Quantum Machine Learning for Electronic Structure Calculations

Rongxin Xia[1] and Sabre Kais *[1,2,3]

[1]*Department of Physics, Purdue University, West Lafayette, IN, 47907 USA*
[2]*Department of Chemistry and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907 USA*
[3]*Santa Fe Institute, 1399 Hyde Park Rd, Santa Fe, NM 87501*

Abstract

Considering recent advancements and successes in the development of efficient quantum algorithms for electronic structure calculations — alongside similarly impressive results using machine learning techniques for computation — hybridizing quantum computing with machine learning for the intent of perform electronic structure calculations is a natural progression. Here we present a hybrid quantum algorithm employing a quantum restricted Boltzmann machine to obtain accurate molecular potential energy surfaces. The Boltzmann machine trains parameters within an Ising-type model which exists in thermal equilibrium. By exploiting a quantum algorithm to optimize the underlying objective function, we obtained an efficient procedure for the calculation of the electronic ground state energy for a system. Our approach achieves high accuracy for the ground state energy of a simple diatomic molecular system such as H2, LiH, H2O at a specific location on its potential energy surface. With the future availability of larger-scale quantum computers and the possible training of some machine units with the simple dimensional scaling results for electronic structure, quantum machine learning techniques are set to become powerful tools to obtain accurate values for ground state energies and electronic structure for molecular systems.

1 Introduction

Machine learning techniques are demonstrably powerful tools displaying remarkable success in compressing high dimensional data [1, 2]. These methods have been applied to a variety of fields in both science and engineering, from computing excitonic dynamics [3], energy transfer in light-harvesting systems [4], molecular electronic properties [5], surface reaction network [6], learning density functional models [7] to classify phases of matter, and the simulation of classical and complex quantum systems [8, 9, 10, 11, 12, 13, 14]. The ability of modern machine learning techniques to classify, identify, and interpret massive data sets showcases their suitability to provide analyses on the exponentially large data sets embodied in the state space of complex condensed-matter systems [9] and to speed-up searches for novel energy generation/storage materials [15, 16]. Quantum machine learning [17] — a hybridization of classical machine learning techniques with quantum computation – is emerging as a powerful approach both allowing speed-ups and improving classical machine learning algorithms [18, 19, 20, 21, 22]. Recently, Wiebe *et. al.* [23] have shown that quantum computing is capable of reducing the time required to train a restricted Boltzmann machine (RBM), while also providing a richer framework for deep learning than its classical analogue. The standard RBM models the probability of a given configuration of visible and hidden units by the Gibbs distribution with interactions restricted between different layers. Here, we focus on an RBM where the visible and hidden units assume binary forms [24]. Accurate electronic structure calculations for large systems continue to be a challenging problem in the field of chemistry and material science. Toward this goal — in addition to the impressive progress in developing classical algorithms based on *ab initio* and density functional methods — quantum computing

*kais@purdue.edu

based simulation have been explored [25, 26, 27, 28, 29, 30]. Recently, Kivlichan *et. al.* [31] show that using a particular arrangement of gates (a fermionic swap network) it is possible to simulate electronic structure Hamiltonian with linear depth and connectivity. These results present significant improvement on the cost of quantum simulation for both variational and phase estimation based quantum chemistry simulation methods.

Recently, Troyer and coworkers proposed using a restricted Boltzmann machine to solve quantum many-body problems, for both stationary states and time evolutions of the quantum Ising and Heisenberg models [32]. However, this simple approach has to be modified for cases where the wave function's phase is required for accurate calculations [33].

Herein, we propose a three-layered RBM structure in addition to visible and hidden layers, and a new term correction for the phase of the wave function. We have shown that this model has the potential to solve complex quantum many-body problems and to obtain very accurate electronic potential energy surfaces for simple molecules.

2 Restricted Boltzmann Machine

We will begin by briefly outlining the original RBM structure as described by [32]. For a given Hamiltonian, H, and a trial state, $|\phi\rangle$, the expected value can be written as:

$$\langle H\rangle = \frac{\langle|\phi|H|\phi\rangle}{\langle\phi|\phi\rangle} = \frac{\sum_{x,x'}\langle\phi|x\rangle\langle x|H|x'\rangle\langle x'|\phi\rangle}{\sum_x\langle\phi|x\rangle\langle x|\phi\rangle} \tag{1}$$

where $\phi(x) = \langle x|\phi\rangle$ will be used throughout this letter to express the overlap of the complete wave function with the basis function, x.

We can map the above to a RBM model with visible layer units $\sigma_z^1,\ \sigma_z^2...\ \sigma_z^n$ and hidden layer units $h_1,\ h_2...\ h_m$ if we define $x = \{\sigma_z^1, \sigma_z^2...\sigma_z^n\}$, where x is taken to be a combination of the Pauli matrix σ_z and $\phi(x) = \sqrt{P(x)}$; where $P(x)$ is the probability of x. The probability of a specific set $x = \sigma_z^1\sigma_z^2...\sigma_z^n$ as:

$$P(x) = \sum_{\{h\}} \frac{exp(\sum_{i=1}^n a_i\sigma_z^i + \sum_{j=1}^m b_j h_j + \sum_{i=1}^n\sum_{j=1}^m w_{ij}\sigma_z^i h_j)}{\sum_{x'} P(x')}. \tag{2}$$

Within the above a_i and b_j are trainable weights for unit σ_z^i and h_j, w_{ij} are trainable weights describing the connections between σ_z^i and h_j (see **Figure 1**.)

