

Variational quantum eigensolver for chemical molecules

Luca Petru Ion

Nottingham

(Dated: May 16, 2023)

Solving multi-particle systems is an important topic in quantum chemistry and condensed matter physics. In this article I focus on finding the ground states and ground state energies for the He-H+ and H2O molecules using the power of quantum computing. I will be using a variational quantum eigensolver algorithm (VQE) which will run on both a simulation of a quantum computer and on the IBM quantum computer. I will compare the results against the exact ground state energy found through other classical means. The H2O simulations were run on Nottingham's High Performance Computer (HPC).

I. INTRODUCTION

The second quantum revolution [1] gave rise to the fields of quantum information and quantum computation. In the present we find ourselves living in the era of Noisy Intermediate-Scale Quantum (NISQ) [2] devices; meaning that our quantum devices are build from a few hundred qubits and suffer from qubit errors (noisy). Nevertheless, this era has provided us with new technologies and disciplines like quantum cryptography [3] and quantum machine learning [4].

An important field that benefits from quantum computation is the quantum chemistry of molecules. The reason for this is that the Hilbert space grows exponentially with the number of qubits N and therefore complex molecules comprising of many particles take a lot of computer resources to store and manipulate on a classical computer using classical methods, whereas, using quantum computation we can work with quantum circuits and quantum gates which simplifies the process.

I focus on the VQE algorithm used to find the ground state and ground state energy of a multi particle Hamiltonian. I will first look at the He-H+ molecule which was studied extensively in [5]. I will introduce the circuit ansatz and discuss how the ansatz differs when a deeper circuit structure is desired. I will then discuss the results of VQE optimization. After, I will look at the more complex H2O molecule to see how VQE performs for a bigger system.

II. EXACT DIAGONALIZATION

First, to find the ground state energy for the Hamiltonian associated with the molecule under inspection one can use a normal eigen-value decomposition. The method used to obtain the Hamiltonian's for various molecules will be discussed in III. Once the Hamiltonian H is known, one can use any classical method to find the ground state energy. I have used used numpy for all mathematical operations, including the task of doing eigen-value decomposition and finding the exact ground state energy.

III. FINDING THE HAMILTONIAN

The first step is to get the Hamiltonian for the molecule under inspection. For He-H+ the Hamiltonian used in [5] is also used here; this is a 2 qubit Hamiltonian and is shown in section VI. For other more complex molecules, H2O in this case, I use the build in PennyLane molecularHamiltonian method [6].

IV. QUANTUM CIRCUIT ANSATZ

The main goal is to carry out optimization on a quantum circuit to reach a state that approximates the ground state of the Hamiltonian in question. The first task is to come up with a quantum circuit structure that is able to approximate the ground state. The circuit structure used in this paper is one that can extend over multiple layers M and over multiple qubits N . The circuit in figure 1 is an example of such a structure. The symbols RX and RZ signify the single quantum rotation gates

$$R_X(\theta_a) = \exp(-i\frac{\theta_a}{2}X), \quad (1)$$

$$R_Z(\theta_b) = \exp(-i\frac{\theta_b}{2}Z) \quad (2)$$

where X and Z are the Pauli gates and θ_a, θ_b are some rotation angles. Note that each rotation gate takes in as an argument an angle. The list of all these angles are precisely the parameters to be optimized so that the output state is approximately the ground state. In addition to these gates, there is a need to make the individual qubits 'communicate'; this is achieved by introducing the CZ (Control-Z) gate which is a 2 qubit gate with the control qubit as the first qubit and the target qubit as the second one. In matrix form Control-Z is

$$CZ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (3)$$

Control-Z simply applies the Z gate to the target qubit if the control qubit is $|1\rangle$, otherwise it lets it unchanged. In figure 1 all the connections between qubits are Control-Z gates.

