon_i$, then we can prepare a more noisy state

$$\rho(\lambda(\epsilon_0 \cdots \epsilon_n)) = \prod_i ((1-\lambda\epsilon_i)\mathcal{I} + \lambda\epsilon_i\mathcal{E})\mathcal{U}_i(|\bar{0}\rangle\langle\bar{0}|). \tag{77}$$

Richardson extrapolation uses measurements at several values of λ to estimate the value of the observable in the presence of noise. We can increase the error rate using the techniques described in Refs. [29, 67].

2. Linear extrapolation

For linear extrapolation, we can either use a first order Richardson extrapolation (which uses two points), or take a linear best fit with several points. For the former case, the estimated value of the observable is given by

$$\bar{O}_0^{\text{est}} = \frac{\lambda\bar{O}(\epsilon_0 \cdots \epsilon_n)) - \bar{O}(\lambda(\epsilon_0 \cdots \epsilon_n)))}{\lambda - 1}. \tag{78}$$

If we consider two qubit gates to be the dominant source of errors in the circuit, and assume that each two qubit gate has the same error rate, ϵ, then we can approximate the total noise rate in the circuit by $n_2\epsilon$ where n_2 is the number of two qubit gates.

While this method can significantly improve the accuracy of our calculations, it requires additional measurements in order to keep the variance of our measurements the same as the non-extrapolated case.

The linear extrapolation method has recently been experimentally demonstrated. Kandala *et al.* dramatically improved the accuracy of their VQE experiments on the molecules H_2 and LiH by using the linear extrapolation method, achieving results close to chemical accuracy [29].

3. Exponential extrapolation

The linear extrapolation method only works well when the total noise rate $n_2\epsilon$ is much less than one [136]. However, the performance of the extrapolation method can be improved by considering an exponential extrapolation. The justification for this is as follows. When our error model is stochastic, and the error rate is the same for each gate, the noisy measurement result can be expanded according to the number of errors which occur in the circuit,

$$\bar{O} = \text{Tr}[\rho O] = p_0\bar{O}_0 + p_1\bar{O}_1 + p_2\bar{O}_2 + \dots, \tag{79}$$

where $p_k = \binom{n_2}{k}\epsilon^k(1-\epsilon)^{n_2-k}$ is the probability of having k errors and $\bar{O}_k$ is the corresponding measurement result. When n_2 is large and $n_2\epsilon \approx 1$, the binomial distribution converges to the Poisson distribution

$$p_k = \binom{n_2}{k}\epsilon^k(1-\epsilon)^{n_2-k} \approx e^{-n_2\epsilon}\frac{(n_2\epsilon)^k}{k!}. \tag{80}$$

Thus, the noisy measurement result decays exponentially with increasing error rate,

$$\bar{O} \approx e^{-n_2\epsilon}(\bar{O}_0 + (n_2\epsilon)\bar{O}_1 + \frac{(n_2\epsilon)^2}{2}\bar{O}_2 + \dots), \tag{81}$$

and exponential fitting is more appropriate than polynomial fitting, as confirmed by the numerical experiments of Ref. [136].

C. Probabilistic error cancellation

The probabilistic error cancellation method [135, 136] works by effectively realising the inverse of the error channel, $\mathcal{N}^{-1}$, such that $\mathcal{N}^{-1}(\mathcal{N}(\rho)) = \rho$. However, as realising the inverse channel is an unphysical process, we use the scheme depicted in Fig. 11 to effectively realise the inverse channel by focusing only on measurement results.

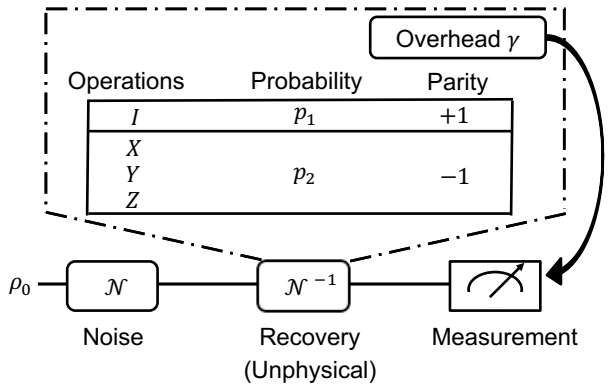

FIG. 11. A schematic of the probabilistic error cancellation method for a depolarising error resulting from a single qubit gate. After the gate is applied, there is a noise channel $\mathcal{N}$. The method works by effectively realising the inverse channel $\mathcal{N}^{-1}$. This is achieved by randomly applying one of the X, Y or Z operators with probability p_2, or the identity gate with p_1. The expectation values resulting from the circuits are combined. If one of the Pauli matrices was applied to realise the inverse channel, the resulting expectation value is subtracted, rather than added (parity -1). The overhead γ determines the number of additional measurements required to keep the variance of the error mitigated result equal to the variance of the noisy result. This can be generalised to multi-qubit gates as described in the main text.

As an example, we consider the case of a depolarising error channel,

$$\mathcal{D}(\rho_0) = \left(1 - \frac{3}{4}p\right)\rho_0 + \frac{p}{4}(X\rho_0 X + Y\rho_0 Y + Z\rho_0 Z). \quad (82)$$

The unphysical inverse channel is

$$\begin{aligned}
&\mathcal{D}^{-1}(\rho) \\
&= \left(1 + \frac{3p}{4(1-p)}\right)\rho - \frac{p}{4(1-p)}(X\rho X + Y\rho Y + Z\rho Z), \\
&= \gamma[p_1\rho - p_2(X\rho X + Y\rho Y + Z\rho Z)],
\end{aligned} \quad (83)$$

where, the overhead coefficient $\gamma = (p+2)/(2-2p) > 1$, $p_1 = (4-p)/(2p+4)$, and $p_2 = p/(2p+4)$ in this case. The overhead coefficient, γ, increases the variance of O_0. For n gates, the variance of O_0 is proportional to γ^n.

We cannot directly realise $\mathcal{N}^{-1}$ due to the minus sign before p_2. However, we can consider and correct its effect on the expectation value. Suppose the expectation value is O, then the corrected measurement outcome is

$$\begin{aligned}
O_0 &= \mathrm{Tr}[O\mathcal{N}^{-1}(\rho)], \\
&= \gamma[p_1 \langle O\rangle_\rho - p_2(\langle O\rangle_{X\rho X} + \langle O\rangle_{Y\rho Y} + \langle O\rangle_{Z\rho Z})], \\
&= \gamma[p_1 \langle O\rangle_\rho - p_2(\langle XOX\rangle_\rho + \langle YOY\rangle_\rho + \langle ZOZ\rangle_\rho)],
\end{aligned} \quad (84)$$

where $\langle O\rangle_\rho = \mathrm{Tr}[O\rho]$. We can therefore measure O, XOX, YOY, ZOZ, and linearly combine the measurement results to effectively realise the inverse channel, thus obtaining the noiseless measurement result O_0.

In practice, it is not possible to exactly measure all of the possible terms resulting from errors if there are many gates in the circuit. Instead, we can consider only the most important terms, which result from a small number of errors occurring. If the error rate is low, then the other terms can be considered negligibly small. After each single qubit gate, we can apply X, Y or Z operators with probability p_2, or the identity gate with p_1. We repeat that circuit variant many times to extract the expectation value, and multiply the expectation value by $(-1)^{N_p}$, where N_p is the number of additional X, Y or Z gates that were applied in that circuit iteration. We then sum up the values for several circuit variants and multiply by γ to obtain the error mitigated result. For two qubit gates in the depolarising noise model, after each two qubit gate we insert the gates: XI, IX, YI, IY, ZI, IZ (parity -1) with probability p_2, the gates $XX, YY, ZZ, XY, YX, XZ, ZX, YZ, ZY$ (parity $+1$) with probability p_2 and II (parity $+1$) with probability p_1.

The probabilistic error cancellation has been shown to work for general Markovian noise, but is not suitable for correlated errors [136]. We note that the probabilistic error cancellation method requires full knowledge of the noise model associated with each gate. This can be obtained from either process tomography, or a combination of process and gate set tomography. The latter approach reduces errors due to state preparation and measurement [136]. As the probabilistic error cancellation method requires tomography, it is practically more difficult to realise than the other error mitigation methods described in this section.

D. Quantum subspace expansion

The quantum subspace expansion [131] described in Sec. V C can mitigate errors in the VQE, in addition to calculating the excited energy eigenstates. This method can suppress stochastic errors such as depolarising and dephasing errors. Suppose that after the VQE, an approximate ground state $|\tilde{E}_0\rangle$ has been discovered. Such a state may deviate from the true ground state $|E_0\rangle$ due to errors in the whole process. For example, when $|\tilde{E}_0\rangle = X_1|E_0\rangle$, we can simply apply an X_1 gate to recover the correct ground state.

However, as we do not know which errors have occurred, we can instead consider an expansion in the subspace $\{|P_i\tilde{E}_0\rangle\}$, where P_i are matrices belonging to the Pauli group. Then, one can measure the matrix representation of the Hamiltonian in the subspace,

$$H_{ij} = \langle\tilde{E}_0| P_i H P_j |\tilde{E}_0\rangle . \tag{85}$$

As the subspace states are not orthogonal to each other, we should also measure the overlap matrix

$$S_{ij} = \langle\tilde{E}_0| P_i P_j |\tilde{E}_0\rangle . \tag{86}$$

By solving the generalised eigenvalue problem

$$HC = SCE, \tag{87}$$

with eigenvectors C and diagonal matrix of eigenvalues E, we can get the error mitigated spectrum of the Hamiltonian. In general, when the chosen subspace $\{|P_i\tilde{E}_0\rangle\}$ can represent the full Hilbert space, we can recover the error free spectrum of the Hamiltonian. This assumes that there is no error in the measurement of the two matrices H_{ij} and S_{ij}. However, if $\{|P_i\tilde{E}_0\rangle\}$ represents the full Hilbert space, we would need to measure an exponential number of terms. As such, we generally consider a limited number of Pauli group matrices P_i.

