Quantum computing for quantum chemistry

Phillip W. K. Jensen, Sumner Alperin-Lea, Robert A. Lang

December 24, 2018

Contents

1 Introduction

1.1 Background of the field

Quantum chemistry is the subfield of chemistry which concerns itself with the quantum mechanical description of molecular systems. Arguably, quantum chemistry became distinct from quantum physics with the first applications of quantum mechanics to molecular systems, at the end of the 1920's. [1] [2] Beginning in the 1960's, computational methods were developed which revolutionized the field and brought it to ascendancy. [3] [4] Increases in computational power and the refinement of computational methods advanced the field to such a degree that theoretical quantum chemistry is often, instead, called computational chemistry, though the term extends to other parts of physical chemistry. Computational methods allow scientists to make predictions about the quantum-mechanically dictated properties of molecules, such as the energy of the electrons in the molecule, the energy stored in its bonds, the physical geometry of the molecule, and so forth – predictions which cannot always be made by analytic mathematical methods.

We intend, here, to give a brief surveillance of the chemical applications of Quantum Computing. This introduction is meant, primarily, for those with a computer science background but without substantial knowledge of chemistry or physics.

1.2 The Fundamental problem of quantum chemistry

Summarized, quantum chemistry is about determining the energy of molecular quantum systems, composed of electrons and nuclei. Little focus, however, is given to the nuclei, which are usually approximated to be stationary – the so-called Born-Oppenheimer approximation. This approximation is easy to justify; even though nuclei and electrons are attracted to each other by the same potential, the mass of even the smallest nucleus is thousands of times greater than the mass of an electron; the force acting on the nuclei is therefore weak by comparison. The vast majority of attention, in quantum chemistry, is paid to determining the energies and dynamics of electrons – the so called Electronic problem. The Electronic problem, therefore, is the fundamental problem of quantum chemistry; the nuclear problem is most often an afterthought.

In its simplest form, quantum mechanics focuses on solving the Schrodinger equation for a system of particles under a specific potential. The time-independent, non-relativistic Schrodinger equation (TISE), which captures the dynamics of a typical quantum system which is not evolving in time, is an eigenvalue equation of the following form:

$$\hat{H}\psi = E\psi \tag{1}$$

where the operator $\hat{H}$ is the Hamiltonian operator of the system, which describes the energies of the electrons in the system, ψ is the wavefunction of the electrons, dependent upon the position of all the electrons in the system, and E is the eigenvalue of the Hamiltonian acting on the wavefunction ψ, which is understood as an energy. For a general K atom , N electron system, the Hamiltonian is of the following form:

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{a=1}^{K}\sum_{i=1}^{N}\frac{Z_a}{|r_i - R_a|} + \sum_{i\neq j}\frac{1}{|r_i - r_j|} + \sum_{a\neq b}\frac{Z_a Z_b}{|R_a - R_b|} \tag{2}$$

where ∇_i^2 is the kinetic energy of electron i, $\frac{Z_a}{|r_i - R_a|}$ is the Coulombic attraction exerted by atom a with Z_a protons on electron i, $\frac{1}{|r_i - r_j|}$ is the Coulombic repulsion that electrons i and j exert upon one another. The final term, $\sum_{a\neq b}\frac{Z_a Z_b}{|R_a - R_b|}$, is the Coulombic repulsion between nuclei, which is not negligible, but is usually treated as some constant when possible. In all cases, variables r and R should be understood as positions of the relevant objects. For ease of understanding, we use atomic units.

At first glance, it appears that this succinct general form should allow easy solutions to arbitrarily large systems with multiple electrons and multiple attractive potentials. Unfortunately, this is not the case in the slightest. Analytic solutions to the TISE are proven to be impossible for systems with two or more electrons, because of the Coulombic repulsion term; this is due to the general unsolvability of the three-body problem. This means that, out of the billions of different molecules in existence, we can solve the Schrodinger equation for exactly one of them: the molecular hydrogen ion, H_2^+. Adding insult to injury, if we were to dispose of the Born-Oppenheimer approximation, even the TISE of molecular hydrogen ion would be rendered impossible to solve.

Clearly, simulatory methods for obtaining the quantum mechanical energies of molecules are necessary, if quantum theory is to produce any quantitative predictions about molecular properties and dynamics.

1.3 Useful mathematics

One of the fundamental postulates of quantum mechanics is the essential *indistinguishability* of particles of the same species. From the indistinguishably of particles it follows that the interchange of particle coordinates in an N particle system has no affect upon the system's state. The resultant state function differs from the original by a simple phase factor λ, where $\lambda = e^{2i\pi S}$. S is the *spin number* of the particles in question. This "spin" is the intrinsic angular momentum of quantum particles along a fixed axis (generally said to be the z axis, the axis of measurement), so named for its seeming resemblance to classical spin.

$$\begin{aligned}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_k,\ldots,\mathbf{r}_N) &= \lambda\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_k,\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_N) \\ &= \lambda^2\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_k,\ldots,\mathbf{r}_N). \end{aligned} \tag{3}$$

If the same two coordinates are interchanged again, we return to the same state function. We can conclude that $\lambda^2 = 1$ or $\lambda = \pm 1$. For particles in real space, only two types of particles are possible – the *bosons*, $e^{2i\pi S} = 1 \rightarrow S = 0, 1, 2, \ldots$ and *fermions*, $e^{2i\pi S} = -1 \rightarrow S = 1/2, 3/2, \ldots$

$$\text{bosons}: \quad \Psi(\mathbf{r}_1,\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_k,\ldots,\mathbf{r}_N) = +\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_k,\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_N) \quad S = \pm 0, 1, 2\ldots \tag{4}$$

$$\text{fermions}: \quad \Psi(\mathbf{r}_1,\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_k,\ldots,\mathbf{r}_N) = -\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_k,\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_N) \quad S = \pm 1/2, 3/2\ldots. \tag{5}$$