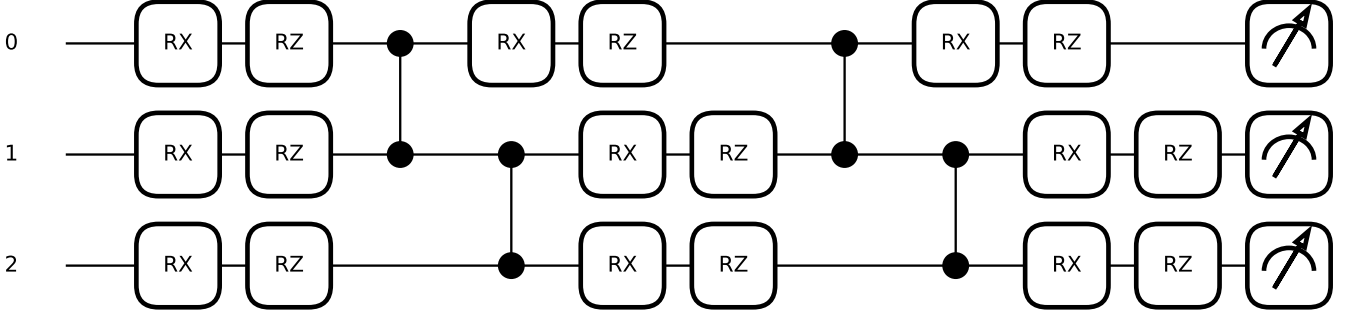


FIG. 1. Circuit ansatz for $N = 3$ and $M = 2$. R_X and R_Z are rotation gates and the connection between the wires is done through Control-Z gates.

V. METHOD

In order to approximate the ground state energy, VQE algorithm will be used. The ansatz circuit will be optimized such that the state produced by the circuit $|\psi\rangle$ minimizes the energy

$$E = \langle \psi | H | \psi \rangle. \quad (4)$$

It is important to note that $|\psi\rangle$, and as a result E , are both dependent on the parameter list $\vec{\theta}$; In machine learning language E is the loss function. The method used here for the optimization is gradient descent [7] and variations of it. For the classical gradient descent method the parameter list is updated over multiple iterations according to

$$\vec{\theta}_{n+1} = \vec{\theta}_n - \eta \nabla E(\vec{\theta}_n), \quad (5)$$

where η is the learning rate. In order to approximate the gradient there are also multiple different methods one can use. In this paper I compare the optimization for the following gradient descent approximation schemes: first order finite difference (FOGD) [8], second order finite difference (SOGD) [8], simultaneous perturbation stochastic approximation (SPSA) [9] and parameter shift (PS) [10].

Choosing a tuning method for η is not an easy task; here a constant η is used for FOGD, SOGD, PS and a decay in η for SPSA. Other more complex methods might be used [11]. The number of parameters to optimize depends on the number of qubits N and on the number of layers in the circuit ansatz M according to the following equation

$$\text{parameters} = 2N(M + 1). \quad (6)$$

The way we initialize these parameters is through a random process; each parameter is initialized to 2π times a random number between 0 and 1. The number of iterations used in the optimization is also fixed, even though one can design a stopping criteria implementation where the optimization stops whenever the stopping criteria is reached. Finally, the number of circuit evaluations (shots) is set to 8192.

VI. HE-H+ RESULTS

The first molecule considered is He-H+ (helium hydride). This molecule's Hamiltonian depends on the distance between the atoms R . The Hamiltonian is a $N = 2$ qubit Hamiltonian and is given by

$$H = \frac{1}{2} [J_x(X\mathbb{1} + \mathbb{1}X) + J_z(Z\mathbb{1} + \mathbb{1}Z) + J_{xx}XX + J_{zz}ZZ + J_{xz}(XZ + ZX) + C] \quad (7)$$

where J_x, J_z, J_{xx}, J_{zz} and C are all parameters dependent on R .

For this molecule I use $\eta = 0.8$ and 20 iterations in the optimization. The optimization plots for each of the methods are shown in figure 2. It is seen that the best convergence is for the parameter shift and SOGD methods. SPSA decreases a lot in very few iterations but then stagnates without decreasing further. Of course these observations depend heavily on the η tuning, therefore as mentioned above it is an important task to find an appropriate tuning for the method used.

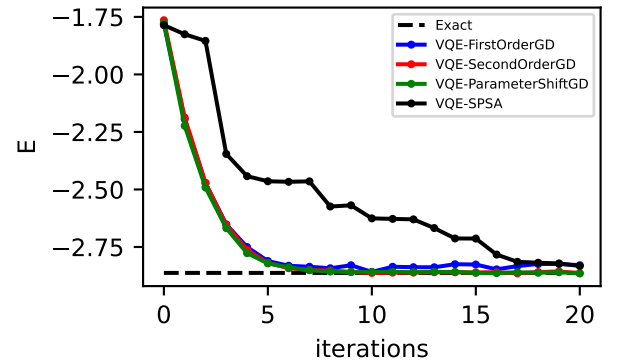


FIG. 2. Simulated He-H+ optimization for $R = 0.9\text{\AA}$, $M = 1$, $N = 2$, $\eta = 0.8$ and for all the different gradient descent variants.