By setting $\langle H\rangle$ as the loss function of this RBM, we can use the standard gradient decent method to update parameters via a backpropagation algorithm, effectively minimizing $\langle H\rangle$ to obtain the ground eigenvalue. To perform the calculations, we expand the wave function $|\phi\rangle = \sum_{i=1}^n \phi(x_i)|x_i\rangle$, where $|x_i\rangle$ is combinations of all values of $|\sigma_z^1\sigma_z^2...\sigma_z^n\rangle$.

However, previous prescriptions considering the use of RBMs for electronic structure problems have found difficulty as $\phi(x_i)$ can only assume non-negative values. We have thus appended an additional layer to the neural network architecture to compensate for sign features specific to electronic structure problems.

We propose an RBM with three layers. The first layer, σ_z, describes the parameters building the wave function. The h's within the second layer are parameters for the coefficients for the wave functions and the third layer s, is to train to represent the signs associated with each basis function.

$$s(\sigma_z^1, \sigma_z^2...\sigma_z^n) = 2\text{Sigmoid}(\sum_h \log(exp(\sum_{j=1}^m c_j h_j + \sum_{j=1}^n d_i\sigma_z^i))) - 1 \tag{3}$$

Sigmoid functions were employed to describe signs, used for their easier during implementation of back-prorogation algorithm through the gradient decent method. Due to the Sigmoid function, the s does not assume the exact values ±1; this issue was solved by application of subsequent $Sign$ function of the signs.

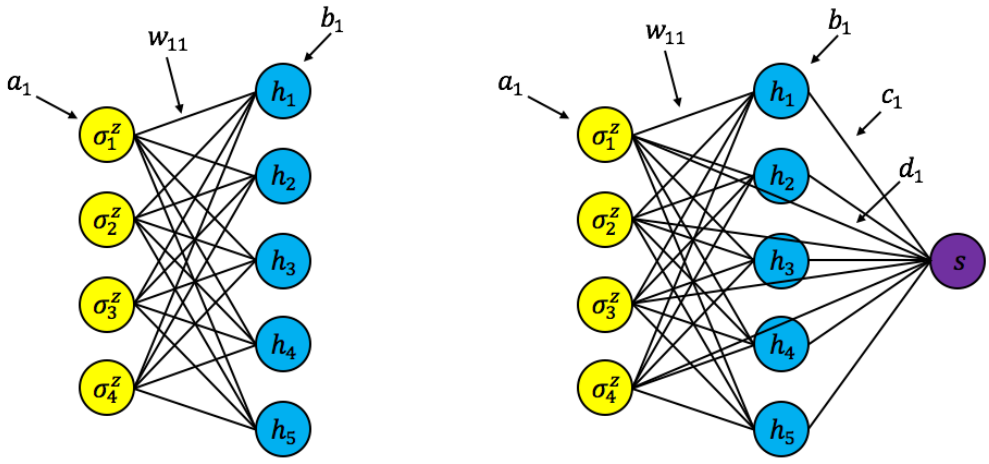


Figure 1: Left: the original restricted Boltzmann machine (RBM) structure with visible $\sigma_i^z$and hidden h_i layers. **Right**: Improved RBM structure with three layers, visible, hidden and sign units. $a_i, w_{ij}, b_i, c_i, d_i$ are trainable weights describing the different connection between layers.

see Supplemental Material for approach. Within this scheme, c_j, d_i are weights of connection between h_j and s and between σ_z^i and s (see **Figure 1**.) Our final wave function should be $|\phi\rangle = \sum_{i=1}^{n} s(x_i)\ \phi(x_i)|x_i\rangle$.

3 Approaches

We employed two quantum approaches to optimize the RBM model: (1) The phase estimation algorithm, and (2) a direct approach. The standard approach to optimize parameters within an RBM is Gibbs Sampling, which is not efficient and likely takes many iterations to converge (details in Supplemental Material).

The first quantum algorithm is based on the problem of thermalizing a quantum system with a quantum computer which has been thoroughly discussed in the literature [34, 35, 36]. Building on this work, Poulin and Wocjan present a quantum algorithm to prepare the thermal Gibbs state of interacting quantum systems and evaluating the relevant partition function [37]. The first approach makes use of the phase estimation algorithm (PEA) to measure the energy of the system according to the Hamiltonian:

$$H = \sum_{i=1}^{n} a_i Z_{\sigma_z,i} + \sum_{j=1}^{m} b_j Z_{h,j} + \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij} Z_{\sigma_z,i} Z_{h,j}, \tag{4}$$

with $Z_{\sigma_z,i}$ is the Pauli matrix σ^z on $|\sigma_z^i\rangle$ and $Z_{h,j}$ is the Pauli matrix σ^z on $|h_j\rangle$. By using this Hamiltonian we can get the energy which can be used to calculate probability distribution.

In **Figure 2** we show the flow chart of our algorithm with two main features: the first, is to perform phase estimation algorithm; and the second is to calculate and measure probability.

In this algorithm we have a system register (with $n+m$ qubits), energy register (with N qubits), scratchpad register comprised of k_1 qubits and one final register, all as shown in **Figure 3**(a). We performed a Hadamard transform on the initial states $|0^{\otimes N}\rangle$ followed by a Quantum phase estimation algorithm (PEA) to record energies within the energy register. For efficient procedure as shown in the supplementary materials, we transform the Hamiltonian to $H' = 2arcsin(e^{(H/2k_1)}e^{-k_2})$, where k_1 k_2 are the constants to control errors.

