The Variational Quantum Eigensolver

Zixin Chen, Shoumik Chowdhury, Shantanu Jha, and Yiming Zhang

Department of Physics, Yale University, New Haven, CT 06510 [a]

In this final project report, we review the basic theory behind the variational quantum eigensolver (VQE) algorithm and discuss various applications to quantum computation and information. We also employ the VQE algorithm to simulate the ground state energies of small molecules, before reviewing a number of practical considerations for implementation of VQE.

1. Introduction

Quantum computers promise to deliver exponential speedups over their classical counterparts for certain classes of computational problems [1–3]. Among the most notable examples are eigenvalue and optimization problems, which have key applications in finance [4, 5], machine learning, and simulations of quantum chemistry [6, 7] — the latter dating back to Richard Feynman's original proposal for quantum computation [8]. Some of the earliest theoretical quantum algorithms providing speedups to such eigenvalue and optimization problems often relied on the use of quantum phase estimation [9, 10] or the quantum Fourier transform [11–13]. These original proposals also assumed the availability universal fault-tolerant quantum computers. Nevertheless, these ideas subsequently inspired the simulation of important physical processes in chemistry, such as electronic structure, chemical dynamics, and protein folding [10, 14–21].

In contrast to the theoretical advances above, most current physical quantum computers are still small-scale machines and largely prone to errors and decoherence beyond the fault-tolerance threshold. These so-called Noisy Intermediate-Scale Quantum (NISQ) devices operate with fewer than 100 qubits and shallow quantum circuit depths, limiting the size and scope of quantum algorithms that can be implemented on them [22]. Despite these hardware limitations, the advent of NISQ-era quantum computing has spurred research into short-depth quantum circuits and hybrid quantum-classical algorithms that make use of quantum computers in conjunction with

classical optimization techniques [23, 24]. Such hybrid quantum-classical algorithms allow for the possibility of performing useful computational tasks, even with NISQ devices. A crucial milestone in this direction was the invention of the Variational Quantum Eigensolver (VQE) algorithm by Peruzzo and McClean et. al. in Ref. [25], which uses hybrid quantum-classical resources to find variational solutions to certain eigenvalue and and optimization problems — in particular, computing molecular ground state energies. This algorithm was initially demonstrated experimentally using quantum photonic chips [25], but has since been implemented on other platforms, such as trapped-ions [26–29] and superconducting circuits [30–33]. In the recent work from the Google AI Quantum group [33], the authors also notably reported variational simulations of H_6, H_8, H_{10}, and H_{12} binding energies to within chemical accuracy. Many of the recent experimental and theoretical results, including Ref. [33], have also crucially relied on powerful error mitigation strategies to improve the performance of the VQE algorithm in the presence of noise [28, 32, 34–36].

In this final paper, we will explore the theoretical background, ongoing research directions, and exciting applications of the Variational Quantum Eigensolver. We review basic theory and notation in Section 2, before discussing notable applications of the VQE algorithm in chemistry and combinatorial optimization in Section 3. Here, we also address questions of runtime and query complexity of VQE; state preparation and Hamiltonian encoding for quantum chemistry; and choice of classical optimizer. In Section 4, we motivate one- and two-qubit variational forms in VQE. We then subsequently demonstrate our simulation results on modelling the ground state energies of small molecules (H_2 and LiH) with varying interatomic distances, reproducing results from Ref. [31].

---

[a] **Electronic addresses:** {zixin.chen, shoumik.chowdhury, shantanu.jha, yiming.zhang}@yale.edu

Here we also briefly discuss a hardware-efficient implementation of VQE circuits and other experimental considerations. Finally, in Section 5, we review novel error mitigation techniques that have played a crucial role in several recent VQE experiments — most notably, the use of Richardson zero-noise extrapolation in Ref. [32].

2. Theoretical Background & Formulation

In this section, we will follow the notation of Ref. [23]. We consider an N-qubit quantum computer $\mathcal{S}$ — and another physical quantum system $\mathcal{Q}$ with Hamiltonian $\hat{H}$, that we wish to simulate. We only require that $\dim(\mathcal{Q}) \leq \dim(\mathcal{S}) = N$.
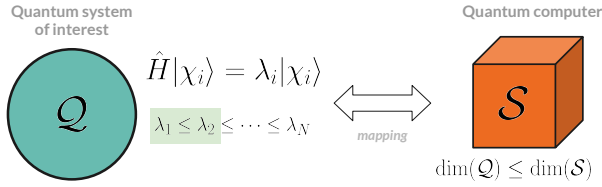

Quantum system of interest

$$\hat{H}|\chi_i\rangle = \lambda_i|\chi_i\rangle$$

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$$

Quantum computer

$$\mathcal{S}$$

mapping

$$\dim(\mathcal{Q}) \leq \dim(\mathcal{S})$$

FIG. 1: We consider a quantum system $\mathcal{Q}$ with Hamiltonian $\hat{H}$, where $|\chi_i\rangle$ and λ_i are the eigenvectors and eigenvalues of $\hat{H}$ respectively. We simulate this on a quantum computer $\mathcal{S}$.

Note that here, $\hat{H}$ could represent the molecular Hamiltonian of some molecule or interacting electronic system of interest; but it could also be an encoding of some other optimization problem whose minimum we wish to compute. Our goal will be to calculate and study eigenvectors $|\chi_i\rangle$ and eigenvalues λ_i of the Hamiltonian $\hat{H}$ using our quantum computer $\mathcal{S}$. Now, without loss of generality, we can order the non-negative eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$. In many cases, however, we will primarily only be concerned with the lowest few eigenvalues — for instance ground state energies in the application to quantum chemistry, since these energy levels often determine the relevant chemical properties of interest. Thus, an important problem we wish to solve is how best to determine λ_1. To this end, let us recall the well-known variational theorem:

$$\langle \hat{H}(\vec{\theta})\rangle = \langle\psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle = \lambda_\theta \geq \lambda_{\min} = \lambda_1 \quad (1)$$

which states that the expectation value of $\hat{H}$ with respect to an arbitrary state $|\psi(\vec{\theta})\rangle$ is bounded below by the smallest eigenvalue of $\hat{H}$ [37]. Note that here, we have parameterized $|\psi(\vec{\theta})\rangle$ by a vector of

real parameters $\vec{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$. The variational bound may be proved by writing $|\psi(\vec{\theta})\rangle$ in the energy eigenbasis:

$$
\begin{aligned}
\langle \hat{H}\rangle &= \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\sum_i |\langle\chi_i|\psi\rangle|^2 \lambda_i}{\sum_i |\langle\chi_i|\psi\rangle|^2} \\
&\geq \lambda_{\min}\frac{\sum_i |\langle\chi_i|\psi\rangle|^2}{\sum_i |\langle\chi_i|\psi\rangle|^2} = \lambda_{\min} = \lambda_1
\end{aligned}
\quad (2)
$$

where we briefly dropped the $\vec{\theta}$ dependence for brevity. The argument above relied on the fact that λ_1 is the smallest eigenvalue. Also note that we explicitly included the normalization in Eq. (2), whereas in Eq. (1) and now subsequently in this paper we simply assume it, i.e. $\langle\psi(\vec{\theta})|\psi(\vec{\theta})\rangle = 1$. Eq. (1) allows us to recast a minimum eigenvalue problem into an optimization problem, where the cost function $\langle\hat{H}(\vec{\theta})\rangle$ is variationally minimized over the parameter space of $\vec{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$ to reach the minimal eigenvalue.

2.1. General Outline of the VQE Algorithm

As we mentioned in the introduction, the variational quantum eigensolver is implemented as a hybrid quantum-classical algorithm. In general, it can be summarized in the following four steps:

1. Apply a unitary $U(\vec{\theta})$ to prepare an ansatz state $|\psi(\vec{\theta})\rangle$ on the quantum computer, where the parameter $\vec{\theta}$ could be controlled experimentally to a desired degree so that the subspace of Hilbert space of interest is accessible.

