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: January 18, 2019)

One of the most promising applications of quantum computing is solving classically intractable chemistry problems. This may enable the design of new materials, medicines, catalysts, or high temperature superconductors. As a result, quantum computational chemistry is rapidly emerging as an interdisciplinary field requiring knowledge of both quantum information and computational chemistry. This review provides a comprehensive introduction to both fields, bridging the current knowledge gap. We review the major developments in this area, with a focus on near-term quantum computation. Illustrations of key methods are provided, 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

arXiv:1808.10402v2 [quant-ph] 17 Jan 2019

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."* (Dirac, 1929). The problem described by Dirac is that the complexity of 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."* (Feynman, 1982). Building a quantum computer has taken over 30 years, but Feynman's vision may soon be fulfilled, following recent developments in quantum hardware including ion traps (Ballance *et al.*, 2016; Gaebler *et al.*, 2016; Harty *et al.*, 2014; Monz *et al.*, 2011), superconducting systems (Barends *et al.*, 2014; Chow *et al.*, 2012; Song *et al.*, 2017; Wendin, 2017), and photonic systems (Chen *et al.*, 2017; Wang *et al.*, 2016). It is believed that using quantum systems as our simulation platform will yield unprecedented developments in chemistry (Aspuru-Guzik *et al.*, 2018), biology (Reiher *et al.*, 2017), medicine (Cao *et al.*, 2018a), and materials science (Babbush *et al.*, 2018c).

To date, several efficient quantum algorithms have been proposed to solve problems in chemistry (Aspuru-Guzik *et al.*, 2005; Huh *et al.*, 2015; Kassal *et al.*, 2008; Lidar and Wang, 1999; Peruzzo *et al.*, 2014). 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 (Colless *et al.*, 2018; Du *et al.*, 2010; Ganzhorn *et al.*, 2018; Hempel *et al.*, 2018; Kandala *et al.*, 2017, 2018; Lanyon *et al.*, 2010; Li *et al.*, 2011; O'Malley *et al.*, 2016; Paesani *et al.*, 2017a; Peruzzo *et al.*, 2014; Santagati *et al.*, 2018; Shen *et al.*, 2018, 2017; Sparrow

et al., 2018; Wang *et al.*, 2015). 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 (Mueck, 2015). New developments are needed to solve classically intractable chemistry problems on a shorter timescale.

These breakthroughs may 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 low lying energy levels of molecules (Aspuru-Guzik *et al.*, 2005). There are three reasons for this restriction of scope. Primarily, this problem is a fundamental one in classical computational chemistry. Knowledge of the energy eigenstates enables the prediction of reaction rates, determination of stable structures, and optical properties. Secondly, it is because the machinery developed to solve this problem on quantum computers is easily applied to other types of problems. Finally, 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.

This review is organised as follows. 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. The methods developed to merge these two fields, including mapping chemistry problems onto a quantum computer, are described in Sec. IV. 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 review, providing worked examples for the reader. We conclude this review 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.

Several other detailed reviews on this topic exist in the literature. Summaries of early theoretical and experimental work in quantum computational chemistry can be found in (Kassal *et al.*, 2011; Lu *et al.*, 2012). More focused discussions of quantum algorithms for chemistry simulation, and the computational complexity of problems in chemistry can be found in (Kais *et al.*, 2014; Veis and Pittner, 2012; Yung *et al.*, 2012). A comprehensive review was recently released by Cao *et al.* (2018b). Said review, and our own, are complementary, and together provide a complete overview of progress to date in quantum computational chemistry.

Despite being a relatively young field, quantum computational chemistry has grown extremely rapidly, and has already evolved beyond the stage that it can be fully summarised by a single review. As such, there are approaches to solving chemistry problems with a quantum computer which we are not able to describe fully in this review, such as: using quantum computers as part of a problem decomposition approach to simulation (Bauer *et al.*, 2016; Dallaire-Demers and Wilhelm, 2016a,b; Kreula *et al.*, 2016; Rubin, 2016), hybrid quantum algorithms for time dependent density functional theory (Whitfield *et al.*, 2014), quantum algorithms for relativistic quantum chemistry (Veis *et al.*, 2012), gate based methods for simulating molecular vibrations (McArdle *et al.*, 2018b; Sawaya and Huh, 2018), analog simulators of molecular vibrations (Huh *et al.*, 2015; Huh and Yung, 2017; Shen *et al.*, 2018; Sparrow *et al.*, 2018), quantum methods for electron-phonon systems (Macridin *et al.*, 2018a,b), protein folding (Babbush *et al.*, 2012; Babej *et al.*, 2018; Fingerhuth *et al.*, 2018; Perdomo *et al.*, 2008; Perdomo-Ortiz *et al.*, 2012), and solving problems in chemistry using a quantum annealer (Babbush *et al.*, 2014; Genin *et al.*, 2019).

II. QUANTUM COMPUTING AND SIMULATION

In this section, we introduce the basic elements of quantum computing and quantum simulation. We refer the reader to Nielsen and Chuang (2002) and Georgescu *et al.* (2014) for more detailed introductions.

A. Quantum computing

In this review, we focus on the qubit-based circuit model of quantum computation (Nielsen and Chuang, 2002). Other paradigms that vary to a greater or lesser extent include: adiabatic quantum computing (Aharonov *et al.*, 2008; Farhi *et al.*, 2000) (including in the context of chemistry simulation (Babbush *et al.*, 2014)), one-way or measurement based quantum computing (Jozsa, 2005; Raussendorf and Briegel, 2001; Raussendorf *et al.*, 2003), and continuous-variable quantum computing (Braunstein and van Loock, 2005; Lloyd and Braunstein, 1999).

The canonical circuit model of quantum computation is so-named because of its resemblance to the logic circuits used in classical computing. In the classical circuit model, logic gates (such as AND, OR and NOT) act on bits of information. In the quantum case, quantum gates are acted upon the basic unit of information, the qubit. The qubit 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 in the computational basis, 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

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \tag{4}$$

In this review, 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 $(|\bar{0}\rangle = |0\rangle^{\otimes n} = |0\rangle \otimes |0\rangle \otimes \cdots \otimes |0\rangle)$. A quantum circuit generally concludes with measurements to extract information. It may also employ additional intermediate measurements, for example, to check for errors. From a mathematical perspective, the qubit gates are unitary matrices. Typical gates include the Pauli gates

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

the single qubit rotation gates

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

the Hadamard and T gates

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}, \quad (7)$$

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

$$I_T \otimes |0\rangle \langle 0|_C + X_T \otimes |1\rangle \langle 1|_C, \quad (8)$$

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 the entangled state of 2 qubits given by Eq. (4), meaning the state cannot be written as a tensor product of individual qubit states.
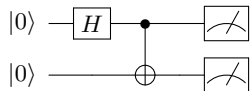

FIG. 1 A quantum circuit that generates the entangled state $(|00\rangle + |11\rangle)/\sqrt{2}$ and measures each qubit in the computational basis. Here, H is the Hadamard gate, defined in the main text. Half of the time both of the qubits will be measured to be 0, and the other half, both 1. Time runs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

With only single qubit operations and the CNOT gate, it is possible to approximate an arbitrary multi-qubit gate to any desired accuracy (DiVincenzo, 1995). 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

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

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

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

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

$$\bar{O} = \langle \psi | O | \psi \rangle, \quad (11)$$

referred to as the expected value of the operator O. Measuring the expectation value of qubit i in the computational basis corresponds to $\langle \psi | Z_i | \psi \rangle$. In practice, this means that we repeatedly measure the state of qubit i, labelling the outcomes $+1$ (for $|0\rangle$) and -1 (for $|1\rangle$). We then take the mean of these measurement outcomes. 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. To measure the expectation value of a product of operators, such as $Z_i Z_j$, we assign the outcomes of a single measurement of $Z_i Z_j$ to be the product of the individual measurement outcomes of Z_i and Z_j. These outcomes are typically correlated for entangled states. For example, the measurement outcome for ZZ on the state in Eq. (4) is always $+1$ ($+1 \times +1$ for $|00\rangle$ and -1×-1 for $|11\rangle$).

The Pauli operators form a complete basis for any Hermitian operator. Therefore any observable can be expanded into strings of Pauli operators, the expectation values of 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 (Landauer, 1995; Unruh, 1995), we can encode n logical qubits into $m > n$ physical qubits. These codes are analogous to classical error correcting codes, but are in general more complex, due to the delicate nature of quantum information. Depending on the code used, we can either detect, or detect *and* correct the errors which occur. Error correcting codes seek to endow the quantum computer with a

property known as 'fault-tolerance' (Aharonov and Ben-Or, 1997; Gottesman, 1998; Shor, 1996). This means that if the physical error rate in the gates is below a certain (code-dependent) threshold value, the error rate in the logical operations can be made arbitrarily low - even if the error checking measurements cause errors (Knill *et al.*, 1996). A more detailed discussion of specific codes can be found in Devitt *et al.* (2013); Raussendorf (2012); and Terhal (2015). One of the most widely studied error correction codes is the surface code (Kitaev, 1997), which is particularly suitable for 2D grids of qubits with nearest-neighbour connectivity. Physical error rates below the surface code threshold of around 1 % (Fowler *et al.*, 2012b; Stephens, 2014; Wang *et al.*, 2011) have recently been achieved for trapped ion (Ballance *et al.*, 2016; Gaebler *et al.*, 2016) and superconducting (Barends *et al.*, 2014) qubits. However, with such error rates, we would require around $10^3 - 10^4$ physical qubits per logical qubit to perform interesting tasks in a fault-tolerant manner (Campbell *et al.*, 2017; Fowler *et al.*, 2012a; O'Gorman and Campbell, 2017).

In contrast, current quantum computers are described in terms of tens of error-prone physical qubits. Quantum computers of this size are too large to exactly 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 (Boixo *et al.*, 2018; Harrow and Montanaro, 2017). Preskill (2018) has referred to these machines as noisy intermediate-scale quantum (NISQ) devices, and observed that it is currently unclear whether they will be able to perform useful tasks better than classical computers. 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 review, 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. However, analog simulators are generally considered more robust to noise, and therefore easier to construct (Georgescu *et al.*, 2014). To date, there have been several proposals for the simulation of chemistry using analog simulators (Argüello-Luengo *et al.*, 2018; Chin and Huh, 2018; Huh *et al.*, 2015; Huh and Yung, 2017; Torrontegui *et al.*, 2011), some of which have been experimentally realised (Clements *et al.*, 2017; Hu *et al.*, 2018; Shen *et al.*, 2018; Smirnov *et al.*, 2007; Sparrow *et al.*, 2018). Nevertheless, to perform accurate simulations of large chemical systems, we will likely require digital quantum simulation, as it is not yet clear how to protect large analog simulations from errors. Digital quantum simulation is more vulnerable to noise and device imperfections than analog simulation. While such imperfections can be addressed via error correction, this requires additional qubits and places stringent requirements on gate fidelities (O'Gorman and Campbell, 2017). In this review we focus solely on digital quantum simulation of chemistry problems. We refer the reader to Aspuru-Guzik and Walther (2012); Blatt and Roos (2012); Georgescu *et al.* (2014); Houck *et al.* (2012); and Schneider *et al.* (2012) for information about digital 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 (Lloyd, 1996) and further developed by Abrams and Lloyd (Abrams and Lloyd, 1997). 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 time evolution operator, e^{-itH_s}, to a series of gates. This can be achieved using a Lie-Trotter-Suzuki decomposition (Trotter, 1959), commonly referred to as Trotterization. This means that if the Hamiltonian of the system, H_s 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. Each of the exponential terms in Eq. (13) can be realised efficiently using a quantum computer. As the dynamics of local Hamiltonians can be efficiently simulated on
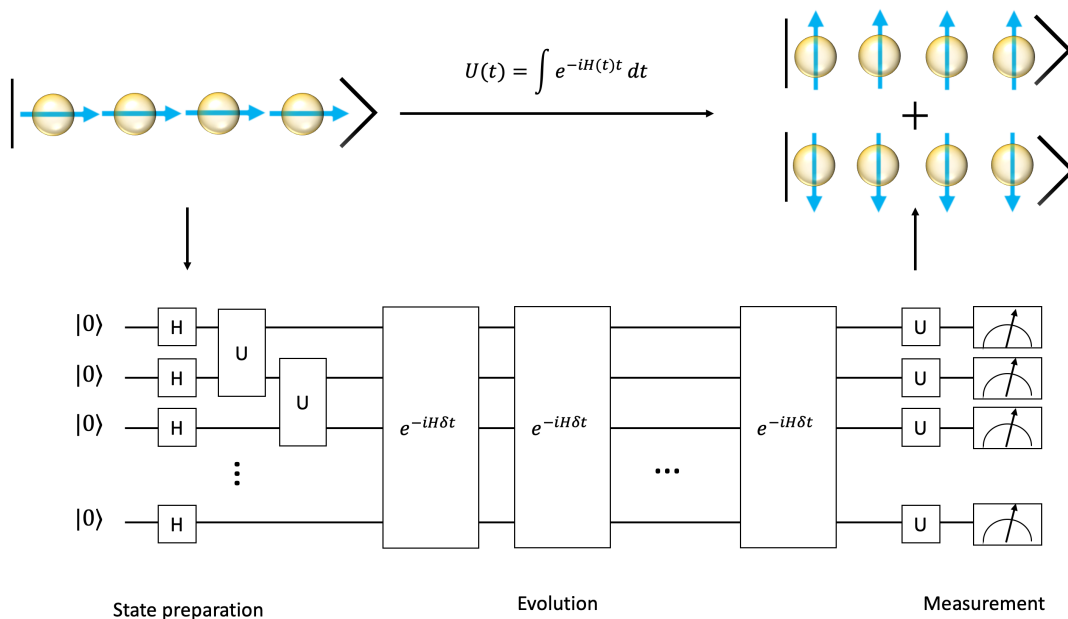
FIG. 2 Digital quantum simulation of time evolution of a spin chain, using a canonical Trotter-type method. We first map the system Hamiltonian, H_s, to a qubit Hamiltonian, H_q. Then the initial system wavefunction $|\psi_s^i\rangle$ is mapped to a qubit wavefunction $|\psi_q^i\rangle$. The time evolution 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.

a quantum computer, but are generally thought to be inefficient to simulate on a classical computer, this problems belongs to the computational complexity class BQP (bounded-error quantum polynomial-time). Further discussion of general Hamiltonian simulation methods is given in Sec. V. It was recently shown that time evolution can also be simulated using variational approaches (Li and Benjamin, 2017). This may enable simulation of time evolution using circuits with fewer gates than Trotterization. However, a variational circuit with fixed parameters is tailored to the time evolution of a specific initial state, in contrast to a Trotter circuit, which can be used to time evolve any valid initial state. Once the system has been time evolved for the desired duration, we can extract useful dynamical quantities from these simulations. Examples of such quantities include the electronic charge density distribution, or two-particle correlation functions (Abrams and Lloyd, 1997).

In chemistry, one is often concerned with determining whether two molecules will react with each other, when brought together with a certain energy. This can be determined by studying the chemical reaction dynamics of the system. One might assume that this could be studied by simply initialising the reactant molecules on the quantum computer, and time evolving under the system Hamiltonian, using the methods described above. However, whether this is possible depends on the method used to map the chemical problem onto the quantum computer. While this will be discussed in more detail in

Sec. III and Sec. IV, we briefly elaborate on this point here, as it is an important result. Kassal *et al.* (2008) showed that if a grid based method is used, then the electrons and nuclei can be treated on an equal footing, and the method described above can be used. This enables the explicit simulation of chemical reactions, using resources which scale polynomially with the system size. In contrast, if the problem is projected onto a finite basis set of electron spin-orbitals, the time evolution method described above cannot be used. This is because the spatial form of the wavefunction is fixed when setting up the problem, and so cannot change sufficiently during the simulation to accommodate molecular bonding. To study chemical reaction dynamics under this mapping, one must instead use a combination of quantum-calculated static properties, and classical methods.

We can obtain static properties by mapping the target wavefunction (such as the ground state wavefunction of the system) onto a qubit wavefunction. We can then use the quantum computer to calculate the expectation value of the desired observable, $\langle \psi_q | O_q | \psi_q \rangle$. In particular, Abrams and Lloyd showed that the phase estimation algorithm (Kitaev, 1995) can be used to find the energy of a quantum system (Abrams and Lloyd, 1999), and collapse it into the desired energy eigenstate. Finding the low lying energy levels of a quantum Hamiltonian is in general an exponentially difficult problem for classical computers. Moreover, it is important to note that solving the ground state problem for a general local Hamiltonian

belongs to the complexity class QMA (quantum Merlin-Arthur), and is in fact, QMA-complete (the quantum analogue of NP-complete) (Cubitt and Montanaro, 2016; Kempe *et al.*, 2006). Problems in this complexity class are not believed to be efficiently solvable with either a classical or quantum computer.

Despite this, the situation is not as bleak as it may initially seem. As stated in the introduction, we focus on finding the low lying energy levels of molecules. This is commonly referred to as the molecular electronic structure problem. It is widely believed that this problem should be efficiently solvable with a quantum computer (Whitfield *et al.*, 2013). The molecular electronic structure problem has received significant attention since it was first introduced in the context of quantum computational chemistry by Aspuru-Guzik *et al.* (2005), and is widely considered to be one of the first applications of quantum computing. Solving the electronic structure problem is often a starting point for more complex calculations in chemistry, including the calculation of reaction rates, the determination of molecular geometries, and calculations of optical properties.