To obtain the probability distribution, we perform a controlled-rotation with angles $2arcsin(e^{(E/2k_1)}e^{-k_2})$ on k_1 qubits initiated in $|0\rangle$ states. Then we use the CNOT gate to sum up amplitudes in k_1 qubits to the

final register followed by measurement of the final register and first n-qubits in the system register to obtain the probability distribution associated with σ^z (details in Supplemental Material).

The algorithm's complexity scales as $O(Nmn)$; qubits requirements are $O(mn/2)$ and the error works out to be $O(k_1 e^{-k_2})$. If we can set restrictions on w in the parameters, the complexity reduces to $O(Nmn)$, and the qubit requirements comes to $O(m+n)$.

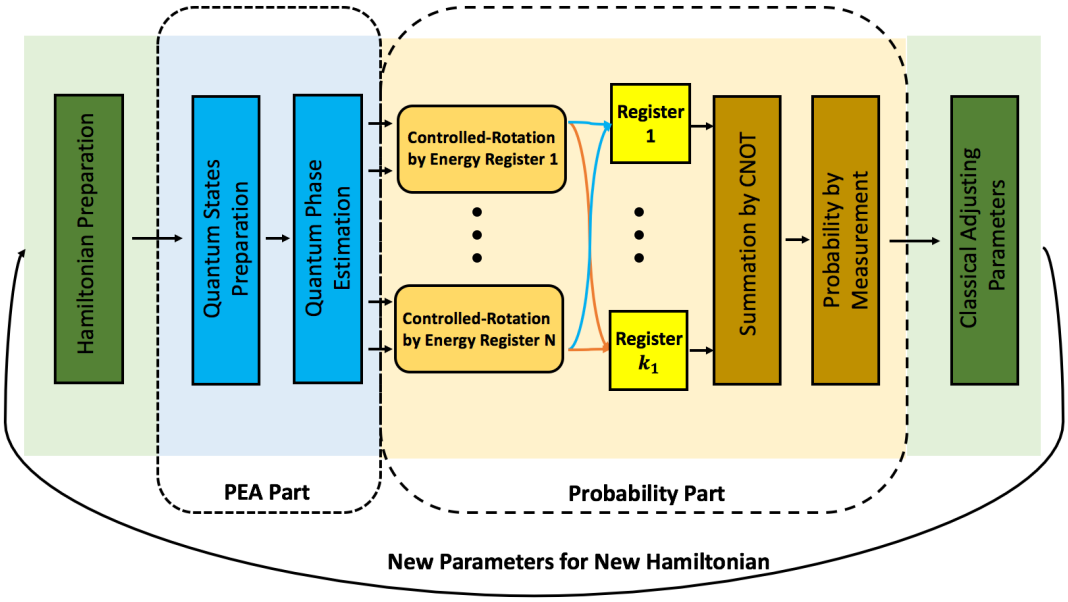
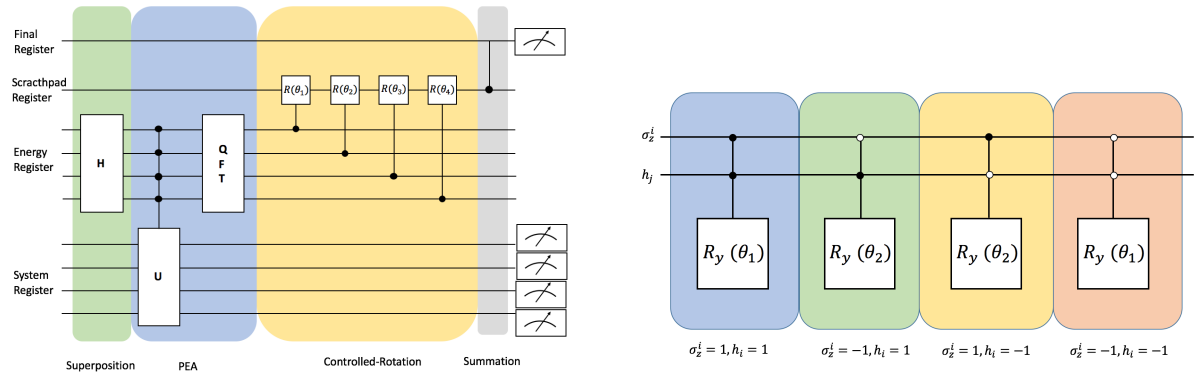


Figure 2: The algorithmic flow chart of our quantum algorithm based on the phase estimation algorithm.

The second quantum algorithm is based on sequential applications of controlled-rotation operations on qubits which represent the visible and hidden units. In **Figure 4** we provide the algorithmic flow chart two main features: (1) application of controlled-rotation gates to calculate amplitudes; and (2) the summation of amplitudes and measure probability.



(a) (b)

Figure 3: (a) The entire quantum circuit for our PEA approach. (b) The complete circuit for our controlled-rotation gate approach.

The idea of sequential controlled-rotation gates is to check whether the target qubit is in state $|0\rangle$ or state $|1\rangle$, then rotate the corresponding angle (see **Figure 3** (b).) In this algorithm is comprised of a system register (with $n+m$ qubits), three different scratchpad registers (with n,m and nm qubits) and a final register, as shown in **Figure 3 (a)**. This algorithm uses controlled-rotation gates on all three scratchpad registers and then sums amplitudes on the final register; followed by measurement of the final register and the first n-qubits of the system register to obtain the probability distribution associated with σ^z. The complexity comes to $O(4mn)$, and the qubits requirement comes to $O(4mn)$ (see Supplemental Material for details.)