2. Apply the encoded Hamiltonian on the quantum computer and measure $\langle\hat{H}(\vec{\theta})\rangle$.

3. Update the value of $\vec{\theta}$ that would decrease $\langle\hat{H}(\vec{\theta})\rangle$ via a classical optimization routine.

4. Repeat 1-3 with a new ansatz parametrized by the updated $\vec{\theta}$ until the final resulting expectation converges to a desired precision. The expectation obtained in the last iteration will be a good approximation of the minimal eigenvalue.

This procedure described by the steps above are succinctly summarized by the schematic in Fig. 2. Note that here, the preparation of the ansatz state $|\psi(\vec{\theta})\rangle$ and the Hamiltonian encoding (which depends on the mapping between $\mathcal{Q}$ and $\mathcal{S}$) are performed on a quantum computer. The classical data generated
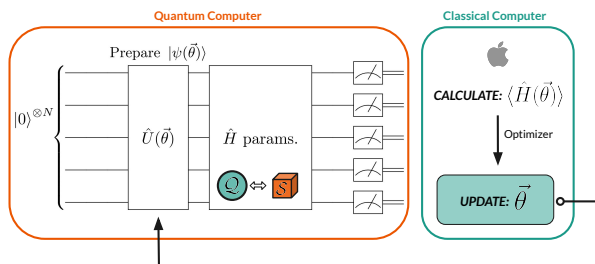
FIG. 2: Schematic of VQE algorithm. Here, the state preparation of $|\psi(\vec{\theta})\rangle$ and encoding of $\hat{H}$ occur in the quantum computer (orange). The evaluation of $\langle \hat{H}(\vec{\theta}) \rangle$ and optimization routine are then run on a classical computer (teal).

from the series of measurements then allows the evaluation of the cost function $\langle \hat{H}(\vec{\theta}) \rangle$ and subsequent optimization — these steps are performed on a classical computer.

Now, before proceeding to a detailed treatment of the algorithm outlined, we must first review some important background on the fermionic structure problem, which is among the most notable proposed applications of the VQE. This will later aid in our analysis of the algorithm in Section 3, and will also be play an important role in the implementation of numerical simulations in Section 4.

2.2. Femionic Structure Problems

The ground state energy of a molecule is of great interest in chemistry since it determines many important chemical properties of a substance. Ab initio calculations of ground state energies to chemical accuracy[1] are known to be difficult due to the large number of interaction terms in the molecular Hamiltonian, which scale factorially with system size. Therefore in practice both purely classical and hybrid quantum-classical approaches to such calculations would require certain approximation schemes. In this section, we review qualitatively the well-known Hartree-Fock approximation, which is central to a number of referenced works on VQE and fermionic problems in this paper.

---

[1] Chemical accuracy: approximately 4 kJ/mol $= 1.6 \times 10^{-3}$ Hartree (Ha), which is the accuracy needed for the calculated reaction rate to be within an order of magnitude of the experimentally measured rate at room temperature [38].

A generic molecular Hamiltonian could be written, in atomic units, as:

$$
\hat{H} = -\sum_i \frac{\nabla_{\vec{R}_i}^2}{2M_i} - \sum_i \frac{\nabla_{\vec{r}_i}^2}{2}
$$
$$
- \sum_{i,\,j} \frac{Z_i}{\left| \vec{R}_i - \vec{r}_j \right|} + \sum_{j>i} \frac{Z_i Z_j}{\left| \vec{R}_i - \vec{R}_j \right|} + \sum_{j>i} \frac{1}{|\vec{r}_i - \vec{r}_j|}
$$
$$(3)$$

where $\vec{R}_i$ is the position of the i−th nucleus, $\vec{r}_i$ the position of the i−th electron, M_i the mass of the i−th nucleus, and Z_i, Z_j the nuclear charges. The first two terms are kinetic energies of nuclei and electrons, while the remaining terms account for nucleus-electron, nucleus-nucleus and electron-electron interaction potentials respectively. The first step in simplifying the eigenvalue problem of equation (3) is to recognize the large difference between masses of electrons and nuclei (a factor of 10^3 to 10^5), which allows us to separate the motion of electrons from that of nuclei. Formally such a treatment is known as the Born-Oppenheimer approximation, which states that the total wave function of the molecule could be approximated by a product of nuclear wave function and electronic wave function [39]:

$$
\Psi_{\mathrm{BO}}(\vec{R}; \vec{r}) = \Psi_{\mathrm{nuc}}(\vec{R}) \Psi_{\mathrm{el}}(\vec{R}, \vec{r}) \tag{4}
$$

where the electronic wave function satisfies $H_{\mathrm{el}}(\vec{r}, \vec{R}) \Psi_{\mathrm{el}}(\vec{r}, \vec{R}) = E_{\mathrm{el}}(\vec{R}) \Psi_{\mathrm{el}}(\vec{r}, \vec{R})$ and H_{el} does not contain derivatives in $\vec{R}$.

To simplify the many-body interaction among electrons, one usually adopts a mean-field description of interacting particles where each electron is treated as only interacting with the average Coulomb potential of other electrons. The total wavefunction therefore could be approximately decoupled into products of single-electron wave functions:

$$
\Psi_{\mathrm{el}}(\vec{x}_1, \vec{x}_2, \cdots) = \psi_1(\vec{x}_1) \psi_2(\vec{x}_2) \cdots \tag{5}
$$

where we have changed our notation from $\vec{r}$ to $\vec{x} \equiv (\vec{r}, \vec{s})$ in order to explicitly denote electronic spin-orbitals. Equation (5) is known as the Hartree product [40]. To capture the antisymmetry of fermions (Pauli exclusion principle), this could be modified by the Hartree-Fock approximation, which states that *a many-body electronic wavefunction could be approximated by a single Slater determinant* [40, 41]. For a two-particle case this is:

$$
\Psi(\vec{x}_1, \vec{x}_2) = \frac{1}{2} \left(\psi_1(\vec{x}_1) \psi_2(\vec{x}_2) - \psi_2(\vec{x}_1) \psi_1(\vec{x}_2) \right) \tag{6}
$$

and for N particles:

$$\Psi_{el}(\vec{x}_1, \vec{x}_2, \cdots) = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} \mathbf{sgn}(\sigma) \prod_{i=1}^{N} \psi_i(\vec{x}_{\sigma(i)})$$

(7)

where σ is a permutation of natural numbers from 1 to N.

The Hartree-Fock approximation in equation (7) is the starting point of what is usually referred to as the Hartree-Fock method, whereby a system of eigenvalue equations (Hartree-Fock equations) for single electron spin-orbitals $\psi(\vec{x})$ are solved iteratively in a fashion known as the Self-Consistent Field Procedure [39, 42] . Alternatively for the purpose of the VQE, it is oftentimes helpful to work in second quantization, where a product state of the following form (Hartree-Fock state) is permissible as solutions to the Hartree-Fock-approximated many-body electron interaction:

$$|\text{HF}\rangle = \prod_{j}^{N} \left(\sum_{i}^{n} c_{ij} a_i^\dagger \right) |\mathbf{vac}\rangle$$

(8)

where N is the total number of electrons, n the number of sites an electron could occupy, and $a^\dagger$, a the usual fermionic creation and annihilation operators. This gives us a good choice for the initial ansatz to be used in the step one of the VQE algorithm outlined in Section 2.1.