The quantum subspace expansion technique has been experimentally demonstrated, using a two qubit superconducting system to measure the ground and excited state energies of H_2. [26]. Using the subspace $\{|\sigma_j\tilde{E}_0\rangle\} = \{|\tilde{E}_0\rangle, X|\tilde{E}_0\rangle, Y|\tilde{E}_0\rangle, Z|\tilde{E}_0\rangle\}$, the spectrum of hydrogen was calculated to approximately chemical accuracy.

An extension of the QSE was recently developed, which considers the case where the excitation operators P_i are symmetry operators which commute with the Hamiltonian [139].

E. QVECTOR

The quantum variational error corrector (QVECTOR) [137] uses a hybrid quantum-classical algorithm to construct device-tailored error corrected quantum memories, in order to dramatically reduce the number of physical qubits needed to encode a logical qubit. As shown in Fig. 12, suppose the k qubit state is initialised in $|\varphi\rangle = U_S|0\rangle^{\otimes k}$ via unitary U_S. To protect the state, it is encoded into $n > k$ qubits using $V(\vec{\theta}_1)$ and can be finally decoded using $V^\dagger(\vec{\theta}_1)$. When errors occur on the encoded state, they can be corrected using $W(\vec{\theta}_2)$.

The QVECTOR circuit is an $[n, k]$ quantum code (with k logical qubits encoded in n physical qubits) that uses r additional (refresh) qubits for the recovery process. In order to optimise the performance of the error correction circuit, the parameters $\vec{\theta}_1$ and $\vec{\theta}_2$ are varied to maximise

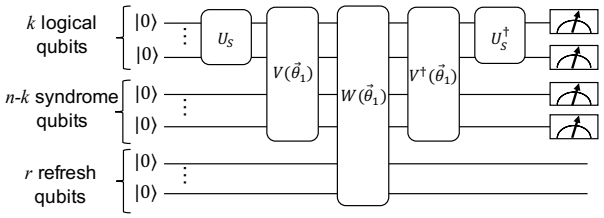

FIG. 12. The schematic of the variational circuit for QVECTOR.

the average code fidelity,

$$\mathcal{F} \equiv \int_{\varphi\in\mathcal{C}} \langle\varphi|\, \mathcal{K}_R(\vec{\theta}_1, \vec{\theta}_2)\big(\,|\varphi\rangle\langle\varphi|\,\big)\,|\varphi\rangle\, d\varphi, \tag{88}$$

where the integration is over all states in the Haar measure, and $\mathcal{K}_R(\vec{\theta}_1, \vec{\theta}_2)$ is the circuit generated by combining the encoding, decoding and recovery operations. In practice, the integration can be efficiently calculated via unitary 2-design [141] and the average code fidelity can be obtained by measuring the final state in Fig. 12 in the computational basis. For different device errors, the average code fidelity $\mathcal{F}$ can be regarded as a cost function and minimised over parameters $\vec{\theta}_1$ and $\vec{\theta}_2$ via a classical optimiser. QVECTOR can thus generate a device-tailored quantum error correction circuit. QVECTOR was found to be highly effective in simulations of small systems, enabling a 6-fold extension in the T_2 time of the system, and was also found to outperform the five qubit stabiliser code for a system experiencing amplitude and phase damping noise [137].

F. Stabiliser based methods

Recently, a new method of error mitigation has been introduced, which makes use of stabiliser checks on a suitably constructed trial state [138, 139]. A key concern for the VQE is preserving particle number, as states with electron number far from the true value appear to have a larger energy variance than those with smaller particle number errors [132]. Consequently, some ansatze, such as the UCC, Hamiltonian variational, and certain types of hardware efficient ansatze [119], construct variational trial states with the same number of electrons and total spin as the Hartree-Fock state. As such, we can perform 'checks' on quantities which should be conserved, discarding the results if the measured value is not as expected.

The most simple stabiliser measurement checks the total electron number parity. This procedure is shown in Fig. 13, and enables the detection of any error which changes the electron number parity by one. Further checks can be performed on the electron spin parities, and on the total number of electrons in the molecule [138]. Ref. [139] derived a method of performing these stabiliser

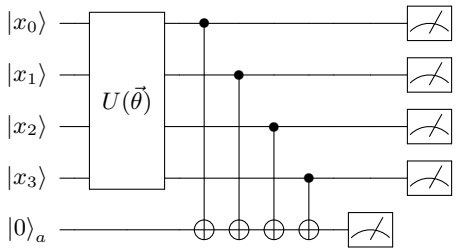

FIG. 13. The stabiliser-VQE circuit to check the electron number parity of a physical trial state. The ancilla should be measured in the state $|\frac{1}{2}(1-(-1)^N)\rangle$, where N is the total number of electrons in the molecule. If errors occur, and the measured value of the ancilla is not correct, the measurement of the Hamiltonian term $\hat{h}_i$ on the main qubit register is not performed, and the circuit is reinitialised. The qubits are initialised in $x_i = 0$ or 1 such that the register is in the mean-field state of the system.

checks efficiently for qubit arrays with limited connectivity, as well as ways to increase the number of errors detected by introducing additional symmetries.

This method of error mitigation can easily be combined with some of the other techniques mentioned above, such as the extrapolation method of error mitigation. Combining the extrapolation and stabiliser methods was shown to be considerably more effective than either method in isolation [138].

VII. ILLUSTRATIVE EXAMPLES

In this section we illustrate many of the techniques described in the previous sections of this work, by explicitly demonstrating how to map molecular ground state problems onto a quantum computer. We do this for the Hydrogen molecule (H_2) in the STO-3G, 6-31G and cc-PVDZ bases, and Lithium Hydride (LiH) in the STO-3G basis (see Sec. III D). Across these examples, we showcase the Jordan–Wigner (JW), Bravyi–Kitaev (BK) and BK tree mappings (Sec. IV B), reduction of active orbitals using the Natural Molecular Orbital (NMO) basis (Sec. III E), reduction of qubits using symmetry conservation (Sec. IV C) and the UCC ansatze (Sec. V B). These examples are designed to familiarise the reader with the key steps of formulating a quantum computational chemistry problem. These steps are applicable to both ground state and general chemical problems.

A. Hydrogen

The first quantised molecular Hamiltonian for H_2 is given by Eq. (16), with two electrons. To convert this Hamiltonian into the second quantised representation, as given by Eq. (25), we need to select a basis set. As discussed in Sec. III D, this is a discrete set of functions which are used to approximate the spin-orbitals of the molecule. By considering a larger number of orbitals, we are able to recover a larger proportion of the correlation energy in a molecule, resulting in a more accurate estimate of the ground state energy. Fig. 14 shows the H_2 ground state dissociation curves in the STO-3G, 6-31G and cc-PVDZ bases. We can see that the differences in energy between the three minima are considerably larger than chemical accuracy. This highlights that working in a suitably large basis set is crucial for obtaining accurate results.

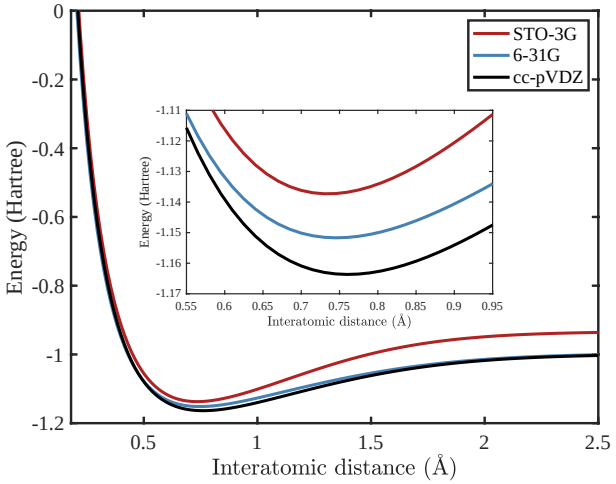

FIG. 14. Comparing the ground state dissociation curves of H_2 for a range of basis sets.

1. *STO-3G basis*

The STO-3G basis for H_2 includes only the $\{1s\}$ orbital for each hydrogen atom. The $1s$ orbital is represented by a linear combination of three Gaussian functions (GTOs). Each hydrogen atom contributes one spin-orbital, and there are two possible spins for each orbital - resulting in a total of 4 orbitals for STO-3G H_2. We denote these orbitals as

$$|1s_{A\uparrow}\rangle, |1s_{A\downarrow}\rangle, |1s_{B\uparrow}\rangle, |1s_{B\downarrow}\rangle, \tag{89}$$

where the subscript A or B denotes which of the two atoms the orbital is centred on, and the $\uparrow$ / $\downarrow$ denotes the spin of the electron in the orbital. For convenience, we work in the molecular orbital basis for H_2, which is simple to construct manually. These single electron molecular

orbitals are given by

$$\begin{aligned}|\sigma_{g\uparrow}\rangle &= \frac{1}{\sqrt{2}}(|1s_{A\uparrow}\rangle + |1s_{B\uparrow}\rangle),\\ |\sigma_{g\downarrow}\rangle &= \frac{1}{\sqrt{2}}(|1s_{A\downarrow}\rangle + |1s_{B\downarrow}\rangle),\\ |\sigma_{u\uparrow}\rangle &= \frac{1}{\sqrt{2}}(|1s_{A\uparrow}\rangle - |1s_{B\uparrow}\rangle),\\ |\sigma_{u\downarrow}\rangle &= \frac{1}{\sqrt{2}}(|1s_{A\downarrow}\rangle - |1s_{B\downarrow}\rangle).\end{aligned} \tag{90}$$