Infinitely many bosons may occupy the same quantum state. This is not true for fermions. This leads to the *Pauli exclusion principle* – that the quantum states of each and every fermion in a system is unique, even though the fermions themselves are not. Electrons, which have spin one-half, are fermions. Thus, the wavefunctions of electronic systems must be *anti-symmetric* – an exchangle of particle indices must lead to a change of sign. To ensure anti-symmetry, quantum chemists often make use of wavefunction constructed from *Slater determinants* to construct anti-symmetric fermion wave functions:

$$\Psi(\mathbf{r}_1,\ldots\mathbf{r}_N) = \begin{vmatrix} f_1(\mathbf{r}_1) & f_2(\mathbf{r}_1) & \cdots & f_N(\mathbf{r}_1) \\ f_1(\mathbf{r}_2) & f_2(\mathbf{r}_2) & \cdots & f_N(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(\mathbf{r}_N) & f_2(\mathbf{r}_N) & \cdots & f_N(\mathbf{r}_N) \end{vmatrix} \quad \text{Slater determinant} \tag{6}$$

where $f(\mathbf{r})$ are arbitrary functions that live in the Hilbert space. For example, a two-particle state is given by,

$$\Psi(\mathbf{r}_1,\mathbf{r}_2) = \begin{vmatrix} f_1(\mathbf{r}_1) & f_2(\mathbf{r}_1) \\ f_1(\mathbf{r}_2) & f_2(\mathbf{r}_2) \end{vmatrix} = f_1(\mathbf{r}_1)f_2(\mathbf{r}_2) - f_1(\mathbf{r}_2)f_2(\mathbf{r}_1) = -\Psi(\mathbf{r}_2,\mathbf{r}_1). \tag{7}$$

1.4 Standard methods in computational chemistry, and their problems

We will focus on four protocols for approximating the wavefunction which are well known throughout the field. Each method is tied to an accompanying set of approximations meant to render the electronic problem tractable – that is, to deal with Coulombic repulsion. We will treat them rather generally.

1.4.1 Hartree-Fock

The dawn of quantum chemistry occurred with the creation of the *Hartree-Fock*[1, 2] method. This method is comparatively elegant and simple. It is a mean-field method, which assumes that the electrons are uncorrelated, with respect to their spins – that is to say, Hartree-Fock ignores the fact that two electrons with the same spin cannot occupy the same physical space simultaneously, though electrons of opposite spins can.

In Hartree-Fock theory, the wavefunction is represented by a single Slater determinant of N electrons, using N functions. These functions are called *spin orbitals*, which depend on three spatial coordinates plus the spin of the electron. In general, the functions which describe the behaviors of electrons are called orbitals, in reference to the analytical solutions of the TISE of the hydrogen atom, which they are assumed to resemble. Spin orbitals are a simple product of a spatial orbital (the hydrogenic orbitals are such orbitals) and a spin function, hence they always come in pairs – one spin "up", one "down". Hartree-Fock is a mean-field method because it assumes that the electron-electron repulsion can be treated in an average manner;we at first ignore electronic interactions, but reintroduce their average in calculating energies. We can re-express the normal electronic Hamiltonian as follows:

$$\hat{H}_{el} = \sum_{i}^{N} h(i) + \sum \sum_{i \neq j} v(i,j) + V_{NN} \tag{8}$$

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_{a=1}^{K} \frac{Z_a}{|r_i - R_a|} \tag{9}$$

$$v(i,j) = \frac{1}{|r_i - r_j|} \tag{10}$$

V_{NN} is a scalar, for fixed nuclear coordinates, which only affects the Hamiltonian's eigenvalues, but not its eigenfunctions. Via mathematics outside the scope of this treatment, calculating energies in the Hartree-Fock method is simply solving a set of one-electron and two-electron integrals:

$$E_{HF} = \sum_{i}^{N} \langle i|\, h \,|i\rangle + \frac{1}{2}\sum_{i \neq j} [ii|jj] - [ij|ji] \tag{11}$$

$$\langle i|\, h \,|j\rangle = \int \chi_i^*(\mathbf{x}_u) h(\mathbf{r}_u) \chi_j(\mathbf{x}_u) d\mathbf{x}_u \tag{12}$$

$$[ij|kl] = \int\int \chi_i^*(\mathbf{x}_u)\chi_j(\mathbf{x}_u) v(u,w) \chi_k^*(\mathbf{x}_w)\chi_l(\mathbf{x}_w) d\mathbf{x}_u d\mathbf{x}_w \tag{13}$$

where $\mathbf{x}_u$ is the three spatial dimensions of the electron u plus its spin, and the χ are the spin orbitals. Hartree-Fock is guided by the variational theorem.[5] This theorem, rendered simply, states that no approximation to the wavefunction of a quantum mechanical system can ever have an expectation value for the ground state energy which is lower, with respect to the system's Hamiltonian, than the ground state energy expectation value of the true wavefunction. Phrased fast and loose: You can't guess a better wavefunction than nature can create. Philosophically, this is grounded in our understanding that physical systems behave in whatever fashion minimizes their energy. The true wavefunction therefore *is* the wavefunction which has the lowest energy in the ground state. In variational methods, the wavefunction ansatz is simply altered iteratively until convergence. In the case of Hartree-Fock, this means iteratively changing the N spin orbitals. The integrals in Hartree-Fock theory are quite efficiently calculated on a computer. Unfortunately, the approximation of uncorrelation means that, in almost all cases, the Hartree-Fock method produces energies which are too far from the true energy of the system to be useful, even though E_{HF} is much larger, Nevertheless, sufficiently good heuristics make Hartree-Fock a powerful starting point for more advanced methods which incorporate correlation. These methods which build off Hartree-Fock are, naturally, called *post-Hartree-Fock* methods, which we shall now address.