Making contact with step two of the VQE outline in Section 2.1, we now move on to see how a molecular Hamiltonian could be encoded on qubits. The second-quantized molecular Hamiltonian is given by:

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

(9)

where the second term counts all two-electron interactions. The coefficients h_{pq}, h_{pqrs} are given by spin-orbital integrals:

$$h_{pq} = \int d\vec{x} \varphi_p^*(\vec{x}) \left(\frac{\nabla_{\vec{r}}^2}{2} - \sum_i \frac{Z_i}{\left| \vec{R}_i - \vec{r} \right|} \right) \varphi_q(\vec{x})$$

(10)

$$h_{pqrs} = \int d\vec{x}_1 d\vec{x}_2 \frac{\varphi_p^*(\vec{x}_1) \varphi_q^*(\vec{x}_2) \varphi_s(\vec{x}_1) \varphi_r(\vec{x}_2)}{|\vec{r}_1 - \vec{r}_2|}$$

(11)

which are usually evaluated in advance on a classical computer [29, 43]

Note the fermionic operators in Eq. (9) satisfy the following algebra (CAR or, more generally, C* algebra):

$$\{a_p^\dagger, a_r\} = \delta_{p,r}$$

(12)

$$\{a_p^\dagger, a_r^\dagger\} = \{a_p, a_r\} = 0$$

(13)

This algebra is distinct from the usual SU(2) algebra (i.e. of Pauli matrices) which gates on a quantum computer obey. To encode the Hamiltonian properly is to find a map between the two algebra. Some of the popular choices include Jordan-Wigner transformation, Bravyi-Kitaev transformation and the parity transformation [15, 44]. We quote here the Jordan-Wigner transformation:

$$a_p^\dagger = \left(\prod_{m<p} \sigma_m^z \right) \sigma_p^+$$

(14)

$$a_p = \left(\prod_{m<p} \sigma_m^z \right) \sigma_p^-$$

(15)

where we defined:

$$\sigma^\pm \equiv (\sigma^x \mp i\sigma^y)/2$$

(16)

3. VQE Analysis & Applications

In this section, we will return to a more detailed analysis of the VQE algorithm, and mention some of its potential applications to both quantum and classical problems.

3.1. Complexity and measurement scale

It can be easily shown that any Hamiltonian can be written as

$$\hat{H} = \sum_{i\alpha} h_\alpha^i \sigma_\alpha^i + \sum_{ij\alpha\beta} h_\alpha^i \sigma_\beta^j \ldots$$

(17)

with real h, and the Greek subscripts identify different Pauli operators while the Roman indices identify the subsystems on which the operator acts. Specifically, we consider Hamiltonians with the number of terms that scale polynomially with the system size in the Pauli basis. Therefore, the measurement of $\langle \hat{H}(\vec{\theta}) \rangle$ will reduce to the evaluation of a polynomial number of expectation values of Pauli operators. It has been shown that a quantum simulation of such Hamiltonians can be done efficiently in polynomial

time [45]. Despite this limitation, the class of Hamiltonians under our consideration can describe a wide range of problems of interest, including the electronic structure in quantum chemistry, the quantum and classical Ising Model, the Heisenberg Model. Later, we will examine both the quantum chemistry and some classically-hard combinatorics problems.

To compare the power of the hybrid VQE to its classical analog, we note that on a classical computer, to evaluate $\langle \hat{H}(\vec{\theta}) \rangle$, one need to optimize all reduced states separately according to the problem Hamiltonian. However, this gives rise to the N-representability problem, which is known to be in the complexity class QMA-hard and cannot be solved efficiently both on the quantum and classical computers [46]. In VQE, however, a global quantum state can be prepared on the quantum circuits with exponentially fewer resources, thus avoiding the N-representability problem.

As mentioned in the introduction, quantum phase estimation (QPE) is also a good candidate for finding the eigenvalues of Hamiltonians were we not in the NISQ era and limited by the physical size of quantum devices. QPE can be used to estimate the eigenvalue of a given eigenstate, and if such state does not exist, QPE can be combined with other primitives such as amplitude amplification to solve for both the state and the corresponding energy. In the recent paper by Ge et. al. [47], it is found that at a given precision ϵ, a ground state with known energy can be approximated using QPE-related algorithms with constant probability in a gate complexity of $O(1/\epsilon)$. In comparison, the performance of VQE is still known as heuristic.

The advantage of VQE comes from its less stringent demand on the circuit depth. VQE can perform independent and parallel measurement of the Pauli operators on local qubits, and only requires a coherent circuit depth of $O(1)$ for a single measurement on the prepared ansatz state. This improvement is traded off with an increase in measurement samples. The sample size for VQE and QPE is $O(1/\epsilon^2)$ and $O(1)$ respectively [48]. This repetition in sampling poses a potential problem for VQE [49], and propels innovations in expectation estimations. But in many cases, the repeated measurement and preparation of states add in a polynomial overheads, confirming the current application value of VQE [25].

3.2. Hamiltonian mapping

One of the most important parts in the VQE algorithm is the mapping of $\langle \hat{H}(\vec{\theta}) \rangle$. The mapping should yield $\langle \hat{H}(\vec{\theta}) \rangle$ with polynomial number of terms and later decides the preparation of the ansatz states. Both quantum and classical problems require the mapping to formulate the problem in the Pauli basis. Especially for the quantum problems, there is a subtlety that the annihilation operators in fermionic basis follow $\{a_i, a_j\} = 0$, $\{a_i^\dagger, a_j^\dagger\} = 0$, $\{a_i, a_j^\dagger\} = \delta_{ij}$. This is different from the Pauli algebraic rules and motivates the mapping.

As introduced in the previous section, a usual way to implement the mapping of fermionic Hamiltonian uses the Jordan-Wigner mapping, shown schematically in Figure 3. Written out more explicitly,

$$a_i^\dagger \rightarrow I^{\otimes i-1} \otimes \sigma^- \otimes (\sigma^z)^{\otimes N-i} \qquad (18)$$

$$a_i \rightarrow I^{\otimes i-1} \otimes \sigma^+ \otimes (\sigma^z)^{\otimes N-i} \qquad (19)$$

where N is the total number of qubits and i is the orbit of the fermion in consideration [50]. In this mapping, there is non-locality restricted by parity calculation, and a fermionic wave function can spread over $O(N)$. This non-locality may lead to fidelity issues in the readout, facilitating the need for a robust classical optimization algorithm, which we shall explain more later.

For classical problems, a simpler mapping of the cost function into the Ising Hamiltonian can be implemented assuming a general solution $\vec{x}$ polynomial in binary components [7]. Following this treatment, we can further shift the variables and transform to Pauli operators as follows,

$$x_i \rightarrow \frac{1-z_i}{2}, \; z_i \in \{-1, 1\} \qquad (20)$$

$$z_i \rightarrow \sigma_z^i \qquad (21)$$

and obtain the cost Hamiltonian.

3.3. Preparation of the ansatz states

After mapping the Hamiltonian, we should decide on the preparation method for the ansatz states. Ideally, the ansatz states should both incorporate the structure of the problem Hamiltonian and can be updated efficiently.

We will explain our claims with the example of the Unitary Coupled Cluster (UCC) approach, which is
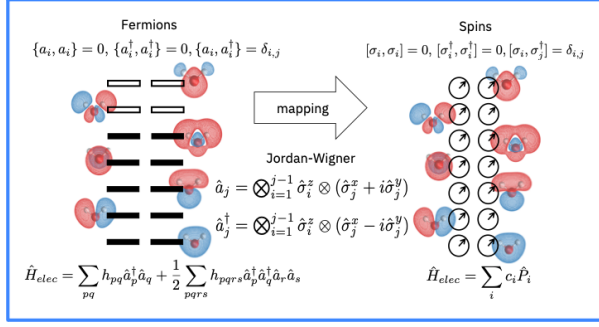
FIG. 3: This is a schematic of the Jordan-Wigner transformation from fermionic creation and annihilation operators to qubit spin operators. Image courtesy IBM Qiskit Electronic Structure tutorial (open access).