Before discussing how the electronic structure 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 solve 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 basis set approaches. 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 Helgaker *et al.* (2014) and Szabo and Ostlund (2012) for further information.

A. The electronic structure problem

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

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

where M_I, $\mathbf{R}_I$, and Z_I denote the mass, position, and atomic number of the Ith nucleus, and $\mathbf{r}_i$ is the position of the ith 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

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

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

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

Our aim is to find energy eigenstates $|E_i\rangle$ and the corresponding energy eigenvalues E_i of the Hamiltonian H_e. In the rest of this review, 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 (Christiansen, 2012).

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 (Evans and Polanyi, 1935; Eyring, 1935)

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

where T is the temperature of the system, and ΔE is the energy difference between the reactant and product states. In computational chemistry, we are typically 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 review 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

It is essential to distinguish between simulations carried out in first quantisation, and those carried out in second quantisation. The distinguishing feature between these representations is how the antisymmetry of the wavefunction is stored. The electronic wavefunction must be antisymmetric under the exchange of any two electrons, as required by the Pauli exclusion principle.

As we will show in Sec. III.B.2, the second quantised method maintains the correct exchange statistics through the properties of the operators which are applied to the wavefunction. In contrast, first quantised methods explicitly retain the antisymmetry in the wavefunction. These differences will become more apparent in the context of quantum computational chemistry mappings, which we discuss in Sec. IV.

It is important to note that whether a simulation is carried out in first or second quantisation is a distinct choice from whether a 'basis set' or 'real space' (grid based) method is used. This will be elaborated on in more detail in the following sections. However, one key difference that is important to note here is as follows. Using a basis set is known as a 'Galerkin discretisation', which ensures that the energy converges to the correct value from above, as the number of basis functions tends to infinity. This property does not hold for grid based methods. A more detailed discussion on the differences between real space and basis set methods can be found in the main text and Appendix A of Babbush et al. (2018c).

1. First quantisation

Here, we focus on classical first quantised simulation methods. Discussion of first quantised chemistry simulation on quantum computers is postponed until Sec. IV.A. We only discuss first quantised real space methods in this section, as these are more common than first quantised basis set calculations in classical computational chemistry. First quantised basis set calculations are discussed in the context of quantum computational chemistry in Sec. IV.A.

a. Real space methods Here we consider the wavefunction in the position representation, $\{|\mathbf{r}\rangle\}$, which must be explicitly anti-symmetrised to enforce exchange symmetry (Abrams and Lloyd, 1997). 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 ith 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 discretised 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 amplitudes, 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 real space using a classical computer.

Grid based methods are very useful when considering chemical reaction 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. 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 may be better to treat the nuclei and electrons on an equal footing, which is best achieved with real space methods. This is discussed further in Kassal et al. (2008).

The real space method directly stores the wavefunction without fully exploiting our knowledge of the molecular orbitals. In contrast, basis set methods exploit our knowledge of the general spatial form of the orbitals. This dramatically reduces the resources needed to simulate molecules. We discuss basis set methods in the context of second quantisation in the next section.

2. Second quantisation

We only discuss second quantised basis set methods in this section, as these are more common than second quantised real space calculations in both classical and quantum computational chemistry.

a. Basis set methods 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 ith 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

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

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 (Szabo and Ostlund, 2012)

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

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 review as Slater determinants. This is common practice in computational chemistry (Szabo

and Ostlund, 2012).

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

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

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

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

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

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

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 (Helgaker *et al.*, 2014; Szabo and Ostlund, 2012).

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

with

$$
h_{pq} = \int d\mathbf{x}\, \phi_p^*(\mathbf{x}) \left(-\frac{\nabla^2}{2} - \sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} \right) \phi_q(\mathbf{x}),
$$
$$
h_{pqrs} = \int d\mathbf{x_1} d\mathbf{x_2} \frac{\phi_p^*(\mathbf{x_1})\phi_q^*(\mathbf{x_2})\phi_s(\mathbf{x_1})\phi_r(\mathbf{x_2})}{|\mathbf{x_1} - \mathbf{x_2}|}. \tag{26}
$$

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. We can define a suitable basis for this fictitious system as the set of molecular orbitals which diagonalise the non-interacting Hamiltonian. These molecular orbitals are typically linear combinations of the atomic orbitals. The energy eigenfunctions are anti-symmetrised tensor products of these single-particle molecular orbitals. This can be achieved by creating Slater determinants from the molecular orbitals.

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

$$|\psi\rangle = \sum_f \alpha_f |f\rangle, \qquad (27)$$

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 (Helgaker *et al.*, 2014; Szabo and Ostlund, 2012). 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 electrons, 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 the following section. Alternatively, one may consider only the most important MOs, which we discuss in Sec. III.D. We note that in practice, the molecular orbitals obtained by diagonalising the non-interacting part of the Hamiltonian will likely form a poor basis for the system. Instead, the Hartree-Fock procedure, described in Sec. III.C.1 can be used to obtain more suitable molecular orbitals.

C. Classical computation methods

In this section, we review four methods for approximating the ground state wavefunction with a restricted number of Slater determinants; the Hartree-Fock (HF), multi-configurational 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.

The methods discussed below are in the context of second quantised basis set calculations, as these translate most easily to the methods used in quantum computational chemistry. However, these methods can also be applied in first quantised or real space simulations. These methods are, in a sense, the most basic in classical computational chemistry. For a review of the more advanced methods developed recently for the electronic structure problem, see Cao *et al.* (2018b).

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 using 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 (Helgaker *et al.*, 2014)

$$\hat{f} = \sum_{i,j}(h_{ij} + V_{ij})a_i^\dagger a_j,$$
$$V_{ij} = \sum_{k \in occ}(h_{ikkj} - h_{ikjk}), \qquad (28)$$

where V_{ij} describes the effective potential, and *occ* is the set of occupied orbitals. We see that the Fock operator depends on the spatial form of the orbitals through h_{ij}, h_{ikkj}, and h_{ikjk} 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 con-

verge (Szabo and Ostlund, 2012). 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 term h_{ikkj} describes the Coulomb interaction of an electron with the charge distribution of the other electrons, while the term h_{ikjk} describes exchange effects (also called Fermi correlation) arising from the required antisymmetrisation. 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 (Helgaker et al., 2014).

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 (Helgaker et al., 2014).

2. Multiconfigurational self-consistent field

As discussed above, the HF method performs poorly for strongly correlated systems. 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 (Wang et al., 2008). 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 (Roos et al., 1980). MCSCF can be considered the best approximation to the exact wavefunction for a given number of determinants (Wang et al., 2008). It is not possible to perform an MCSCF calculation on systems with more than a few electrons, as the number of determinants scales exponentially with the number of electrons. 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 (Roos et al., 1980). 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 (Wang et al., 2008). However, they are the most effective methods at treating systems with strong static correlation (Helgaker et al., 2014).

3. Configuration interaction

The configuration interaction (CI) method generates a correlated wavefunction by considering excitations above a reference state, typically the HF state. The CI method is effective at recovering dynamic correlation, but less effective at recovering static correlation (Helgaker et al., 2014). If all determinants are included, we recover the full configuration interaction (FCI) wavefunction, generated by considering all excitations above the HF wavefunction

$$|\psi_{\mathrm{FCI}}\rangle$$
$$= \left(I + \sum_{i,\alpha} C_{i\alpha} a_i^\dagger a_\alpha + \sum_{i>j,\alpha>\beta} C_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta + ... \right) |\psi_{\mathrm{HF}}\rangle,$$
(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 (Helgaker et al., 2014; Szabo and Ostlund, 2012). 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') (Helgaker et al., 2014).

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. Like the CI method, the CC method is effective at recovering dynamic correlation, but not static correlation (Helgaker

et al., 2014). The CC wavefunction is given by

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

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

$$|\psi_{\text{CC}}\rangle = e^T |\psi_{\text{HF}}\rangle, \quad (31)$$

where $T = \sum_i T_i$,

$$T_1 = \sum_{i\in virt, \alpha\in occ} t_{i\alpha} a_i^\dagger a_\alpha,$$
$$T_2 = \sum_{i,j\in virt, \alpha,\beta\in occ} t_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta, \quad (32)$$
$$...,$$

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 intractable. As a result, the method is normally truncated at a lower excitation level; often single and double excitations (CCSD). Because of its product parametrisation, the CC 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 does not obey the Rayleigh-Ritz variational principle (Helgaker *et al.*, 2014). Furthermore, the CC method cannot be used with MCSCF states, and so struggles with systems displaying strong static correlation (Romero *et al.*, 2019). 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 (Romero *et al.*, 2019).

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 a Galerkin discretisation, 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 real space methods with a very fine grid spacing, which is only possible for small atoms or simple molecules. The orbital functions introduced in this section are approximations of these true orbitals.

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

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

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 (Griffiths, 2016). While these solutions are exact for one electron atoms, they perform poorly for many-electron atoms. They rapidly become diffuse, so cannot 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.

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)

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

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 (Helgaker *et al.*, 2014). Unlike the true atomic orbitals, these functions do not display oscillatory behaviour. Consequently, linear combinations of STOs 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)
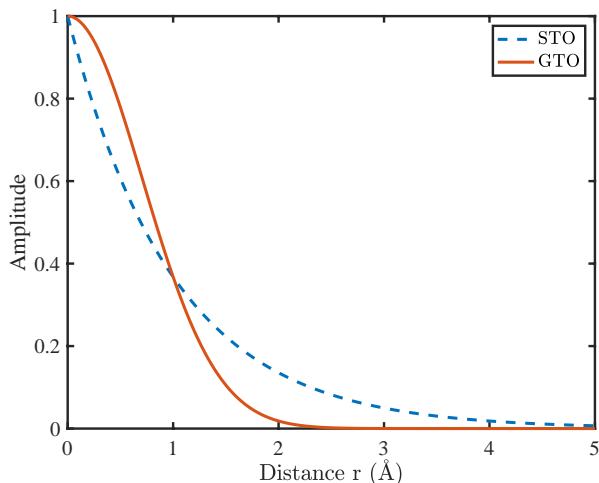
FIG. 3 Comparing the shapes of Slater (STO) and Gaussian (GTO) type orbitals. Normalisation factors are neglected.

computationally difficult. As a result, they are not used as basis functions in practice.

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 a Gaussian-type orbital (GTO) is given by

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

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 (Helgaker *et al.*, 2014).

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.

1. STO-nG basis sets

Some of the most simple bases are the STO-nG basis sets (Slater Type Orbital-n Gaussians) (Hehre *et al.*, 1969). In an STO-nG basis, each atomic orbital is considered to be an approximate STO. The STOs are approximated using n GTOs. STO-nG 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 (i.e. the energy obtained by performing a HF calculation using a grid based method, on an infinitely precise grid) 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.

2. Split-valence basis sets

Split-valence (or Pople (Ditchfield *et al.*, 1971)) 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 orbitals (the orbitals of the highest occupied energy level). 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 polar-

isation of electronic charge. Nevertheless, as Pople basis sets were designed for HF calculations, even large, polarised Pople basis sets are not well suited for correlated post-HF calculations (Helgaker *et al.*, 2014).

3. Correlation-consistent basis sets

Additional accuracy can be obtained by using cc-PVnZ basis sets (correlation consistent polarised valence n zeta), introduced by Dunning (Dunning Jr., 1989). 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 (Helgaker *et al.*, 2014). 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 energy level (the valence level) is the first 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.

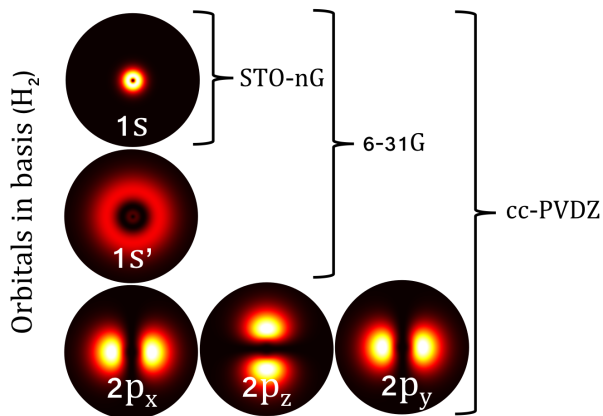For lithium in the cc-PVTZ basis ($n = 3$, T = triple),

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.

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 $n^3/3$ (Helgaker *et al.*, 2014). 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 (Helgaker *et al.*, 2014). 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.

4. 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 by Babbush *et al.* (2018c). The

plane wave basis functions, $\phi_\nu(r)$, are given by

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

for a wave with wavevector corresponding to the νth 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. This reduces the number of Hamiltonian terms from $O(M^4)$ to $O(M^2)$. Multiple benefits are obtained from this reduction – most notably, improved 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 (Babbush *et al.*, 2018c). A similar improvement in algorithmic scaling might be obtained using the recently proposed gausslet basis sets (White, 2017), 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 (Helgaker *et al.*, 2014)

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

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

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). The RDMS are defined with respect to a state which is an approximation of the ground state, which could be the results of a classically tractable configuration interaction or coupled cluster calculation. These RDMs contain all of the information required to evaluate $\langle O \rangle_\psi$. From the definition above, we can see that the diagonal elements of ρ^1 are the expectation values 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 Hempel *et al.* (2018) 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 set.

This part of the review 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 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. As discussed in Sec. III.B the distinguishing feature between first and second quantised methods is whether antisymmetry is enforced in the wavefunction directly (first quantised), or in the behaviour of the operators which act on the wavefunction (second quantised). As in the previous section, we consider a molecule with M spin-orbitals (when discussing basis set approaches) and N electrons.

We note that most of the work to date in quantum computational chemistry has focused on second quantised simulation methods. This is because, while first quantised simulations require asymptotically fewer qubits than second quantised simulations, for the smallest simulable systems (such as small molecules in minimal

basis sets), second quantised methods require either fewer qubits and/or shorter gate sequences. This has caused second quantised methods to become the *de facto* option for experimental demonstrations of quantum computational chemistry algorithms, due to the limits of current quantum hardware.

A. First quantised encoding methods

Here we give an overview of first quantised quantum simulation, which can be carried out using either a discrete single-particle basis, or real space methods.

1. Basis set methods

The original algorithm for simulating quantum systems in the first quantisation using a discrete basis was given by Abrams and Lloyd (1997). If we consider M single-particle basis functions (the molecular orbitals described in the previous section), we can enumerate these from 0 to $M - 1$. We can store these orbitals using $\log_2(M)$ qubits, denoting orbital 0 as $|0...0\rangle$, orbital 1 as $|0...01\rangle$ and so on, such that orbital $M - 1$ is represented as $|1...1\rangle$. We then use N registers of these $\log_2(M)$ qubits (one register for each electron) to describe the states of all of the electrons in the molecule. As a result, it requires $O(N\log_2(M))$ qubits to store the wavefunction.

If we consider a product state generated by each electron being in a single orbital, we observe that the wavefunction does not have the correct antisymmetry. As such, it must be antisymmetrised. The original algorithm (Abrams and Lloyd, 1997) accomplishes this in $O(N^2)$ gates. This has recently been dramatically improved upon by Berry *et al.* (2018), who used a circuit of depth $O(\log_2^c(N)\log_2\log_2(M))$, where $c \geqslant 1$ and depends on the choice of sorting network used.

The Hamiltonian can be obtained by projecting it onto the single-particle basis functions, and the wavefunction can then be time evolved under the Hamiltonian, which will maintain the correct antisymmetry (Abrams and Lloyd, 1997). A first quantised representation of the plane wave basis has recently been used to achieve the best scaling chemistry algorithm to date (Babbush *et al.*, 2018a). It requires $O(N\log_2(M))$ qubits to store the wavefunction, and can perform time evolution under the Hamiltonian with a gate count of $O(N^{8/3}M^{1/3}t)$, neglecting logarithmic factors. This is a substantial development, as it is the first quantum chemistry algorithm scaling sublinearly with the number of basis functions. This can be used to mitigate the drawback of plane wave basis functions; that more are required than gaussian basis functions to achieve accurate results.

2. Real space methods

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

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

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 by Eq. (38). This makes it efficient to represent quantum systems on quantum computers in real space. As discussed in Sec. III.B, simulation of reaction dynamics is more efficient in real space, without making the Born-Oppenheimer approximation, which necessitates costly classical precomputation (Kassal *et al.*, 2008).

Real space methods were first introduced for the quantum simulation of general quantum systems by Wiesner (1996) and Zalka (1998). They were then adapted for simulating problems in chemistry by Lidar and Wang (1999) and Kassal *et al.* (2008). The algorithm given by Kassal *et al.* (2008) proceeds as follows. The qubits are used to create a discretised grid, as described above. Physically relevant states can then be prepared using the algorithms outlined by Kassal *et al.* (2008) and Ward *et al.* (2009). The state can be propagated in time by repeatedly using the quantum Fourier transform to move between the position and momentum bases, so that the potential and kinetic terms are diagonal (respectively) and so are simple to apply. Finally, the relevant observables can be measured (Kassal *et al.*, 2008; Whitfield, 2015). A thorough investigation of the resources required to perform these simulations in a fault-tolerant manner was carried out by Jones *et al.* (2012).