Additionally one can carry out the optimization for other layers M and compare how the final energy com-

compares when varying M . For this molecule, it is clear that $M = 1$ is sufficient for optimization to converge, so increasing M doesn't give much better approximations.

To run the optimization on a real IBM machine, it is impractical to do so for all the different methods because of the very long queue times one needs to wait for the jobs to actually run on the machine, therefore SOGD was used since it did well in the simulation. Figure 3 shows the results of the optimization. At first it may be surprising that at the end of optimization the approximated energy is not very close to the exact energy; this however is due to the fact that the real devices are currently noisy and suffer from quantum errors.

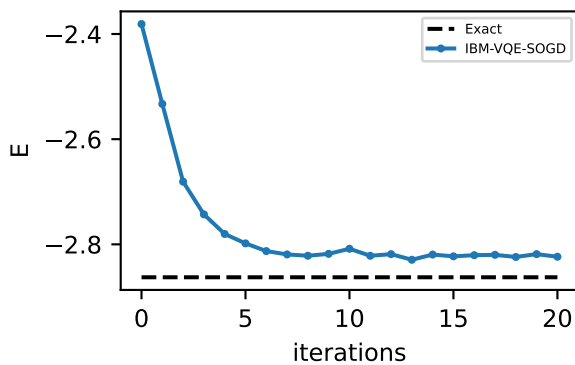


FIG. 3. IBM He-H+ optimization for $R = 0.9\text{\AA}$, $M = 1$, $N = 2$, $\eta = 0.8$ and SOGD.

The optimization was also carried out for $R = 0.75\text{\AA}$ and $R = 1.05\text{\AA}$ on the real IBM machine and it was also run on the simulation for every $R \in [0.5, 2.5]$ where R always increases by 0.05 . In figure 4 the approximated energy is plotted for each R and is compared against the exact energy found through exact diagonalization. The shape of the energy curve clearly shows that the energy is a minimum at $R = 0.9\text{\AA}$; this is the main reason we picked this value of R for the other plots in this section.

VII. H2O RESULTS

The next molecule considered is the water molecule H_2O . Here I consider a simplified model [12] where the active orbitals and active electrons are both set to 4. With these constraints, PennyLane returns an 8 qubit Hamiltonian for this system; not limiting the active orbitals would return a 14 qubit Hamiltonian. Since this system is so much larger than the He-H+ system the optimization takes considerably longer time to complete, even for a $M = 1$ ansatz. As a result, the HPC was used to carry out the optimization in order to benefit from its parallelism capabilities.

The water molecule is more complex than the He-H+ molecule and requires 2 parameters to describe its position in space. These are R and ϕ where the former is

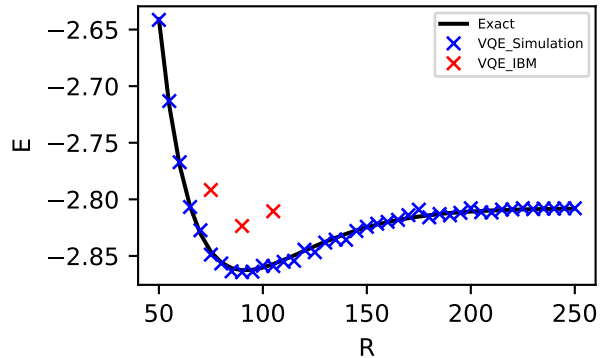


FIG. 4. He-H+: Energy(E) against R in picometres plot for the exact, VQE-simulation(SOGD) and IBM results using $M = 1$ ansatz and $\eta = 0.8$.

the bond length between the O-H bond and the latter is the angle between the two H atoms in the molecule; these are shown in figure 5. As in the He-H+, different values for these parameters give different Hamiltonian's and thus have different ground state energies. When examining the Hamiltonian's returned by PennyLane it was found that the lowest ground state energy occurs for the $R = 1.9\text{\AA}$ and $\phi = 1.75$ rad configuration which is different from the typical values found in the literature of $R = 0.96\text{\AA}$ and $\phi = 1.8$ rad. The reason for this discrepancy is that our model is an approximation to the full model, which comes from limiting the active orbitals and active electrons allowed to take part in the interactions.