(a) A signal with noise

(b) Corresponding derivative

FIG. 4: Noisy signal and its derivative

widely used for fermionic Hamiltonians. As shown in the previous section, if we try to generate all possible excited Slater determinant, the number of trial states will scale factorially with the number of electrons. The UCC approach improves the efficiency by characterizing the ansatz states with a polynomial number of terms, lowering the time overheads from ansatz updates [51]. The unitary operator $U(\vec{\theta})$ to generate the trial wavefunctions $|\psi(\vec{\theta})\rangle$ from the reference state $|\phi\rangle$ is given by the cluster operators $T(\vec{\theta})$ as follows

$$|\psi(\vec{\theta})\rangle = U(\vec{\theta}) = e^{T(\vec{\theta}) - T^\dagger(\vec{\theta})} |\phi\rangle \qquad (22)$$

$$T(\vec{\theta}) = \sum_k T_k(\vec{\theta}) \qquad (23)$$

$$T_1(\vec{\theta}) = \sum_{\substack{i \in \text{occ} \\ j \in \text{unocc}}} \theta_i^j a_j^\dagger a_i \qquad (24)$$

$$T_2(\vec{\theta}) = \sum_{\substack{i,j \in \text{occ} \\ l,k \in \text{unocc}}} \theta_{i,j}^{k,l} a_l^\dagger a_k^\dagger a_j a_i \qquad (25)$$

Only the first two terms of the cluster operator are shown, but in general, the anti-Hermitian $T(\vec{\theta}) - T^\dagger(\vec{\theta})$ cluster is the sum of polynomial number of terms. In practice, we can truncate the series up to the double or triple level without significant loss of accuracy and remove some time cost from higher uncorrelated excitations at each level.

3.4. Classical optimization algorithm

The classical component of the VQE is also important. Specifically, to complete our VQE algorithm, we want a robust classical optimization algorithm
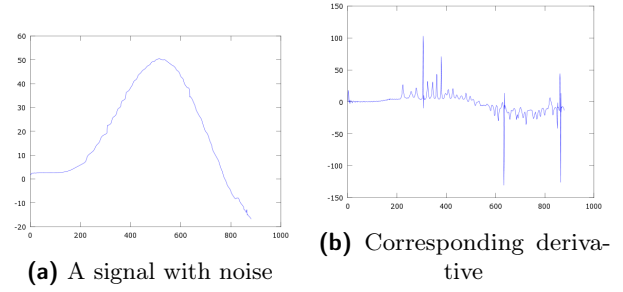
that requires a few measurement samples to reduce time overheads from repetitions. The robustness demand mainly comes from the statistical noise from the limited number of samples and possible errors in our circuit. We will introduce the Nelder-Mead method (NM) and the simultaneous perturbation stochastic approximation (SPSA) that fit our need. NM is a derivative-free algorithm based on simplex direct search (DS) [52]. Shown in Fig. 3 are a noisy signal and its corresponding derivative. The sharp peaks in the derivative due to noise imply the potential problem that noise poses on gradient-based algorithms. Since NM is derivative-free, it can be more robust compared to some other gradient-based DS methods. The NM method has been tested for He^+–H molecules and some other energy optimizations, yielding robust performance [25].

In NM, parameters and the cost function are mapped to a simplex S as shown in Fig. 4(a). It typically begins with a set of $n + 1$ points (x_i) as the vertices of the working simplex S, where n is the dimension of the problem. And a set of function values $f(x_i)$ will be assigned to the vertices. The initial simplex must be non-degenerate (i.e., the points do not lie in the same hyperplane). The method performs transformations on S to decrease the function values at each vertex. The process is terminated when S geometrically becomes sufficiently small (or algebraically, the function values are sufficiently close). Typically, NM needs only one or two function evaluations at each step, which is much smaller than the n evaluations required by some other DS methods.

Another important algorithm is the SPSA algorithm. SPSA is not derivative-free, but it approximates the gradient of the objective function with only two measurements regardless of the dimension of the optimization problem, thus saving us time from sampling and is suitable for high-dimensional problems
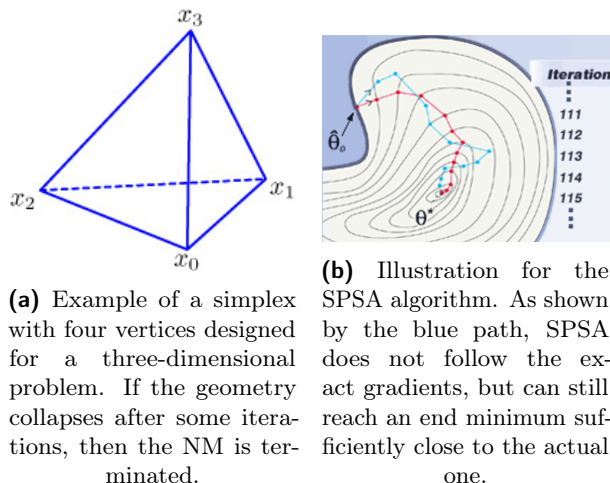
(a) Example of a simplex with four vertices designed for a three-dimensional problem. If the geometry collapses after some iterations, then the NM is terminated.

(b) Illustration for the SPSA algorithm. As shown by the blue path, SPSA does not follow the exact gradients, but can still reach an end minimum sufficiently close to the actual one.

FIG. 5: Schemes for the classical optimization algorithms

like fermionic problems. It does so by concurrently perturbing all of the parameters in a random fashion; in contrast to gradient-descent where each parameter is perturbed independently. Although the SPSA random search does not follow exactly the gradient path, this also allows it to escape from the local minima while provides sufficient information on the gradient of the function for a final convergence [53, 54], as illustrated in Fig. 4(b). Since the algorithm itself uses noisy gradient approximation, we also see that the noisy measurement input is acceptable for this algorithm, showing its robustness against noise. When utilizing VQE in either a noisy simulator or on real hardware, SPSA is a recommended classical optimizer.

3.5. Applications

Now that we have completed our analysis of the VQE algorithm, we now proceed to discuss some notable applications.

Quantum chemistry problems are of specific interest here because the electronic structure problems in quantum chemistry scales exponentially, and therefore cannot be solved efficiently on a classical computer. However, the understanding of the electronic structures for different molecules can be crucial to the development of new materials. As mentioned earlier, we can use VQE in the electronic structure problem, and estimate the bond-energy of the many-electron Hamiltonian or the ground state energy of some small

molecules. There have been various successful experimental demonstrations in the area [25, 28–31]. In our project, we also simulate the ground state energy for H_2 and LiH to show a working example on our own.

As shown earlier, VQE is also suitable for optimizing classical problems by mapping the cost function to the Ising-type Hamiltonian solvable on VQE. Here, we consider two NP-complete combinatorics problems.

The MaxCut problem is ubiquitously present in many areas, including clustering, networks, and statistical physics. The problem consists of grouping the nodes of a graph into two subgroups by cutting across the links between them, and the cut should be made in such a way that the added weights on the edges that were cut are maximized [55]. On an ideal quantum simulator, the MaxCut problem can be solve with a probability higher than 95% within 100 trial steps [7].

3-SAT is another prime NP-compete problem. It can be translated into an eigenvalue problem by searching for the maximum number of clauses in $C(\vec{x})$ that can be satisfied by the bit string $\vec{x}$. If the number at least equal to the total number of clauses, we can claim that we have found a solution. Recently, some innovative algorithm using imaginary time evolution [56] has produced the energy spectrum for some simple 3-SAT problems with up to 18 qubits.

4. VQE Circuits & Simulations

The following section will apply the ideas thus developed in this review article. In particular, we will implement the general outline of the VQE algorithm detailed in Section 2.1 seeded with a Hartree-Fock inspired ansatz state described in Section 2.2. Then, we will utilize the parity map, which is intimately related to the Jordan-Wigner mapping described in Section 3.2, to describe the Hamiltonian of small molecules. Finally, we will use classical optimization algorithms, such as those in Section 3.5, to find the ground state energy of these small molecules, as proposed in Section 3.5. This section will also closely follow the tutorial we have created using IBM's Qiskit SDK to supplement this paper, available at: www.github.com/yaleqc/vqe-error-mitigation.