We can write a Slater determinant in the occupation number basis as

$$|\psi\rangle = |f_{\sigma_{u\downarrow}}, f_{\sigma_{u\uparrow}}, f_{\sigma_{g\downarrow}}, f_{\sigma_{g\uparrow}}\rangle\,, \tag{91}$$

where $f_i = 1$ if spin-orbital i is occupied, and $f_i = 0$ if spin-orbital i is unoccupied. We can now calculate the integrals given in Eq. (26) using these molecular orbitals. These integrals have been calculated for a large number of basis sets, and the results can be obtained by using a computational chemistry package [142–145]. The resulting second quantised Hamiltonian is [88]

$$\begin{aligned}H = {} & h_{00}a_0^\dagger a_0 + h_{11}a_1^\dagger a_1 + h_{22}a_2^\dagger a_2 + h_{33}a_3^\dagger a_3\\ & + h_{0110}a_0^\dagger a_1^\dagger a_1 a_0 + h_{2332}a_2^\dagger a_3^\dagger a_3 a_2 + h_{0330}a_0^\dagger a_3^\dagger a_3 a_0\\ & + h_{1221}a_1^\dagger a_2^\dagger a_2 a_1 + (h_{0220} - h_{0202})a_0^\dagger a_2^\dagger a_2 a_0\\ & + (h_{1331} - h_{1313})a_1^\dagger a_3^\dagger a_3 a_1 + h_{0132}(a_0^\dagger a_1^\dagger a_3 a_2 + a_2^\dagger a_3^\dagger a_1 a_0)\\ & + h_{0312}(a_0^\dagger a_3^\dagger a_1 a_2 + a_2^\dagger a_1^\dagger a_3 a_0),\end{aligned} \tag{92}$$

where the coefficients are given by the electron integrals. We must map the problem Hamiltonian from being written in terms of creation and annihilation operators, to being written in terms of qubit operators. Using the JW encoding, we can obtain the 4 qubit Hamiltonian for H_2, given by [88]

$$\begin{aligned}H = {} & h_0 I + h_1 Z_0 + h_2 Z_1 + h_3 Z_2 + h_4 Z_3 +\\ & h_5 Z_0 Z_1 + h_6 Z_0 Z_2 +\\ & h_7 Z_1 Z_2 + h_8 Z_0 Z_3 +\\ & h_9 Z_1 Z_3 + h_{10} Z_2 Z_3 +\\ & h_{11} Y_0 Y_1 X_2 X_3 + h_{12} X_0 Y_1 Y_2 X_3 +\\ & h_{13} Y_0 X_1 X_2 Y_3 + h_{14} X_0 X_1 Y_2 Y_3.\end{aligned} \tag{93}$$

While it is important to understand this procedure, every step from selecting a basis to producing an encoded qubit Hamiltonian can be carried out using a quantum computational chemistry package such as OpenFermion [146] or Qiskit Aqua [147].

In the JW encoding, it is simple to construct the Hartree-Fock (HF) state for the H_2 molecule. The HF state for H_2 is given by

$$\psi^{\mathrm{H_2}}_{\mathrm{HF}}(\mathbf{r_1}, \mathbf{r_2}) = \frac{1}{\sqrt{2}}(\sigma_{g\uparrow}(\mathbf{r_1})\sigma_{g\downarrow}(\mathbf{r_2}) - \sigma_{g\uparrow}(\mathbf{r_2})\sigma_{g\downarrow}(\mathbf{r_1})) \tag{94}$$

where $\mathbf{r_i}$ is the position of electron i. In the occupation number basis, we can write this as

$$|\psi^{\mathrm{H_2}}_{\mathrm{HF}}\rangle = |0011\rangle\,. \tag{95}$$

The most general state for H_2 (with the same spin and electron number as the HF state) is given by

$$|\psi^{\mathrm{H_2}}\rangle = \alpha\,|0011\rangle + \beta\,|1100\rangle + \gamma\,|1001\rangle + \delta\,|0110\rangle\,, \tag{96}$$

and the ground state of the H_2 molecule at its equilibrium bond distance is given by [72]

$$|\psi^{\mathrm{H_2}}_g\rangle = 0.9939\,|0011\rangle - 0.1106\,|1100\rangle\,. \tag{97}$$

The first determinant in the ground state wavefunction is the HF state for H_2, showing that a mean-field solution is a good approximation for this molecule at this interatomic distance. The second determinant represents the antibonding state, and accounts for dynamical correlation between the electrons due to their electrostatic repulsion. While the HF determinant dominates at the equilibrium separation, at large separation the two determinants contribute equally to the wavefunction. This is because the bonding and antibonding configurations become degenerate. We require both determinants to accurately describe the state, ensuring that only one electron locates around each atom. This is an example of static correlation, which can also be dealt with using multiconfigurational self-consistent field (MCSCF) methods, as described in Sec. III C 1.

As discussed previously, in order to find the ground state of the H_2 molecule (using either the VQE or PEA), we need to construct the state on the quantum computer. This can be done using adiabatic state preparation [14], or using an ansatz. Here we explicitly derive the UCCSD ansatz for H_2. As discussed in Sec. V B, the UCCSD operator we seek to realise is given by

$$U = e^{(T_1 - T_1^\dagger) + (T_2 - T_2^\dagger)}, \tag{98}$$

where

$$\begin{aligned}T_1 &= \sum_{i\in virt, \alpha\in occ} t_{i\alpha} a_i^\dagger a_\alpha,\\ T_2 &= \sum_{i,j\in virt, \alpha,\beta\in occ} t_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta,\end{aligned} \tag{99}$$

and *occ* are initially occupied orbitals in the HF state, *virt* are initially unoccupied orbitals in the HF state, and $t_{i\alpha}$ and $t_{ij\alpha\beta}$ are variational parameters to be optimised. For H_2, the only operators which don't change the spin of the molecule when acting upon the HF state are: $a_2^\dagger a_0$, $a_3^\dagger a_1$, $a_3^\dagger a_2^\dagger a_1 a_0$. Other valid operators are equivalent to these operators, and can be combined with them, such as $a_3^\dagger a_0^\dagger a_1 a_0 = -a_3^\dagger a_1$. As a result, the UCCSD operator takes the form

$$U = e^{t_{02}(a_2^\dagger a_0 - a_0^\dagger a_2) + t_{13}(a_3^\dagger a_1 - a_1^\dagger a_3) + t_{0123}(a_3^\dagger a_2^\dagger a_1 a_0 - a_0^\dagger a_1^\dagger a_2 a_3)}. \tag{100}$$

We can split this operator using Trotterization with a single Trotter step

$$\begin{aligned} U = \; & e^{t_{02}(a_2^\dagger a_0 - a_0^\dagger a_2)} \\ & \times e^{t_{13}(a_3^\dagger a_1 - a_1^\dagger a_3)} \\ & \times e^{t_{0123}(a_3^\dagger a_2^\dagger a_1 a_0 - a_0^\dagger a_1^\dagger a_2 a_3)}. \end{aligned} \tag{101}$$

Using the JW encoding, we find that

$$\begin{aligned} (a_2^\dagger a_0 - a_0^\dagger a_2) &= \frac{i}{2}(X_2 Z_1 Y_0 - Y_2 Z_1 X_0) \\ (a_3^\dagger a_1 - a_1^\dagger a_3) &= \frac{i}{2}(X_3 Z_2 Y_1 - Y_3 Z_2 X_1) \\ (a_3^\dagger a_2^\dagger a_1 a_0 - a_0^\dagger a_1^\dagger a_2 a_3) &= \\ & \frac{i}{8}(X_3 Y_2 X_1 X_0 + Y_3 X_2 X_1 X_0 + Y_3 Y_2 Y_1 X_0 + Y_3 Y_2 X_1 Y_0 \\ & - X_3 X_2 Y_1 X_0 - X_3 X_2 X_1 Y_0 - Y_3 X_2 Y_1 Y_0 - X_3 Y_2 Y_1 Y_0). \end{aligned} \tag{102}$$

It was shown in Ref. [71] that all Pauli terms arising from the same excitation operators commute. As a result, each of the exponentials in Eq. (101) can be separated into a product of exponentials of a single Pauli string. For example

$$e^{t_{02}(a_2^\dagger a_0 - a_0^\dagger a_2)} = e^{\frac{it_{02}}{2} X_2 Z_1 Y_0} \times e^{\frac{-it_{02}}{2} Y_2 Z_1 X_0}. \tag{103}$$

In Ref. [28] the UCCSD operator for H_2 was simplified by implementing the single excitation terms as basis rotations, and combining terms in the double excitation operator (by considering the effect of each term on the HF state). We note that this latter technique is only possible because there is only one double excitation operator for this molecule, and so is not a scalable technique in general. The UCCSD operator is simplified to

$$U = e^{-i\theta X_3 X_2 X_1 Y_0}. \tag{104}$$

This can be implemented using the circuit [148] shown in Fig. 15.

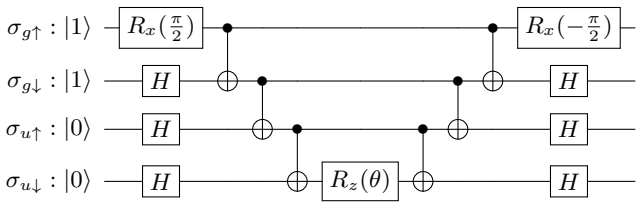

FIG. 15. The circuit for implementing the UCCSD operator for H_2 in the STO-3G basis, as given by Eq. 104. The $R_x(\frac{\pi}{2})$ and H gates rotate the basis such that the exponentiated operator applied to the corresponding qubit is either Y or X, respectively. Single excitation terms are implemented with a change of basis [28].

2. 6-31G basis

As discussed in Sec. III D, H_2 in the 6-31G basis has a double-zeta representation of the valence electrons. This means we have 8 orbitals to consider in total; $\{1s_\uparrow, 1s_\downarrow, 1s'_\uparrow, 1s'_\downarrow\}$ from each atom. Working in the canonical orbital basis, we show how to construct Bravyi-Kitaev encoded states of 6-31G H_2. The BK transform matrix for an 8 orbital system is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{105}$$

We order the orbitals such that the first $M/2$ orbitals are spin up, and the final $M/2$ orbitals are spin down. When the orbitals are ordered in this way, the 4^{th} entry in the BK encoded vector is the sum (mod 2) of the spin up occupancies, which sums to the number of spin up electrons. Moreover, the 8^{th} entry is the sum (mod 2) of all of the orbital occupanies, which sums to the number of electrons. As these quantities are conserved, we can remove these two qubits from the simulation, following the procedure of Sec. IV C. We note that if the orbitals are arranged 'up-down, up-down', then while the 8^{th} entry is still equal to the number of electrons, the 4^{th} entry no longer necessarily equal to a conserved quantity.