1.4.2 Full configuratio interaction

The first post-Hartree-Fock method we will address is called *Full Configuration Interaction*[6, 7] full configuration interaction (FCI or full CI). In full CI, the wavefunction ansatz is a linear combination of all Slater determinants formed by all $1, 2..N$-fold *excitations* of electrons in the Hartree-Fock reference state $|HF\rangle$, barring those with incorrect symmetries. Full CI includes not only the N occupied spin orbitals of Hartree Fock, but instead draws from a basis set $\{\chi_i\}$. The FCI wavefunction is, then, as follows:

$$|\Psi\rangle = c_0 |HF\rangle + \sum_{a,r} C_a^r |\Phi_a^r\rangle + \sum_{a<b,r<s} C_{ab}^{rs} |\Phi_{ab}^{rs}\rangle + \sum_{a<b<c,r<s<t} C_{abc}^{rst} |\Phi_{abc}^{rst}\rangle \ldots \tag{14}$$

where $|\Phi_a^r\rangle$ is the Slater determinant by replacing the occupied spin orbital a with the unoccupied orbital r, and so on for higher excitations. All C are expansion coefficients. In the limit of an infinite basis set, full CI converges to an exact solution of the Schrodinger equation. In practice, full CI uses a finite basis set. FCI will, however, converge to the best representation of the wavefunction in the subspace spanned by the chosen basis set; in this sense, full CI is locally exact. To express the full CI hamiltonian, we write it as a matrix, where each element of the matrix is defined by its effect on functions in the basis set. Thus, if the basis set forms K determinants Φ_1 through Φ_k, the Hamiltonian is:

$$\begin{bmatrix} \langle\Phi_1| H |\Phi_1\rangle & \ldots & \langle\Phi_1| H |\Phi_k\rangle \\ \vdots & \ddots & \vdots \\ \langle\Phi_k| H |\Phi_1\rangle & \ldots & \langle\Phi_k| H |\Phi_k\rangle \end{bmatrix}$$

There are a well known set of rules for evaluating these matrix elements, called the Slater-Condon rules [6]. For identical determinants:

$$\langle\Phi_1| H |\Phi_1\rangle = \sum_i^K \langle i| h |i\rangle + \sum_{i>j} [ii|jj] - [ij|ji] \tag{15}$$

For determinants which differ by a single spin orbital:

$$|\Phi_1\rangle = |...ab...\rangle \tag{16}$$

$$|\Phi_2\rangle = |...cb...\rangle \tag{17}$$

$$\langle\Phi_1| H |\Phi_2\rangle = \langle a| h |c\rangle + \sum_i^K [ac|ii] - [ai|ic] \tag{18}$$

if the determinants differ by two spin orbitals:

$$|\Phi_1\rangle = |...ab...\rangle \tag{19}$$
$$|\Phi_2\rangle = |...cd...\rangle \tag{20}$$
$$\langle\Phi_1| H |\Phi_2\rangle = [ac|bd] - [ad|bc] \tag{21}$$

If the determinants differ by more than two orbitals, the corresponding matrix element is zero. After constructing the matrix in this fashion, one can diagonalize it and find its eigenvalues and eigenvectors. These eigenvalues are the energies of the system, with the lowest eigenvalue being the ground state energy; corresponding eigenvectors reveal the states of the system. By this method, FCI obtains not only a ground state energy, but an entire spectrum of energies.

1.4.3 Density Functional Theory

Because Full CI is a (locally) exact method, it is also a costly one. Computationally cheap correlated methods can produce surprisingly accurate approximations. One approximate post-HF method rose to usefulness well in advance of the others. This method is called *Density Functional Theory*, or DFT.[8, 3]

Amusingly, it is easier and more succinct to show the mathematics of DFT than to describe them. We do so for illustrative purposes. In DFT, the TISE is replaced by the Kohn-Sham equations, a set of eigenequations for the approximate system.

These equations are of the form:

$$(-\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r}))\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{22}$$

where ϕ_i is called a Kohn-Sham orbital, and ϵ_i is its corresponding energy. These Kohn-Sham orbitals describe individual electrons in the system and are chosen from a basis set; these basis sets are meant to be easy to numerically manipulate, whilst still "resembling" the hydrogenic orbitals. To progress further, we must define the electron density $\rho(\mathbf{r})$:

$$\rho(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \tag{23}$$

The total energy is a functional of this electron density – hence the name of the method. For convenience we will simplify notation, dropping $\mathbf{r}$ when possible.

$$E[\rho] = T[\rho] + \int v_{ext}\rho d\mathbf{r} + E_H[\rho] + E_{xc}[\rho] \tag{24}$$

Here,T[ρ] is the kinetic energy of the system, v_{ext} is the external potential – the nucleus-electron repulsion, E_H is the Hartree Energy, and E_{xc} is the exchange correlation energy. The exchange correlation energy is unknown and must be approximated – this is the crux of DFT. Kinetic energy and Hartree energy are defined as follows:

$$T[\rho] = \sum_i^N \int \phi_i^*(-\frac{1}{2}\nabla^2)\phi_i dr \tag{25}$$
$$E_H = \frac{1}{2}\int\int \frac{\rho(r)\rho(r')}{|r-r'|}drdr' \tag{26}$$

These definitions yield a definition for the effective potential:

$$v_{eff} = v_{ext} + \int \frac{\rho(r')}{|r - r'|} + \frac{\delta E_{xc}[\rho]}{\delta \rho(r)} \tag{27}$$

This definition allows us to relate the total energy of the system to the orbital energies:

$$E[\rho] = \sum_{i}^{N} \epsilon_i - E_H[\rho] + E_{xc}[\rho] - \int \frac{\delta E_{xc}[\rho]}{\delta \rho(r)} \rho(r) dr \tag{28}$$

Note that, in general, the individual orbital energies are not meaningful on their own.

Given the mathematics shown herein, the procedure of DFT is simply the variation of the chosen orbitals until the total energy is minimized, which occurs when the exchange correlation energy and the electron density are self-consistent – that is to say, when the chosen density function is minimal with respect to the exchange correlation it produces.

The choice of basis set in DFT is highly nontrivial. There are thousands of commonly used basis sets, varying in size and the functions chosen, and no (known) basis set uniformly outperforms the others. The *a priori* choice of what basis set produces the best solution for a given problem is an open question in DFT; heuristics based on experience predominate.

1.4.4 Coupled Cluster

Coupled cluster is the more recently refined of the two dominant methods; it can offer significant improvements over DFT, though it is more computationally costly. [9, 10] In coupled cluster theory, the wave-function of the system is described in terms of *correlating operators* acting on an uncorrelated wavefunction – usually, the Hartree-Fock wave function. These correlating operators are excitation operators. The effects of these correlating operators are, e.g, combined into a single operator, which is an exponential of the combined correlating operators.