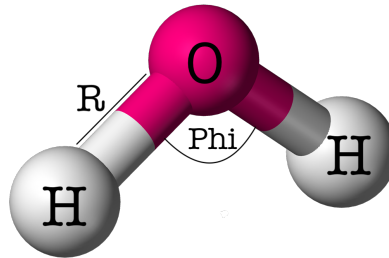


FIG. 5. The geometry of the H_2O molecule. R is the O-H bond length and ϕ is the HOH angle.

Optimization was run for $M = 1, 2, 3, 4, 5, 8$ layered ansatz for 100 iterations with $\eta = 0.8$. It was generally found that even $M = 1$ ansatz produces good approximations. Figure 6 shows the optimization for each method. We can see that all methods successfully converge apart from SPSA, this is in part due to the learning rate needing a better tuning other than a simple decay.

In order to show the dependence of the ground state energy on the angle ϕ , the bond length is fixed to $R = 1.9\text{\AA}$ and the angle is varied from 5° to 180° by 5° increments. Optimization is then run for each of these angles and both the exact energy and the approximated one are plotted. By running optimization multiple times, it was observed

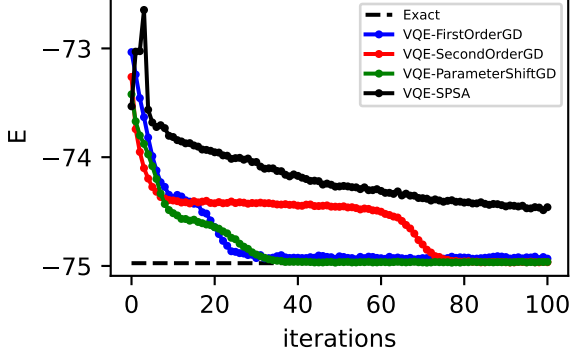


FIG. 6. Simulated H2O optimization for $\phi = 1.75$ rad, $R = 1.9\text{\AA}$, $M = 1$, $N = 8$, $\eta = 0.8$ and for all the different gradient descent variants.

that PS method generally did the best, therefore this is the method used here. Figure 7 shows the results for $M = 1, 2, 3$ ansatz circuits. We can see that there is a dip in the energy at $\phi = 1.75$ rad which is reminiscent to the dip seen for the He-H+ case in figure 4. This dip is expected since it occurs precisely at the value of ϕ which gives a minimum for the ground state energy. Moreover, we see that for some values of ϕ optimization is getting stuck and is not converging in the required number of sweeps. To fix this, one will need to find an appropriate η tuning scheme and run the optimization for longer and multiple times.

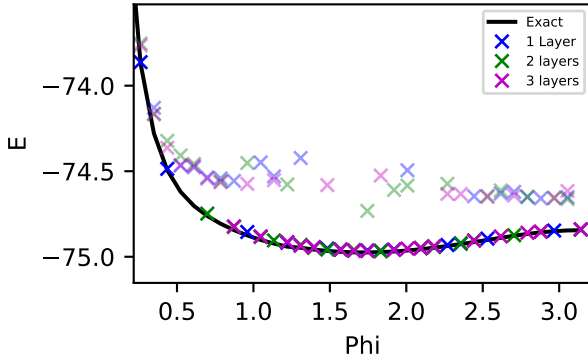


FIG. 7. H2O: Energy(E) against ϕ plot for the exact and VQE-simulation(PS) results using $M = 1, 2, 3$ layered ansatz and $\eta = 0.8$. The more transparent data is the one which didn't converge in the optimization

VIII. DISCUSSION

In this work I have carried out VQE for He-H+ and H2O using different ansatz circuits and for multiple gradient descent methods; it is found that VQE can success-

fully approximate the ground state energy even for a complex molecule like H2O. For the simpler He-H+ molecule the algorithm was also performed on a real IBM machine and the noisy nature of current quantum computers was exposed in the optimization plot.