3. cc-PVDZ basis

As discussed in Sec. III D, the cc-PVDZ basis for H_2 includes a double-zeta representation of the valence shell, and additional polarisation orbitals. Each atom contributes $\{1s, 1s', 2p_x, 2p_y, 2p_z\}$ orbitals, resulting in 20 orbitals in total. In order to reduce our active space, we first change to the natural molecular orbital (NMO) basis, using the single particle reduced density matrix (1-RDM), as discussed in Sec. III E. There are only 10 rows and columns in the 1-RDM shown below because the spin-up and spin-down entries have been combined. The diagonal elements of the 1-RDM are the occupation numbers of the corresponding canonical orbitals (the orbitals produced from a HF calculation on H_2 in this basis). The 1-RDM obtained from a classically tractable CISD calculation (at an internuclear separation of 0.75 $\dot{A}$) on cc-PVDZ H_2 is shown in Eq. (106)

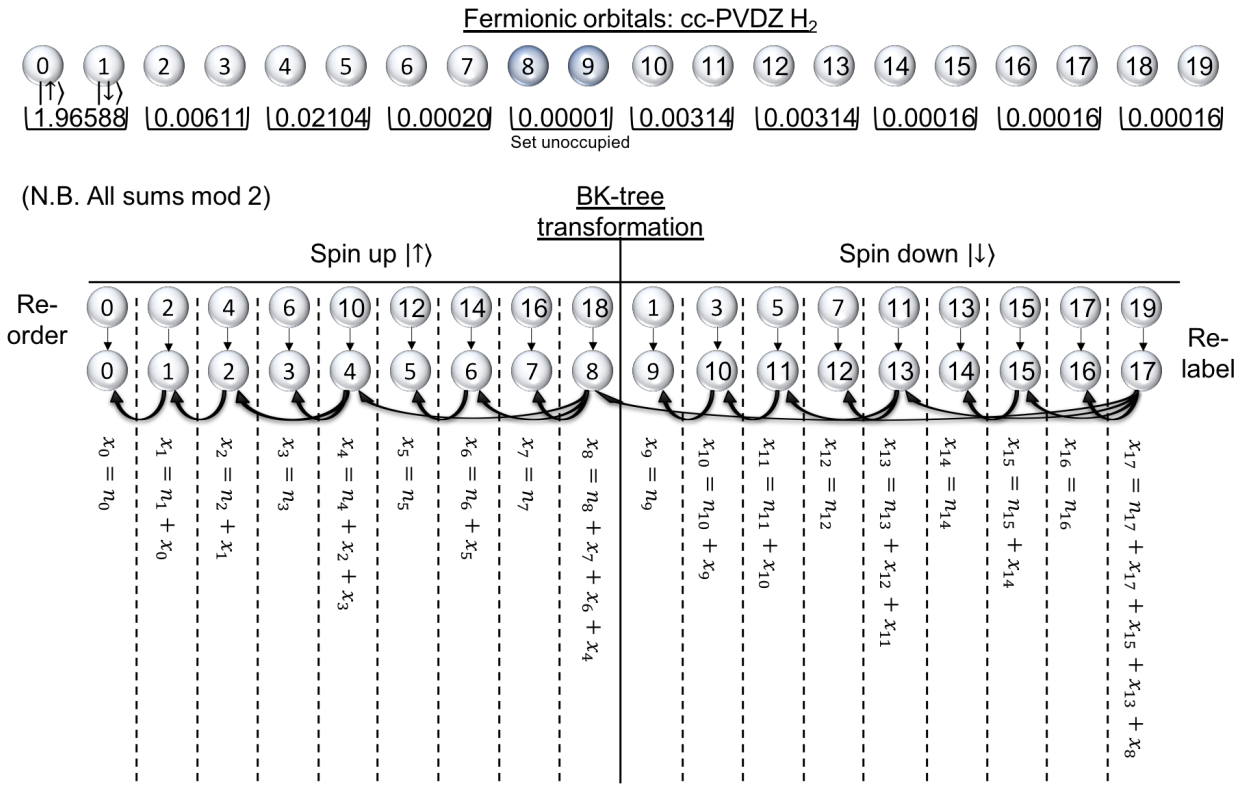

FIG. 16. A pictorial representation of the fermion-to-qubit mapping procedure for H_2 in the cc-PVDZ basis. The fermionic natural molecular orbitals (NMO) are initially arranged 'spin up, spin down, spin up, spin down, ...', and have their corresponding natural orbital occupation number (NOON) below. As the NOON of orbitals 8 and 9 is so small, they can be assumed unfilled, and removed from the Hamiltonian. We then rearrange the remaining orbitals to be 'all spin up, all spin down', and re-label them from 0 to 17. We then perform the BK-tree mapping by constructing the Fenwick tree, Fen(0,17). The value x_i is the value of the i^{th} qubit under the BK-tree mapping, while n_i is the value of the i^{th} qubit under the JW mapping. We see that qubit 8 stores the sum $\sum_{i=0}^{8} n_i$, and qubit 17 stores the sum $\sum_{i=0}^{17} n_i$. As these sums are conserved quantities, these qubits do not flip throughout the simulation, and so can be removed from the Hamiltonian as described in Sec. IV C.

$$\begin{bmatrix}
\mathbf{1.96578} & 0.00000 & \mathbf{-0.01174} & 0.00000 & 0.00000 & 0.00000 & \mathbf{-0.00844} & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & \mathbf{0.01052} & 0.00000 & \mathbf{0.01032} & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{-0.00174} \\
\mathbf{-0.01174} & 0.00000 & \mathbf{0.00553} & 0.00000 & 0.00000 & 0.00000 & \mathbf{-0.00179} & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & \mathbf{0.01032} & 0.00000 & \mathbf{0.01031} & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{-0.00183} \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00314} & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00314} & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
\mathbf{-0.00844} & 0.00000 & \mathbf{-0.00179} & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00088} & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00016} & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00016} & 0.00000 \\
0.00000 & \mathbf{-0.00174} & 0.00000 & \mathbf{-0.00183} & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00038}
\end{bmatrix}. \tag{106}$$

We perform a unitary diagonalisation of this matrix, and rotate the orbitals by the same unitary matrix. This constitutes a change of basis to the NMO's of the molecule. The diagonalised 1-RDM is shown in Eq. (107)

$$\begin{bmatrix}
\mathbf{1.96588} & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & \mathbf{0.00611} & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & \mathbf{0.02104} & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & \mathbf{0.00020} & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00001} & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00314} & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00314} & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00016} & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00016} & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00016}
\end{bmatrix}. \tag{107}$$

As discussed in Sec. III E, the diagonal entries are the natural orbital occupation numbers (NOONs). We can see that the 5^{th} orbital has a NOON that is 20 times smaller than the next smallest NOON. As a result, we consider this orbital to always be empty, and so remove all terms involving it from the Hamiltonian. This leaves a Hamiltonian acting on $M = 18$ spin-orbitals. We now map these into qubits using the BK-tree method, as M is not a power of 2, so the standard BK method will leave us unable to remove 2 qubits using symmetries. A visual representation of the fermion-to-qubit mapping process is shown in Fig. 16. We can see that the 9^{th} and 18^{th} orbitals store the number of spin up electrons and total number of electrons, respectively. As a result, they can be removed. This reduces the problem to one of 16 qubits. In Fig. 16 x_i is the value of the i^{th} qubit under the BK-tree mapping, while n_i is the value of the i^{th} qubit in the JW mapping. The HF state of cc-PVDZ H_2 has $n_0, n_8 = 1, n_{i\neq 0,8} = 0$, written as $|0000000010000001\rangle$.Using Fig. 16, we can see that the corresponding BK-tree mapped HF state is given by $|0001011100010111\rangle$.

B. Lithium Hydride STO-3G basis

For LiH in the STO-3G basis, we consider $1s, 2s, 2p_x, 2p_y, 2p_z$ functions for lithium, and a single $1s$ orbital for hydrogen. This gives a total of 12 spin-orbitals, when spin degeneracy is included. We can reduce this problem to one of six qubits, following a similar procedure to that of cc-PVDZ H_2 above, visually illustrated in Fig. 17. The canonical orbital 1-RDM from a CISD calculation on LiH (at an internuclear separation of 1.45 $\dot{A}$) is given by

$$\begin{bmatrix}
\mathbf{1.9999} & \mathbf{-0.0005} & \mathbf{0.0006} & 0.0000 & 0.0000 & \mathbf{-0.0010} \\
\mathbf{-0.0005} & \mathbf{1.9598} & \mathbf{0.0668} & 0.0000 & 0.0000 & \mathbf{0.0084} \\
\mathbf{0.0006} & \mathbf{0.0668} & \mathbf{0.0097} & 0.0000 & 0.0000 & \mathbf{-0.0138} \\
0.0000 & 0.0000 & 0.0000 & \mathbf{0.0017} & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & \mathbf{0.0017} & 0.0000 \\
\mathbf{-0.0010} & \mathbf{0.0084} & \mathbf{-0.0138} & 0.0000 & 0.0000 & \mathbf{0.0273}
\end{bmatrix}. \tag{108}$$

We diagonalise this 1-RDM, moving to the NMO basis. The NMO 1-RDM is given by

$$\begin{bmatrix}
\mathbf{1.99992} & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & \mathbf{1.96206} & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & \mathbf{0.03454} & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & \mathbf{0.00005} & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00171} & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & \mathbf{0.00171}
\end{bmatrix}. \tag{109}$$