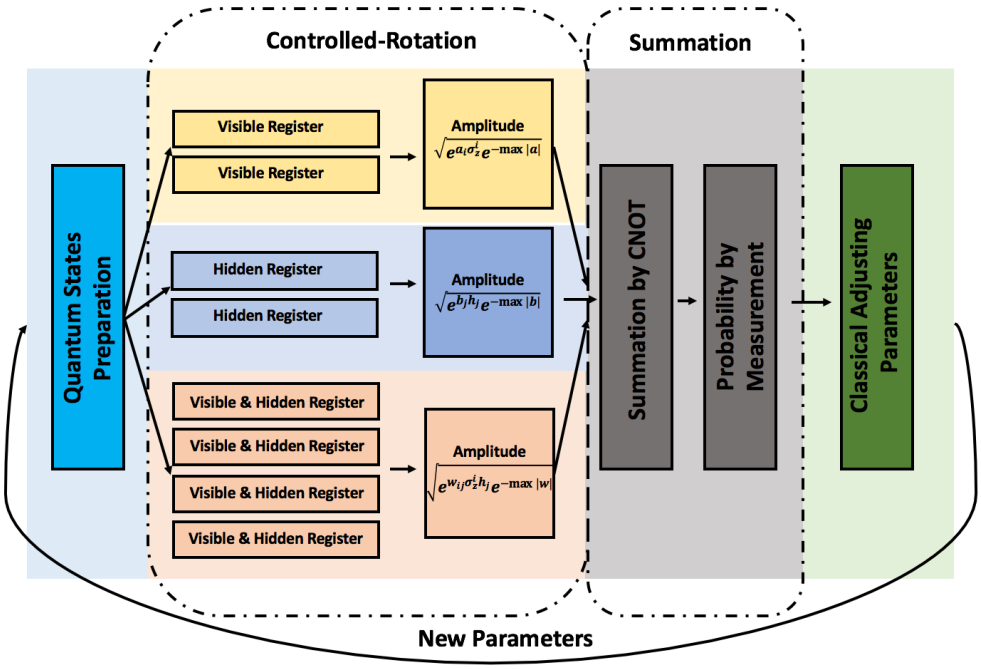


Figure 4: The algorithmic flow chart of the quantum algorithm based on sequential controlled-rotations gates.

4 Result and Discussion

We now present the results derived from our RBM for H_2, LiH and H_2O diatomic. It can clearly be seen that our three layer RBM yields a very accurate results. Points deviating from the ideal curve are likely due to local minima trapping during the optimization procedure. This can be avoided in the future by implementing optimization methods which include momentum or exication, increasing the escape probability from any local features of the potential energy surface.

Further discussion about our results should mention instances of transfer learning. Transfer learning is a unique facet of neural network machine learning algorithms describing an instance (engineered or otherwise) where the solution to a problem can inform or assist in the solution to another similar subsequent problem. Given a diatomic Hamiltonian at a specific intermolecular separation, the solution yielding the variational parameters — which are the weighting coefficients of the basis functions —- are adequate first approximations to those parameters at a subsequent calculation where the intermolecular separation is a small perturbation to the previous value.

Also noteworthy, to examine the transfer learning facility for these systems we count the number of iterations necessary to determine each point along the ground state energy curve for the LiH system (see **Figure 5** (d)). Except the first point in the **Figure 5** (d), calculations initiated with transferred parameters from previous iterations require 1/10 the iterations for each points and still achieve a good result. We also see

that the local minimum is avoided if the starting point achieve global minimum.