Coupled cluster, it must be noted, is NOT a variational method in this formulation. In optimizing coupled cluster, one optimizes the combined correlating operator. A variational formulation of coupled cluster, called Unitary coupled cluster, does exist, but is beyond the scope of this work.

Because of the exponential ansatz operator, coupled cluster is size-extensive. This is particularly convenient – subsystems of large systems can be treated independently,when such a treatment is reflective of the underlying physics of the system. Commonly, coupled cluster is truncated to 1,2,3,or 4 levels of correlating operators; that is, single, double, triple, and quadruple excitations. Most often, the variant of coupled cluster theory employed is CCSD(T) – coupled cluster Single Double (Triple), where the parentheses indicate that the triple excitations are actually handled by perturbation theory, a powerful method which we will not address here. In the case that one includes N-fold excitation, where N is the number of electrons of the system, coupled cluster converges to FCI. [4, 9]

1.4.5 The shortcomings of current methods

Each of the methods above has its own particular weaknesses. Notably, their worst-case complexity is severe: Hartree-Fock itself is worst-case NP-complete[11], so methods which extend beyond it or include it as a preliminary subroutine have the potential to become seriously intractable.

This is, however, the *worst case* complexity, not the typical complexity-otherwise our general prospects would be grim indeed.

In the case of FCI, computational costs are severe – usually $O(N!)$. This rapidly renders it unusable, even with a truncated basis set.

The complexity class of DFT is dependent on the Hamiltonian of the system. It has been shown that the problem of finding a universal functional for DFT is QMA complete[12]. QMA is the Quantum Merlin-Arthur class, which can be thought of as the NP-Analogue to BQP. However, DFT rarely has greater than polynomial cost. Being a mean-field method, it typically has mean-field computational cost – $O(N^3)$. The computational cost of DFT is greatly reduced by the heuristics employed, particularly basis set choice, taking advantage of additional structure in the system.

Coupled cluster is currently the most accurate of the approximate methods in computational chemistry. When computational resources are sufficient, CC methods tend to be preferred over any other post-HF method. In general, CC methods scale as $O(N^7)$. [10] This average cost remains polynomial, like the typical DFT cost – but is prohibitive nonetheless. Additionally, CC methods can have serious convergence issues. Increasing the degree of excitation does not monotonically improve convergence, and when exponential-ansatz coupled cluster fails to converge, it can generate physically meaningless results, like infinite negative energies. In this sense, "when it rains, it pours" – When coupled cluster fails, it is because the equations it attempts to solve are unsolveable. Poor results from DFT are still physically *possible*, with a more smooth boundary between the regimes of success and failure.

Both DFT and coupled cluster fail to capture the dynamics of certain systems. Both methods fail in the event of strong correlation in general. This includes any system with a large number of un-coupled electrons – such as those containing transition metals, which includes the vast majority of important catalysts – are difficult for both DFT and coupled cluster methods. Coupled cluster also fares poorly for systems with a large number of degenerate – that is, equal energy – electrons, and may produce a number of solutions, with no criteria for choosing amongst them.

1.5 Quantum chemistry on a quantum computer

So far we have discussed the problems of specific methods in classical quantum chemistry. At this point, the more optimistic quantum Luddite may claim that we simply have not yet found the best classical algorithms. However, there are two problems which all classical methods suffer from which require address.

First: the cost of representation of quantum states on classical computers scales exponentially. On a quantum computer, state representation scales linearly. Systems which are impossible to represent classically (such as highly entangled systems) are cheap to represent on a quantum computer.

The more interesting failure of classical methods relates to time evolution. We return to the Schrodinger equation, this time its time dependent variant:

$$i\frac{\partial\psi(r,t)}{\partial t} = \hat{H}\psi(r,t) \tag{29}$$

where ψ depends on both position and time and $\hat{H}$ depends on position at least.For time-independent Hamiltonians:

$$e^{-i\hat{H}t}\psi(r,0) = \psi(r,t) \tag{30}$$

calculating the exponential term above is in general intractable on a classical computer. On (fault tolerant) quantum computers, such calculations can be performed efficiently for arbitrary t and a broad class of Hamiltonians.[13, 14]

These universal quantum advantages are the deepest source of the desire to use quantum computers for quantum chemistry – they can efficiently and exactly (with high probability) perform a wide range of Hamiltonian simulations which classical computers can at best approximate at high cost, with no guarantee of success or accuracy.

2 Quantum computational chemistry mapping to qubits

Before we can calculate molecular properties on a quantum computer, we need to express/map wave functions and Hamiltonians as qubits and gates, respectively. First, we will give a brief overview of second quantization, and finally show difference types of mapping methods.

2.1 Second quantization

We may describe a system of fermions in the formalism of second quantization, where N single particle states (i.e spin orbitals) are either occupied or unoccupied by a fermion, and consequently either existing or not existing in the first quantized Slater determinant appearing in the wavefunction describing the fermionic system. In first quantization, states are associated with (wave)functions, and observable quantities are associated with operators. In the formalism of second quantization, all observables *and* states are represented by operators, and the abstract construct of the vacuum state. Before transforming the electronic structure problem to the qubit space, we first wish to reformulate it in second quantization, for motivations which will become clear within this section.

2.1.1 The occupation number representation

The N-particle fermionic or bosonic wavefunction may be cast in the *occupation number representation* using second quantization, described by occupation vector

$$|k_1, k_2, k_3, \ldots, k_M\rangle, \tag{31}$$

where we have for bosons and fermions,

$$k_P = \begin{cases} 0, 1 & \text{fermions} \\ 0, 1, 2, \ldots & \text{bosons} \end{cases}$$

where k_P describes the occupation number of spin orbital P in the case of a molecular electronic wavefunction. If $k_P = 1$ we say the spin orbital is occupied, and if $k_P = 0$, the spin orbital is "virtual". The space spanned by the occupation number basis is the Fock space, defined as $\mathcal{F} = \mathcal{F}_0 \oplus \mathcal{F}_1 \oplus \mathcal{F}_3 \oplus \cdots$, where $\mathcal{F}_N = \text{span}\{|k_1, k_2 \ldots\rangle \,|\, \sum_P k_P = N\}$. The all-virtual occupation state $|0_1, 0_2, \ldots\rangle$ is referred to as the vacuum state, corresponding to a 0×0 determinant in first quantization. We will see how to populate and perform operations on occupation vectors while keeping the Slater determinant it represents physical within the subsection below.