The algorithm described above was improved upon considerably by Kivlichan *et al.* (2017), who also performed a more thorough analysis of both gate counts and errors. Their method discretises the kinetic and potential terms of the Hamiltonian, separates them into linear combinations of unitary operators, and applies the Taylor expansion method for simulating time evolution (Berry *et al.*, 2015a). The number of gates required scales as $O((N/h^2 + N^2)t)$, where h is the grid spacing and t is the simulation time. Although the spatial resolution of the grid increases exponentially with the number of qubits

used, it is not possible to use this as a route to exponentially improving the accuracy of the calculation. To see this, note that the gate count scales polynomially with the inverse grid spacing, so any attempts to exponentially increase the simulation accuracy by exponentially reducing the grid spacing will cause the gate count to increase exponentially. Kivlichan *et al.* (2017) also show that there exist systems where the grid spacing must decrease exponentially with the number of particles in the system to maintain constant accuracy. Consequently, these systems are not efficient to simulate using this method. However, those authors note that such pathological cases can also exist for basis set methods, but are typically dealt with efficiently using a clever choice of basis function.

The simulation of molecules in real space requires considerably more qubits than in basis set approaches (Kassal *et al.*, 2008, 2011; Kivlichan *et al.*, 2017; Ward *et al.*, 2009). For example, it would require around 100 logical qubits to simulate the lithium atom in real space (a task beyond classical computers using numerical grid based methods) (Kassal *et al.*, 2008). Consequently, real space approaches are typically considered unsuitable for near-term quantum computers, which will have small numbers of qubits.

B. Second quantised encoding methods

To simulate chemical systems in the second 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 Hilbert space of qubits, such that every fermionic state can be represented by a qubit state. There are multiple methods of encoding, which we describe below. In the following section, we only discuss second quantised basis set methods, as second quantised grid based methods have only briefly been discussed in the context of second quantised quantum computational chemistry (see Babbush *et al.* (2018c) Appendix A).

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 formally,

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

The fermionic creation and annihilation 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 preserve 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 or $Q^\dagger$ 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.

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 a qubit mapped fermionic operator generally has a weight of $O(M)$ Pauli operators, each acting on a different qubit.

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 electrons, $N < M$. 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 (Aspuru-Guzik *et al.*, 2005). 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 second quantised encodings.

2. Parity encoding

In the parity encoding, instead of directly storing the occupation number, one can instead use the pth qubit to store the parity of the first p modes (Seeley *et al.*, 2012),

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

The creation and annihilation operators are

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

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

TABLE I Example mappings of a fermionic Fock state and its fermionic operators onto the corresponding qubit state, and qubit operators. $\hat{n}_i$ is the fermionic number operator.

Fermion	Jordan-Wigner	Parity	Bravyi-Kitaev
$a\left\|0001\right\rangle + b\left\|0010\right\rangle$ $+c\left\|0100\right\rangle + d\left\|1000\right\rangle$	$a\left\|0001\right\rangle + b\left\|0010\right\rangle$ $+c\left\|0100\right\rangle + d\left\|1000\right\rangle$	$a\left\|1111\right\rangle + b\left\|1110\right\rangle$ $+c\left\|1100\right\rangle + d\left\|1000\right\rangle$	$a\left\|1011\right\rangle + b\left\|1010\right\rangle$ $+c\left\|1100\right\rangle + d\left\|1000\right\rangle$
a_0	Q_0	$X_3 X_2 X_1 Q_0$	$X_3 X_1 Q_0$
a_1	$Q_1 Z_0$	$X_3 X_2 \left(Q_1 \left\|0\right\rangle \left\langle 0\right\|_0 - Q_1^\dagger \left\|1\right\rangle \left\langle 1\right\|_0 \right)$	$X_3 \left(Q_1 \left\|0\right\rangle \left\langle 0\right\|_0 - Q_1^\dagger \left\|1\right\rangle \left\langle 1\right\|_0 \right)$
a_2	$Q_2 Z_1 Z_0$	$X_3 \left(Q_2 \left\|0\right\rangle \left\langle 0\right\|_1 - Q_2^\dagger \left\|1\right\rangle \left\langle 1\right\|_1 \right)$	$X_3 Q_2 Z_1$
a_3	$Q_3 Z_2 Z_1 Z_0$	$Q_3 \left\|0\right\rangle \left\langle 0\right\|_2 - Q_3^\dagger \left\|1\right\rangle \left\langle 1\right\|_2$	$\frac{1}{2} \left(Q_3 (\mathbf{1} + Z_2 Z_1) - Q_3^\dagger (\mathbf{1} - Z_2 Z_1) \right)$
$\hat{n}_i = a_i^\dagger a_i$	$\left\|1\right\rangle \left\langle 1\right\|_i$	$\left\|1\right\rangle \left\langle 1\right\|_{i=0}$, $\frac{1}{2} \left(\mathbf{1} - Z_i Z_{i-1} \right)_{i=1,2,3}$	$\left\|1\right\rangle \left\langle 1\right\|_{i=0,2}$, $\frac{1}{2} (\mathbf{1} - Z_1 Z_0)_{i=1}$, $\frac{1}{2} (\mathbf{1} - Z_3 Z_2 Z_1)_{i=3}$

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).

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.

$$P \cdot (a\left|0001\right\rangle + b\left|0010\right\rangle + c\left|0100\right\rangle + d\left|1000\right\rangle)$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \cdot \left(a\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + d\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right),$$

$$= \left(a\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + c\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} + d\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right),$$

$$= a\left|1111\right\rangle + b\left|1110\right\rangle + c\left|1100\right\rangle + d\left|1000\right\rangle. \tag{43}$$

In contrast to the JW encoding, the parity mapping stores the parity information locally, and the occupation number non-locally. Like the JW mapping, each mapped fermionic operator has a weight of $O(M)$ Pauli operators, each acting on a different qubit.

3. Bravyi–Kitaev encoding

The Bravyi–Kitaev (BK) encoding (Bravyi and Kitaev, 2002) 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}.

$$\left|f_{M-1}, f_{M-2}, \dots, f_0\right\rangle \to \left|q_{M-1}, q_{M-2}, \dots, q_0\right\rangle,$$

$$q_p = \left[\sum_{q=0}^{p} \beta_{pq} f_q \right] \pmod{2}. \tag{44}$$

It is defined recursively (Bravyi and Kitaev, 2002; Seeley et al., 2012) via

$$\beta_1 = [1],$$

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

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 $M = 4$ ($x = 1$), the matrix β_{pq} is

$$\beta_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}. \tag{46}$$

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 Table I for an example and works by Seeley et al. (2012) and Tranter et al. (2015) for a complete presentation 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 in general scales

as $O(\log_2 M)$. A thorough comparison of the BK and JW mappings was recently performed by Tranter *et al.* (2018) for 86 molecular systems. They found that the BK transform was in general at least as efficient as the JW transform, and was in many cases considerably more so.

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 (Havlicek *et al.*, 2017). 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, in general. 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 (Sung, 2018). 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 (Steudtner and Wehner, 2018a). In some cases, the number of qubits saved can be exponential in M, although this introduces the need for $O(M)$-controlled gates.

In contrast to these non-linear encodings, other mappings have been developed which exchange an increased number of qubits for a lower gate count. Verstaete and Cirac (2005) 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. This is achieved by doubling the number of qubits. Similar ideas were introduced by Ball (2005) and Farrelly and Short (2014). These ideas were generalised by Whitfield *et al.* (2016), who refer to these mappings as 'auxiliary fermion transforms'. They too introduce ad-

ditional qubits (at most doubling the number of qubits) to map local fermionic operators to local qubit operators. Related techniques have recently been introduced by Steudtner and Wehner (2018b).

There is also another variant of the BK transform, known as the Bravyi-Kitaev superfast transform (Bravyi and Kitaev, 2002). 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. We refer the reader to (Setia and Whitfield, 2018) for a detailed discussion of the BKSF transform. Recently, generalised superfast encodings (GSE) have been proposed (Setia *et al.*, 2018), which introduce additional benefits to the mapping. The first GSE reduces the weight of each of the Pauli operators in the Hamiltonian to $O(\log d)$, while the second GSE provides the possibility of correcting any single qubit error, as long as $d \geq 6$. A related mapping, known as the majorana loop stabilizer code, was recently proposed by Jiang *et al.* (2018a). It is capable of correcting any single qubit error on a square lattice (for which $d = 4$), and has lower weight operators than those of the second GSE.

C. Hamiltonian reduction

In this section, we focus on methods used to reduce the number of qubits required for the second quantised approach, using $\mathbb{Z}_2$ symmetries. More general qubit reduction schemes have also been developed (Bravyi *et al.*, 2017), but these have yet to be numerically or experimentally investigated. Alternatively, one may make use of quantum autoencoders, which can compress the wavefunction into a smaller Hilbert space (Romero *et al.*, 2017).

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. Here, we will describe the method by Bravyi *et al.* (2017), 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 (with the caveat that the number of orbitals must be a power of two), 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. As can be seen for the BK matrix given in Eq. IV.B.3, every element in the final row is one, and the first half of the elements in the $M/2^{\text{th}}$ row are also one.

Consequently, the final element of the vector encoded by this matrix, q_{M-1}, is equal to the number of electrons (mod 2). Similarly, the $M/2^{\text{th}}$ element in the 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 parity 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 that can be found. 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 (Wang *et al.*, 2018; Yung *et al.*, 2014) which occupy the intermediate region in the quantum computational chemistry timeline. Moreover, algorithms to find the ground state using methods which differ from both QPE and VQE have also been devised (Ge *et al.*, 2017).

A. Quantum phase estimation

1. Implementation

Phase estimation (Kitaev, 1995) can be used to find the lowest energy eigenstate, $|E_0\rangle$, of a Hamiltonian (Abrams and Lloyd, 1999). 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 (Nielsen and Chuang, 2002), and shown in Fig. 5. The number of ancilla qubits required for phase estimation is determined by the desired success probability and precision in the energy estimate. Nielsen and Chuang (2002) show that to obtain a binary estimate of the energy, precise to n bits, with success probability p, requires

$$\omega = n + \lceil \log_2(2 + \frac{1}{2p}) \rceil \qquad (47)$$

ancilla qubits.

1. We initialise the qubit register in state $|\psi\rangle$, which has non-zero overlap with the true FCI ground state of the molecule. We require an additional register of ω ancilla qubits. We can expand the state $|\psi\rangle$ in terms of energy eigenstates of the molecular Hamiltonian, writing that $|\psi\rangle = \sum_i c_i |E_i\rangle$, where c_i are complex coefficients.

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

$$\frac{1}{\sqrt{2^\omega}} \sum_i \sum_x |x\rangle c_i |E_i\rangle \rightarrow \frac{1}{\sqrt{2^\omega}} \sum_i \sum_x e^{-2\pi i E_i x} c_i |x\rangle |E_i\rangle. \qquad (48)$$

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

$$\frac{1}{\sqrt{2^\omega}} \sum_i \sum_x e^{-2\pi i E_i x} c_i |x\rangle |E_i\rangle \xrightarrow{QFT^{-1}} \sum_i c_i |\text{bin}(E_i)\rangle |E_i\rangle. \qquad (49)$$

4. We measure the ancilla qubits in the Z basis, which gives the ground state energy eigenvalue as a binary bit-string $\text{bin}(E_0)$, accurate to n bits with probability $p \times |c_0|^2$. This collapses the main register into the corresponding energy eigenstate, $|E_0\rangle$.
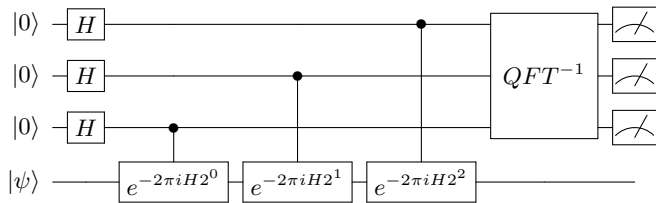
FIG. 5 The canonical quantum phase estimation circuit with three ancilla qubits. When the ancilla qubits are in state $|x\rangle$, a control rotation $e^{-2\pi i H x}$ is applied to the target state $|\psi\rangle$. QFT denotes the quantum Fourier transform (Nielsen and Chuang, 2002). By measuring the ancilla qubits in the computational basis, they collapse to an eigenvalue of H and the register qubits collapse to the corresponding energy eigenstate.

To realise the standard phase estimation algorithm given above, we sequentially need to time evolve the main register under the Hamiltonian H for times $t_0 = 2\pi, t_1 = 4\pi, ..., t_{\omega-1} = 2^\omega \pi$. The total coherent time evolution, T, is then given by approximately $T = 2^{\omega+1}\pi$. For a success probability of $p = 0.5$, we require $\omega = n + 2$ ancilla qubits. The total evolution time can be related to the binary precision $\epsilon = 1/2^n$, to show that $T = 8\pi/\epsilon$. Given that our success probability for this estimate is 0.5, we expect to have to repeat the procedure twice to obtain a good estimate of the ground state. This is equivalent to a total of $16\pi/\epsilon$ calls to the unitary e^{-iH} (Reiher *et al.*, 2017).

The basic phase estimation algorithm described above can be improved in many ways. It can be modified to use only a single ancilla qubit, which is used to measure each bit in the energy eigenvalue sequentially (Aspuru-Guzik *et al.*, 2005). It can also be made more efficient (Svore *et al.*, 2013), parallelised (Knill *et al.*, 2007; Reiher *et al.*, 2017), or made more resilient to noise using time-series analysis (O'Brien *et al.*, 2018). Recent work has further improved upon the asymptotic scaling of phase estimation by using classically obtainable knowledge about the energy gap between the ground and first excited states (Berry *et al.*, 2018). The ultimate limit for the number of calls required to e^{-iH} is π/ϵ, which is approximately obtained using Bayesian approaches (Berry *et al.*, 2009; Paesani *et al.*, 2017b; Wiebe and Granade, 2016). For the case of molecular simulation Reiher *et al.* (2017) show that a number of calls scaling as $\pi/2\epsilon$ will suffice.

Regardless of which version of phase estimation is used, there are two universal features. Firstly, it is necessary for the register to initially be in a state with a non-zero overlap with the ground state. Secondly, we must have a way to coherently implement time evolution under the Hamiltonian H. We will discuss ways to satisfy both of these requirements in the following two sections.

2. State preparation

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 a randomly chosen state would have vanishing probability of collapsing to the desired ground state, as the system size increases. Moreover, 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 (McClean *et al.*, 2014). Several techniques have been proposed for state preparation, including: specific chemical state preparation routines (Babbush *et al.*, 2015; Sugisaki *et al.*, 2016, 2018; Tubman *et al.*, 2018; Wang *et al.*, 2008), using the variational methods discussed in the next section (Yung *et al.*, 2014), quantum cooling (Xu *et al.*, 2014), and most commonly, adiabatic state preparation (Aspuru-Guzik *et al.*, 2005; Veis and Pittner, 2014). 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 time evolve the system under a Hamiltonian that changes slowly 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 may be 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 $H(t) = (1-t/T)H_0 + (t/T)H_s$, where t is the time and T is the maximum desired simulation time (Aspuru-Guzik *et al.*, 2005; Veis and Pittner, 2014). The maximum annealing time, T, is given by

$$T \approx O(\frac{M^4}{\min_s \Delta(s)}), \qquad (50)$$

where $\Delta(s) = E_1(s) - E_0(s)$ and M is the number of spin-orbitals in the molecule, although the true scaling may be closer to $O(M^2/\min_s \Delta(s))$, using physical arguments (Reiher *et al.*, 2017). It is difficult to know the size of the gap along the entire adiabatic path *a priori*, which is potential problem for applying ASP.

There are a variety of methods that can be used to evolve the system under this time-dependent Hamiltonian, which are discussed below.

TABLE II Asymptotic query complexities of quantum simulation algorithms as a function of the simulation time t, simulation error ε, and the Hamiltonian sparsity d.

Algorithm	Query complexity (d, t, ε)
Product formulae	$O\left(\text{poly}(d)\frac{t^2}{\varepsilon}\right)$
Taylor series expansion	$O\left(d^2 t \frac{\log(d^2 t/\varepsilon)}{\log\log(d^2 t/\varepsilon)}\right)$
Qubitization + Quantum signal processing	$O\left(dt + \frac{\log(1/\varepsilon)}{\log\log(1/\varepsilon)}\right)$

3. Hamiltonian simulation

As discussed above, both QPE and adiabatic state preparation require implementation of the time evolution operator, e^{-iHt}, where H may or may not be time dependent. There are several ways to do this, each with their own advantages and disadvantages.

a. Product formulae

The most simple method has already been described in Sec. II; Trotterization. For a time-independent Hamiltonian H, if it can be decomposed as $H = \sum_i h_i$, where h_i are local Hamiltonians, then a first order Lie-Trotter-Suzuki approximation (Trotter, 1959) of the time evolution is

$$e^{-iHt} = \left(\prod_i e^{-ih_i t/S}\right)^S + O(t^2/S). \quad (51)$$

This approach is also referred to as the 'product formulae' method. In practice, to achieve accuracy ε, the number of Trotter steps $S = O(t^2/\varepsilon)$ should be large in order to suppress the errors in the approximation. This is effectively a stroboscopic evolution under time evolution operators corresponding to each of the terms in the Hamiltonian. It is also possible to use higher order product formulae (Berry *et al.*, 2007; Dür *et al.*, 2008; Suzuki, 1976), which scale better with respect to the simulation error than the first order method. Recently, randomisation procedures have been shown to improve the accuracy obtained using product formulae (Campbell, 2018; Childs *et al.*, 2018).