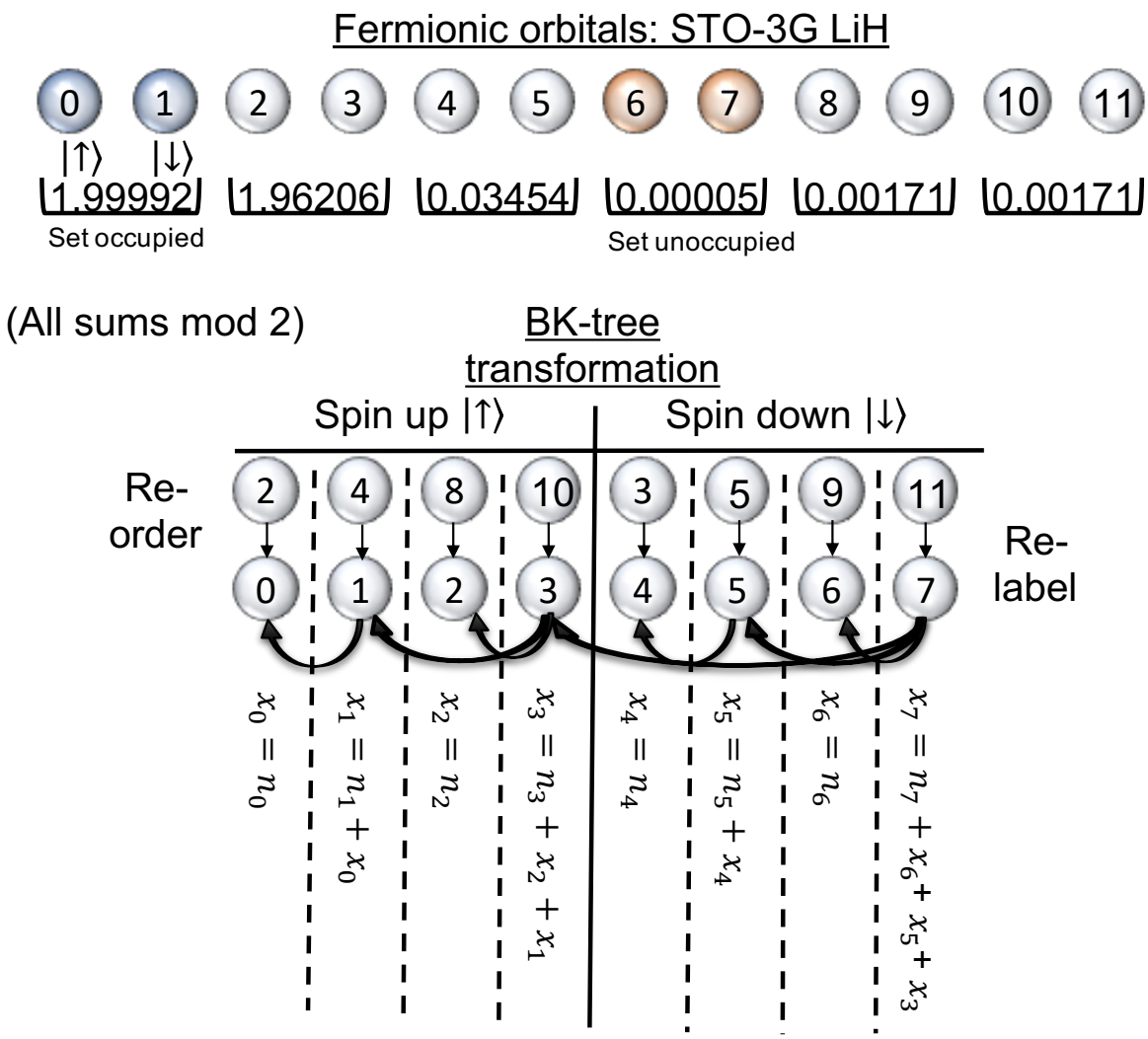

FIG. 17. A pictorial representation of the fermion-to-qubit mapping procedure for LiH in the STO-3G basis. The fermionic natural molecular orbitals (NMO) are initially arranged 'spin up, spin down, spin up, spin down, ...', and have their corresponding natural orbital occupation number (NOON) below. As the NOON of orbitals 6 and 7 is so small, they can be assumed unfilled, and removed from the Hamiltonian. As the combined NOON of orbitals 0 and 1 is close to 2, they can be assumed filled, and removed from the Hamiltonian. We then rearrange the remaining orbitals to be 'all spin up, all spin down', and re-label them from 0 to 7. We then perform the BK-tree mapping by constructing the Fenwick tree, Fen(0,7), as described in Fig. 18. The value x_i is the value of the i^{th} qubit under the BK-tree mapping, while n_i is the value of the i^{th} qubit under the JW mapping. We see that qubit 3 stores the sum $\sum_{i=0}^{3} n_i$, and qubit 7 stores the sum $\sum_{i=0}^{7} n_i$. As these sums are conserved quantities, these qubits do not flip throughout the simulation, and so can be removed from the Hamiltonian as described in Sec. IV C.

The first orbital has a NOON close to two, and so we consider it to always be doubly occupied. We can then remove any terms containing $a_0^\dagger, a_0, a_1^\dagger, a_1$ from the Hamiltonian. In contrast, the fourth orbital has a very small NOON. As a result, we assume that this orbital is never occupied, and so remove the two corresponding fermion operators from the Hamiltonian. This leaves a Hamiltonian acting on 8 spin-orbitals. As the number of orbitals is now a power of 2, we can use either the BK or BK-tree mappings to remove the 2 qubits associated with conservation symmetries. We use the BK-tree mapping in order to provide an explicit example of Fenwick tree construction. The Fenwick tree tells us which qubits store which orbitals in the BK-tree mapping. We denote the Fenwick tree for the M orbitals as $\text{Fen}(0, M-1)$. We can obtain this data structure using an iterative algorithm, which we reproduce from Ref. [91] below. The generation of the Fenwick tree for the LiH molecule using this algorithm is shown in Fig. 18.

Our final Hamiltonian acts on 6 qubits, but differs in energy from the full 12 qubit Hamiltonian by only 0.2 mHartree. A similar procedure is described in

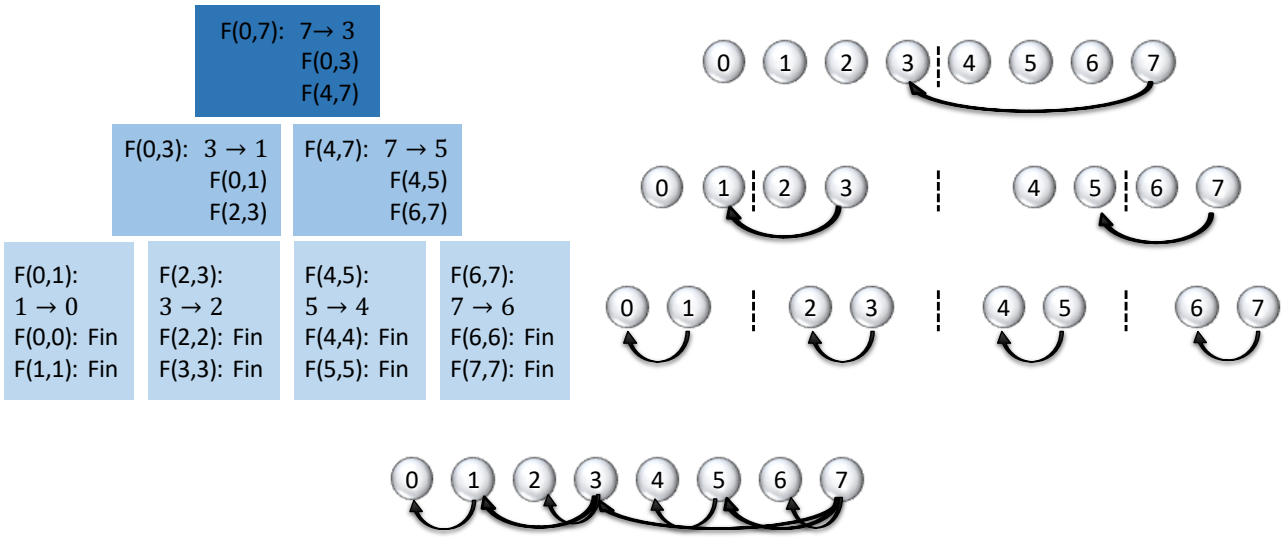

FIG. 18. A pictorial representation of the Fenwick tree construction for LiH, shown in Fig. 17. We carry out the BK-tree mapping by constructing the Fenwick tree, Fen(0,7), as described in Algorithm. 1. The algorithmic steps are shown on the left hand side of the figure, while the corresponding actions are shown on the right hand side. The notation $X \rightarrow Y$ means connect orbital X to orbital Y with an arrow. 'Fin' means that the corresponding branch of the Fenwick tree is finished. The finished Fenwick tree Fen(0,7) is shown at the bottom of the figure.

Algorithm.1 : Fenwick tree generation

Define Fen(L, R)

If $L \neq R$:

- Connect R to Floor $\left(\frac{R+L}{2}\right)$;
- Fen $\left(L,\ \text{Floor}\left(\frac{R+L}{2}\right)\right)$;
- Fen $\left(\text{Floor}\left(\frac{R+L}{2}\right)+1, R\right)$.

Else:

- End the current Fenwick tree.

Refs. [28, 126].

VIII. DISCUSSION AND CONCLUSIONS

In order to draw conclusions about the outlook for the nascent field of quantum computational chemistry, it is necessary to first consider the limitations of classical computational chemistry. We must also consider the resources required by the different quantum techniques, and the timeframe over which these resources may become available. As in previous sections, M denotes the number of spin-orbitals considered, and N denotes the number of electrons in the molecule.

A. Classical limits

Classical computational chemistry techniques can be broadly divided into two categories: qualititative methods, such as Hartree-Fock and density functional theory, which can be applied to systems with hundreds of atoms [72], and quantitative techniques such as full configuration interaction (and to a lesser extent, coupled cluster based methods), which are limited to molecules containing a small number of atoms.

In the case of FCI, classically storing the wavefunction requires an amount of memory which scales exponentially with the number of spin-orbitals considered for the molecule. Consequently, it is only possible to classically store and calculate the FCI ground state wavefunction of single atoms or small molecules, such as: the nitrogen molecule (N_2) in an accurate basis set using true FCI [149], or the Cr_2 molecule with 24 active electrons in 30 spin-orbitals [150], or the fluorine atom in a cc-PV5Z basis with additional basis functions [151]. The latter two examples used variations of quantum Monte Carlo full configuration interaction; a powerful approximation of the FCI wavefunction. However, this method is not without its own limitations, including the infamous 'sign problem' [152].

It is important to note that being able to accurately predict the ground state energy of small molecules leaves us far from our desired goal of designing new drugs and materials *in silico*. For example, as noted by Ref. [153], over 95 % of the approved drug molecules in DrugBank 5.0 are larger than is classically simulable using FCI methods. However, in practice it is not necessary to perform highly accurate calculations on the

entirety of a large molecule or enzyme. Instead, problem decomposition approaches can be utilised, whereby the most important part of the system is accurately simulated, and then integrated with a potentially less accurate simulation of the less challenging parts of the system [12, 153, 154].