4.1. Motivating variational forms

To implement the algorithm detailed in Section 2.1 for some Hamiltonian H, we first map the Hamiltonian onto the operator basis of interest and develop a variational form that can produce arbitrary trial states $|\psi(\vec{\theta})\rangle = U(\vec{\theta})|\psi_0\rangle$ in the relevant Hilbert of the Hamiltonian H. Finding a universal variational form that lends to tractable classical optimizations is often challenging. In an effort to illuminate the creation of such variational forms, we begin with the simplest systems of a single qubit and two qubits.

4.1.1. Single Qubit VQE

To demonstrate single qubit VQE, we must first choose a single qubit Hamiltonian. In the linked tutorial, we choose to randomly generate a single qubit Hamiltonian using the Pauli gates.

$$\hat{H} = \vec{\varepsilon} \cdot \vec{\sigma} = \varepsilon_0 I + \varepsilon_1 \sigma_x + \varepsilon_2 \sigma_y + \varepsilon_3 \sigma_z$$

where $\vec{\varepsilon} \in [0,1]^4$. Our goal is to now find the ground state energy of this Hamiltonian $\hat{H}$. Using the variational inequality stated in Equation 1, we quickly see that we must minimize $\langle\psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle$ over the space of all possible valid quantum states $|\psi(\vec{\theta})\rangle$.
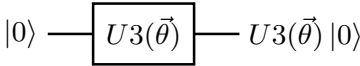
In the case of a single qubit system, we use the simple universal variational form detailed in Figure 6 to explore the space of arbitrary trial states $|\psi(\vec{\theta})\rangle$ by varying $\vec{\theta}$ classically.

Notably, the calculation of $\langle\psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle$ can quickly become a computationally intractable operation. As the number of qubits in our Hilbert space increases linearly, the size of $\hat{H}$ increases exponentially. Since matrix-vector multiplication has polynomial time complexity, the calculation of $\langle\psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle$ also takes exponentially longer as the number of qubits in our system increase linearly.

This is where VQE shines. Using our QPU, we are able to *approximate* $\langle\psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle$ using shot-based sampling. There is a section of our tutorial devoted to exploring different methods of calculating $\langle\psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle$ using IBM Qiskit.

Upon minimizing $\langle\psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle$ using the *Constrained Optimization by Linear Approximation* (COBYLA) classical optimizer, we are left with the smallest upper bound $\lambda_{\min}$ on the true ground state

eigenvalue λ_1 of $\hat{H}$. For the case of the 1-qubit system, we can also mathematically diagonalize $\hat{H}$ and calculate λ_1 to compare to our VQE optimizer found $\lambda_{\min}$. By simply averaging across multiple random $\hat{H}$, we find a crude average error rate of $\frac{\lambda_{\min}-\lambda_1}{|\lambda_1|} * 100\% \sim 10^{-6}\%$. These *noiseless* simulations suggest that the 1-qubit VQE algorithm works as intended. Moreover, the variational form expressed in Figure 6 is able explore the single qubit Hilbert space well enough to find an incredible tight bound on λ_1. This numerically suggests that the variational form is universal, although this can be proven rigorously by calculating the Lie group of reachable states given our variational form [57].

$$|0\rangle \quad\longrightarrow\quad \boxed{U3(\vec{\theta})} \quad\longrightarrow\quad U3(\vec{\theta})|0\rangle$$

$$U3(\vec{\theta}) = \begin{pmatrix} \cos\left(\frac{\theta_1}{2}\right) & -e^{i\theta_3}\sin\left(\frac{\theta_1}{2}\right) \\ e^{i\theta_2}\sin\left(\frac{\theta_1}{2}\right) & e^{i\theta_3+i\theta_2}\cos\left(\frac{\theta_1}{2}\right) \end{pmatrix}$$

FIG. 6: The U3 gate is enough to produce arbitrary $|\psi(\vec{\theta})\rangle = U3(\vec{\theta})|0\rangle$ trial states with which to minimize $\langle\psi|\hat{H}|\psi\rangle$.

4.1.2. Two Qubit VQE

Next, we repeat the process detailed in Section 4.1.1 for a two qubit Hilbert space. Instead of choosing a random Hamiltonian, as done in the previous section, we instead choose to tackle the problem of finding the ground state of the hydrogen molecule H_2. To do this, we must first express H_2 in spin operator basis using the Parity mapping [58]. The Parity mapping is a mapping from fermionic states onto qubit states that exploits symmetries to lower the number of qubits required to represent the Hamiltonian of H_2. By numerically solving for the mapping, we find the following Hamiltonian for H_2.

$$\begin{aligned} \hat{H} = &(-1.0523732 * I \otimes I) + (0.39793742 * I \otimes Z) \\ &+ (-0.3979374 * Z \otimes I) + (-0.0112801 * Z \otimes Z) \\ &+ (0.18093119 * X \otimes X) \end{aligned}$$

$$(26)$$

Next, we cite another universal variational form to produce arbitrary trial states $|\psi(\vec{\theta})\rangle$ [59]. This minimal universal variational form is noticeably longer

with 24 optimization parameters, as shown in Figure 7. Already, we see how quickly universal variational forms may scale in VQE problems. This is because a fixed variational form with a polynomial number of parameters can only generate transformations to a polynomial sized subspace in the larger exponentially sized Hilbert space. Thus, the study of approximate variational forms for VQE problems is an important step towards realizing practical and scalable variation quantum eigensolvers.
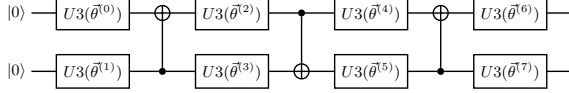

FIG. 7: The minimal universal variational form for a two-qubit system must consider two-body interactions and thus it involves entanglement through CNOT gates [59].

Using the variational form depicted in Figure 7 along with the H_2 Hamiltonian mapping found in Equation 26, we can again use the COBYLA classical optimizer to minimize $\langle \psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle$. By mathematically diagonalize $\hat{H}$ to find λ_1, we can then benchmark the performance of the VQE algorithm on this two-qubit system. Through noiseless simulations, we find that $\frac{\lambda_{\min}-\lambda_1}{|\lambda_1|} * 100\% \sim 10^{-7}\%$. This tight bound again shows the power of both the VQE algorithm and the minimal universal variational form.

4.2. Using VQE to learn about small molecules

Next, we apply the ideas developed in the previous section to the problem of solving for the ground state energy of Hydrogen H_2 and Lithium Hydride LiH, with varying interatomic distances. Again, we begin by mapping each small molecules Hamiltonian form fermionic operators to spin operators using the Parity map. Then, we use the Unitary Coupled-Cluster Single and Double excitations variational form [60]. Finally, we then seed the Sequential Least Squares Programming (SLSQP) optimizer with the Hartree-Fock state ansatz. The results of this optimization are shown in Figure 8.