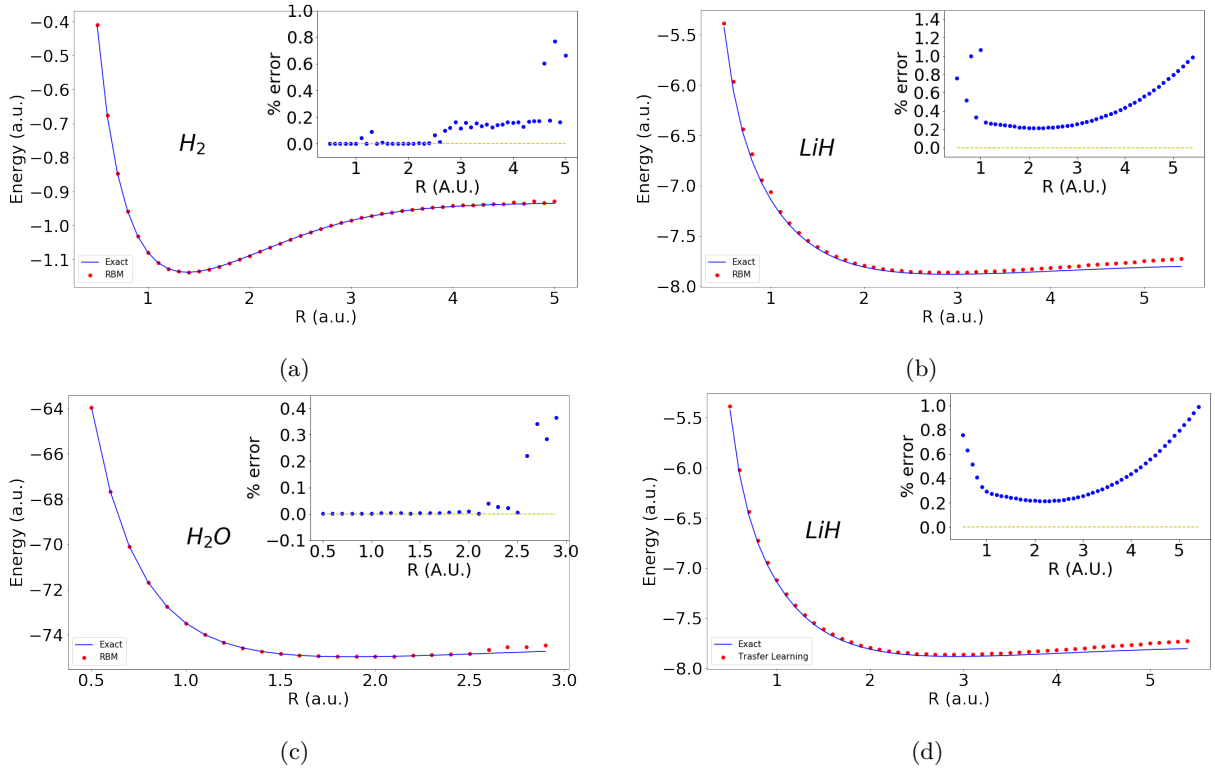


Figure 5: (a), (b), (c) are the results of H_2 ($n = 2,\ m = 4\ N = 10$), LiH ($n = 4,\ m = 8\ N = 10$) and H_2O ($n = 4,\ m = 8\ N = 10$) calculated by our three layer RBM compared with exact diagonalized results. (d) is the result of LiH ($n = 4,\ m = 8\ N = 10$) calculated by the Transfer Learning method.

Herein, we have demonstrated the ability of our three-layer RBM at solving a class of complex quantum many-body problems. Additionally, the coupling of machine learning and the phase estimation algorithm can give an excellent approximate solution for the ground state energy in molecular systems.

For training some of the units in the quantum machine, one can rely on dimensional scaling results for electronic structure calculations as a zeroth approximation. This approach pioneered by Herschbach [38], treats the dimensionality of the system as a free parametric and solve the problem at the large-dimensional limit. At this limit the problem becomes simpler as one could treat it classically by finding the minimum energy structure of the localized electrons. With the rapid development of larger-scale quantum computers and the possible training of some machine units with the simple dimensional scaling results for electronic structure, quantum machine learning techniques are set to become powerful tools to perform electronic structure calculations and assist in designing new materials for specific applications.

References

[1] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.

[2] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[3] Florian Häse, Stéphanie Valleau, Edward Pyzer-Knapp, and Alán Aspuru-Guzik. Machine learning exciton dynamics. *Chemical Science*, 7(8):5139–5147, 2016.

[4] Florian Häse, Christoph Kreisbeck, and Alán Aspuru-Guzik. Machine learning for quantum dynamics: deep learning of excitation energy transfer properties. *Chemical Science*, 8(12):8419–8426, 2017.

[5] Grégoire Montavon, Matthias Rupp, Vivekanand Gobre, Alvaro Vazquez-Mayagoitia, Katja Hansen, Alexandre Tkatchenko, Klaus-Robert Müller, and O Anatole Von Lilienfeld. Machine learning of molecular electronic properties in chemical compound space. *New Journal of Physics*, 15(9):095003, 2013.

[6] Zachary W Ulissi, Andrew J Medford, Thomas Bligaard, and Jens K Nørskov. To address surface reaction network complexity using scaling relations machine learning and dft calculations. *Nature communications*, 8:14621, 2017.

[7] Felix Brockherde, Leslie Vogt, Li Li, Mark E Tuckerman, Kieron Burke, and Klaus-Robert Müller. Bypassing the kohn-sham equations with machine learning. *Nature communications*, 8(1):872, 2017.

[8] Lei Wang. Discovering phase transitions with unsupervised learning. *Physical Review B*, 94(19):195105, 2016.

[9] Juan Carrasquilla and Roger G Melko. Machine learning phases of matter. *Nature Physics*, 2017.

[10] Peter Broecker, Juan Carrasquilla, Roger G Melko, and Simon Trebst. Machine learning quantum phases of matter beyond the fermion sign problem. *Scientific reports*, 7(1):8823, 2017.

[11] Kelvin Ch'ng, Juan Carrasquilla, Roger G Melko, and Ehsan Khatami. Machine learning phases of strongly correlated fermions. *Physical Review X*, 7(3):031038, 2017.

[12] Evert PL van Nieuwenburg, Ye-Hua Liu, and Sebastian D Huber. Learning phase transitions by confusion. *Nature Physics*, 13(5):435–439, 2017.

[13] Louis-François Arsenault, Alejandro Lopez-Bezanilla, O Anatole von Lilienfeld, and Andrew J Millis. Machine learning for many-body physics: the case of the anderson impurity model. *Physical Review B*, 90(15):155136, 2014.

[14] Aaron Gilad Kusne, Tieren Gao, Apurva Mehta, Liqin Ke, Manh Cuong Nguyen, Kai-Ming Ho, Vladimir Antropov, Cai-Zhuang Wang, Matthew J Kramer, Christian Long, et al. On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets. *Scientific reports*, 4, 2014.

[15] Phil De Luna, Jennifer Wei, Yoshua Bengio, Alán Aspuru-Guzik, and Edward Sargent. Use machine learning to find energy materials. *Nature*, 552(7683):23, 2017.

[16] Jennifer N Wei, David Duvenaud, and Alán Aspuru-Guzik. Neural networks for the prediction of organic chemistry reactions. *ACS central science*, 2(10):725–732, 2016.

[17] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195, 2017.

[18] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum algorithms for supervised and unsupervised machine learning. *arXiv preprint arXiv:1307.0411*, 2013.

[19] Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. *Physical review letters*, 113(13):130503, 2014.