2.1.2 The fermion creation and annihilation operators

Given the occupation number operator, starting as the vacuum state, it is natural to introduce the *creation operator*, denoted as $a_j^\dagger$, which adds a particle to virtual state j.

$$a_P^\dagger |k_1, k_2, \ldots, 0_P, \ldots, k_M\rangle = \Gamma_p |k_1, k_2, \ldots, 1_P, \ldots, k_M\rangle, \tag{32}$$

where $\Gamma_p = \prod_{Q=1}^{p-1}(-1)^{k_Q}$ is the phasefactor which equal to +1 if there are an even number of electrons in the spin orbitals $Q < P$ and equal to *-1* if there are an odd number of electrons in the spin orbitals. Consequently, one defines the *annihilation operator* $\langle k_P + 1|a_P^\dagger|k_P\rangle = \langle (a_P^\dagger)^\dagger k_P + 1|k_P\rangle = \langle a_P k_P + 1|k_P\rangle$ thus a_P removes a particle i.e. making the state virtual. Removing a particle from a empty state "kills" the state $a\,|0\rangle = 0$ and $a^\dagger\,|1\rangle = 0$. Due to the anti-symmetry requirement for fermions, we have the following *anti-commutation relations*,

$$\{a_i, a_j^\dagger\} = \delta_{ij}, \quad \{a_i, a_j\} = 0, \quad \{a_i^\dagger, a_j^\dagger\} = 0, \tag{33}$$

where $\{A, B\} = AB + BA$.

For example, consider the Hydrogen atom (1 proton and 1 electron) with its first five energy eigenstates, $|1s\rangle$, $|2s\rangle$, $|2p_x\rangle$, $|2p_y\rangle$, $|2p_z\rangle$, illustrated in Fig. 1(a). In the case of the Hydrogen atom, the energy eigenstates are exact solutions to the time-independent Schrödinger equation, $\hat{H}\,|\psi\rangle = E\,|\psi\rangle$, where $|1s\rangle$ has the lowest eigenvalue (lowest energy) known as the ground state, and the other states have same energy known as the 1st excited states. The shape of the energy eigenstates represents the electron probability distribution. That is, $|1s\rangle$ and $|2s\rangle$ have a finite probability to find the electron infinite close to the nucleus, whereas $|2p_x\rangle$, $|2p_y\rangle$, and $|2p_z\rangle$ have zero probability of finding the electron at the nucleus, but increasing probability to find electron more away from the nucleus. In first quantization, we would write the ground state as $\Psi(\mathbf{r}) = |1s\rangle$. In second quantization, we write the ground state as $|1_1 0_2 0_3 0_4 0_5\rangle$ meaning only the lowest energy eigenstate is occupied and all other are virtual states, illustrated with a white background.

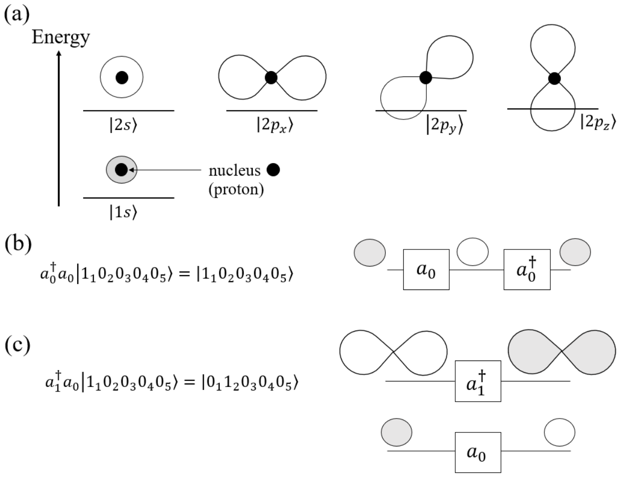


Figure 1: (a) The first five energy eigenstates of the Hydrogen atom. The $|1s\rangle$ is occupied indicated with a shaded grey background. (b) The occupation number operator, $a_0^\dagger a_0$, acting on the ground state $|1_1 0_2 0_3 0_4 0_5\rangle$. (c) An excitation operator, $a_1^\dagger a_0$, it takes the electron from the ground state and add it to $|2p_z\rangle$.

Fig. 1(b) shows a schematic representation of the occupation number operator, $a_0^\dagger a_0$, acting on the ground state, $|1_1 0_2 0_3 0_4 0_5\rangle$. The occupation number operator simple counts the number of occupied particles in the system, that is it first takes an electron, a_0, and then add it back to the same state $a_0^\dagger$. Likewise, one can make an excitation operator, $a_1^\dagger a_0$, which take an electron from the ground state and add it to a higher energy eigenstate as illustrated Fig. 1(c)

2.1.3 The molecular electronic Hamiltonian

From the creation and annihilation operators, we can now construct the full second quantization non-relativistic and spin-free molecular electronic Hamiltonian in the Born-Oppenheimer approximation, given by,

$$\hat{H} = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2}\sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s + h_{\text{nuc}} \tag{34}$$

where the first term is the one-electron operator and the second term is the two-electron term. The integrals h_{pq} and h_{pqrs} (in atomic units) are given as,

$$h_{pq} = \int d\mathbf{r}\chi_p^*(\mathbf{r})\bigg(-\frac{1}{2}\nabla^2 - \sum_\alpha \frac{Z_\alpha}{r_{\alpha\mathbf{r}}}\bigg)\chi_q(\mathbf{x}), \quad h_{pqrs} = \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\chi_p^*(\mathbf{r}_1)\chi_q^*(\mathbf{r}_2)\chi_r(\mathbf{r}_2)\chi_s(\mathbf{r}_1)}{r_{12}}, \tag{35}$$

where $\chi_p(\mathbf{r})$ are a selected single-particle basis. The h_{nuc} term is the nuclear-nuclear repulsion energy.