Notably, the Hartree-Fock ansatz is a loose upper bound to the much tighter and visually overlapping VQE upper bound on the exact ground state energies of the tested small molecules. Moreover, through this experiment, we are able to verify the most stable

interatomic distance for each small molecule. As an added bonus, these simulation results match closely with the results presented in this Refs. [31] and [32].
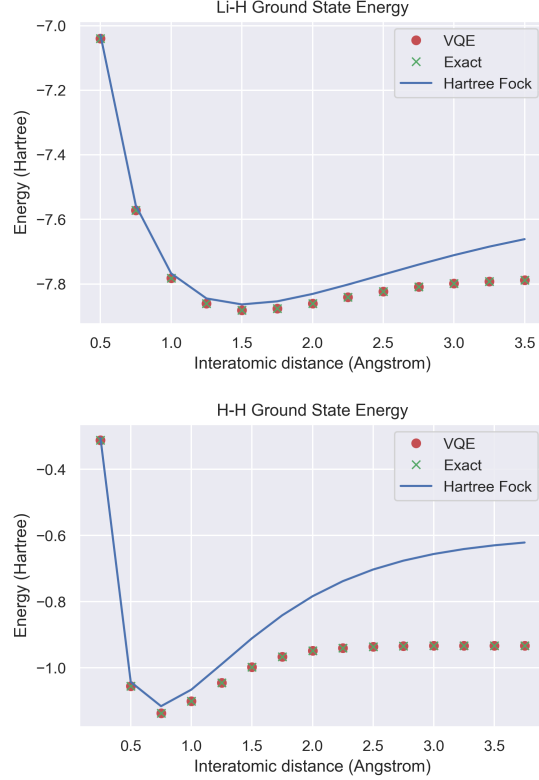

FIG. 8: (Top) Using the Hartree-Fock state as an ansatz to our classical optimization, we compare the performance of our VQE simulation results with the exact values of the ground state energies of LiH for varying interatomic distances. (Bottom) We do the same for H_2. These simulated results closely match the equivalent results found in Figure 10.

4.3. Hardware-efficient VQE for small molecules

To end this discussion of VQE implementations and simulations, we briefly mention the hardware-efficient VQE circuits built and experimentally probed in Ref. [31]. An overview of the hardware-efficient six-qubit quantum circuit has been detailed in the caption of Figure 9.

Most notably, the hardware-efficient trial states prepared using superconducting quantum processing units are composed of entangling unitary operators of the form $U_{\mathrm{ent}} = \exp\{-iH_0\tau\}$, which entangle all the transmon qubits in the circuit by letting the drift
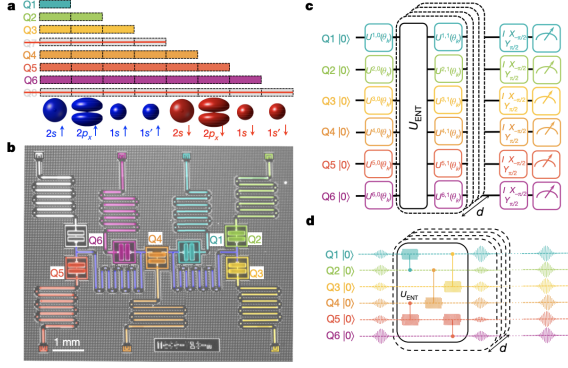
FIG. 9: (a) Parity map of spin orbitals to eight qubits, reduced to six using fermionic spin and parity symmetries. (b) Optical micrograph of the physical superconducting quantum processor with seven transmon qubits, coupled using two co-planar waveguide resonators shown in violet. (c) Variational form of this six qubit Hilbert space, depicted schematically. Single-qubit rotations $U^{q,d}(\vec{\theta_k})$ interleaved with entangling unitary operations U_{ent} are enough to prepare arbitrary trial states. Once the trial state is prepared, the final set of post-rotations before readout are used to estimate the energy of the trial state. (d) Here, U_{ent} is realized as a sequence of two-qubit cross-resonance gates, and depicted along side an example pulse sequence used to prepare some trial state. Adapted from Ref. [31].

Hamiltonian H_0 evolve the system for some time τ. These entanglers, however, are not enough to create a universal variation form. To accomplish this, the entanglers are interleaved with single-qubit Euler rotations,

$$U^{q,i}(\vec{\theta}) = Z^q_{\theta_1^{q,i}} X^q_{\theta_2^{q,i}} Z^q_{\theta_3^{q,i}}$$

where q identifies the qubit and $i = 0, \cdots, d$ specifies the depth position of the gate, and $\vec{\theta}$ represents the Euler angles associated with that gate. The explicit form of the trial state can be written as follows

$$|\Phi(\vec{\theta})\rangle = \prod_{q=1}^{N}[U^{q,d}(\vec{\theta})] \times U_{ENT} \times \prod_{q=1}^{N}[U^{q,d-1}(\vec{\theta})] \times \cdots$$

$$\times U_{ENT} \times \prod_{q=1}^{N}[U^{q,0}(\vec{\theta})] |00\cdots0\rangle$$

Since the first set of Z rotations of $U^{q,0}(\vec{\theta})$ are not implemented on $|\vec{0}\rangle$, this leaves $p = N(3d + 2)$ independent parameters to control. While there is control over the phase (evolution time τ) of U_{ENT},

the entanglign gates are kept constant across the d iterations in this experiment. In particular, U_{ENT} is composed of a sequence of two-qubit cross-resonance gates.

As we did in simulation, after the preparation of each trial state the associated energy is estimated by measuring the expectation values of the individual Pauli terms in the Parity mapped Hamiltonian of the small molecules of interest, which include H_2, LiH, and BeH_2. Then this energy estimate is used in a gradient descent based classical optimization algorithm called simultaneous perturbation stochastic approximation (SPSA) to update the p control parameters and repeat the cycle. Finally, the optimized experimental results are shown in Figure 10.

It should be mentioned that the numerical simulations shown in the density plots of Figure 10 account for decoherence effects simulated by adding amplitude damping and dephasing channels after each iteration of qubit gates in the trial state circuit. Thus, we quickly see that substantial noise and error mitigation strategies are needed to even numerically approach chemical accuracy. These techniques are explored in depth in Section 5.
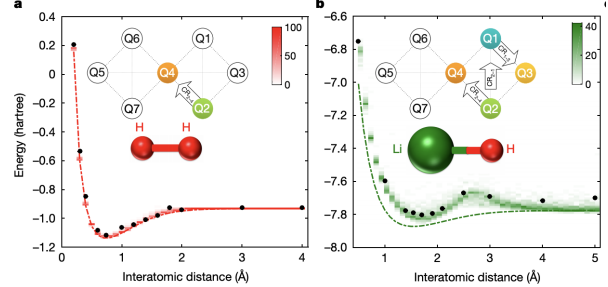

FIG. 10: (a-b) Experimental data (black dots), exact energy surfaces (dotted lines), and density plots (shaded) from numerical simulations are all shown.

5. Error Mitigation Strategies for VQE

As the results of Ref. [31] demonstrate, VQE experiments on real hardware are unfortunately susceptible to various sources of noise and decoherence. Furthermore, fully error corrected fault-tolerant architectures remains out of reach in the NISQ-era. Thus, we must resort to other techniques for mitigating incoherent errors in short-depth quantum circuits that may improve the performance of NISQ-era quantum algorithms.

5.1. Error Mitigation via Richardson Extrapolation

In this subsection, we will describe a relatively novel strategy for performing error mitigation based on Richardson's deferred approach to the limit [61]. This was first proposed by the IBM Research group in Ref. [62] and later implemented experimentally in [32] to improve the performance of the VQE algorithm among other quantum circuits by suppressing incoherent errors.