[20] Hartmut Neven, Geordie Rose, and William G Macready. Image recognition with an adiabatic quantum computer i. mapping to quadratic unconstrained binary optimization. *arXiv preprint arXiv:0804.4457*, 2008.

[21] Hartmut Neven, Vasil S Denchev, Geordie Rose, and William G Macready. Training a binary classifier with the quantum adiabatic algorithm. *arXiv preprint arXiv:0811.0416*, 2008.

[22] Hartmut Neven, Vasil S Denchev, Geordie Rose, and William G Macready. Training a large scale classifier with the quantum adiabatic algorithm. *arXiv preprint arXiv:0912.0779*, 2009.

[23] Nathan Wiebe, Ashish Kapoor, and Krysta M Svore. Quantum deep learning. *arXiv preprint arXiv:1412.3489*, 2014.

[24] Mohammad H Amin, Evgeny Andriyash, Jason Rolfe, Bohdan Kulchytskyy, and Roger Melko. Quantum boltzmann machine. *arXiv preprint arXiv:1601.02036*, 2016.

[25] S. Kais. *Quantum Information and Computation for Chemistry: Advances in Chemical Physics*, volume 154. Wiley Online Library, 2014.

[26] Ammar Daskin and Sabre Kais. Direct application of the phase estimation algorithm to find the eigenvalues of the hamiltonians. *Chemical Physics*, 2018.

[27] Alán Aspuru-Guzik, Anthony D Dutoi, Peter J Love, and Martin Head-Gordon. Simulated quantum computation of molecular energies. *Science*, 309(5741):1704–1707, 2005.

[28] PJJ O'Malley, Ryan Babbush, ID Kivlichan, Jonathan Romero, JR McClean, Rami Barends, Julian Kelly, Pedram Roushan, Andrew Tranter, Nan Ding, et al. Scalable quantum simulation of molecular energies. *Physical Review X*, 6(3):031007, 2016.

[29] Ivan Kassal, James D Whitfield, Alejandro Perdomo-Ortiz, Man-Hong Yung, and Alán Aspuru-Guzik. Simulating chemistry using quantum computers. *Annual review of physical chemistry*, 62:185–207, 2011.

[30] Ryan Babbush, Peter J Love, and Alán Aspuru-Guzik. Adiabatic quantum simulation of quantum chemistry. *Scientific reports*, 4:6603, 2014.

[31] Ian D Kivlichan, Jarrod McClean, Nathan Wiebe, Craig Gidney, Alán Aspuru-Guzik, Garnet Kin-Lic Chan, and Ryan Babbush. Quantum simulation of electronic structure with linear depth and connectivity. *Physical Review Letters*, 120(11):110501, 2018.

[32] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

[33] Giacomo Torlai, Guglielmo Mazzola, Juan Carrasquilla, Matthias Troyer, Roger Melko, and Giuseppe Carleo. Many-body quantum state tomography with neural networks. *arXiv preprint arXiv:1703.05334*, 2017.

[34] Barbara M Terhal and David P DiVincenzo. Problem of equilibration and the computation of correlation functions on a quantum computer. *Physical Review A*, 61(2):022301, 2000.

[35] RD Somma, S Boixo, Howard Barnum, and E Knill. Quantum simulations of classical annealing processes. *Physical review letters*, 101(13):130504, 2008.

[36] Pawel Wocjan, Chen-Fu Chiang, Daniel Nagaj, and Anura Abeyesinghe. Quantum algorithm for approximating partition functions. *Physical Review A*, 80(2):022340, 2009.

[37] David Poulin and Pawel Wocjan. Sampling from the thermal quantum gibbs state and evaluating partition functions with a quantum computer. *Physical review letters*, 103(22):220502, 2009.

[38] Dudley R Herschbach, John S Avery, and Osvaldo Goscinski. *Dimensional scaling in chemical physics*. Springer Science & Business Media, 2012.

[39] David Poulin and Pawel Wocjan. Sampling from the thermal quantum gibbs state and evaluating partition functions with a quantum computer. *Physical review letters*, 103(22):220502, 2009.

[40] Dominic W Berry, Andrew M Childs, Richard Cleve, Robin Kothari, and Rolando D Somma. Exponential improvement in precision for simulating sparse hamiltonians. In *Forum of Mathematics, Sigma*, volume 5. Cambridge University Press, 2017.

[41] Jacob T Seeley, Martin J Richard, and Peter J Love. The bravyi-kitaev transformation for quantum computation of electronic structure. *The Journal of chemical physics*, 137(22):224109, 2012.

[42] Sergey Bravyi, Jay M Gambetta, Antonio Mezzacapo, and Kristan Temme. Tapering off qubits to simulate fermionic hamiltonians. *arXiv preprint arXiv:1701.08213*, 2017.

[43] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Jerry M Chow, and Jay M Gambetta. Hardware-efficient quantum optimizer for small molecules and quantum magnets. *arXiv preprint arXiv:1704.05018*, 2017.

[44] Rongxin Xia, Teng Bian, and Sabre Kais. Electronic structure calculations and the ising hamiltonian. *The Journal of Physical Chemistry B*.

Acknowledgement

We would thank to Dr. Ross Hoehn for critical reading and useful discussions.