Product formulae can also be used to simulate dynamics under a time dependent Hamiltonian, $H(t)$. Wiebe *et al.* (2011) showed that the accuracy of such simulations depends on the derivatives of the Hamiltonian (although this dependence may be alleviated by incorporating randomisation procedures (Poulin *et al.*, 2011)).

As discussed above, the error in the simulation is determined by the Trotter order used, and the number of Trotter steps. However, it is important to note that the gate counts of all product formula based methods

scale as $O(\text{poly}(1/\epsilon))$. Recent work (Endo *et al.*, 2018b) investigated suppressing the error arising from using a finite number of Trotter steps, by using extrapolation (using extrapolation for physical error mitigation will be described in more detail in Sec. VI). Those authors performed simulations with several different numbers of Trotter steps, and then extrapolated their values to obtain an estimate of the result in the limit of an infinite number of Trotter steps.

b. Advanced Hamiltonian simulation methods

Alternative methods have been introduced which can realise the time evolution operator more efficiently than Trotterization, including; quantum walk based methods (Berry *et al.*, 2015b; Childs and Kothari, 2011), Taylor series expansions (Berry and Childs, 2012; Berry *et al.*, 2015a,c) or Chebyshev polynomial approximations (Subramanian *et al.*, 2018), and qubitization (Low, 2018; Low and Chuang, 2016) in conjunction with quantum signal processing (Low and Chuang, 2017; Low *et al.*, 2016). These methods make use of oracle access to the Hamiltonian and provide improved efficiency when simulating Hamiltonians which are d-sparse (they have at most d non-zero elements in each row and column, where d is a polylogarithmic function of the matrix dimension). The Taylor series and qubitization based methods also scale exponentially better with regards to the accuracy of the simulation than product formula based methods. We refer the readers to work by Childs *et al.* (2017) and the review of Cao *et al.* (2018b) for a summary of recent progress in Hamiltonian simulation and a comparison of the different methods.

4. Implementing time evolution for chemistry simulation

Both product formulae and more advanced methods of Hamiltonian simulation have been considered in the context of solving problems in quantum computational chemistry. Once again, the number of spin-orbitals in the molecule, or spin-sites in a lattice is given by M, and the number of electrons is given by N.

Early works on finding the ground state of molecular systems using phase estimation used first and second order product formalae (Aspuru-Guzik *et al.*, 2005; Babbush *et al.*, 2015; Hastings *et al.*, 2015; Kassal *et al.*, 2008; Poulin *et al.*, 2015; Reiher *et al.*, 2017; Wecker *et al.*, 2014; Whitfield *et al.*, 2011). A series of improvements throughout these papers reduced the scaling of phase estimation for molecules from $O(M^{11})$ (Wecker *et al.*, 2014) to (empirically) $O(M^5)$ (Babbush *et al.*, 2015). Additional benefits have recently been obtained using plane wave based basis sets, which reduces the scaling to at most $O(M^{7/2}t^{3/2})$ (Babbush *et al.*, 2018c).

As would be expected, the advanced Hamiltonian simulation algorithms have led to a reduction in the asymptotic scaling of chemistry algorithms. Babbush *et al.* made use of the Taylor series method (Berry *et al.*, 2015a) to produce algorithms which simulate time evolution under a molecular Hamiltonian scaling as $O(M^5t)$ (Babbush *et al.*, 2016) and $O(N^2M^3t)$ (Babbush *et al.*, 2017) (both excluding logarithmic factors). The algorithms scale exponentially better in the simulation error than Trotter based algorithms. The first result uses the second quantised representation of the Hamiltonian. The second, more efficient result is obtained by writing the Hamiltonian in the basis of Slater determinants, where it is known as the CI matrix, which was initially suggested as the basis for a quantum algorithm by Toloui and Love (2013). While this matrix has an exponential number of elements, it has a sparsity of $O(N^2M^2)$ when the number of orbitals is much larger than the number of electrons (the continuum limit, required for high accuracy calculations). The Slater rules (Helgaker *et al.*, 2014) can be used to determine where the non-zero elements are located. The benefit obtained from using the CI matrix is both a reduced gate count, and fewer qubits (a reduced spatial complexity). Because the particle number in each Slater determinant is fixed, only $O(N\log_2(M))$ qubits are required (plus additional ancilla qubits). By constructing oracle circuits to find the non-zero elements, the reported algorithmic scalings can be obtained. These algorithms also require calculation of the molecular integrals on the quantum computer, which are performed efficiently by exploiting an analogy between the discretisation of space in Riemannian integration and the discretisation of time in the Taylor series method to simulate a time dependent Hamiltonian (Babbush *et al.*, 2016). Once again, the introduction of the plane wave and plane wave dual basis sets made both of these algorithms more efficient to perform, resulting in an asymptotic depth scaling of $O(M^{8/3}t)$ (Babbush *et al.*, 2018c).

The Taylor series method is also used in the first quantised real space algorithm of Kivlichan *et al.* (2017) to obtain a gate scaling of $O((N/h^2 + N^2)t)$ (excluding logarithmic factors), where h is the grid spacing. It was also used in its time dependent form for simulation in the 'interaction picture', which enables more efficient time evolution in a plane wave basis, scaling as $O(M^2t)$ (Low and Wiebe, 2018) (although this comes at a cost of $O(M\log_2(M))$ spatial complexity). A similar interaction picture method was used in the first quantised, plane wave basis method discussed in Sec. IV.A, which achieved a gate scaling of $O(N^{8/3}M^{1/3}t)$ and a spatial complexity of $O(N\log_2(M))$ (Babbush *et al.*, 2018a).

The recently proposed quantum signal processing and qubitization methods have also been applied to quantum computational chemistry simulation. Berry *et al.* (2018) showed how qubitization can be used to perform the time evolution required for phase estimation with zero error. Recent work has shown that by implementing $e^{-if(H)t}$ in phase estimation (where $f(H)$ is an efficiently invertible function of the Hamiltonian), the errors that arise from trying to approximate e^{-iHt} can be circumvented (Berry *et al.*, 2018; Poulin *et al.*, 2018). These techniques, together with qubitization and the plane wave dual basis, were used by Babbush *et al.* (2018b), who developed new techniques and performed a resource analysis for fault-tolerant implementations of phase estimation. When performing a fault-tolerant resource estimation, the quantity of interest is the number of T gates, as these are the most costly gates to implement for error correcting schemes such as the surface code. As such, the scaling reported below is not comparable to the other scalings reported in this review, which describe primitive gate scalings, without considering the overhead of error correction. As part of that work, Babbush *et al.* (2018b) have carried out T gate estimations for many of the papers described in this review. Their algorithm results in a T gate scaling of $O((M^3 + M^2log(1/\epsilon))/\epsilon)$, which is asymptotically better than all previous results. We will discuss the cost (in terms of wall time and number of physical qubits) of applying this method to problems of interest in Sec. VIII.

Recently, new procedures have been developed for time evolution under lattice Hamiltonians. These Hamiltonians have geometrically local interactions of qubits that are laid out on a lattice. They relate to systems of interest in condensed matter physics and chemistry, such as the Fermi-Hubbard model, a prototypical model for high temperature superconductivity. For example, a qubit lattice Hamiltonian is obtained for the Fermi-Hubbard model when using a locality preserving mapping (Jiang *et al.*, 2018a; Verstaete and Cirac, 2005). Haah *et al.* (2018) made use of arguments about the speed of information propagation in lattice systems (Lieb and Robinson, 1972; Osborne, 2006) to obtain a simulation algorithm requiring $O(Mt\,\text{polylog}(Mt/\varepsilon))$ gates to simulate time evolution under lattice Hamiltonians. Similar results were recently obtained by Childs and Su (2019), who proved that a kth order product formula can simulate time evolution of an M qubit lattice Hamiltonian using $O((Mt)^{1+\frac{1}{k}}/\epsilon^{\frac{1}{k}})$ elementary operations. These algorithms are almost optimal in terms of gate complexity.

While the advanced Hamiltonian simulation methods described above are asymptotically more efficient than Trotterization, the Trotter error bounds appear to be loose by several orders of magnitude (Babbush *et al.*, 2015; Poulin *et al.*, 2015). A recent study of spin Hamiltonians (Childs *et al.*, 2017) found that while the asymptotic scaling of Trotter methods was much worse than the

qubitization and Taylor series methods, when numerical simulations were performed, Trotter methods required lower gate counts than the other methods. Moreover, the Taylor series method requires elementary logic operations, resulting in a large T gate count when considering fault-tolerant approaches (although recent work by Sanders et al. (2019) may help to alleviate this problem). However, the Taylor series method is required to implement the most efficient time dependent techniques, such as the interaction picture. As such, it is not yet possible to say which method will perform best for chemical systems.

Despite this progress, all of the methods discussed above require circuits with a large number of gates. As a result, these methods are typically assumed to require fault-tolerance (Jones et al., 2012). As near-term quantum computers will not have enough physical qubits for error correction, the long gate sequences required by these algorithms make them unsuitable for near-term hardware. Consequently, alternative methods are required for near-future chemistry simulation.

B. Variational algorithms

A highly promising chemistry algorithm for the NISQ era is the variational quantum eigensolver (VQE), first proposed and experimentally realised by Peruzzo et al. (2014), and elaborated on by McClean et al. (2016). 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 (Colless et al., 2018; Ganzhorn et al., 2018; Hempel et al., 2018; Kandala et al., 2017; O'Malley et al., 2016; Peruzzo et al., 2014; Santagati et al., 2018; Shen et al., 2017), and shows many desirable features. It appears to be robust against errors (McClean et al., 2016; O'Malley et al., 2016), and capable of finding the ground state energies of small molecules using low depth circuits (Kandala et al., 2017). Despite the successes of the VQE, several challenges remain – most notably the difficulty of classical optimisation, the large number of measurements required, the construction of suitable quantum circuits, and mitigating the effects of noise. To date, the VQE has only been applied to second quantised basis set simulations, and so our discussion of it will only be concerned with that scenario.
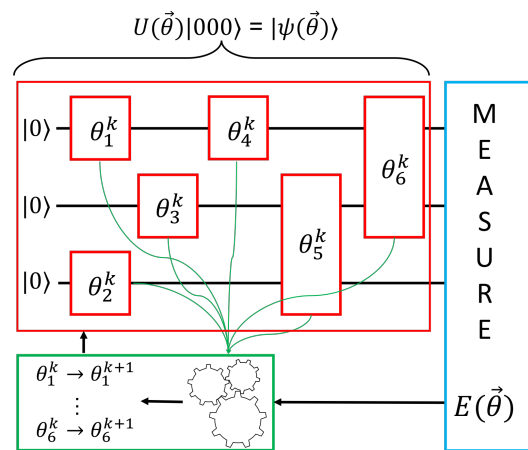

FIG. 6 A schematic of the variational quantum eigensolver (VQE). The classically intractable state preparation and measurement subroutines (red and blue) are performed on the small quantum computer. The current energy and parameter values are fed into a classical optimisation routine (green), which outputs 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 (Sakurai and Napolitano, 2017). This states that for a parametrised trial wavefunction $|\psi(\vec{\theta})\rangle$

$$\langle\psi(\vec{\theta})| H |\psi(\vec{\theta})\rangle \geq E_0, \tag{52}$$

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 (Babbush et al., 2015; Dallaire-Demers et al., 2018; Sugisaki et al., 2016, 2018; Tubman et al., 2018; Wang et al., 2008) describing the chemical system of interest

$$|\psi_{\text{ref}}\rangle = U_{\text{prep}}|\bar{0}\rangle. \tag{53}$$

A series of parametrised gates $U(\vec{\theta}) = U_N(\theta_N)\ldots U_k(\theta_k)\ldots U_1(\theta_1)$ are then applied to the

qubits. Here, $U_k(\theta_k)$ is the kth single or two qubit unitary gate, controlled by parameter θ_k. This circuit generates the trial wavefunction

$$|\psi(\vec{\theta})\rangle = U(\vec{\theta})|\psi_{\text{ref}}\rangle. \quad (54)$$

We refer to $|\psi(\vec{\theta})\rangle$ as the ansatz state, and $U(\vec{\theta})$ as the ansatz circuit. However, the reader will find that the word ansatz is typically used interchangeably to describe both. The set of all possible states that can 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 in the second quantised basis set approach can be mapped to a linear combination of products of local Pauli operators, using the transforms introduced in Sec. IV.B,

$$H = \sum_j h_j \prod_i \sigma_i^j, \quad (55)$$

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

$$E(\vec{\theta}_k) = \sum_j^N h_j \langle\psi(\vec{\theta}_k)| \prod_i \sigma_i^j |\psi(\vec{\theta}_k)\rangle. \quad (56)$$

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 (McClean *et al.*, 2014), and requires $O(1/\epsilon^2)$ measurements to determine the energy to a precision ϵ (McClean *et al.*, 2016; Romero *et al.*, 2019). Specifically, it can be shown that the number of measurements required scales as $O(M^6/\epsilon^2)$ in a gaussian orbital basis, and $O(M^4/\epsilon^2)$ for a plane wave dual basis (Babbush *et al.*, 2018c; Cao *et al.*, 2018b; Mc-Clean *et al.*, 2014). The cost of measurement can be reduced by a suitable grouping of Hamiltonian into terms that can be simultaneously measured (Izmaylov *et al.*, 2018; Kandala *et al.*, 2017; McClean *et al.*, 2016). As the quantum computer is reinitialised for each repetition, the required coherence time is dramatically reduced compared to quantum phase estimation.

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 have been in use since the first VQE experiment by Peruzzo *et al.* (2014). They were also independently introduced by Farhi *et al.* (2014) for the Quantum Approximate Optimisation Algorithm (QAOA), 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.* (2017).

Recent work by McClean *et al.* (2018) 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.

Recent work (Barkoutsos *et al.*, 2018) has improved the suitably of hardware efficient ansatze for chemistry problems, by considering gates which conserve particle number, and so permit the use of a Hartree–Fock initial state. This proposal has been experimentally demonstrated (Ganzhorn *et al.*, 2018) on a superconducting system for which the two-qubit number conserving entangling gates were natural operations.

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 can be extended to produce the unitary coupled cluster (UCC) ansatz (Bartlett *et al.*, 1989; Hoffmann and Simons, 1988). The UCC method creates a parametrised trial state by considering excitations above the initial reference state, and can be written as

$$U(\vec{\theta}) = e^{T-T^\dagger}, \qquad (57)$$

where $T = \sum_i T_i$, and

$$
\begin{aligned}
T_1 &= \sum_{i\in virt, \alpha\in occ} t_{i\alpha} a_i^\dagger a_\alpha, \\
T_2 &= \sum_{i,j\in virt, \alpha,\beta\in occ} t_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta, \qquad (58) \\
&...
\end{aligned}
$$

and *occ* are occupied orbitals in the reference state, and *virt* are orbitals that are initially unoccupied in the reference state. The UCC method is intractable on classical computers, but can be efficiently implemented on a quantum computer. It was originally proposed for quantum computational chemistry by Yung *et al.* (2014) and Peruzzo *et al.* (2014). A comprehensive review of the UCC method for quantum computational chemistry is given by Romero *et al.* (2019).

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 multireference ground states. As with the CC method, the UCC ansatz is typically truncated at a given excitation level – usually single and double excitations (known as UCCSD). We show a canonical implementation of the UCC ansatz (as described by (Romero *et al.*, 2019)) in Sec. VII. The formal gate scaling of this implementation of the UCC ansatz is approximately $O(M^3N^2)$ (where M is the number of spin-orbitals, and N is the number of electrons) when using the Jordan-Wigner mapping (Romero *et al.*, 2019). This assumes that a single Trotter step can be used, which appears to yield accurate results (Barkoutsos *et al.*, 2018; Romero *et al.*, 2019). The number of parameters required scales formally as $O(M^2N^2)$. 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 (Hempel *et al.*, 2018). Moreover, we can use classically tractable methods to get initial approximations for the remaining

non-zero parameters (O'Malley *et al.*, 2016; Romero *et al.*, 2019), which makes the classical optimisation step of the VQE easier. In addition, recent work by Motta *et al.* (2018) has reduced the gate scaling of the UCC ansatz to $O(M^4)$ with increasing molecular size, and $O(M^3)$ for a fixed molecular size and increasing basis set size.

Alternative (often heuristic) variants of the UCC ansatz have also been proposed for solving problems in quantum computational chemistry. These include: the Bogoliubov-UCC ansatz (Dallaire-Demers *et al.*, 2018) (a quasiparticle variant of UCC that is suitable for more general Hamiltonians than the UCC ansatz, potentially including pairing terms (superconductivity) or three body terms (nuclear physics)), the 'low-depth circuit ansatz' (Dallaire-Demers *et al.*, 2018) (which attempts to mimic the aforementioned BUCC ansatz using a circuit scaling linearly in the number of qubits), qubit coupled-cluster (Ryabinkin *et al.*, 0) (which produced far lower gate counts than the UCC ansatz when applied to several small molecules), and generalised-UCC (Lee *et al.*, 0) (which constructs a powerful heuristic ansatz out of a coupled-cluster ansatz which considers pairwise electron excitations).

c. Hamiltonian variational

There are also variational ansatze that lie between the two extremes described above. One important example is the Hamiltonian variational ansatz (also commonly referred to as a Trotterized adiabatic state preparation (TASP) ansatz), proposed by Wecker *et al.* (2015a). This ansatz was inspired by adiabatic state preparation and the QAOA. The idea is to Trotterize an adiabatic path to the ground state, using a number of Trotter steps that would be insufficient for accurate results. One can then variationally optimise the Trotter evolution times to create a variational ansatz for the ground state. The number of parameters in this ansatz scales linearly with the number of Trotter steps, S.