If we limit ourselves to a reduced accuracy, and instead consider a classical coupled cluster singles and doubles (CCSD) approach, we are able to simulate larger molecules. The current state of the art for classical CCSD calculations is around 400-600 spin-orbitals [155], which corresponds to the DNA base guanine ($C_5H_5N_5O$) in a cc-PVTZ basis [156], or the hydrocarbon octane (C_8H_{18}) in a cc-PVTZ basis [153]. The classical implementation of CCSD does not store the wavefunction, as this would again be exponentially costly (as the CCSD wavefunction has support on all possible Slater determinants). Instead, coupled non-linear equations to solve the CCSD ground state problem can be derived [72, 157]. The time taken to solve these equations scales as $O((M-N)^4N^2)$ [157], while the memory needed to store the molecular integrals needed scales as $O(M^4)$. As discussed in Sec. III C 4, the two main limitations of the CC method are that it is not fully variational, and that it does not work well when applied to initial states which have support on multiple Slater determinants (states with strong static correlation). Moreover, the CCSD method itself is typically not considered accurate enough for truly quantitative calculations [72]. Instead, the CCSD(T) method can be used, which treats the triple excitations pertubatively, and scales in time approximately as $O(M^7)$.

B. Quantum resources

As discussed in Sec. IV B, quantum computers can store the FCI wavefunction of M spin-orbitals using only M qubits. As such, it seems relatively simple to surpass our current classical capabilities by constructing a device with 100 qubits, only slightly larger than the devices currently available. However, as discussed in Sec. II, we must also take into consideration the qubit overhead of error correction, which is determined by the error rates of the gates. It is not feasible to generate the FCI wavefunction by including all possible excitations in the UCC ansatz, as this would lead to a number of gates which scales exponentially with the number of orbitals considered. Instead, we can use adiabatic state preparation, as described in Sec. V A, with the assumption that the gap is well behaved [36].

Initial work showed that around 10^{12} gates would be necessary to simulate a system of around 100 spin-orbitals (excluding the overhead of error correction) [158–160], although this can be somewhat reduced by considering a more localised orbital basis set [101]. Reiher *et al.* used this as the starting point to find the resources required to carry out a transformative chemistry calculation [12]. They considered the problem of biological dinitrogen fixation. Currently, fertiliser is produced from nitrogen using the energy intensive Haber-Bosch process, which consumes up to 2% of the world's energy output [12]. However, bacteria containing the nitrogenase enzyme can convert nitrogen into ammonia under ambient conditions. The crux of understanding this enzyme is a small molecule, an iron molybdenum cofactor (FeMoco), which could be modelled with an active space of around 60 electrons in 110 spin-orbitals [12]. However, as a transition metal compound, the ground state of FeMoco likely shows strong static correlation, and so cannot be solved with CC methods. Moreover, the active space required is too large to be addressed with classical MCSCF approaches. The authors calculated the resources required to perform an FCI calculation on this active space using an error corrected quantum computer with around 100 logical qubits. They found that this would require around 200 million physical qubits, and take on the order of months (assuming current best error rates) [12]. Recent work has dramatically improved upon these estimates (for the Hubbard model or periodic systems such as the homogeneous electron gas), reducing the requirements to around one million physical qubits running for a few hours [113]. However, it may be many years before we possess quantum devices with one million physical qubits.

Consequently, in order to achieve transformative chemistry simulations within the next decade, different approaches are required. For example, we could use variational algorithms, together with heuristic methods. Alternatively, we can target good approximations of the FCI ground state using methods like UCCSD. We refer to these approaches as 'near-future' approaches.

Heuristic methods include the VQE with the Hamiltonian variational or LDCA ansatze. As discussed in Sec. V B, the depth of each block in these ansatze scales linearly with the number of orbitals considered. Moreover, both methods target the FCI ground state, so can be considered more accurate than the UCCSD ansatz (if they are successful). However, as heuristic methods, it is unclear as to how the number of ansatz blocks required for convergence scales with the system size.

As discussed above, the scaling of classical CCSD is approximately $O(M^6)$. The time required to carry out a UCCSD calculation on a hybrid quantum classical system in the Jordan-Wigner encoding scales as $O(M^3N^2)$, multiplied by the overhead of classical optimisation. The total scaling can thus be assumed to be roughly equivalent to classical methods. Moreover, as discussed above, classical CCSD calculations can be applied to systems with 400-600 orbitals, meaning that we would need a similar number of qubits to be competitive on a memory front.

As such, quantum UCCSD calculations are best suited to small system sizes for which classical CCSD methods are not applicable. As discussed previously, these

are systems with strong static correlation, for which a multireference initial wavefunction is required. This includes transition metal complexes, and systems undergoing bonding or bond-breaking [12].

One regularly cited limitation of the UCCSD ansatz is the large number of gates required. The number of gates scales as $O(M^3N^2)$ in the Jordan-Wigner encoding (without employing gate count reduction methods [71, 120]), which would imply over a billion gates for interesting transition metal complexes with around 100 orbitals, such as FeMoco. Recent work [120] has shown that the number of gates required to implement the UCCSD operator for a 100 orbital system can be reduced to around one million. Moreover, this does not take into account the reduction of excitation terms due to molecular point group symmetries. For example, the LiH molecule in an STO-3G basis naively has 200 excitation operators to consider. However, taking into account symmetries and a reduced active space, one can achieve accurate results while considering only around 12 excitation operators [28]. If the number of excitation operators required for FeMoco can be reduced in a similar manner, then a UCCSD simulation on a near-term quantum computer would become much more feasible.

C. Outlook for near-future approaches

For near-future approaches, the greatest remaining difficulties are constructing good ansatze, reducing the number of measurements required, and mitigating the effects of errors.

While good, chemically motivated state preparation routines, such as the UCC ansatz, have been developed, the gate count requirements may be too large for near-term quantum computers. This makes low depth, chemically motivated methods such as the LDCA and Hamiltonian variational ansatze particularly valuable in the near-term. However, most of these heuristic ansatze have only been tested on small molecules, making their scaling unclear. As such, while they may provide a low depth route to the FCI ground state, further work is needed to estimate their scaling for problems of interest. Furthermore, recent work [120] has shown that it is still possible to find improvements in the gate count of both heuristic (Hamiltonian variational) and chemically motivated (UCC) ansatze. Clearly, any work which substantially reduces the gate counts required for near-term chemistry simulation will be invaluable in achieving our goal of transformative chemistry simulation.

One of the largest problems with variational approaches is the number of measurements required, particularly when considering molecules. For example, Ref. [121] found that around 10^8 measurements were required for each energy evaluation for small molecules. This rose to 10^{13} samples per energy evaluation for a 112 spin-orbital molecule such as Fe_2S_2. This results in a total gate count for the algorithm of around 10^{26}, which would take billions of years (even with a gate time of 10ns) or an enormous cluster of small quantum computers working in parallel. While recent work has reduced the number of measurements required by several orders of magnitude [13, 100, 119, 161], the number of measurements is still dauntingly high for molecular simulations.

Overcoming errors is another critical challenge facing near-future approaches. Recent experimental demonstrations have shown that noise can corrupt the results of quantum computational chemistry calculations [25, 28]. Even with an optimistic two qubit gate error rate of 0.01 %, we could only carry out around 10,000 gates before we would always expect an error to occur in the circuit. While the error mitigation techniques presented in Sec. VI may increase this to approximately 30,000-50,000 gates, this is still lower than the gate counts required for most problems of interest. As such, it is crucial to find new methods of error mitigation or low-resource error correction.

Despite the challenges discussed above, there are several potential avenues for which these problems look tractable in the near-term.

D. Target problems

One possible target is calculating the energy density of the 2D uniform electron gas (jellium), which could then be used in DFT calculations – as suggested by Ref. [13]. This may benefit many research areas in computational chemistry. The plane wave basis discussed in Sec. III D 5 is the natural basis set for jellium, and as such, simulations could be performed using shallow circuits, by making use of the Hamiltonian variational ansatz [13]. A simulation requiring around 100 physical qubits would suffice to surpass classical methods, if the circuit depth could be kept low enough to prevent significant error accumulation [13].

An alterative simulation target is the Fermi-Hubbard model. The Fermi-Hubbard model is the prototypical system for many areas of materials chemistry and condensed matter physics, including high temperature superconductivity [162]. However, classical methods to solve the Fermi-Hubbard model exactly are limited to around 20 lattice sites [163]. As such, if we were able to solve a classically intractable Fermi-Hubbard model problem on a near-term quantum computer, it would signal a clear quantum advantage over classical methods.

It has been shown possible to prepare initial states of the Hubbard model using $O(N^{1.5})$ gates, and per-

form Trotter steps of the Fermi-Hubbard Hamiltonian using $O(N)$ gates for each Trotter step [163]. Previous work has shown that the Hamiltonian variational ansatz performs well for the Fermi-Hubbard model, achieving good convergence for a 12 site problem with 20 Trotter steps [121]. As such, near-future quantum computers may be able to solve a classically intractable problem with around 100 lattice sites (200 qubits). The time required to solve this problem is only on the order of days [121], and could be reduced further by taking advantage of the inherent parallelisability of the VQE. As such, the Fermi-Hubbard model provides an interesting and computationally feasible goal to aim for in the near-future.

E. Conclusions

This work has sought to bridge the gap between those scientists working on quantum information, and those working on computational chemistry. We have discussed the key methods used in classical computational chemistry, and how these have been incorporated into quantum algorithms. We have highlighted the key differences between quantum and classical methods of chemistry simulation, and the resulting benefits that quantum computing is widely predicted to bring to the field of computational chemistry.

However, we have also shown that quantum methods still face many challenges, not least the high error rates and low qubit counts of current hardware. As such, it is important to continue to develop new algorithms, mappings, basis sets, and error mitigation techniques. Doing so will reduce the resources required for transformative chemistry simulations, enabling us to reap their benefits on a much shorter timescale.

ACKNOWLEDGEMENTS

This work was supported by BP plc and by the EPSRC National Quantum Technology Hub in Networked Quantum Information Technology (EP/M013243/1). AAG acknowledges Anders G Froseth for his generous support, as well as the Vannevar Bush Faculty Fellowship program of the US Department of Defense. S.M. and X.Y. thank L. Lindoy for helpful discussions on basis sets.