Supplemental Material

Rongxin Xia[1] and Sabre Kais*[1,2,3]

[1]*Department of Physics, Purdue University, West Lafayette, IN, 47907 USA*
[2]*Department of Chemistry and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907 USA*
[3]*Santa Fe Institute, 1399 Hyde Park Rd, Santa Fe, NM 87501*

1 Restricted Boltzmann Machine (RBM)

1.1 Original RBM

We briefly introduce the RBM structure [32] as shown in Figure 1 of the main text. For a Hamiltonian, H, and a trial wave function, $|\phi\rangle$, the expectation value can be written as:

$$\langle H\rangle = \frac{\langle\phi|H|\phi\rangle}{\langle\phi|\phi\rangle} = \frac{\sum_{x,x'}\langle\phi|x\rangle\langle x|H|x'\rangle\langle x'|\phi\rangle}{\sum_x\langle\phi|x\rangle\langle x|\phi\rangle} = \frac{\sum_x|\phi(x)|^2E_{loc}(x)}{\sum_x|\phi(x)|^2}. \tag{1}$$

Where $E_{loc}(x) = \sum_{x'}H_{xx'}\frac{\phi(x')}{\phi(x)}$, $H_{xx'} = \langle x|H|x'\rangle$ and $\phi(x) = \langle x|\phi\rangle$

If we set $x = \{\sigma_z^1, \sigma_z^2...\sigma_z^n\}$, where x is a combination of σ_z and $\phi(x) = \sqrt{P(x)}$ where $P(x)$ is the probability of x; we can map the above to an RBM with visible layer units $\sigma_z^1,\ \sigma_z^2...\ \sigma_z^n$ and hidden layer units $h_1,\ h_2...\ h_m$ with the probability of specific basis expansion $x = \sigma_z^1\sigma_z^2...\sigma_z^n$ as:

$$P(x) = \frac{exp(\sum_{i=1}^n a_i\sigma_z^i + \sum_{j=1}^m b_jh_j + \sum_{i=1}^n\sum_{j=1}^m w_{ij}\sigma_z^ih_j)}{\sum_{x'}P(x')}. \tag{2}$$

Where a_i and b_j are weights for the units σ_z^i and h_j, w_{ij} is the weight for the connection between σ_z^i and h_j (see Figure 1 in the main text)

If we select $\langle H\rangle$ as the loss function of this RBM. We have:

$$\partial_{p_k}\langle H\rangle \approx \langle\langle G_k(x)\rangle\rangle, \tag{3}$$

$$G_k(x) = 2Re((E_{loc}(x) - \langle\langle E_loc(x)\rangle\rangle)D_k^*(x), \tag{4}$$

$$\text{and}D_k^*(x) = \frac{\partial_{p_k}\phi(x)}{\phi(x)}, \tag{5}$$

where p_k is the parameters a_i, b_j, w_{ij} for k_{th} iterations.

We can use the gradient decent method to update the parameters and decrease $\langle H\rangle$ to obtain the desired ground state eigenvalue.

*kais@purdue.edu

1.2 Improved RBM

Based on the original RBM, which has one visible layer for the wave function and a hidden layer, we have appended an additional layer describing the signs of the basis functions used in the wave function expansion. Thus, this RBM can be used to perform electronic structure calculations.

The main idea of the RBM above is to calculate the parameters $\phi(x_i)$ for a state which can be written as:

$$|\phi\rangle = \sum_{i=1}^{n} \phi(x_i)|x_i\rangle, \tag{6}$$

where $|x_i\rangle$ is combinations of all values of $|\sigma_z^1\sigma_z^2...\sigma_z^n\rangle$. Continual optimization of the RBM yields k_i, which achieves the minimum of the optimization function, $\langle H\rangle$. However, when considering electronic structure problems, the coefficients are insufficient as $\phi(x_i)$ in the above constructions can only assume a non-negative values. We, thus, add an additional layer to correct for the sign.

Our RBM has three layers. The first layer, σ_z, constructs the wavefunction, Eq. 7 below. The third layer, s, generates the signs associated with the basis functions, and the second layer, h, are parameters for the wave function and corrected signs.

$$\phi(\sigma_z^1, \sigma_z^2...\sigma_z^n) = \sqrt{P(\sigma_z^1, \sigma_z^2...\sigma_z^n)} \tag{7}$$

$$P(x_i) = \sum_{\{h\}} \frac{\sum_h exp(\sum_{i=1}^n \sum_{j=1}^m w_{ij}\sigma_z^i h_j + \sum_{j=1}^m b_j h_j + \sum_{j=1}^n a_i\sigma_z^i)}{\sum_{x'} P(x')} \tag{8}$$

We use a Sigmoid function for signs as it is simpler during the application of the back-prorogation algorithm based-on the gradient decent method.

$$s(\sigma_z^1, \sigma_z^2...\sigma_z^n) = 2Sigmoid(\sum_h \log(exp(\sum_{j=1}^m c_j h_j + \sum_{j=1}^n d_i\sigma_z^i))) - 1 \tag{9}$$

In our RBM construction, s does not exactly span ± 1; we transform s to be the sign of s when we finally evaluate the ground state energy, where c_j, d_i are weights of connection between h_j and s and between σ_z^i and s (see Figure 1 in the main text).

In our construction, the expectation value can be written as:

$$\langle H\rangle = \frac{\sum_{x,x'} \langle\phi|x\rangle\langle x|Hs(x)s(x')|x'\rangle\langle x'|\phi\rangle}{\sum_x \langle\phi|x\rangle\langle x|\phi\rangle}. \tag{10}$$

1.3 Gradient Descent Method

We use the back propagation algorithm with the gradient decent method to optimize our RBM, yielding the global minimum corresponding to the ground energy.

$$p_{k+1} = p_k - \alpha\partial_{p_k}\langle H\rangle \tag{11}$$

Where α is the learning rate, controlling the convergence rate.