The introduction of the plane wave based basis sets (Babbush *et al.*, 2018c) has made the Hamiltonian variational algorithm more efficient to implement. The recent work of Kivlichan *et al.* (2018) showed that for Hamiltonians in this basis, we can implement Trotter steps in depth $O(M)$, using $O(M^2)$ two qubit gates (where M is again the number of orbitals or plane waves in the basis set). This is a significant improvement over the scaling found by Wecker *et al.* (2015a), which was approximately $O(M^4)$. Similar improvements in asymptotic scaling were obtained using the nested decomposition method introduced by Motta *et al.* (2018). 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.

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 (Kolda *et al.*, 2006).

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.* (2014), 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. Nelder-Mead was also used successfully (Shen *et al.*, 2017) and partially successfully (Hempel *et al.*, 2018) 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 by Colless *et al.* (2018) and Santagati *et al.* (2018), who 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.* (2017). They used the SPSA algorithm, because of its purported resilience to noise and low number of required function evaluations (Spall, 1992). This same algorithm was also used by Ganzhorn *et al.* (2018).

To date, there have been several numerical studies comparing different optimisation algorithms for the VQE. McClean *et al.* (2016) compared four direct search algorithms: Nelder-Mead simplex, and TOMLAB's glc-Cluster, LGO, and multiMin algorithms. 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 glc-Cluster to converge to the most accurate result, especially as the measurement precision was increased.

Romero *et al.* (2019) 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). 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. These authors found that they were able to markedly improve the performance of the algorithms by using a chemically motivated (Möller-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.* (2015a) introduced the 'global variational' method for optimisation, alongside their Hamiltonian variational ansatz. 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 (Dallaire-Demers *et al.*, 2018; Romero *et al.*, 2019). 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 parameter 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.

Annealed and morphed Hamiltonians

Several works have used concepts from adiabatic quantum computing to aid the classical optimisation procedure. Wecker *et al.* (2015a) proposed an 'annealed variational' method alongside their Hamiltonian variational ansatz. They first decompose the molecular Hamiltonian of interest into

$$H_s = H_0 + sH_1, \tag{59}$$

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 Trotter steps. 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
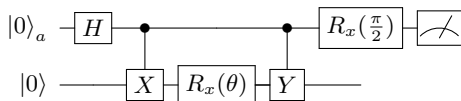

FIG. 7 A quantum circuit to calculate the gradient of a toy VQE simulation. In this toy problem, the ansatz used is $|\psi\rangle = R_x(\theta)|0\rangle$, and the Hamiltonian is $H = Y$. The energy is given by $E(\theta) = \langle\psi(\theta)|H|\psi(\theta)\rangle = \langle 0|R_x^\dagger(\theta)YR_x(\theta)|0\rangle$. 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.

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 by Garcia-Saez and Latorre (2018).

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 non-zero 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 (McArdle *et al.*, 2018a). 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. Recent work by Mitarai and Fujii (2018) has increased the experimental feasibility of the variational imaginary time algorithm, by reducing the degree of qubit connectivity required.

Particle-hole transform

Recent work by Barkoutsos *et al.* (2018) 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 (Barkoutsos *et al.*, 2018).

C. Evaluation of excited states

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

1. WAVES

The witness-assisted variational eigenspectra solver (WAVES) (Santagati *et al.*, 2018) 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(a) 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.

$$F_{\text{cost}} = E(\vec{\theta}) + aS(\vec{\theta}), \tag{60}$$

where a is a constant which determines the trade-off between minimising energy and entropy. Minimising the
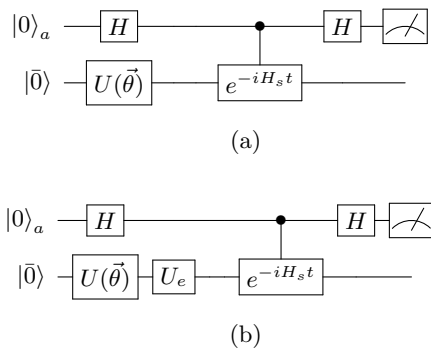
FIG. 8 Quantum circuits in the WAVES protocol. (a) The circuit can generate an eigenstate of the Hamiltonian, H_s. (b) The operator U_e excites the system from its ground state to a state that is close to an excited state.

entropy forces the system into the ground state of the Hamiltonian, H_s and we obtain the parameter $\vec{\theta}_{\min}$ which gives the ground state.

We then apply the circuit shown in Fig. 8(b), where the state $U_e U(\vec{\theta}_{\min}) |0\rangle$ approximates an excited state of the Hamiltonian. By varying $\vec{\theta}$, we can minimise the cost function in the limit of $a \to \infty$, 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.

The WAVES method has been experimentally demonstrated using a silicon photonics device (Santagati *et al.*, 2018). 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. Overlap-based methods

It is possible to use variational algorithms and state overlap calculations to find excited states (Endo *et al.*, 2018a; Higgott *et al.*, 2018).

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

$$H' = H + \alpha |E_0\rangle \langle E_0| , \tag{61}$$

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 circuits to calculate the overlap between the states. This includes using the SWAP test, which has recently been modified to use a more shallow circuit (Cincio *et al.*, 2018; Garcia-Escartin and Chamorro-Posada, 2013).

As this method uses only low depth circuits, has the potential for error mitigation (Endo *et al.*, 2018a; Higgott *et al.*, 2018) and uses much of the machinery underlying the VQE, it is suitable for use with near-future hardware. This method has recently been numerically investigated by Lee *et al.* (0), who also considered the propagation of errors resulting from only obtaining an approximation of lower lying eigenstates, rather than the true eigenstates.

3. The folded spectrum method

The folded spectrum method can be used for finding the excited states of molecular Hamiltonians (McClean *et al.*, 2016). 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 (Wecker *et al.*, 2015a). 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 (McClean *et al.*, 2017a). 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.

Those authors consider the linear response subspace around the fermionic ground state. This subspace is spanned by the states $a_i^\dagger a_j |E_0\rangle$ 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

$$H^{\text{QSE}}C = S^{\text{QSE}}CE, \tag{62}$$

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

$$H^{\text{QSE}}_{ij,kl} = \langle E_0 | a_i a_j^\dagger H a_k^\dagger a_l | E_0 \rangle. \tag{63}$$

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

$$S^{\text{QSE}}_{ij,kl} = \langle E_0 | a_i a_j^\dagger a_k^\dagger a_l | E_0 \rangle. \tag{64}$$

We provide more information on the QSE method in Sec. VI, where we discuss how it can also 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 (Sawaya et al., 2016). Circuits with a large number of gates are presumed to require the full machinery of error correction for protection. 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 been developed which seek to mitigate errors, rather than correct them (Bonet-Monroig et al., 2018; Endo et al., 2017; Huo and Li, 2018; Johnson et al., 2017; Li and Benjamin, 2017; McArdle et al., 2018c; McClean et al., 2017a; Otten and Gray, 2018a,b; Temme et al., 2017). 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. We note that many of these techniques were introduced for use in general NISQ algorithms, and so can be applied to problems beyond chemistry simulation.

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 (Nielsen and Chuang, 2002).

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

$$\rho_0 = \mathcal{U}_N \ldots \mathcal{U}_2 \mathcal{U}_1 (|\bar{0}\rangle \langle \bar{0}|), \tag{65}$$

where for a density matrix ρ, $\mathcal{U}(\rho) = U \rho U^\dagger$. We extract information from the circuit by measuring a Hermitian observable, O

$$\bar{O}_0 = \text{Tr}[\rho_0 O]. \tag{66}$$

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

$$\rho = \prod_i \mathcal{N}_i(\mathcal{U}_i(|\bar{0}\rangle \langle \bar{0}|)), \tag{67}$$

and the measurement result becomes $\bar{O} = \text{Tr}[\rho O]$. 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 (McClean et al., 2016).

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 O'Malley et al. (2016).

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

$$L = H + \sum_j \beta_j (Q_j - q_j I)^2, \tag{68}$$

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 (Li and Benjamin, 2017; Temme et al., 2017) works by intentionally increasing the error rate by a factor λ, and inferring the error free result by extrapolation. We can increase the error rate using the techniques described by Kandala et al. (2018) and Li and Benjamin (2017). The technique is based on Richardson extrapolation (Richardson et al., 1927), which to first order corresponds to linear extrapolation
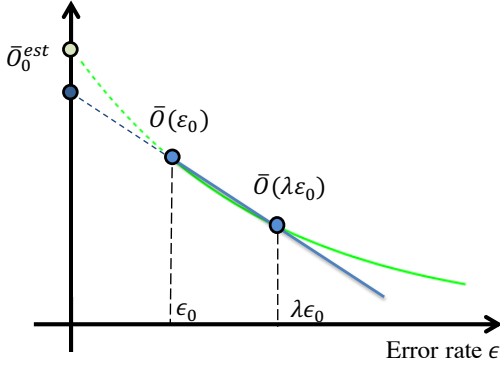
FIG. 9 A comparison of linear (blue) and exponential (green) extrapolation. The horizontal axis is the error rate of each gate and the vertical axis is the expectation value of the measured observable, $\bar{O}$.

using two points. Alternatively, we can take a linear best fit with several points. For the former case, the estimated value of the observable is given by

$$\bar{O}_0^{\text{est}} = \frac{\lambda \bar{O}(\epsilon_0 \cdots \epsilon_n)) - \bar{O}(\lambda(\epsilon_0 \cdots \epsilon_n)))}{\lambda - 1}. \quad (69)$$

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. (2018) 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.

Exponential extrapolation was introduced by Endo et al. (2017) as a more appropriate extrapolation technique for large quantum circuits. A comparison between the two extrapolation methods is shown in Fig. 9. Otten and Gray (2018b) have also extended the extrapolation method to the scenario where the error rates of different gates are increased with different factors.

C. Probabilistic error cancellation

The probabilistic error cancellation method introduced by Temme et al. (2017) 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. 10 to effectively realise the inverse channel by focusing only on measurement results.

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

$$\mathcal{D}(\rho_0) = \left(1 - \frac{3}{4}p\right)\rho_0 + \frac{p}{4}(X\rho_0 X + Y\rho_0 Y + Z\rho_0 Z). \quad (70)$$
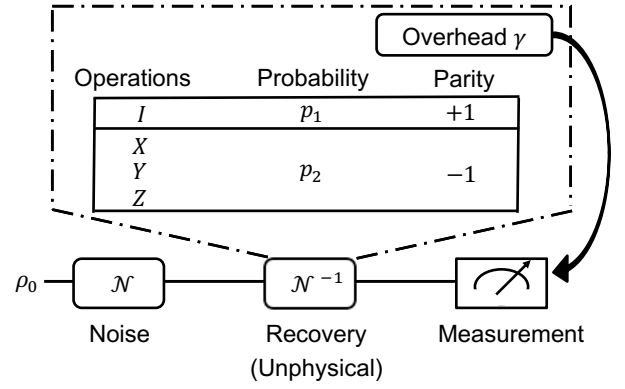

FIG. 10 A schematic of the probabilistic error cancellation method for a depolarising error resulting from a single qubit gate. After the gate is applied, there is a noise channel $\mathcal{N}$. The method works by effectively realising the inverse channel $\mathcal{N}^{-1}$. This is achieved by randomly 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.

The unphysical inverse channel is

$$\mathcal{D}^{-1}(\rho)$$
$$= \left(1 + \frac{3p}{4(1-p)}\right)\rho - \frac{p}{4(1-p)}(X\rho X + Y\rho Y + Z\rho Z),$$
$$= \gamma[p_1 \rho - p_2(X\rho X + Y\rho Y + Z\rho Z)], \quad (71)$$

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 variance in our measurement of O_0 is increased by a factor of γ^n, where γ is the overhead coefficient, and n is the number of gates in the circuit.

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

$$O_0 = \text{Tr}[O\mathcal{N}^{-1}(\rho)],$$
$$= \gamma[p_1 \langle O \rangle_\rho - p_2(\langle O \rangle_{X\rho X} + \langle O \rangle_{Y\rho Y} + \langle O \rangle_{Z\rho Z})],$$
$$= \gamma[p_1 \langle O \rangle_\rho - p_2(\langle XOX \rangle_\rho + \langle YOY \rangle_\rho + \langle ZOZ \rangle_\rho)], \quad (72)$$

where $\langle O \rangle_\rho = \text{Tr}[O\rho]$. 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 example, for two qubit gates in the depolarising noise model, after each two qubit gate we insert one of the gates: XI, IX, YI, IY, ZI, IZ (parity -1) with probability p_2, one of 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 method has been shown to work for general Markovian noise, but is not suitable for correlated errors (Endo *et al.*, 2017). 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 the effect of errors due to state preparation and measurement (Endo *et al.*, 2017). The probabilistic error cancellation method has recently been experimentally demonstrated on a superconducting processor (Song *et al.*, 2018).

D. Quantum subspace expansion

The quantum subspace expansion (McClean *et al.*, 2017a) described in Sec. V.C can mitigate errors in the VQE, in addition to calculating the excited energy eigenstates. This method is most effective at correcting systematic errors, but can also suppress stochastic 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,

$$H_{ij} = \langle\tilde{E}_0| P_i H P_j |\tilde{E}_0\rangle. \tag{73}$$

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

$$S_{ij} = \langle\tilde{E}_0| P_i P_j |\tilde{E}_0\rangle. \tag{74}$$

By solving the generalised eigenvalue problem

$$HC = SCE, \tag{75}$$

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 (Colless *et al.*, 2018). 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 H_2 was calculated to near chemical accuracy.

E. Stabiliser based methods

Recently, a new method of error mitigation has been introduced, which makes use of stabiliser checks on a suitably constructed second quantised state (Bonet-Monroig *et al.*, 2018; McArdle *et al.*, 2018c). 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 (Sawaya *et al.*, 2016). Consequently, we can perform 'checks' on quantities which should be conserved, discarding the results if the measured value is not as expected. This can be achieved by conducting suitably constructed stabiliser checks on the register using additional ancilla qubits (Bonet-Monroig *et al.*, 2018; McArdle *et al.*, 2018c), or by taking additional measurements and performing some postprocessing (Bonet-Monroig *et al.*, 2018). This method of error mitigation can easily be combined with some of the other techniques mentioned above, such as extrapolation. Combining the extrapolation and stabiliser methods was shown to be considerably more effective than either method in isolation (McArdle *et al.*, 2018c). Bonet-Monroig *et al.* (2018) also derived a way to increase the number of errors detected by introducing additional symmetries, or modifying the ansatz circuit.

F. Other methods of error mitigation

QVECTOR

The quantum variational error corrector (QVECTOR), introduced by Johnson *et al.* (2017), 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. It was found to be effective in simulations of a three qubit system, enabling a 6-fold extension in the T_2 time of the system. It was also found to outperform the five qubit stabiliser code for a system experiencing amplitude and phase damping noise (Johnson *et al.*, 2017).

Individual Error Reduction

Otten and Gray (2018a) showed error mitigation is possible when there are a number of error sources, which may be reduced individually. They use error correction to protect a single qubit, while leaving the rest of the physical qubits subject to noise. They repeat the process several times, with each physical qubit being protected in turn. Linearly combining the results produces a more accurate expectation value than would be obtained without the mitigation technique.

Error mitigation for temporally correlated noise

Recent work by Huo and Li (2018) has extended the probabilistic error cancellation method to the case of temporally correlated errors and low frequency noise. They introduced a method known as 'linear operator tomography' (LOT), in which quantum gates with correlated errors are perfectly characterised by a set of linear operators. The information obtained from LOT is then used to construct the inverse operation of the error, so that correlated noise can be suppressed.

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 in second quantisation 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 ansatz (Sec. V.B). These examples are designed to familiarise the reader with the key steps of formulating a quantum computational chemistry problem. Many of these techniques are applicable to both ground state and general chemical problems.
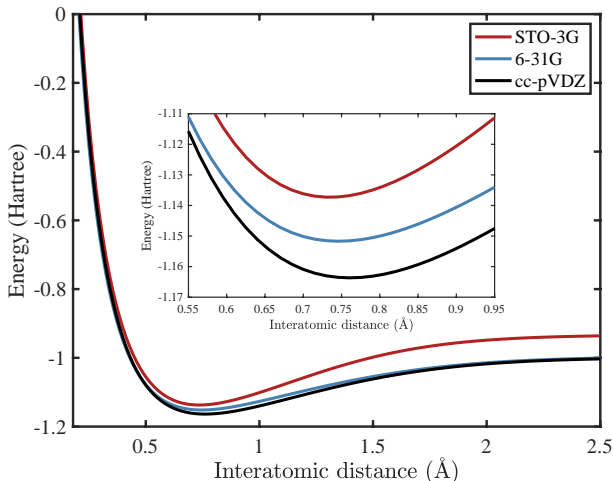

FIG. 11 Comparing the ground state dissociation curves of H_2 for a range of basis sets.