From the optimization plots it is seen that generally PS does the best while SPSA the worst. PS has the advantage that it computes the exact gradient whereas SPSA has some stochastic nature to it. On the other hand SPSA is the fastest to complete optimization due to the considerably less calls to the quantum circuit needed compared to all the other methods. For the He-H+ this speed up is not noticed due to the fact that the circuit only has 2 qubits, but for the H2O molecule this is a very noticeable difference. Moreover, due to this reduction in calls to the quantum circuit, SPSA is also highly favoured when running optimization on the real IBM device. This is because for each call in the code, the request is put in a queue on the IBM servers and one has to wait the queue for the call to be executed; since SPSA has less calls this means that the IBM results are obtained a lot faster. SPSA however, contains more hyper parameters and therefore the challenge of fine tuning them is also higher.

More investigation is needed in the study of the H2O molecule. Firstly one needs to find a better tuning of the hyper parameters and to run optimization multiple times in an attempt to get convergence for as many of the points in figure 7. Another change one can try in order to make optimization easier is to change how the initial gates are initialized in the ansatz circuit. Instead of having fully random rotation gates, one can initialise them to very small rotations $\vec{\theta}_0 \approx \vec{0}$, so that the initial output state is very close to the first basis state and therefore carries low entanglement. This is done with the idea in mind that the actual ground state might exhibit low entanglement and then the initial ansatz output state would be closer to the ground state in the Hilbert space, and as a result convergence would be easier. Next, it would be interesting to keep ϕ constant and vary R and see how VQE can approximate the E-R plot. Finally, changing the active orbitals parameter to get a more realistic model of H2O (which will inevitably give a higher N for the circuit) and carrying optimization for that more complex model would be a next step.

-
- [1] J. P. Dowling and G. J. Milburn, Quantum technology: The second quantum revolution (2002).
 - [2] J. Preskill, Quantum Computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
 - [3] D. J. Bernstein, J. Buchmann, and E. Dahmén, *Post-quantum cryptography* (Springer, 2009).
 - [4] X. Gao, Z. Zhang, and L. Duan, An efficient quantum algorithm for generative machine learning (2017).
 - [5] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, A variational eigenvalue solver on a photonic quantum processor, *Nature Communications* **5**, 10.1038/ncomms5213 (2014).
 - [6] V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. Akash-Narayanan, A. Asadi, J. M. Arrazola, U. Azad, S. Banning, C. Blank, T. R. Bromley, B. A. Cordier, J. Ceroni, A. Delgado, O. D. Matteo, A. Dusko, T. Garg, D. Guala, A. Hayes, R. Hill, A. Ijaz, T. Isacsson, D. Ittah, S. Jangiri, P. Jain, E. Jiang, A. Khandelwal, K. Kottmann, R. A. Lang, C. Lee, T. Loke, A. Lowe, K. McKiernan, J. J. Meyer, J. A. Montañez-Barrera, R. Moyard, Z. Niu, L. J. O’Riordan, S. Oud, A. Panigrahi, C.-Y. Park, D. Polatajko, N. Quesada, C. Roberts, N. Sá, I. Schoch, B. Shi, S. Shu, S. Sim, A. Singh, I. Strandberg, J. Soni, A. Száva, S. Thabet, R. A. Vargas-Hernández, T. Vincent, N. Vitucci, M. Weber, D. Wierichs, R. Wiersema, M. Willmann, V. Wong, S. Zhang, and N. Killoran, PennyLane: Automatic differentiation of hybrid quantum-classical computations (2022), arXiv:1811.04968 [quant-ph].
 - [7] C. Lemaréchal, Cauchy and the gradient method, *Optimization Stories*, 251–254 (2012).
 - [8] S. Venkateshan and P. Swaminathan, Numerical differentiation, in *Computational Methods in Engineering* (Elsevier, 2014) pp. 295–315.
 - [9] J. Spall, Multivariate stochastic approximation using a simultaneous perturbation gradient approximation, *IEEE Transactions on Automatic Control* **37**, 332–341 (1992).
 - [10] PennyLane, Parameter-shift rules (2022).
 - [11] L. N. Smith, Cyclical learning rates for training neural networks, 2017 IEEE Winter Conference on Applications of Computer Vision (WACV) 10.1109/wacv.2017.58 (2017).
 - [12] A. Delgado, Building molecular hamiltonians - pennylane (2022).