[1] P. Dirac, Proceedings of the Royal Society of London **123** (1929), 10.1098/rspa.1929.0094.

[2] R. P. Feynman, International Journal of Theoretical Physics **21**, 467 (1982).

[3] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, Phys. Rev. Lett. **106**, 130506 (2011).

[4] T. Harty, D. Allcock, C. J. Ballance, L. Guidoni, H. Janacek, N. Linke, D. Stacey, and D. Lucas, Physical review letters **113**, 220501 (2014).

[5] C. J. Ballance, T. P. Harty, N. M. Linke, M. A. Sepiol, and D. M. Lucas, Phys. Rev. Lett. **117**, 060504 (2016).

[6] J. P. Gaebler, T. R. Tan, Y. Lin, Y. Wan, R. Bowler, A. C. Keith, S. Glancy, K. Coakley, E. Knill, D. Leibfried, *et al.*, Physical review letters **117**, 060505 (2016).

[7] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, *et al.*, Nature **508**, 500 (2014).

[8] C. Song, K. Xu, W. Liu, C.-p. Yang, S.-B. Zheng, H. Deng, Q. Xie, K. Huang, Q. Guo, L. Zhang, P. Zhang, D. Xu, D. Zheng, X. Zhu, H. Wang, Y.-A. Chen, C.-Y. Lu, S. Han, and J.-W. Pan, Phys. Rev. Lett. **119**, 180511 (2017).

[9] X.-L. Wang, L.-K. Chen, W. Li, H.-L. Huang, C. Liu, C. Chen, Y.-H. Luo, Z.-E. Su, D. Wu, Z.-D. Li, H. Lu, Y. Hu, X. Jiang, C.-Z. Peng, L. Li, N.-L. Liu, Y.-A. Chen, C.-Y. Lu, and J.-W. Pan, Phys. Rev. Lett. **117**, 210502 (2016).

[10] L.-K. Chen, Z.-D. Li, X.-C. Yao, M. Huang, W. Li, H. Lu, X. Yuan, Y.-B. Zhang, X. Jiang, C.-Z. Peng, L. Li, N.-L. Liu, X. Ma, C.-Y. Lu, Y.-A. Chen, and J.-W. Pan, Optica **4**, 77 (2017).

[11] A. Aspuru-Guzik, R. Lindh, and M. Reiher, ACS Central Science **4**, 144 (2018).

[12] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, Proceedings of the National Academy of Sciences (2017).

[13] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and G. K.-L. Chan, Phys. Rev. X **8**, 011044 (2018).

[14] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Science **309**, 1704 (2005).

[15] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, Proceedings of the National Academy of Sciences **105**, 18681 (2008).

[16] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New Journal of Physics **18**, 023023 (2016).

[17] J. Huh, G. G. Guerreschi, B. Peropadre, J. R. McClean, and A. Aspuru-Guzik, Nature Photonics **9**, 615 (2015).

[18] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, *et al.*, Nature chemistry **2**, 106 (2010).

[19] J. Du, N. Xu, X. Peng, P. Wang, S. Wu, and D. Lu, Phys. Rev. Lett. **104**, 030502 (2010).

[20] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. Obrien, Nature communications **5** (2014).

[21] Y. Wang, F. Dolde, J. Biamonte, R. Babbush, V. Bergholm, S. Yang, I. Jakobi, P. Neumann, A. Aspuru-Guzik, J. D. Whitfield, *et al.*, ACS nano **9**, 7769 (2015).

[22] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Phys. Rev. X **6**, 031007 (2016).

[23] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Phys. Rev. A **95**, 020501 (2017).

[24] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O'Brien, and M. G. Thompson, Phys. Rev. Lett. **118**, 100503 (2017).

[25] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017).

[26] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi, Phys. Rev. X **8**, 011021 (2018).

[27] R. Santagati, J. Wang, A. A. Gentile, S. Paesani, N. Wiebe, J. R. McClean, S. Morley-Short, P. J. Shadbolt, D. Bonneau, J. W. Silverstone, D. P. Tew, X. Zhou, J. L. O'Brien, and M. G. Thompson, Science Advances **4** (2018), 10.1126/sciadv.aap9646.

[28] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. Roos, ArXiv e-prints (2018), arXiv:1803.10238 [quant-ph].

[29] A. Kandala, K. Temme, A. D. Corcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, arXiv:1805.04492 (2018).

[30] Y. Shen, Y. Lu, K. Zhang, J. Zhang, S. Zhang, J. Huh, and K. Kim, Chem. Sci. **9**, 836 (2018).

[31] C. Sparrow, E. Martín-López, N. Maraviglia, A. Neville, C. Harrold, J. Carolan, Y. N. Joglekar, T. Hashimoto, N. Matsuda, J. L. OBrien, *et al.*, Nature **557**, 660 (2018).

[32] M. A. Nielsen and I. Chuang, "Quantum computation and quantum information," (2002).

[33] I. M. Georgescu, S. Ashhab, and F. Nori, Rev. Mod. Phys. **86**, 153 (2014).

[34] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, arXiv:quant-ph/0001106 (2000).

[35] D. Aharonov, W. Van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, SIAM review **50**, 755 (2008).

[36] R. Babbush, P. J. Love, and A. Aspuru-Guzik, Scientific reports **4**, 6603 (2014).

[37] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[38] R. Raussendorf, D. E. Browne, and H. J. Briegel, Phys. Rev. A **68**, 022312 (2003).

[39] R. Jozsa, arXiv:quant-ph/0508124 (2005).

[40] S. Lloyd and S. L. Braunstein, Phys. Rev. Lett. **82**, 1784 (1999).

[41] S. L. Braunstein and P. van Loock, Rev. Mod. Phys. **77**, 513 (2005).

[42] D. P. DiVincenzo, Phys. Rev. A **51**, 1015 (1995).
[43] R. Landauer, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **353**, 367 (1995).
[44] W. G. Unruh, Phys. Rev. A **51**, 992 (1995).
[45] P. W. Shor, Physical review A **52**, R2493 (1995).
[46] A. M. Steane, Physical Review Letters **77**, 793 (1996).
[47] E. Knill, R. Laflamme, and W. Zurek, (1996).
[48] D. Gottesman, Physical Review A **57**, 127 (1998).
[49] P. W. Shor, in *Foundations of Computer Science, 1996. Proceedings., 37th Annual Symposium on* (IEEE, 1996) pp. 56–65.
[50] A. Y. Kitaev, Russian Mathematical Surveys **52**, 1191 (1997).
[51] A. Y. Kitaev, Annals of Physics **303**, 2 (2003).
[52] D. S. Wang, A. G. Fowler, A. M. Stephens, and L. C. L. Hollenberg, arXiv preprint arXiv:0905.0531 (2009).
[53] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Physical Review A **86**, 032324 (2012).
[54] D. S. Wang, A. G. Fowler, and L. C. Hollenberg, Physical Review A **83**, 020302 (2011).
[55] E. T. Campbell, B. M. Terhal, and C. Vuillot, Nature **549**, 172 (2017).
[56] X. Chen, H. Chung, A. W. Cross, B. Zeng, and I. L. Chuang, Phys. Rev. A **78**, 012353 (2008).
[57] B. Eastin and E. Knill, Phys. Rev. Lett. **102**, 110502 (2009).
[58] S. Bravyi and A. Kitaev, Physical Review A **71**, 022316 (2005).
[59] J. O'Gorman and E. T. Campbell, Phys. Rev. A **95**, 032338 (2017).
[60] J. Wooten, arXiv preprint 1806.02736 (2018).
[61] A. W. Harrow and A. Montanaro, Nature **549**, 203 (2017).
[62] J. Huh and M.-H. Yung, Scientific reports **7**, 7462 (2017).
[63] J. Argello-Luengo, A. Gonzlez-Tudela, T. Shi, P. Zoller, and J. I. Cirac, arXiv:1807.09228 (2018).
[64] S. Lloyd, Science **273**, 1073 (1996).
[65] D. S. Abrams and S. Lloyd, Phys. Rev. Lett. **79**, 2586 (1997).
[66] H. F. Trotter, Proceedings of the American Mathematical Society **10**, 545 (1959).
[67] Y. Li and S. C. Benjamin, Phys. Rev. X **7**, 021050 (2017).
[68] A. Y. Kitaev, Preprint at http://arxiv. org/abs/quant-ph/9511026 (1995).
[69] D. S. Abrams and S. Lloyd, Phys. Rev. Lett. **83**, 5162 (1999).
[70] I. Kassal and A. Aspuru-Guzik, The Journal of Chemical Physics **131**, 224102 (2009), https://doi.org/10.1063/1.3266959.
[71] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. Love, and A. Aspuru-Guzik, arXiv preprint arXiv:1701.02691 (2017).
[72] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular electronic-structure theory* (John Wiley & Sons, 2014).
[73] A. Szabo and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory* (Courier Corporation, 2012).
[74] O. Christiansen, Physical Chemistry Chemical Physics **14**, 6672 (2012).
[75] H. Wang, S. Kais, A. Aspuru-Guzik, and M. R. Hoffmann, Phys. Chem. Chem. Phys. **10**, 5388 (2008).
[76] B. O. Roos, P. R. Taylor, and P. E. Sigbahn, Chemical Physics **48**, 157 (1980).
[77] D. J. Griffiths, *Introduction to quantum mechanics* (Cambridge University Press, 2016).
[78] R. Ditchfield, W. J. Hehre, and J. A. Pople, The Journal of Chemical Physics **54**, 724 (1971).
[79] T. H. D. Jr., The Journal of Chemical Physics **90**, 1007 (1989).
[80] S. R. White, The Journal of Chemical Physics **147**, 244102 (2017).
[81] S. Wiesner, arXiv preprint quant-ph/9603028 (1996).
[82] C. Zalka, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **454**, 313 (1998).
[83] D. A. Lidar and H. Wang, Phys. Rev. E **59**, 2429 (1999).
[84] N. J. Ward, I. Kassal, and A. Aspuru-Guzik, The Journal of Chemical Physics **130**, 194105 (2009).
[85] J. D. Whitfield, arXiv preprint arXiv:1502.03771 (2015).
[86] I. Kassal, J. D. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik, Annual review of physical chemistry **62**, 185 (2011).
[87] K. I. D., W. N., B. R., and A.-G. A., Journal of Physics A: Mathematical and Theoretical **50** (2017).
[88] J. T. Seeley, M. J. Richard, and P. J. Love, The Journal of Chemical Physics **137**, 224109 (2012).
[89] S. B. Bravyi and A. Y. Kitaev, Annals of Physics **298**, 210 (2002).
[90] T. Andrew, S. Sarah, S. Jake, K. Michael, M. Jarrod, B. Ryan, C. P. V., M. Florian, W. Frank, and L. P. J., International Journal of Quantum Chemistry **115**, 1431.
[91] V. Havlicek, M. Troyer, and J. D. Whitfield, Phys. Rev. A **95**, 032332 (2017).
[92] "`https://github.com/quantumlib/OpenFermion/issues/259`," Accessed: 24.7.2018.
[93] M. Steudtner and S. Wehner, ArXiv e-prints (2017), arXiv:1712.07067 ["quant-ph"].
[94] V. F. and C. J. I., Journal of Statistical Mechanics: Theory and Experiment **2005** (2005).
[95] W. J.D., H. V., and T. M., Physical Review A **94** (2016).
[96] S. K. and W. J. D., The Journal of Chemical Physics **148**, 164104 (2018).
[97] S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, arXiv preprint arXiv:1701.08213 (2017).
[98] J. Romero, J. P. Olson, and A. Aspuru-Guzik, Quantum Science and Technology **2**, 045001 (2017).
[99] M.-H. Yung, J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano, Scientific reports **4**, 3589 (2014).
[100] D. Wang, O. Higgott, and S. Brierley, (2018), arXiv:1802.00171.
[101] J. R. McClean, R. Babbush, P. J. Love, and A. Aspuru-Guzik, The Journal of Physical Chemistry Letters **5**, 4368 (2014), pMID: 26273989, https://doi.org/10.1021/jz501649m.
[102] J.-S. Xu, M.-H. Yung, X.-Y. Xu, S. Boixo, Z.-W. Zhou, C.-F. Li, A. Aspuru-Guzik, and G.-C. Guo, Nature Photonics **8**, 113 (2014).
[103] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, Communications in Mathematical Physics **270**, 359 (2007).