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. 11 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.

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

$$|1s_{A\uparrow}\rangle, |1s_{A\downarrow}\rangle, |1s_{B\uparrow}\rangle, |1s_{B\downarrow}\rangle, \qquad (76)$$

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

$$|\sigma_{g\uparrow}\rangle = \frac{1}{\sqrt{2}}(|1s_{A\uparrow}\rangle + |1s_{B\uparrow}\rangle),$$

$$|\sigma_{g\downarrow}\rangle = \frac{1}{\sqrt{2}}(|1s_{A\downarrow}\rangle + |1s_{B\downarrow}\rangle),$$

$$|\sigma_{u\uparrow}\rangle = \frac{1}{\sqrt{2}}(|1s_{A\uparrow}\rangle - |1s_{B\uparrow}\rangle), \qquad (77)$$

$$|\sigma_{u\downarrow}\rangle = \frac{1}{\sqrt{2}}(|1s_{A\downarrow}\rangle - |1s_{B\downarrow}\rangle).$$

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

$$|\psi\rangle = |f_{\sigma_{u\downarrow}}, f_{\sigma_{u\uparrow}}, f_{\sigma_{g\downarrow}}, f_{\sigma_{g\uparrow}}\rangle, \qquad (78)$$

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 (Frisch *et al.*, 2016; Muller, 2004; Parrish *et al.*, 2017; Sun *et al.*, 2017). The resulting second quantised Hamiltonian is (Seeley *et al.*, 2012)

$$\begin{aligned} H &= h_{00}a_0^\dagger a_0 + h_{11}a_1^\dagger a_1 + h_{22}a_2^\dagger a_2 + h_{33}a_3^\dagger a_3 \\ &+ h_{0110}a_0^\dagger a_1^\dagger a_1 a_0 + h_{2332}a_2^\dagger a_3^\dagger a_3 a_2 + h_{0330}a_0^\dagger a_3^\dagger a_3 a_0 \\ &+ h_{1221}a_1^\dagger a_2^\dagger a_2 a_1 + (h_{0220} - h_{0202})a_0^\dagger a_2^\dagger a_2 a_0 \\ &+ (h_{1331} - h_{1313})a_1^\dagger a_3^\dagger a_3 a_1 + h_{0132}(a_0^\dagger a_1^\dagger a_3 a_2 + a_2^\dagger a_3^\dagger a_1 a_0) \\ &+ h_{0312}(a_0^\dagger a_3^\dagger a_1 a_2 + a_2^\dagger a_1^\dagger a_3 a_0), \end{aligned}$$
$$(79)$$

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 (Seeley *et al.*, 2012)

$$\begin{aligned} H &= h_0 I + h_1 Z_0 + h_2 Z_1 + h_3 Z_2 + h_4 Z_3 + \\ &\quad h_5 Z_0 Z_1 + h_6 Z_0 Z_2 + \\ &\quad h_7 Z_1 Z_2 + h_8 Z_0 Z_3 + \\ &\quad h_9 Z_1 Z_3 + h_{10} Z_2 Z_3 + \\ &\quad h_{11} Y_0 Y_1 X_2 X_3 + h_{12} X_0 Y_1 Y_2 X_3 + \\ &\quad h_{13} Y_0 X_1 X_2 Y_3 + h_{14} X_0 X_1 Y_2 Y_3. \end{aligned}$$
$$(80)$$

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 (McClean *et al.*, 2017b) or Qiskit Aqua (IBM, 2018).

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

$$\psi_{\mathrm{HF}}^{\mathrm{H_2}}(\mathbf{r_1}, \mathbf{r_2}) = \frac{1}{\sqrt{2}}(\sigma_{g\uparrow}(\mathbf{r_1})\sigma_{g\downarrow}(\mathbf{r_2}) - \sigma_{g\uparrow}(\mathbf{r_2})\sigma_{g\downarrow}(\mathbf{r_1})),$$
$$(81)$$

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

$$|\psi_{\mathrm{HF}}^{\mathrm{H_2}}\rangle = |0011\rangle. \qquad (82)$$

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

$$|\psi^{\mathrm{H_2}}\rangle = \alpha|0011\rangle + \beta|1100\rangle + \gamma|1001\rangle + \delta|0110\rangle, \quad (83)$$

and the ground state of the H_2 molecule at its equilibrium bond distance is given by (Helgaker *et al.*, 2014)

$$|\psi_g^{\mathrm{H_2}}\rangle = 0.9939|0011\rangle - 0.1106|1100\rangle. \qquad (84)$$

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, 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

$$U = e^{(T_1 - T_1^\dagger) + (T_2 - T_2^\dagger)}, \qquad (85)$$

where

$$\begin{aligned} T_1 &= \sum_{i \in virt, \alpha \in occ} t_{i\alpha} a_i^\dagger a_\alpha, \\ T_2 &= \sum_{i,j \in virt, \alpha, \beta \in occ} t_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta, \end{aligned}$$
$$(86)$$

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 do not 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

$$U = e^{t_{02}(a_2^\dagger a_0 - a_0^\dagger a_2) + t_{13}(a_3^\dagger a_1 - a_1^\dagger a_3) + t_{0123}(a_3^\dagger a_2^\dagger a_1 a_0 - a_0^\dagger a_1^\dagger a_2 a_3)}. \tag{87}$$

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

$$\begin{aligned} U = \; & e^{t_{02}(a_2^\dagger a_0 - a_0^\dagger a_2)} \\ & \times e^{t_{13}(a_3^\dagger a_1 - a_1^\dagger a_3)} \\ & \times e^{t_{0123}(a_3^\dagger a_2^\dagger a_1 a_0 - a_0^\dagger a_1^\dagger a_2 a_3)}. \end{aligned} \tag{88}$$

Using the JW encoding, we find that

$$\begin{aligned} (a_2^\dagger a_0 - a_0^\dagger a_2) &= \frac{i}{2}(X_2 Z_1 Y_0 - Y_2 Z_1 X_0) \\ (a_3^\dagger a_1 - a_1^\dagger a_3) &= \frac{i}{2}(X_3 Z_2 Y_1 - Y_3 Z_2 X_1) \\ (a_3^\dagger a_2^\dagger a_1 a_0 - a_0^\dagger a_1^\dagger a_2 a_3) &= \\ & \frac{i}{8}(X_3 Y_2 X_1 X_0 + Y_3 X_2 X_1 X_0 + Y_3 Y_2 Y_1 X_0 + Y_3 Y_2 X_1 Y_0 \\ & - X_3 X_2 Y_1 X_0 - X_3 X_2 X_1 Y_0 - Y_3 X_2 Y_1 Y_0 - X_3 Y_2 Y_1 Y_0). \end{aligned} \tag{89}$$

It was shown in Romero et al. (2019) that all Pauli terms arising from the same excitation operators commute. As a result, each of the exponentials in Eq. (88) can be separated into a product of exponentials of a single Pauli string. For example

$$e^{t_{02}(a_2^\dagger a_0 - a_0^\dagger a_2)} = e^{\frac{i t_{02}}{2} X_2 Z_1 Y_0} \times e^{\frac{-i t_{02}}{2} Y_2 Z_1 X_0}. \tag{90}$$

In Hempel et al. (2018) 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

$$U = e^{-i\theta X_3 X_2 X_1 Y_0}. \tag{91}$$

This can be implemented using the circuit (Whitfield et al., 2011) shown in Fig. 12.

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
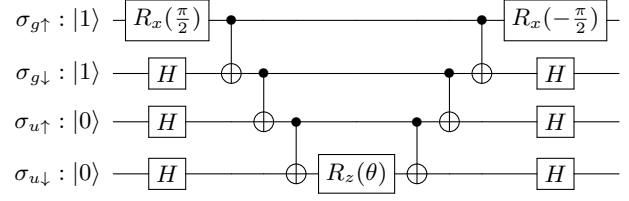

FIG. 12 The circuit for implementing the UCCSD operator for H$_2$ in the STO-3G basis, as given by Eq. (91). The $R_x(\frac{\pi}{2})$ and H gates rotate the basis such that the exponentiated operator applied to the corresponding qubit is either Y or X, respectively. Single excitation terms are implemented with a change of basis (Hempel et al., 2018).

total; $\{1s_\uparrow, 1s_\downarrow, 1s'_\uparrow, 1s'_\downarrow\}$ from each atom. Working in the canonical orbital basis, (obtained by performing a Hartree–Fock calculation) 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

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{92}$$

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$^{\text{th}}$ 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$^{\text{th}}$ entry is the sum (mod 2) of all of the orbital occupancies, 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$^{\text{th}}$ entry is still equal to the number of electrons, the 4$^{\text{th}}$ entry no longer necessarily equal to a conserved quantity. The JW mapped HF state (8 qubits) is given by $|00010001\rangle$. Using the matrix given above, the BK mapped HF state (8 qubits) is $|00111011\rangle$. When the two conserved qubits are removed, the BK mapped HF state (6 qubits) is $|011011\rangle$.

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. We first obtain the 1-RDM for H_2 in the cc-PVDZ basis with a classically tractable configuration interaction singles and double calculation.

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 given by $diag(1.96588, 0.00611, 0.02104, 0.0002, 0.00001, 0.00314, 0.00314, 0.00016, 0.00016, 0.00016)$. There are only 10 diagonal entries in this 1-RDM because the spin-up and spin-down entries for the same spatial orbitals have been combined. As discussed in Sec. III.E, the diagonal entries are the natural orbital occupation numbers (NOONs). We can see that the 5$^{\text{th}}$ orbital has a NOON that is 20 times smaller than the next smallest NOON. As a result, we consider this spatial 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. We follow a similar procedure to that described for the LiH molecule in Fig. 13 in order to map the fermionic orbitals to qubits. The reader will find that the 9$^{\text{th}}$ and 18$^{\text{th}}$ 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. The lowest energy computational basis state of cc-PVDZ H_2 in the Jordan-Wigner encoding (18 qubits) is $|000000001000000001\rangle$. The corresponding BK-tree mapped state (16 qubits) 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, as illustrated in Fig. 13. The canonical orbital 1-RDM from a CISD calculation on LiH (at an internuclear separation of 1.45 $\dot{A}$) is given by

$$
\begin{bmatrix}
1.9999 & -0.0005 & 0.0006 & 0.0000 & 0.0000 & -0.0010 \\
-0.0005 & 1.9598 & 0.0668 & 0.0000 & 0.0000 & 0.0084 \\
0.0006 & 0.0668 & 0.0097 & 0.0000 & 0.0000 & -0.0138 \\
0.0000 & 0.0000 & 0.0000 & 0.0017 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0017 & 0.0000 \\
-0.0010 & 0.0084 & -0.0138 & 0.0000 & 0.0000 & 0.0273
\end{bmatrix}. \tag{93}
$$

There are only 6 diagonal entries in this 1-RDM because the spin-up and spin-down entries for the same

spatial orbitals have been combined. We diagonalise this 1-RDM, moving to the NMO basis. The NMO 1-RDM is given by

$$
\begin{bmatrix}
1.99992 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 1.96206 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.03454 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00005 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00171 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00171
\end{bmatrix}. \tag{94}
$$

The first orbital has a NOON close to two, and so

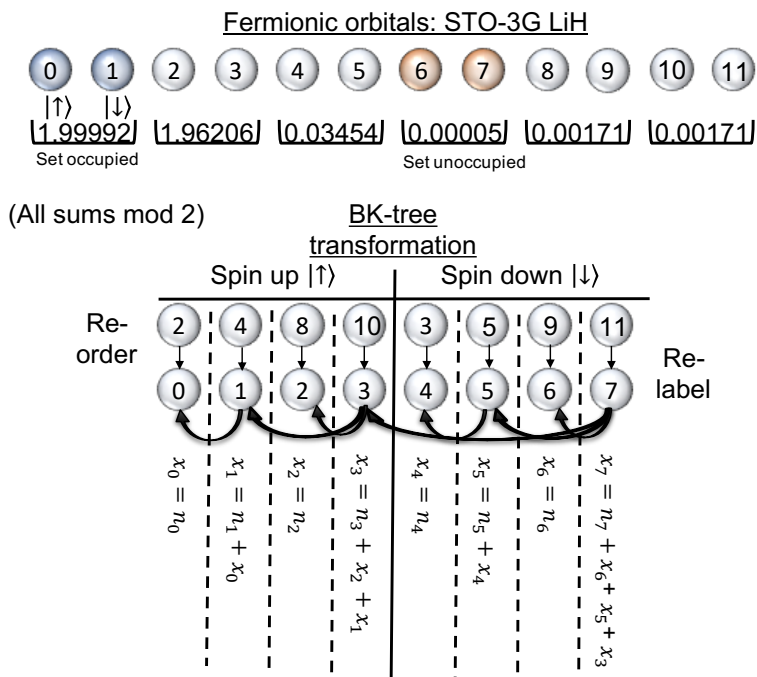we consider it to always be doubly occupied. We can

FIG. 13 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. 14. The value x_i is the value of the ith qubit under the BK-tree mapping, while n_i is the value of the ith 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.

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 Fen$(0, M-1)$. We can obtain this data structure using an iterative algorithm, which we reproduce from Havlicek *et al.* (2017) below. The generation of the Fenwick tree for the LiH molecule using this algorithm is shown in Fig. 14.

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 Hempel *et al.* (2018).

Algorithm.1 : Fenwick tree generation

```
Define Fen(L, R)

If L ≠ R:
    • Connect R to Floor (R+L/2);
    • Fen (L, Floor (R+L/2));
    • Fen (Floor (R+L/2) + 1, R).
Else:
    • End the current Fenwick tree.
```

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.
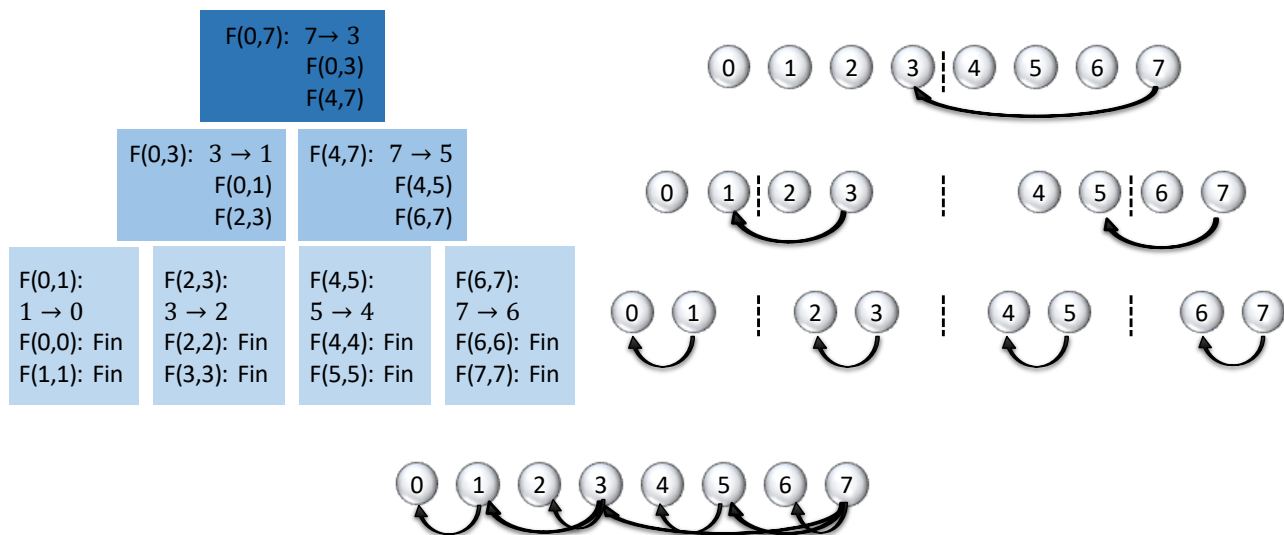
FIG. 14 A pictorial representation of the Fenwick tree construction for LiH, shown in Fig. 13. We carry out the BK-tree mapping by constructing the Fenwick tree, Fen(0,7), as described in Algorithm. 1. The algorithmic steps are shown on the left hand side of the figure, while the corresponding actions are shown on the right hand side. The notation $X \to Y$ 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.

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 (Helgaker *et al.*, 2014), 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 electrons in 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 (Rossi *et al.*, 1999), or the Cr_2 molecule with 24 active electrons in 30 spin-orbitals (Tubman *et al.*, 2016), or the fluorine atom in a cc-PV5Z basis with additional basis functions (Booth and Alavi, 2010). 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' (Ortiz *et al.*, 2001). Tensor product methods, such as density matrix renormalisation group (DMRG) have also proven effective for treating systems with large active spaces, including some metalloenzyme complexes with active spaces

of over 70 spin-orbitals (Kurashige *et al.*, 2013; Sharma *et al.*, 2014). Olivares-Amaya *et al.* (2015) have produced a comprehensive review of progress in DMRG for quantum chemistry.

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 medicines and materials *in silico*. For example, as noted by Yamazaki *et al.* (2018), 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 (Bauer *et al.*, 2016; Kreula *et al.*, 2016; Reiher *et al.*, 2017; Rubin, 2016; Yamazaki *et al.*, 2018).