2.2 Jordan-Wigner encoding

The Jordan-Wigner (JW) transformation is used to map a system of fermions to a system of qubits, and vice versa. The occupation vector considering m orbitals used in second quantization lends itself to a natural mapping to an m-qubit register, which is referred to as the *Jordan-Wigner basis* of qubits.

$$|k_1, k_2, \ldots, k_m\rangle \quad \rightarrow \quad |q_1\rangle \otimes |q_2\rangle \otimes \ldots \otimes |q_m\rangle\,, \quad k_i = q_i \in \{0, 1\} \tag{36}$$

The remaining step is to transform the second quantized annihilation and creation operators to the qubit space in form of gate primitives such that the transformation 1) is isospectral and 2) retains the proper anticommutation relations. Our first (naive) attempt: $a_j^\dagger \rightarrow 1^{\otimes N-j-1} \otimes \sigma_-$ and $a_j \rightarrow 1^{\otimes N-j-1} \otimes \sigma_+$ where $\sigma_- \equiv (\sigma_x - i\sigma_y)/2 = |1\rangle\langle 0|$ and $\sigma_+ \equiv (\sigma_x + i\sigma_y)/2 = |0\rangle\langle 1|$, gives the correct same-site fermionic relations $\{a_j, a_j^\dagger\} = 1$, however, on different sites, we have the relation $[a_i, a_j^\dagger] = 0$, and so spins on different sites commute unlike fermions which anti-commute. A transformation which recovers the true fermion commutation relations from spin-operators was performed in 1928 by Jordan and Wigner, in the original context of spin-lattice models, given by,

$$a_j^\dagger \rightarrow 1^{\otimes N-j-1} \otimes \sigma_- \otimes \sigma_z^{\otimes j} = 1^{\otimes N-j-1} \otimes \frac{1}{2}\big(\sigma_X - i\sigma_Y\big) \otimes \sigma_z^{\otimes j}. \tag{37}$$

$$a_j \rightarrow 1^{\otimes N-j-1} \otimes \sigma_+ \otimes \sigma_z^{\otimes j} = 1^{\otimes N-j-1} \otimes \frac{1}{2}\big(\sigma_X + i\sigma_Y\big) \otimes \sigma_z^{\otimes j}. \tag{38}$$

The σ_-/σ_+ operator changes the occupation number of the target orbital, while the string of σ_z operators recovers the anti-commutation relations by introducing the appropriate phase change.

2.3 Alternative encodings

An undesirable property of the Jordan-Wigner transformation is the number of additional qubit operations required to implement a single annihilation or creation operator scales as $O(m)$ for an m qubit register and correspondingly consideration of m spin orbitals. This is a result of the trailing σ_z operations in Eq. 37 and 38 required to ensure anticommutation relations are satisfied. A potential remedy is to use a qubit basis which does not store occupation numbers locally, but rather information about all orbitals above or below it.[15] Non-local storing of orbital information is used in the *parity* encoding, where individual qubits store the *parity* of all orbitals below it, i.e) qubit i stores p_i,

$$p_i = \sum_{j=0}^{i} k_j \tag{39}$$

where the sum is taken $\mod 2$. The occupation vector $|f\rangle$ can be mapped to vector $|p\rangle$ in the parity basis by transformation

$$p_i = \sum_j [\pi_n]_{ij} k_j \tag{40}$$

where π is upper triangular matrix defined

$$[\pi_n]_{ij} = \begin{cases} 1 & \text{for } i < j \\ 0 & \text{for } i \geq j \end{cases} \tag{41}$$

The transformation equivalent to the Jordan-Wigner transformation in the parity basis is defined

$$a_j^\dagger \equiv \frac{1}{2}\sigma_x^{\otimes N-j-1} \otimes (\sigma_{x,j} \otimes \sigma_{z,j-1} - \sigma_{y,j} \otimes 1_{j-1}) \otimes 1^{\otimes j} \tag{42}$$

$$a_j \equiv \frac{1}{2}\sigma_x^{\otimes N-j-1} \otimes (\sigma_{x,j} \otimes \sigma_{z,j-1} + \sigma_{y,j} \otimes 1_{j-1}) \otimes 1^{\otimes j} \tag{43}$$

One can see by inspection that the parity transformation does not improve on the gate complexity of the field operators when compared to the JW transformation. Instead of the $O(m)$ trailing σ_z gates, we now have $O(m)$ leading σ_x gates required to update the parity of all orbitals above the index acted on by the creation or annihilation operator, since these indices store the sum of all occupations below it. Effectively we have traded off the non-locality of parity for the non-locality of occupation going from the JW to parity basis.

A third transformation exists, named after Bravyi and Kitaev. The Bravyi-Kitaev (BK) transformation achieves $O(\log m)$ gate complexity for implementation of a single field operator. [15, 16] The inclusion of the mathematics of the BK transformation are beyond the scope of this survey, but whose general scheme may be summarized as follows. Within the BK basis, for which a transformation from occupation basis exists, index i stores occupation number of orbital i. For even i, this is the only occupation stored at index i, however for odd i, index i stores the partial sum over a certain set of adjacent orbitals of index less than i. In effect, the BK transformation achieves a middle ground between the non-locality of occupation and parity, resulting in asymptotically fewer gates required to implement the creation and annihilation operators. Since then, increasingly sophisticated fermion-to-qubit transformations have been proposed. [17]

3 Quantum algorithms for chemistry

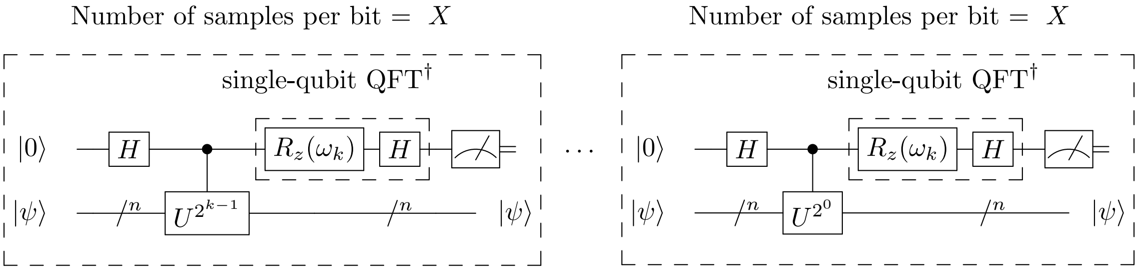
Within this section we outline the two main methodologies of solving the eigenvalue problem for the molecular electronic Hamiltonian with exponential speedup over traditional methods, following one of the transformations to qubit space summarized in Section 2. We first describe the iterative quantum phase estimation, followed by the variational quantum eigenvalue solver.