In this method, we consider the time-dependent Hamiltonian dynamics used to implement any given quantum circuit. Specifically, we consider a time-dependent drive Hamiltonian $K(t) = \sum_\alpha J_\alpha(t) P_\alpha$ applied over a time $[0, T]$. Here each P_α represents an N-qubit Pauli operator with a time-dependent coupling coefficient $J_\alpha(t) \in \mathbb{R}$. Our goal will then be to compute the value of some observable A with respect to an evolved state $\rho_\lambda(T)$, where the evolution is subject to weak noise characterized by a parameter λ. The condition of weak noise simply means that we consider the limit $\lambda \to 0$. We can represent the noisy evolution via the Lindblad equation:

$$\frac{\partial}{\partial t}\rho(t) = -i[K(t), \rho(t)] + \lambda\mathcal{L}(\rho(t)) \qquad (27)$$

Now, we only make two assumptions here: that $\lambda \ll 1$ and that the Lindblad term $\mathcal{L}(\bullet)$ is invariant under time-translation and time rescaling, and independent of the drive parameters $J_\alpha(t)$. Besides this, there is no other requirement on $\mathcal{L}(\bullet)$. Now, we aim to compute $E(\lambda) = \text{Tr}[A\rho_\lambda(T)]$ in the final noisy state $\rho_\lambda(T)$. However, since λ is small, it turns out that we can actually expand $E(\lambda)$ as a power series in λ. Here the constant term $\propto \lambda^0$ gives the noiseless evolution, i.e. what would result from setting $\lambda = 0$ in Eq. (27). Specifically, if we denote the 'ideal' noiseless value of the desired expectation value by $E^* = \text{Tr}[A\rho_0(T)]$, then we can write

$$E(\lambda) = E^* + \sum_{k=1}^n c_k \lambda^k + \mathcal{O}(\lambda^{n+1}) \qquad (28)$$

Note that the coefficients a_k here can be computed directly from the chosen noise model [i.e. the $\mathcal{L}(\bullet)$ that we use to model noise in our system]. Therefore, from a purely theoretical standpoint it is relatively straightforward to compute the power series $E(\lambda)$. The challenge with real quantum devices, however, is that we often do not know the exact noise model that leads to errors. Therefore, in actual experiments, we

can only ever estimate $E(\lambda)$; this estimate[2] is denoted by $\hat{E}(\lambda)$. Even for a good estimate, however, $\hat{E}(\lambda)$ will still deviate from the ideal value E^* by $\mathcal{O}(\lambda)$, as shown in Eq. (28).

To improve this estimate, however, we can employ the trick of repeating the given quantum circuit, i.e. Eq. (27), a total of $n + 1$ times, using rescaled noise parameters $\lambda_i = c_i\lambda$ where $i \in 0, \ldots, n$ and $c_i > 0$. As described in Ref. [62], we can compute the improved estimate:

$$\hat{E}_n(\lambda) = \sum_{j=0}^n \gamma_j \hat{E}(c_j\lambda) \qquad (29)$$

where the γ_j coefficients are chosen to satisfy a system of $n + 1$ linear equations

$$\sum_{j=0}^n \gamma_j = 1 \quad \text{and} \quad \sum_{j=0}^n \gamma_j c_j^k = 0 \qquad (30)$$

for $k = 1, 2, \ldots, n$. This calculated value $\hat{E}_n(\lambda)$ will then be a much better estimate of E^* with error of order only $\mathcal{O}(\lambda^{n+1})$. This technique in Eqs. (29-30) is precisely the Richardson extrapolation technique.

Now, the abstract technique above relies on the ability to amplify the noise parameter λ by successive factors, i.e. we need to rerun the quantum circuit with progressively higher noise strengths $c_1\lambda, c_2\lambda, \ldots$, and so forth. However, *since the noise is time translation-invariant*, it turns out that we can instead achieve the same outcome by rescaling the pulses in the drive Hamiltonian $K(t)$. Specifically, the expectation value $E(c_i\lambda)$ under rescaled noise $c_i\lambda$ is equivalent to the expectation value under a rescaled drive:

$$K^i(t) = \sum_\alpha J_\alpha^i(t) P_\alpha \quad \text{with} \quad J_\alpha^i(t) = c_i^{-1} J_\alpha(c_i^{-1}t) \qquad (31)$$

for total time c_iT, with noise λ (see Ref. [62]). In superconducting circuits, this strategy is very meaningful: we increase the pulse time while simultaneously decreasing the pulse amplitude for each gate pulse (e.g. π pulse enacting an X gate). By simultaneously adjusting both the pulse length and amplitude, we

---

[2] Here the hat denotes an estimate, not a quantum operator. Also we have $\hat{E}(\lambda) = E(\lambda) + \delta$, where δ arises from other experimental inaccuracies not accounted for.

can enact the same logical gate — however, the increased evolution time leads to a higher exposure to noise and thus a greater accumulation of incoherent errors. We note however, that this rescaling is only possible within the characteristic lifetimes T_1 and T_2 of the qubits being used in the computation.
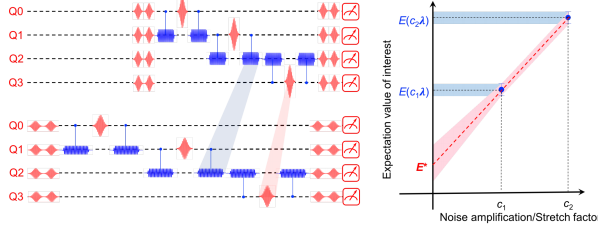

FIG. 11: Measurement of expectation values at higher noise strengths is equivalent to using rescaled pulses in the quantum circuit (left). By making successive measurements at varying noise strengths, we can use Richardson extrapolation to estimate the zero noise value E^* (right). Reprinted and adapted from [32].

This particular technique for error mitigation has been used to improve the performance of the VQE algorithm on a superconducting quantum processor [32]. Here the authors use the variational algorithm to compute the ground state energies for H_2 and LiH by mapping their fermionic Hamiltonians onto two and four qubits respectively. As in Ref. [31], the fermion to qubit mapping employed notable spin-parity symmetries and the freezing of inner shell orbitals for LiH. Furthermore, the device used in this experiment had similar coherence times to that used in Ref. [31] (cf. Section 4.3). Consequently, it is no surprise that the unmitigated ($c = 1$) bare expectation values lie relatively far from the exact ground state energies in Fig 12. By contrast, however, the mitigated results demonstrate significant improvements in accuracy and even visibly lie much closer to the exact molecular energies. Thus, this zero-noise extrapolation technique is able to deliver critical performance gains, even on noisy near-term hardware. Meanwhile, further improvements in error rates and coherence times will allow for longer stretch factors and thus more precise (and accurate) mitigated estimates.

5.2. Other Mitigation Strategies

fFollowing its proposal in Ref. [62], the Richardson zero-noise extrapolation technique described above
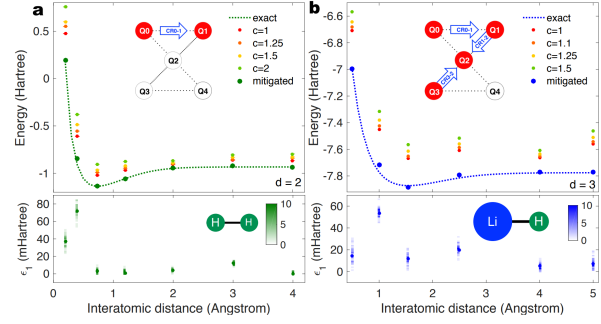

FIG. 12: Experimental results of VQE applied to the calculation of molecular ground state energies for H_2 and LiH. The use of error mitigation enables much more accurate estimates of the true molecular energies. Reprinted from Ref. [32].

has grown to become one of the more popular error mitigation techniques that exist today for improving algorithmic performance on noisy hardware. However, it is far from the only possible mitigation strategy, and various other schemes have since been developed. One technique also proposed in Ref. [62] cancels errors by resampling randomized circuits according to a quasi-probability distribution. Post-selection methods are another popular choice. For instance, in the recent result from the Google AI Quantum group, the authors simulate binding energies of H_6, H_8, H_{10}, and H_{12} molecules to within chemical accuracy — this is made possible using a combination post-selection and pure-state projection to improve the performance of VQE [33]. Indeed, it is fair to say that many of the most notable recent experiments using VQE for chemical simulation have leveraged some form of error mitigation or post-processing to significantly enhance the accuracy of their results [28, 32–36, 63]. The only exception we are aware of is the recent work of Ref. [29], performed on a trapped-ion quantum computer. Here the systematic and statistical errors were comparable to chemical accuracy without the need for mitigation. As quantum devices continue to develop lower hardware error rates, the bare performance of algorithms such as VQE will no doubt improve. Nevertheless, mitigation techniques can enable large performance improvements with little to no additional experimental sophistication, and will thus likely continue to be widely used. These methods can thus be thought of as a free resource to maximize the performance of NISQ devices and algorithms such as VQE in the near term.