[104] D. W. Berry and A. M. Childs, arXiv preprint arXiv:0910.4157 (2009).

[105] D. W. Berry, A. M. Childs, and R. Kothari, in *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on* (IEEE, 2015) pp. 792–809.

[106] R. Babbush, D. W. Berry, I. D. Kivlichan, A. Y. Wei, P. J. Love, and A. Aspuru-Guzik, New Journal of Physics **18**, 033032 (2016).

[107] R. Babbush, D. W. Berry, Y. R. Sanders, I. D. Kivlichan, A. Scherer, A. Y. Wei, P. J. Love, and A. Aspuru-Guzik, Quantum Science and Technology **3**, 015006 (2017).

[108] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, Phys. Rev. Lett. **114**, 090502 (2015).

[109] G. H. Low and I. L. Chuang, arXiv preprint arXiv:1610.06546 (2016).

[110] G. H. Low and I. L. Chuang, Phys. Rev. Lett. **118**, 010501 (2017).

[111] D. W. Berry, M. Kieferová, A. Scherer, Y. R. Sanders, G. H. Low, N. Wiebe, C. Gidney, and R. Babbush, npj Quantum Information **4**, 22 (2018).

[112] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, arXiv preprint arXiv:1711.10980 (2017).

[113] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven, arXiv preprint arXiv:1805.03662 (2018).

[114] J. Preskill, arXiv preprint arXiv:1801.00862 (2018).

[115] J. J. Sakurai and J. Napolitano, *Modern quantum mechanics* (Cambridge University Press, 2017).

[116] P.-L. Dallaire-Demers, J. Romero, L. Veis, S. Sim, and A. Aspuru-Guzik, arXiv preprint arXiv:1801.01053 (2018).

[117] E. Farhi, J. Goldstone, and S. Gutmann, arXiv preprint arXiv:1411.4028 (2014).

[118] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, arXiv preprint arXiv:1803.11173 (2018).

[119] P. K. Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, G. Salis, A. Fuhrer, M. Ganzhorn, D. J. Egger, M. Troyer, A. Mezzacapo, S. Filipp, and I. Tavernelli, ArXiv e-prints (2018), arXiv:1805.04340 ["quant-ph"].

[120] M. Motta, E. Ye, J. R. McClean, Z. Li, A. J. Minnich, R. Babbush, and G. K.-L. Chan, arXiv:1808.02625 (2018).

[121] D. Wecker, M. B. Hastings, and M. Troyer, Phys. Rev. A **92**, 042303 (2015).

[122] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Phys. Rev. Lett. **120**, 110501 (2018).

[123] T. G. Kolda, R. M. Lewis, and V. Torczon, SIAM Rev. **45**, 385482 (2006).

[124] J. C. Spall, IEEE Transactions on Automatic Control **37**, 332341 (1992).

[125] A. Garcia-Saez and J. I. Latorre, arXiv:1806.02287 (2018).

[126] S. McArdle, S. Endo, Y. Li, S. Benjamin, and X. Yuan, arXiv preprint arXiv:1804.03023 (2018).

[127] O. Higgott, D. Wang, and S. Brierley, arXiv preprint arXiv:1805.08138 (2018).

[128] S. Endo, T. Jones, S. McArdle, X. Yuan, and S. Benjamin, arXiv preprint arXiv:1806.05707 (2018).

[129] J. C. Garcia-Escartin and P. Chamorro-Posada, Phys. Rev. A **87**, 052330 (2013).

[130] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles, arXiv preprint arXiv:1803.04114 (2018).

[131] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, Phys. Rev. A **95**, 042308 (2017).

[132] N. P. D. Sawaya, M. Smelyanskiy, J. R. McClean, and A. Aspuru-Guzik, Journal of Chemical Theory and Computation **12**, 3097 (2016), pMID: 27254482, https://doi.org/10.1021/acs.jctc.6b00220.

[133] D. Gottesman, (2009), arXiv:0904.2557.

[134] S. J. Devitt, W. J. Munro, and K. Nemoto, Reports on Progress in Physics **76**, 076001 (2013).

[135] K. Temme, S. Bravyi, and J. M. Gambetta, Phys. Rev. Lett. **119**, 180509 (2017).

[136] S. Endo, S. C. Benjamin, and Y. Li, arXiv preprint arXiv:1712.09271 (2017).

[137] P. D. Johnson, J. Romero, J. Olson, Y. Cao, and A. Aspuru-Guzik, arXiv preprint arXiv:1711.02249 (2017).

[138] S. McArdle, X. Yuan, and S. Benjamin, arXiv:1807.02467 (2018).

[139] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, arXiv:1807.10050 (2018).

[140] L. F. Richardson, B. J Arthur Gaunt, *et al.*, Phil. Trans. R. Soc. Lond. A **226**, 299 (1927).

[141] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Phys. Rev. A **80**, 012304 (2009).

[142] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, "Gaussian16 Revision B.01," (2016), gaussian Inc. Wallingford CT.

[143] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, Journal of Chemical Theory and Computation **13**, 3185 (2017), pMID: 28489372, https://doi.org/10.1021/acs.jctc.7b00174.

[144] "Pyquante," `http://pyquante.sourceforge.net/`, accessed: 24.7.2018.

[145] S. Qiming, B. T. C., B. N. S., B. G. H., G. Sheng, L. Zhendong, L. Junzi, M. J. D., S. E. R., S. Sandeep, W. Sebastian, and C. G. Kin-Lic, Wiley Interdisciplinary Reviews: Computational Molecular Science **8**, e1340,

https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1340.
[146] J. R. McClean, I. D. Kivlichan, K. J. Sung, D. S. Steiger, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby, P. Gokhale, T. Hner, T. Hardikar, V. Havlek, C. Huang, J. Izaac, Z. Jiang, X. Liu, M. Neeley, T. O'Brien, I. Ozfidan, M. D. Radin, J. Romero, N. Rubin, N. P. D. Sawaya, K. Setia, S. Sim, M. Steudtner, Q. Sun, W. Sun, F. Zhang, and R. Babbush, arXiv:1710.07629 (2017).
[147] "Ibm qiskit," https://qiskit.org/, accessed: 24.7.2018.
[148] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Molecular Physics **109**, 735 (2011).
[149] E. Rossi, G. L. Bendazzoli, S. Evangelisti, and D. Maynau, Chemical Physics Letters **310**, 530 (1999).
[150] N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley, The Journal of Chemical Physics **145**, 044112 (2016), https://doi.org/10.1063/1.4955109.
[151] G. H. Booth and A. Alavi, The Journal of Chemical Physics **132**, 174104 (2010), https://doi.org/10.1063/1.3407895.
[152] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Phys. Rev. A **64**, 022319 (2001).
[153] T. Yamazaki, S. Matsuura, A. Narimani, A. Saidmuradov, and A. Zaribafiyan, arXiv:1806.01305 (2018).
[154] N. C. Rubin, (2016), arXiv:1610.06910.
[155] C. Hattig and F. Weigend, The Journal of Chemical Physics **113**, 5154 (2000), https://aip.scitation.org/doi/pdf/10.1063/1.1290013.
[156] P. Hobza and J. poner, Journal of the American Chemical Society **124**, 11802 (2002), pMID: 12296748, https://doi.org/10.1021/ja026759n.
[157] G. D. Purvis and R. J. Bartlett, The Journal of Chemical Physics **76**, 1910 (1982), https://doi.org/10.1063/1.443164.
[158] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, Phys. Rev. A **90**, 022305 (2014).
[159] M. B. Hastings, D. Wecker, B. Bauer, and M. Troyer, QIC **15** (2015).
[160] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doherty, and M. Troyer, QIC **15** (2015).
[161] N. C. Rubin, R. Babbush, and J. McClean, New Journal of Physics **20**, 053020 (2018).
[162] D. Wecker, M. B. Hastings, N. Wiebe, B. K. Clark, C. Nayak, and M. Troyer, Phys. Rev. A **92**, 062318 (2015).
[163] Z. Jiang, K. J. Sung, K. Kechedzhi, V. N. Smelyanskiy, and S. Boixo, Phys. Rev. Applied **9**, 044036 (2018).