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 (Hattig and Weigend, 2000), which corresponds to the DNA base guanine ($C_5H_5N_5O$) in a cc-PVTZ basis (Hobza and poner, 2002), or the hydrocarbon octane (C_8H_{18}) in a cc-PVTZ basis (Yamazaki *et al.*, 2018). The classical implementation of CCSD does not store the wavefunction, as this would again be exponentially costly (as the CCSD wave-

function has support on all possible Slater determinants). Instead, coupled non-linear equations can be derived. The solution to these equations is the CCSD approximation to the ground state (Helgaker *et al.*, 2014; Purvis and Bartlett, 1982). The time taken to solve these equations scales as $O((M-N)^4 N^2)$ (Purvis and Bartlett, 1982), 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 significant 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 (Helgaker *et al.*, 2014). 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: medium to long term

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 electrons in the molecule. Instead, we can use adiabatic state preparation, as described in Sec. V.A, with the assumption that the gap is well behaved (Babbush *et al.*, 2014).

Initial work showed that up to around 10^{18} gates would be necessary to simulate a system of around 100 spin-orbitals (excluding the overhead of error correction) (Wecker *et al.*, 2014), although this estimate was subsequently reduced through a series of algorithmic optimisations (Babbush *et al.*, 2015; Hastings *et al.*, 2015; McClean *et al.*, 2014; Poulin *et al.*, 2015). Reiher *et al.* (2017) used this as the starting point to find the resources required to carry out a transformative chemistry calculation. 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 (Reiher *et al.*, 2017). 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 those authors model with an active spaces of (54 electrons, 108 spin-orbitals) and

(65 electrons, 114 spin-orbitals) (Reiher *et al.*, 2017) (although recent work has shown that an alternative active space of 113 electrons in 152 spin-orbitals would produce more realistic results (Li *et al.*, 2018)). As a bioinorganic transition metal compound, the ground state of FeMoco will likely be difficult to identify with existing classical methods (Podewitz *et al.*, 2011). The authors calculated the resources required to perform an FCI calculation on the aforementioned active spaces with an error corrected quantum computer possessing around 100 logical qubits. They found that this would require around 200 million physical qubits, and take on the order of weeks (10 ns to implement a T gate, including surface code decoding) or months (100 ns per T gate) (both assuming current best error rates) (Reiher *et al.*, 2017).

Recent work has carried out similar resource analysis for other problems (the Fermi-Hubbard model or periodic systems such as the homogeneous electron gas), using state of the art techniques for chemistry simulation and fault-tolerance protocals (Babbush *et al.*, 2018b). The required resources were on the order of one million physical qubits, running for a few hours (assuming 1 μs to implement a T gate, including surface code decoding) (Babbush *et al.*, 2018b).

We note that none of the resource estimation papers described above consider the problem of initial state preparation, and instead focus on the phase estimation aspect of the electronic structure problem. Recent work by Tubman *et al.* (2018) has sought to rectify this. Those authors provide a quantum algorithm for preparation of arbitrary Slater determinants. They then show that for many systems of interest in chemistry, physics and materials science, states that can be efficiently prepared using their algorithm have a large overlap with the true ground state, and thus can be tackled using phase estimation.

Despite these promising results, we note that it may still be many years before we possess a quantum computer with the resources required to implement these algorithms. Consequently, in order to achieve transformative chemistry simulations before that time, different approaches are required. One possible approach is to use the variational algorithms discussed in Sec. V.B. We discuss the challenges facing these methods below.

C. Outlook for near-future approaches

For near-future approaches, the greatest remaining difficulties are designing good ansatze, reducing the number of measurements required, and mitigating the effects of errors. In particular, we will discuss the close relationship between the ansatz circuit and error rates. As stated previously, near-future approaches are those suitable for the quantum computers that will become

available within the next 5-10 years. These machines will likely possess little or no error correction, rather than full fault-tolerance.

Both heuristic and chemically motivated ansatze face a range of challenges, if they are to provide a route to classically intractable calculations on near-term quantum computers. To date, most heuristic ansatze have only been numerically tested on small systems, 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.

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. Moreover, it is not immediately clear where the advantage will lie for UCCSD calculations. As discussed above, the scaling of classical CCSD is approximately $O(M^6)$. The number of gates required to implement a UCCSD ansatz on a quantum computer using the best current methods scales as $O(M^4)$ with increasing molecular size, and $O(M^3)$ for a fixed molecular size and increasing basis set size (Motta *et al.*, 2018). However, this naive comparison does not take into account the cost of classical optimisation in the VQE, constant prefactors in the asymptotic scaling, or the slow clock speed of quantum computers, compared to classical hardware. 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 directly competitive.

As such, quantum UCCSD calculations are likely 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 (Reiher *et al.*, 2017).

It may require on the order of one million gates to implement a UCCSD ansatz for a system with 100 spin-orbitals (Motta *et al.*, 2018). As this does not take into account the reduction of excitation terms due to molecular point group symmetries, the number of gates required may in fact be less than this estimate. 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 (Hempel *et al.*, 2018). Nevertheless, a simple calculation demonstrates the necessity of making ansatz circuits as shallow as possible. If we assume a discrete error model for our circuit, such that error events happen probabilistically and independently following each gate in the circuit, then 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. Recent experimental demonstrations have shown that environmental noise corrupts the results of quantum computational chemistry calculations (Hempel *et al.*, 2018; Kandala *et al.*, 2017). While the error mitigation techniques discussed in Sec. VI may enable us to recover accurate results from a circuit deeper than 10,000 gates, it seems unlikely that current methods alone would enable more than a small multiplicative increase in the circuit depth. There are a few possible ways that one may be able to tackle this problem.

The first is to develop additional methods of error mitigation, which ideally could be combined with existing methods. An alternative proposal is to use a small error correcting code. While such a code would not enable fault-tolerant computation, the idea would be to suppress the error rate to a value low enough to enable the ansatz circuit to be implemented in an error-free manner. For example, recent work by Setia *et al.* (2018) has shown that the Bravyi-Kitaev superfast encoding can be modified such that it provides the ability to correct single qubit errors. A third possibility is to develop alternative ansatze, or improve existing ansatze. This could include heuristic ansatze with a shorter circuit depth (Lee *et al.*, 0), methods to reduce the depth of existing ansatze (Motta *et al.*, 2018), or ansatze which are intrinsically resilient to noise (Kim, 2017; Kim and Swingle, 2017).

Another key concern for variational approaches is the number of measurements required, particularly when considering molecules. For example, Wecker *et al.* (2015a) 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 many 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 (Babbush *et al.*, 2018c; Barkoutsos *et al.*, 2018; Izmaylov *et al.*, 2018; Rubin *et al.*, 2018; Wang *et al.*, 2018), the number of measurements is still dauntingly high for molecular simulations.

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

Babbush *et al.* (2018c). This may benefit many research areas in computational chemistry. The plane wave basis discussed in Sec. III.D.4 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 (Babbush *et al.*, 2018c). 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 (Babbush *et al.*, 2018c).

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 (Wecker *et al.*, 2015b). However, classical methods to solve the Fermi-Hubbard model exactly are limited to around 20 lattice sites (Jiang *et al.*, 2018b). 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 perform Trotter steps of the Fermi-Hubbard Hamiltonian using $O(N)$ gates for each Trotter step (Jiang *et al.*, 2018b). 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 (Wecker *et al.*, 2015a). Recent work has shown promising results for both finding the ground state of the Fermi-Hubbard model with the 'Hamiltonian variational + VQE' approach (Reiner *et al.*, 2018a), and for simulating its dynamics (Reiner *et al.*, 2018b), both in the presence of realistic noise rates. We note that these results were obtained using less efficiently scaling circuits than those described above. There have also been proposals for simulating the Fermi-Hubbard model on near-term quantum computers using a hybrid quantum-classical problem decomposition approach (Bauer *et al.*, 2016; Dallaire-Demers and Wilhelm, 2016a,b).

All of these results give cause for optimism, and raise the question of if near-future quantum computers would 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 (assuming a 1 μs gate time) (Wecker *et al.*, 2015a), 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. Summary

This review has sought to be accessible to both 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. This review has 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.

ACKNOWLEDGMENTS

REFERENCES

Abrams, D. S., and S. Lloyd (1997), Phys. Rev. Lett. **79**, 2586.

Abrams, D. S., and S. Lloyd (1999), Phys. Rev. Lett. **83**, 5162.

Aharonov, D., and M. Ben-Or (1997), in *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing* (ACM) pp. 176–188.

Aharonov, D., W. Van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev (2008), SIAM review **50** (4), 755.

Argüello-Luengo, J., A. González-Tudela, T. Shi, P. Zoller, and J. I. Cirac (2018), arXiv:1807.09228.

Aspuru-Guzik, A., A. D. Dutoi, P. J. Love, and M. Head-Gordon (2005), Science **309** (5741), 1704.

Aspuru-Guzik, A., R. Lindh, and M. Reiher (2018), ACS Central Science **4** (2), 144.

Aspuru-Guzik, A., and P. Walther (2012), Nature Physics **8** (4), 285.

Babbush, R., D. W. Berry, I. D. Kivlichan, A. Y. Wei, P. J. Love, and A. Aspuru-Guzik (2016), New Journal of Physics **18** (3), 033032.

Babbush, R., D. W. Berry, J. R. McClean, and H. Neven (2018a), arXiv:1807.09802.

Babbush, R., D. W. Berry, Y. R. Sanders, I. D. Kivlichan, A. Scherer, A. Y. Wei, P. J. Love, and A. Aspuru-Guzik (2017), Quantum Science and Technology **3** (1), 015006.

Babbush, R., C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven (2018b), Phys. Rev. X **8**, 041015.

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

Babbush, R., J. McClean, D. Wecker, A. Aspuru-Guzik, and N. Wiebe (2015), Phys. Rev. A **91**, 022311.

Babbush, R., A. Perdomo-Ortiz, B. O'Gorman, W. Macready, and A. Aspuru-Guzik (2012), 10.1002/9781118755815.ch05, arXiv:1211.3422.

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

Babej, T., C. Ing, and M. Fingerhuth (2018), arXiv:1811.00713.

Ball, R. C. (2005), Phys. Rev. Lett. **95**, 176407.

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

Barends, R., J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, *et al.* (2014), Nature **508** (7497), 500.

Barkoutsos, P. K., 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 (2018), Phys. Rev. A **98**, 022322.

Bartlett, R. J., S. A. Kucharski, and J. Noga (1989), Chemical Physics Letters **155** (1), 133 .

Bauer, B., D. Wecker, A. J. Millis, M. B. Hastings, and M. Troyer (2016), Phys. Rev. X **6**, 031045.

Berry, D. W., G. Ahokas, R. Cleve, and B. C. Sanders (2007), Communications in Mathematical Physics **270** (2), 359.

Berry, D. W., and A. M. Childs (2012), Quantum Info. Comput. **12** (1-2), 29.

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

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

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

Berry, D. W., B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman (2009), Phys. Rev. A **80**, 052114.

Berry, D. W., M. Kieferová, A. Scherer, Y. R. Sanders, G. H. Low, N. Wiebe, C. Gidney, and R. Babbush (2018), npj Quantum Information **4** (1), 22.

Blatt, R., and C. F. Roos (2012), Nature Physics **8** (4), 277.

Boixo, S., S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven (2018), Nature Physics **14** (6), 595.

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

Booth, G. H., and A. Alavi (2010), The Journal of Chemical Physics **132** (17), 174104, https://doi.org/10.1063/1.3407895.

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

Bravyi, S., J. M. Gambetta, A. Mezzacapo, and K. Temme (2017), arXiv preprint arXiv:1701.08213.

Bravyi, S. B., and A. Y. Kitaev (2002), Annals of Physics **298** (1), 210 .

Campbell, E. (2018), arXiv preprint arXiv:1811.08017.

Campbell, E. T., B. M. Terhal, and C. Vuillot (2017), Nature **549** (7671), 172.

Cao, Y., J. R. Fontalvo, and A. Aspuru-Guzik (2018a), IBM Journal of Research and Development , 1.

Cao, Y., J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferov, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik (2018b), arXiv:1812.09976.

Chen, L.-K., 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 (2017), Optica **4** (1), 77.

Childs, A. M., and R. Kothari (2011), in *Theory of Quantum Computation, Communication, and Cryptography*, edited by W. van Dam, V. M. Kendon, and S. Severini (Springer Berlin Heidelberg, Berlin, Heidelberg) pp. 94–103.

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

Childs, A. M., A. Ostrander, and Y. Su (2018), arXiv preprint arXiv:1805.08385.

Childs, A. M., and Y. Su (2019), arXiv:1901.00564.

Chin, S., and J. Huh (2018), arXiv preprint arXiv:1803.10002.

Chow, J. M., J. M. Gambetta, A. D. Córcoles, S. T. Merkel, J. A. Smolin, C. Rigetti, S. Poletto, G. A. Keefe, M. B. Rothwell, J. R. Rozen, M. B. Ketchen, and M. Steffen (2012), Phys. Rev. Lett. **109**, 060501.

Christiansen, O. (2012), Physical Chemistry Chemical Physics **14** (19), 6672.

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

Clements, W. R., J. J. Renema, A. Eckstein, A. A. Valido, A. Lita, T. Gerrits, S. W. Nam, W. S. Kolthammer, J. Huh, and I. A. Walmsley (2017), arXiv preprint arXiv:1710.08655.

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

Cubitt, T., and A. Montanaro (2016), SIAM Journal on Computing **45** (2), 268, https://doi.org/10.1137/140998287.

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

Dallaire-Demers, P.-L., and F. K. Wilhelm (2016a), Phys. Rev. A **93**, 032303.

Dallaire-Demers, P.-L., and F. K. Wilhelm (2016b), Phys. Rev. A **94**, 062304.

Devitt, S. J., W. J. Munro, and K. Nemoto (2013), Reports on Progress in Physics **76** (7), 076001.

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

Ditchfield, R., W. J. Hehre, and J. A. Pople (1971), The Journal of Chemical Physics **54** (2), 724.

DiVincenzo, D. P. (1995), Phys. Rev. A **51**, 1015.

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

Dunning Jr., T. H. (1989), The Journal of Chemical Physics **90** (2), 1007.

Dür, W., M. J. Bremner, and H. J. Briegel (2008), Phys. Rev. A **78**, 052325.

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

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

Endo, S., Q. Zhao, Y. Li, S. Benjamin, and X. Yuan (2018b), arXiv preprint arXiv:1808.03623.

Evans, M. G., and M. Polanyi (1935), Trans. Faraday Soc. **31**, 875.

Eyring, H. (1935), The Journal of Chemical Physics **3** (2), 107, https://doi.org/10.1063/1.1749604.

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

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

Farrelly, T. C., and A. J. Short (2014), Phys. Rev. A **89**, 012302.

Feynman, R. P. (1982), International Journal of Theoretical Physics **21** (6), 467.

Fingerhuth, M., T. Babej, and C. Ing (2018), "A quantum alternating operator ansatz with hard and soft constraints for lattice protein folding," arXiv:1810.13411.

Fowler, A. G., M. Mariantoni, J. M. Martinis, and A. N. Cleland (2012a), Physical Review A **86** (3), 032324.

Fowler, A. G., A. C. Whiteside, and L. C. L. Hollenberg (2012b), Phys. Rev. Lett. **108**, 180501.

Frisch, M. J., 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 (2016), "Gaussian16 Revision B.01," Gaussian Inc. Wallingford CT.

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

Ganzhorn, M., D. J. Egger, P. K. Barkoutsos, P. Ollitrault, G. Salis, N. Moll, A. Fuhrer, P. Mueller, S. Woerner, I. Tavernelli, and S. Filipp (2018), arXiv:1809.05057.

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

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

Ge, Y., J. Tura, and J. I. Cirac (2017), arXiv:1712.03193.

Genin, S. N., I. G. Ryabinkin, and A. F. Izmaylov (2019), arXiv:1901.04715.

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

Gottesman, D. (1998), Physical Review A **57** (1), 127.

Griffiths, D. J. (2016), *Introduction to quantum mechanics* (Cambridge University Press).

Haah, J., M. B. Hastings, R. Kothari, and G. H. Low (2018), arXiv preprint arXiv:1801.03922.

Harrow, A. W., and A. Montanaro (2017), Nature **549** (7671), 203.

Harty, T., D. Allcock, C. J. Ballance, L. Guidoni, H. Janacek, N. Linke, D. Stacey, and D. Lucas (2014), Physical review letters **113** (22), 220501.

Hastings, M. B., D. Wecker, B. Bauer, and M. Troyer (2015), QIC **15**.

Hattig, C., and F. Weigend (2000), The Journal of Chemical Physics **113** (13), 5154, https://aip.scitation.org/doi/pdf/10.1063/1.1290013.

Havlicek, V., M. Troyer, and J. D. Whitfield (2017), Phys. Rev. A **95**, 032332.

Hehre, W. J., R. F. Stewart, and J. A. Pople (1969), The Journal of Chemical Physics **51** (6), 2657, https://doi.org/10.1063/1.1672392.

Helgaker, T., P. Jorgensen, and J. Olsen (2014), *Molecular electronic-structure theory* (John Wiley & Sons).

Hempel, C., C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos (2018), Phys. Rev. X **8**, 031022.

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

Hobza, P., and J. poner (2002), Journal of the American Chemical Society **124** (39), 11802, pMID: 12296748, https://doi.org/10.1021/ja026759n.

Hoffmann, M. R., and J. Simons (1988), The Journal of Chemical Physics **88** (2), 993, https://doi.org/10.1063/1.454125.