3.1 Iterative quantum phase estimation

The quantum phase estimation (QPE) algorithm is used in quantum chemistry to calculate molecular energies. Given the time-dependent Schrödinger equation (in atomic units $\hbar \equiv 1$) $i\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \hat{H}(\mathbf{r})\Psi(\mathbf{r},t)$

which has the solution $\Psi(\mathbf{r},t) = e^{-i\hat{H}(\mathbf{r})t}\Psi(\mathbf{r},0) = e^{-iEt}\Psi(\mathbf{r},0)$ where $\hat{H}(\mathbf{r})\Psi(\mathbf{r},0) = E\Psi(\mathbf{r},0)$ assuming the Hamiltonian is time-independent. Thus, we have the unitary operator $e^{-i\hat{H}(\mathbf{r})t}\Psi(\mathbf{r},0) = e^{-iEt}\Psi(\mathbf{r},0)$ suitable for the QPE. The QPE algorithm determines the phase i.e., $\theta = -Et \rightarrow E = -\frac{\theta}{t}$, hence the energy of the molecule.

The iterative QPE circuit shown below is implemented on a optical quantum computer[18].



The iterative QPE only has a single ancilla qubit, $|0\rangle$, hence the circuit produces only a single bit output. Due to the probability that the measurement gives the wrong bit, we iterative run the circuit X times until a pattern show the correct bit. We choose the bit by counting the number of 0 and 1 e.g. if number of samples per bit is 31 and say we obtain 10 0's and 21 1's, then we will choose the bit 1. We encode the eigenstate, $|\psi\rangle$, with n qubits. Thus, we have to encode the eigenstate prior to the QPE.

3.2 Variational quantum eigensolver

QPE offers an exponential speedup over traditional methods to finding the ground state eigenvalue with arbitrary precision given a molecular electronic Hamiltonian. However, consecutive applications of implementing the exponentiated Hamiltonian e^{iHt} can require on the order of millions or billions of quantum gates. An alternative method proposed to be more feasible for current and near term devices was the variational quantum eigenvalue solver, or variational quantum eigensolver (VQE) by Perruzo *et al.* [19], in attempt to reduce the required coherent evolution times to efficiently solve the electronic structure problem and potentially other interesting eigenvalue problems.

The VQE is a quantum-classical hybrid technique which can evaluate the expectation energy $\langle H\rangle$ of a $2^n \times 2^n$ qubit-space Hamiltonian using n qubits, providing an exponential speedup over exact diagonalization. Within the VQE, a quantum processing unit (QPU) performs a quantum expectation estimation (QEE) algorithm. The QEE algorithm capitalizes on the linearity of the qubit-space Hamiltonian. For qubit-space Hamiltonian $\hat{H}$ mapped from second quantized form via JW or alternative encoding method,

$$\hat{H} = \sum_i C_i \hat{P}_i \tag{44}$$

. $\hat{P}_i$ are n-fold tensor products of simple Pauli operations, C_i are real coefficients (precomputed as a classical overhead), and $\hat{H}$ has a polynomial number of terms. It follows from linearity that the expectation energy is

$$\langle H\rangle = \sum_i C_i \langle P_i\rangle \ . \tag{45}$$

Given an input state $|\psi\rangle$, the QEE efficiently evaluates the expectation value of simple Pauli words $\hat{P}_i$ on n qubits, which can then be classically substituted into Eq.45 to obtain the total expectation energy. In essence, the QEE replaces the long evolution times of QPE with many short evolutions. However, the QEE, similarly to QPE, suffers from being dependent on being given a wavefunction ansatz sufficiently close to the exact ground state wavefunction to output an expectation energy approaching the lowest eigenvalue. As a remedy to this, the VQE framework introduces a classical optimization scheme, ran on a classical computer, to variationally optimize the wavefunction ansatz entering the QEE.

In essence, the VQE solves a parametrized minimization problem, where the expectation energy may be reformulated as

$$E(\Omega, \tau) = \left\langle \psi(\Omega, \tau) \middle| \hat{H} \middle| \psi(\Omega, \tau) \right\rangle = \sum_i C_i \left\langle \psi(\Omega, \tau) \middle| \hat{P}_i \middle| \psi(\Omega, \tau) \right\rangle \tag{46}$$

where Ω and τ are wavefunction parameters which are chosen by classical optimization algorithm to minimize E,

$$E = \min_{\Omega, \tau} E(\Omega, \tau). \tag{47}$$

Regular convention for parametrization of wavefunction $|\psi(\Omega, \tau)\rangle$ is:

$$|\psi(\Omega, \tau)\rangle = U_{ENT}(\tau) U_{MF}(\Omega) |0\rangle \tag{48}$$

where $|0\rangle$ is the all zero state. $U_{MF}(\Omega)$ are mean-field single-qubit rotation unitaries, analogous to the orbital rotations conducted in the Hartree-Fock procedure. $U_{ENT}(\tau)$ are multi-qubit entangling qubits, which introduce electron-electron correlation. [20]

The VQE method may be summarized as follows. Given an initially prepared wavefunction ansatz, the expectation energy for individual Pauli words appearing in the Hamiltonian are efficiently calculated using a QPU, then classically multiplied by their corresponding real coefficients and summed to give the total electronic expectation energy $E(\Omega, \tau)$. $E(\Omega, \tau)$ is then the input to the classical optimization algorithm which then tunes wavefunction parameters Ω and τ. Wavefunction parameters are used to characterize the mean-field and entangling unitaries in Eq. 48 to produce an improved wavefunction ansatz, whose expectation energy is then calculated as in iteration before. This process is repeated until the expectation energy has sufficiently converged.