6. Conclusions and Outlook

In this project, we carried out a comprehensive study of the variational quantum eigensolver (VQE). We began by reviewing the motivation for hybrid quantum-classical algorithms for tackling important eigenvalue and optimization problems using noisy near-term quantum hardware, before describing the VQE algorithm itself. We addressed the important questions of algorithm runtime, state preparation, implementation of Hamiltonians for quantum chemistry, and the choice of classical optimizer. We also developed a suite of numerical simulations using IBM Qiskit to demonstrate the VQE algorithm in practice[3] — we applied this to simulate the ground state energies of H_2 and LiH. Finally, we reviewed notable experiments and discussed important practical considerations for running the VQE algorithm, including the use of novel noise mitigation strategies to suppress certain errors.

The VQE is a powerful hybrid algorithm for exploring the performance of early quantum computers — especially so since it is largely hardware agnostic and can be run on any quantum device. While fully fault-tolerant error corrected quantum computers remain many years, if not decades, away, current NISQ devices may still be able to deliver a quantum advantage using a combination of hybrid quantum and classical computing resources. While the VQE was developed firmly in the NISQ-era, it has proved to have a surprising number of key near-term applications in quantum chemistry and optimization, and will likely remain an important tool even after the first practical error-corrected quantum devices come into existence. In the meantime, however, variational hybrid algorithms such as VQE hold the promise of performing novel computationally useful tasks, and offer an exciting framework for leveraging the respective strengths of classical and quantum machines.

---

[3] We also began preliminary work on developing molecular simulations to incorporate noise, though have not included these early-stage results in this report. Still, this is something we hope to address soon as a follow-up to this project.

Contributions

References

[1] M. A. Nielsen and I. Chuang, Quantum Computation and Quantum Information (2000).

[2] P. Kaye, R. Laflamme, and M. Mosca, An Introduction to Quantum Computing (2007).

[3] N. D. Mermin, Quantum Computer Science: An Introduction (2007).

[4] N. Elsokkary, F. S. Khan, D. La Torre, T. S. Humble, and J. Gottlieb, Financial portfolio management using D-wave quantum optimizer: The case of Abu Dhabi securities exchange (2017).

[5] R. Orus, S. Mugel, and E. Lizaso, Reviews in Physics **4**, 100028 (2019).

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

[7] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn, et al., Quantum Science and Technology **3**, 030503 (2018).

[8] R. P. Feynman, Int. J. Theor. Phys **21** (1982).

[9] A. Y. Kitaev, arXiv preprint quant-ph/9511026 (1995).

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

[11] D. Coppersmith, An approximate fourier transform useful in quantum factoring, Technical Report RC19642, IBM. (1994).

[12] D. S. Abrams and S. Lloyd, Physical Review Letters **79**, 2586 (1997).

[13] D. S. Abrams and S. Lloyd, Physical Review Letters **83**, 5162 (1999).

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

[15] J. T. Seeley, M. J. Richard, and P. J. Love, The Journal of chemical physics **137**, 224109 (2012).

[16] B. Toloui and P. J. Love, arXiv preprint arXiv:1312.2579 (2013).

[17] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, Physical Review A **90**, 022305 (2014).

[18] J. R. McClean, R. Babbush, P. J. Love, and A. Aspuru-Guzik, The journal of physical chemistry letters **5**, 4368 (2014).

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

[20] C. J. Trout and K. R. Brown, International Journal of Quantum Chemistry **115**, 1296 (2015).

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

[22] J. Preskill, Quantum **2**, 79 (2018).

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

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

[25] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'brien, Nature communications **5**, 4213 (2014).

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

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

[28] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, et al., Physical Review X **8**, 031022 (2018).

[29] Y. Nam, J.-S. Chen, N. C. Pisenti, K. Wright, C. Delaney, D. Maslov, K. R. Brown, S. Allen, J. M. Amini, J. Apisdorf, et al., npj Quantum Information **6**, 1 (2020).

[30] P. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, et al., Physical Review X **6**, 031007 (2016).

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

[32] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Nature **567**, 491 (2019).

[33] Google AI Quantum and others, Science **369**, 1084 (2020).

[34] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. Kimchi-Schwartz, J. McClean, J. Carter, W. De Jong, and I. Siddiqi, Physical Review X **8**, 011021 (2018).

[35] S. E. Smart and D. A. Mazziotti, Physical Review A **100**, 022517 (2019).

[36] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. Bultink, X. Fu, C. Price, V. Ostroukh, N. Muthusubramanian, A. Bruno, et al., Physical Review A **100**, 010302 (2019).

[37] J. Sakurai and J. Napolitano, Modern Quantum Mechanics (2017).

[38] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Rev. Mod. Phys. **92**, 015003 (2020).

[39] J. Almlöf, Lecture Notes in Quantum Chemistry II: European Summer School in Quantum Chemistry (1994).

[40] R. H. Landau, Quantum Mechanics II: A Second Course in Quantum Theory (1995).

[41] A. L. Fetter and J. D. Walecka, Quantum Theory of Many-particle Systems (2003).

[42] J. Slalter, The Self-Consistent Field for Molecules and Solids: Quantum Theory of Molecules and SOlids (1974).

[43] J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, et al., Wiley Interdisciplinary Reviews: Computational Molecular Science **2**, 556 (2012).

[44] S. Bravyi and A. Y. Kitaev, Annals of Physics **298**, 210 (2002).

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

[46] Y.-K. Liu, M. Christandl, and F. Verstraete, Phys. Rev. Lett. **98**, 110503 (2007).

[47] Y. Ge, J. Tura, and J. I. Cirac, Faster ground state preparation and high-precision ground energy estimation with fewer qubits (2018).

[48] D. Wang, O. Higgott, and S. Brierley, Physical Review Letters **122** (2019).

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

[50] P. Jordan and E. P. Wigner, Z. Phys. **47**, 631 (1928).

[51] B. R. J. Taube A G, Int. J. Quantum Chem. **106** (2006).

[52] J. A. Nelder and R. Mead, The Computer Journal **7**, 308 (1965).

[53] J. C. Spall, IEEE Transactions on Automatic Control **37**, 332 (1992).

[54] J. L. Maryak and D. C. Chin, Proceeding of the 2001 Winter Simulation Conference (Cat. No.01CH37304) **1**, 307 (2001).

[55] A. Lucas, Frontiers in Physics **2** (2014).

[56] T. Jones, S. Endo, S. McArdle, X. Yuan, and S. C. Benjamin, Physical Review A **99** (2019).

[57] D. D'Alessandro, Introduction to quantum control and dynamics (2007).

[58] M. Steudtner and S. Wehner, New Journal of Physics **20**, 063010 (2018).

[59] V. V. Shende, I. L. Markov, and S. S. Bullock, Physical Review A **69**, 10.1103/physreva.69.062321 (2004).

[60] P. K. Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, G. Salis, A. Fuhrer, M. Ganzhorn, D. J. Egger, M. Troyer, A. Mezzacapo, and et al., Physical Review A **98**, 10.1103/physreva.98.022322 (2018).

[61] L. F. Richardson and J. A. Gaunt, Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character **226**, 299 (1927).

[62] K. Temme, S. Bravyi, and J. M. Gambetta, Physical review letters **119**, 180509 (2017).

[63] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, Physical review letters **120**, 210501 (2018).

[64] M. Schuld, I. Sinayskiy, and F. Petruccione, Contemporary Physics **56**, 172 (2015).

[65] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature **549**, 195 (2017).