Houck, A. A., H. E. Türeci, and J. Koch (2012), Nature Physics **8** (4), 292.

Hu, L., Y.-C. Ma, Y. Xu, W.-T. Wang, Y.-W. Ma, K. Liu, H.-Y. Wang, Y.-P. Song, M.-H. Yung, and L.-Y. Sun (2018), Science Bulletin **63** (5), 293 .

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

Huh, J., and M.-H. Yung (2017), Scientific reports **7** (1), 7462.

Huo, M., and Y. Li (2018), arXiv:1811.02734.

IBM, (2018), "Ibm qiskit," https://qiskit.org/.

Izmaylov, A. F., T.-C. Yen, and I. G. Ryabinkin (2018), arXiv preprint arXiv:1810.11602.

Jiang, Z., J. McClean, R. Babbush, and H. Neven (2018a), arXiv:1812.08190.

Jiang, Z., K. J. Sung, K. Kechedzhi, V. N. Smelyanskiy, and S. Boixo (2018b), Phys. Rev. Applied **9**, 044036.

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

Jones, N. C., J. D. Whitfield, P. L. McMahon, M.-H. Yung, R. Van Meter, A. Aspuru-Guzik, and Y. Yamamoto (2012), New Journal of Physics **14** (11), 115023.

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

Kais, S., S. A. Rice, and A. R. Dinner (2014), *Quantum information and computation for chemistry* (John Wiley & Sons).

Kandala, A., A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta (2017), Nature **549** (7671), 242.

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

Kassal, I., S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik (2008), Proceedings of the National Academy of Sciences **105** (48), 18681.

Kassal, I., J. D. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik (2011), Annual review of physical chemistry **62**, 185.

Kempe, J., A. Kitaev, and O. Regev (2006),

SIAM Journal on Computing **35** (5), 1070, https://doi.org/10.1137/S0097539704445226.

Kim, I. H. (2017), arXiv:1703.00032.

Kim, I. H., and B. Swingle (2017), arXiv:1711.07500.

Kitaev, A. Y. (1995), Preprint at http://arxiv.org/abs/quant-ph/9511026.

Kitaev, A. Y. (1997), Russian Mathematical Surveys **52** (6), 1191.

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

Kivlichan, I. D., N. Wiebe, R. Babbush, and A. Aspuru-Guzik (2017), Journal of Physics A: Mathematical and Theoretical **50** (30).

Knill, E., R. Laflamme, and W. Zurek (1996), .

Knill, E., G. Ortiz, and R. D. Somma (2007), Phys. Rev. A **75**, 012328.

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

Kreula, J., S. R. Clark, and D. Jaksch (2016), Scientific reports **6**, 32940.

Kurashige, Y., G. K.-L. Chan, and T. Yanai (2013), Nature chemistry **5** (8), 660.

Landauer, R. (1995), Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **353** (1703), 367.

Lanyon, B. P., 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.* (2010), Nature chemistry **2** (2), 106.

Lee, J., W. J. Huggins, M. Head-Gordon, and K. B. Whaley (0), Journal of Chemical Theory and Computation **0** (ja), null, https://doi.org/10.1021/acs.jctc.8b01004.

Li, Y., and S. C. Benjamin (2017), Phys. Rev. X **7**, 021050.

Li, Z., J. Li, N. S. Dattani, C. J. Umrigar, and G. K.-L. Chan (2018), arXiv:1809.10307.

Li, Z., M.-H. Yung, H. Chen, D. Lu, J. D. Whitfield, X. Peng, A. Aspuru-Guzik, and J. Du (2011), Scientific reports **1**, 88.

Lidar, D. A., and H. Wang (1999), Phys. Rev. E **59**, 2429.

Lieb, E. H., and D. W. Robinson (1972), Comm. Math. Phys. **28** (3), 251.

Lloyd, S. (1996), Science **273** (5278), 1073.

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

Low, G. H. (2018), arXiv preprint arXiv:1807.03967.

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

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

Low, G. H., and N. Wiebe (2018), arXiv:1805.00675.

Low, G. H., T. J. Yoder, and I. L. Chuang (2016), Phys. Rev. X **6**, 041067.

Lu, D., B. Xu, N. Xu, Z. Li, H. Chen, X. Peng, R. Xu, and J. Du (2012), Phys. Chem. Chem. Phys. **14**, 9411.

Macridin, A., P. Spentzouris, J. Amundson, and R. Harnik (2018a), Phys. Rev. A **98**, 042312.

Macridin, A., P. Spentzouris, J. Amundson, and R. Harnik (2018b), Phys. Rev. Lett. **121**, 110504.

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

McArdle, S., A. Mayorov, X. Shan, S. Benjamin, and X. Yuan (2018b), arXiv:1811.04069.

McArdle, S., X. Yuan, and S. Benjamin (2018c),

arXiv:1807.02467.

McClean, J., S. Boixo, V. Smelyanskiy, R. Babbush, and H. Neven (2018), Nature Communications **9**, 4812.

McClean, J. R., R. Babbush, P. J. Love, and A. Aspuru-Guzik (2014), The Journal of Physical Chemistry Letters **5** (24), 4368, pMID: 26273989, https://doi.org/10.1021/jz501649m.

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

McClean, J. R., 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 (2017b), arXiv:1710.07629.

McClean, J. R., J. Romero, R. Babbush, and A. Aspuru-Guzik (2016), New Journal of Physics **18** (2), 023023.

Mitarai, K., and K. Fujii (2018), arXiv:1901.00015.

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

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

Mueck, L. (2015), Nature chemistry **7** (5), 361.

Muller, R. P. (2004), "Pyquante," http://pyquante.sourceforge.net/, accessed: 24.7.2018.

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

O'Brien, T. E., B. Tarasinski, and B. M. Terhal (2018), arXiv:1809.09697.

O'Gorman, J., and E. T. Campbell (2017), Phys. Rev. A **95**, 032338.

Olivares-Amaya, R., W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. K.-L. Chan (2015), The Journal of Chemical Physics **142** (3), 034102, https://doi.org/10.1063/1.4905329.

O'Malley, P. J. J., 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 (2016), Phys. Rev. X **6**, 031007.

Ortiz, G., J. E. Gubernatis, E. Knill, and R. Laflamme (2001), Phys. Rev. A **64**, 022319.

Osborne, T. J. (2006), Phys. Rev. Lett. **97**, 157202.

Otten, M., and S. Gray (2018a), arXiv:1804.06969.

Otten, M., and S. Gray (2018b), arXiv:1806.07860.

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

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

Parrish, R. M., 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 (2017), Journal of Chemical Theory and Computation **13** (7), 3185, pMID: 28489372,

https://doi.org/10.1021/acs.jctc.7b00174.

Perdomo, A., C. Truncik, I. Tubert-Brohman, G. Rose, and A. Aspuru-Guzik (2008), Phys. Rev. A **78**, 012320.

Perdomo-Ortiz, A., N. Dickson, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik (2012), Scientific reports **2**, 571.

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

Podewitz, M., M. T. Stiebritz, and M. Reiher (2011), Faraday Discuss. **148**, 119.

Poulin, D., M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doherty, and M. Troyer (2015), QIC **15**.

Poulin, D., A. Kitaev, D. S. Steiger, M. B. Hastings, and M. Troyer (2018), Phys. Rev. Lett. **121**, 010501.

Poulin, D., A. Qarry, R. Somma, and F. Verstraete (2011), Phys. Rev. Lett. **106**, 170501.

Preskill, J. (2018), arXiv preprint arXiv:1801.00862.

Purvis, G. D., and R. J. Bartlett (1982), The Journal of Chemical Physics **76** (4), 1910, https://doi.org/10.1063/1.443164.

Raussendorf, R. (2012), Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **370** (1975), 4541, http://rsta.royalsocietypublishing.org/content/370/1975/4541.full.pdf.

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

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

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

Reiner, J.-M., F. Wilhelm-Mauch, G. Schn, and M. Marthaler (2018a), arXiv:1811.04476.

Reiner, J.-M., S. Zanker, I. Schwenk, J. Leppkangas, F. Wilhelm-Mauch, G. Schn, and M. Marthaler (2018b), Quantum Science and Technology **3** (4), 045008.

Richardson, L. F., B. J Arthur Gaunt, et al. (1927), Phil. Trans. R. Soc. Lond. A **226** (636-646), 299.

Romero, J., R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik (2019), Quantum Science and Technology **4** (1), 014008.

Romero, J., J. P. Olson, and A. Aspuru-Guzik (2017), Quantum Science and Technology **2** (4), 045001.

Roos, B. O., P. R. Taylor, and P. E. Sigbahn (1980), Chemical Physics **48** (2), 157 .

Rossi, E., G. L. Bendazzoli, S. Evangelisti, and D. Maynau (1999), Chemical Physics Letters **310** (5), 530 .

Rubin, N. C. (2016), arXiv:1610.06910.

Rubin, N. C., R. Babbush, and J. McClean (2018), New Journal of Physics **20** (5), 053020.

Ryabinkin, I. G., T.-C. Yen, S. N. Genin, and A. F. Izmaylov (0), Journal of Chemical Theory and Computation **0** (0), null, pMID: 30427679, https://doi.org/10.1021/acs.jctc.8b00932.

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

Sanders, Y. R., G. H. Low, A. Scherer, and D. W. Berry (2019), Phys. Rev. Lett. **122**, 020502.

Santagati, R., 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 (2018), Science Advances **4** (1), 10.1126/sciadv.aap9646.

Sawaya, N. P. D., and J. Huh (2018), arXiv:1812.10495.

Sawaya, N. P. D., M. Smelyanskiy, J. R. McClean,

and A. Aspuru-Guzik (2016), Journal of Chemical Theory and Computation **12** (7), 3097, pMID: 27254482, https://doi.org/10.1021/acs.jctc.6b00220.

Schneider, C., D. Porras, and T. Schaetz (2012), Reports on Progress in Physics **75** (2), 024401.

Seeley, J. T., M. J. Richard, and P. J. Love (2012), The Journal of Chemical Physics **137** (22), 224109.

Setia, K., S. Bravyi, A. Mezzacapo, and J. D. Whitfield (2018), arXiv:1810.05274.

Setia, K., and J. D. Whitfield (2018), The Journal of Chemical Physics **148** (16), 164104.

Sharma, S., K. Sivalingam, F. Neese, and G. K.-L. Chan (2014), Nature chemistry **6** (10), 927.

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

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

Shor, P. W. (1996), in *Foundations of Computer Science, 1996. Proceedings., 37th Annual Symposium on* (IEEE) pp. 56–65.

Smirnov, A. Y., S. Savel'ev, L. G. Mourokh, and F. Nori (2007), EPL (Europhysics Letters) **80** (6), 67008.

Song, C., J. Cui, H. Wang, J. Hao, H. Feng, and Y. Li (2018), arXiv:1812.10903.

Song, C., 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 (2017), Phys. Rev. Lett. **119**, 180511.

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

Sparrow, C., E. Martín-López, N. Maraviglia, A. Neville, C. Harrold, J. Carolan, Y. N. Joglekar, T. Hashimoto, N. Matsuda, J. L. OBrien, et al. (2018), Nature **557** (7707), 660.

Stephens, A. M. (2014), Phys. Rev. A **89**, 022321.

Steudtner, M., and S. Wehner (2018a), New Journal of Physics **20** (6), 063010.

Steudtner, M., and S. Wehner (2018b), arXiv:1810.02681.

Subramanian, S., S. Brierley, and R. Jozsa (2018), arXiv preprint arXiv:1806.06885.

Sugisaki, K., S. Yamamoto, S. Nakazawa, K. Toyota, K. Sato, D. Shiomi, and T. Takui (2016), The Journal of Physical Chemistry A **120** (32), 6459, pMID: 27499026, https://doi.org/10.1021/acs.jpca.6b04932.

Sugisaki, K., S. Yamamoto, S. Nakazawa, K. Toyota, K. Sato, D. Shiomi, and T. Takui (2018), Chemical Physics Letters: X , 100002.

Sun, Q., 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 (2017), Wiley Interdisciplinary Reviews: Computational Molecular Science **8** (1), e1340, https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1340.

Sung, K. (2018), "https://github.com/quantumlib/OpenFermion/issues/259," Accessed: 24.7.2018.

Suzuki, M. (1976), Communications in Mathematical Physics **51** (2), 183.

Svore, K. M., M. B. Hastings, and M. Freedman (2013), arXiv:1304.0741.

Szabo, A., and N. S. Ostlund (2012), *Modern quantum chemistry: introduction to advanced electronic structure theory* (Courier Corporation).

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

Terhal, B. M. (2015), Rev. Mod. Phys. **87**, 307.

Toloui, B., and P. J. Love (2013), arXiv:1312.2579.

Torrontegui, E., A. Ruschhaupt, D. Gury-Odelin, and J. G. Muga (2011), Journal of Physics B: Atomic, Molecular and Optical Physics **44** (19), 195302.

Tranter, A., P. J. Love, F. Mintert, and P. V. Coveney (2018), Journal of Chemical Theory and Computation **14** (11), 5617, pMID: 30189144, https://doi.org/10.1021/acs.jctc.8b00450.

Tranter, A., S. Sofia, J. Seeley, M. Kaicher, J. McClean, R. Babbush, P. V. Coveney, F. Mintert, F. Wilhelm, and P. J. Love (2015), International Journal of Quantum Chemistry **115** (19), 1431.

Trotter, H. F. (1959), Proceedings of the American Mathematical Society **10** (4), 545.

Tubman, N. M., J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley (2016), The Journal of Chemical Physics **145** (4), 044112, https://doi.org/10.1063/1.4955109.

Tubman, N. M., C. Mejuto-Zaera, J. M. Epstein, D. Hait, D. S. Levine, W. Huggins, Z. Jiang, J. R. McClean, R. Babbush, M. Head-Gordon, and K. B. Whaley (2018), arXiv:1809.05523.

Unruh, W. G. (1995), Phys. Rev. A **51**, 992.

Veis, L., and J. Pittner (2012), arXiv:1203.6204.

Veis, L., and J. Pittner (2014), The Journal of Chemical Physics **140** (21), 214111, https://doi.org/10.1063/1.4880755.

Veis, L., J. Višňák, T. Fleig, S. Knecht, T. Saue, L. Visscher, and J. c. v. Pittner (2012), Phys. Rev. A **85**, 030304.

Verstaete, F., and J. I. Cirac (2005), Journal of Statistical Mechanics: Theory and Experiment **2005**.

Wang, D., O. Higgott, and S. Brierley (2018), arXiv:1802.00171.

Wang, D. S., A. G. Fowler, and L. C. L. Hollenberg (2011), Phys. Rev. A **83**, 020302.

Wang, H., S. Kais, A. Aspuru-Guzik, and M. R. Hoffmann (2008), Phys. Chem. Chem. Phys. **10**, 5388.

Wang, X.-L., 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 (2016), Phys. Rev. Lett. **117**, 210502.

Wang, Y., F. Dolde, J. Biamonte, R. Babbush, V. Bergholm, S. Yang, I. Jakobi, P. Neumann, A. Aspuru-Guzik, J. D. Whitfield, *et al.* (2015), ACS nano **9** (8), 7769.

Ward, N. J., I. Kassal, and A. Aspuru-Guzik (2009), The Journal of Chemical Physics **130** (19), 194105.

Wecker, D., B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer (2014), Phys. Rev. A **90**, 022305.

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

Wecker, D., M. B. Hastings, N. Wiebe, B. K. Clark, C. Nayak, and M. Troyer (2015b), Phys. Rev. A **92**, 062318.

Wendin, G. (2017), Reports on Progress in Physics **80** (10), 106001.

White, S. R. (2017), The Journal of Chemical Physics **147** (24), 244102.

Whitfield, J., V. Havlicek, and M. Troyer (2016), Physical Review A **94** (3).

Whitfield, J. D. (2015), arXiv preprint arXiv:1502.03771.

Whitfield, J. D., J. Biamonte, and A. Aspuru-Guzik (2011), Molecular Physics **109** (5), 735.

Whitfield, J. D., P. J. Love, and A. Aspuru-Guzik (2013), Phys. Chem. Chem. Phys. **15**, 397.

Whitfield, J. D., M.-H. Yung, D. G. Tempel, S. Boixo, and A. Aspuru-Guzik (2014), New Journal of Physics **16** (8), 083035.

Wiebe, N., D. W. Berry, P. Høyer, and B. C. Sanders (2011), Journal of Physics A: Mathematical and Theoretical **44** (44), 445308.

Wiebe, N., and C. Granade (2016), Phys. Rev. Lett. **117**, 010503.

Wiesner, S. (1996), arXiv preprint quant-ph/9603028.

Xu, J.-S., M.-H. Yung, X.-Y. Xu, S. Boixo, Z.-W. Zhou, C.-F. Li, A. Aspuru-Guzik, and G.-C. Guo (2014), Nature Photonics **8**, 113.

Yamazaki, T., S. Matsuura, A. Narimani, A. Saidmuradov, and A. Zaribafiyan (2018), arXiv:1806.01305.

Yung, M.-H., J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano (2014), Scientific reports **4**, 3589.

Yung, M.-H., J. D. Whitfield, S. Boixo, D. G. Tempel, and A. Aspuru-Guzik (2012), arXiv:1203.1331.

Zalka, C. (1998), *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **454** (1969), 313.