4 Current issues

Currently, NISQ architectures have limited quality and quantity of quantum resources with $\sim 5-10$ fully interacting qubits and short coherent times. This implies that for quantum chemistry simulations, we can include at best $\sim m = 5-10$ spin orbitals. In minimal basis, this covers atoms and molecules with up m electrons. Examples of such molecules with $\sim 5-10$ valence electrons are H_2O, CH_4, NH_3, etc. These molecules, while potentially chemically interesting, are small enough to be simulated on classical computers to satisfactory accuracy by methods such as coupled cluster or full CI. The largest molecule to be simulated experimentally using the VQE method to date is BeH_2 (four valence electrons) on a superconducting qubit quantum computer with limited accuracy. [21]

Most chemically interesting molecules are much larger than $5-10$ electrons, and chemical accuracy requires more sophisticated and larger basis sets. As a pharmaceutical example, Ibuprofen, $C_{13}H_{18}O_2$, contains 82 valence electrons requiring $m = 82$ qubits at the very least to be efficiently simulated on a quantum computer. For simulating 100 spin orbitals, approximately 10^{12} gates are required excluding overhead of error correction[10]. Thus, until larger-scale and noise-resistant quantum computers are accessible, techniques and algorithms for using quantum computers to solve electronic structure problems must be made as frugal as possible with regard to qubits and gate requirements. [22]

5 Conclusion

Within this survey we provide a general overview of quantum chemistry in the context of an introduction to quantum computing for quantum chemistry. The purpose is to provide a literature review of quantum computational chemistry, and bridge the gap between those scientists working on quantum information, and those working on computational chemistry.

The motivation of using quantum algorithms for quantum chemistry is to solve the electronic structure problem i.e., the time-independent Schrödinger equation, faster and more efficiently which could enable investigation of systems, that are impossible, or limited, with classical algorithms. Even for the simplest methods as Hatree-Fock, the complexity in the worst-case is NP-complete behavior[11].

Using a quantum computer to calculate molecular properties, the first step is to obtain the second quantized Hamiltonian. The occupation vector used in second quantization lends ifself to a natural mapping to an qubit register. The second quantized Hamiltonian is map into qubit gate operators using either 1) the Jordan-Wigner transformation scales as $O(m)$ 2) the parity coding scales as $O(m)$ 3) the Bravyi-Kitaev scales as $O(\log m)$. Further, we demonstrate two quantum algorithms, the quantum phase estimation and variational quantum eigensolver, to solve the eigenvalue problem.

Still, there is a long way to calculate molecular properties for larger molecules on a quantum computer. The high error rates and low qubit counts of current hardware make it impossible, at this state, to focus on larger molecules.

References

[1] D. R. Hartree, "The wave mechanics of an atom with a non-coulomb central field. part i. theory and methods," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 24, no. 1, p. 89110, 1928.

[2] V. Fock, "Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems," *Zeitschrift für Physik*, vol. 61, pp. 126–148, Jan 1930.

[3] A. D. Becke, "Perspective: Fifty years of density-functional theory in chemical physics," *The Journal of Chemical Physics*, vol. 140, no. 18, p. 18A301, 2014.

[4] R. J. Bartlett and M. Musiał, "Coupled-cluster theory in quantum chemistry," *Rev. Mod. Phys.*, vol. 79, pp. 291–352, Feb 2007.

[5] J. J. Sakurai, *Modern quantum mechanics; rev. ed.* Reading, MA: Addison-Wesley, 1994.

[6] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory.* Mineola: Dover Publications, Inc., first ed., 1996.

[7] C. D. Sherrill, "An introduction to configuration interaction theory," *School of Chemistry and Biochemistry, Georgia Institute of Technology*, 1995.

[8] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.*, vol. 140, pp. A1133–A1138, Nov 1965.

[9] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory.* Chichester: John Wiley & Sons, LTD, 2000.

[10] S. McArdle, S. Endo, A. Aspuru-Guzik, S. Benjamin, and X. Yuan, "Quantum computational chemistry," *ArXiv:1808.10402*, 2018.

[11] J. D. Whitfield and Z. Zimborá, "On the np-completeness of the hartree-fock method for translationally invariant systems," *The Journal of Chemical Physics*, vol. 141, no. 23, p. 234103, 2014.

[12] J. D. Whitfield, P. J. Love, and A. Aspuru-Guzik, "Computational complexity in electronic structure," *Phys. Chem. Chem. Phys.*, vol. 15, pp. 397–411, 2013.

[13] S. Lloyd, "Universal quantum simulators," *Science*, vol. 273, no. 5278, pp. 1073–1078, 1996.

[14] I. D. Kivlichan, N. Wiebe, R. Babbush, and A. Aspuru-Guzik, "Bounding the costs of quantum simulation of many-body physics in real space," *Journal of Physics A: Mathematical and Theoretical*, vol. 50, no. 30, p. 305301, 2017.

[15] J. T. Seeley, M. J. Richard, and P. J. Love, "The bravyi-kitaev transformation for quantum computation of electronic structure," *The Journal of Chemical Physics*, vol. 137, no. 22, p. 224109, 2012.

[16] S. B. Bravyi and A. Y. Kitaev, "Fermionic quantum computation," *Annals of Physics*, vol. 298, no. 1, pp. 210 – 226, 2002.

[17] K. Setia and J. D. Whitfield, "Bravyi-kitaev superfast simulation of fermions on a quantum computer," *arXiv:1712.00446 [quant-ph]*, 2017.

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

[19] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. OâEUR(tm)Brien, "A variational eigenvalue solver on a photonic quantum processor," *Nature Communications*, vol. 5, pp. 4213 EP –, Jul 2014. Article.

[20] I. G. Ryabinkin, S. N. Genin, and A. F. Izmaylov, "Constrained variational quantum eigensolver: Quantum computer search engine in the fock space," *arXiv:1806.00461 [physics.chem-ph]*, 2018.

[21] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," *Nature*, Sep 2017.

[22] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and G. K.-L. Chan, "Low-depth quantum simulation of materials," *Phys. Rev. X*, vol. 8, p. 011044, Mar 2018.