As stated before, we have:

$$
\begin{aligned}
D_{a_i}(\sigma_z) &= \frac{1}{2}\sigma_z^i, \\
D_{b_i}(\sigma_z) &= \frac{1}{2}tanh(\theta_i), \\
D_{w_{ij}}(\sigma_z) &= \frac{1}{2}tanh(\theta_i)\sigma_z^j, \\
D_{c_i}(\sigma_z) &= \frac{1}{2}tanh(c_i)(1 - s(\sigma_z)^2), \\
D_{d_i}(\sigma_z) &= \frac{1}{2}\sigma_z^i(1 - s(\sigma_z)^2),
\end{aligned}
\tag{12}
$$

where $\theta_j = \sum_i w_{ij}\sigma_i^z + b_j$.

We can continue iterating until we achieve a minimum value of the optimization function, $\langle H \rangle$.

2 Classical Approach

Here, we consider an RBM with m visible units and n hidden units. The updating procedure follows the normal RBM procedure as before $p_k = p_k - \alpha\partial_{p_k}\langle H \rangle$.

Using the Gibbs Sampling method we can simulate $P(x)$, which delivers $O(mn)$ for each time sampling. We then continue through the stepwise procedure outlined below:

- Generate n random number $r_j \in [0, 1)$.
- Calculate probability for j_{th} h as $P(h_j = 1|\sigma^z) = Sigmoid(2\theta_j)$. If $P(h_j = 1|\sigma^z) > r_j$, change h_j to 1 otherwise unchanged.
- Generate m random number $r_i \in [0, 1)$.
- Calculate probability for $P(\sigma_i^z|h) = Sigmoid(2\gamma_i)$. If $P(\sigma_i^z = 1|h) > r_i$, change σ_i^z to 1 otherwise unchanged.

where $\theta_j = \sum_i w_{ij}\sigma_i^z + b_j$ and $\gamma_i = \sum_j w_{ij}h_j + a_i$. And do N_s times, each time is to just sample one σ^z.

3 Quantum Approach

Here, we have to obtain $P(\sigma^z)$ by using a quantum computation approach to get an error-controlled or even exact result for the probability distribution.

3.1 Phase Estimation Algorithm

Our work — based on the previous work [23, 39] — tries to calculate the probability through application of controlled-rotation gates $R_y(arcsin(e^E))$ controlled by qubits in the state $|E\rangle$. We will not employ the mean-field approximation in [23], as the parameter w_{ij} can not converge to 0 which makes the mean-field approximation non-convergent.

The standard approach to constructing the controlled-rotation gate of $(R_y(arcsin(e^E)))$ cost $O(2^N)$, where N is the number of qubits holding $arcsin(e^E)$. Here we present a construction which is able to write the probability with only $O(Nmn)$ gates and $O(mn)$ qubits requirements and a controllable degree of error.

The probability for each combination of $(\sigma_z^i,\ h_j)$ can be written as:

$$P(x) = \frac{e^{\sum_{i=1}^{n} a_i \sigma_z^i + \sum_{j=1}^{m} b_j h_j + \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\sigma_z^i h_j}}{\sum_{x'} P(x')}, \tag{13}$$

and we can construct a Hamiltonian:

$$H = \sum_{i=1}^{n} a_i Z_{\sigma_z,i} + \sum_{j=1}^{n} b_j Z_{h,j} + \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij} Z_{\sigma_z,i} Z_{h,j}. \tag{14}$$

Within the above, the Hilbert space is constructed as $\otimes_{i=1}^{n}|\sigma_z^i\rangle \otimes_{j=1}^{m} |h_j\rangle$, $Z_{\sigma_z,i}$ is the Pauli matrix σ^z on $|\sigma_z^i\rangle$ and $Z_{h,j}$ is the Pauli matrix σ^z on $|h_j\rangle$. For specific state $|\phi\rangle = \otimes_{i=1}^{n}|\sigma_z^i\rangle \otimes_{j=1}^{m} |h_j\rangle$ we have $H|\phi\rangle = \sum_{i=1}^{n} a_i\sigma_z^i + \sum_{j=1}^{m} b_j h_j + \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\sigma_z^i h_j$.

We constructed a new Hamiltonian, H', from the original Hamiltonian, H, as:

$$H' = 2arcsin(e^{(H/2k_1)}e^{-k_2}). \tag{15}$$

Where k_1, k_2 are the constants to control the eigenvalue spectrum. $arcsin(\ )$ and $e^{(\ )}$ can be achieved by Taylor series expansion in terms of matrices:

$$\begin{aligned}
H_1 &= e^{\frac{1}{2k_1}(\sum_{i=1}^{n} a_i Z_{\sigma_z,i} + \sum_{j=1}^{m} b_j Z_{h,j} + \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij} Z_{\sigma_z,i} Z_{h,j})} \times e^{-k_2} \\
&= e^{-k_2}(1 + (\frac{1}{2k_1}(\sum_{i=1}^{n} a_i Z_{\sigma_z,i} + \sum_{j=1}^{m} b_j Z_{h,j} + \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij} Z_{\sigma_z,i} Z_{h,j}))) + O(e^{-k_2}(\frac{H}{2k_1})^2) \\
H' &= 2arcsin(H_1) = 2H_1 + O(H_1^3) \\
&= e^{-k_2}(2 + (\frac{1}{k_1}(\sum_{i=1}^{n} a_i Z_{\sigma_z,i} + \sum_{j=1}^{m} b_j Z_{h,j} + \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij} Z_{\sigma_z,i} Z_{h,j}))) + O(e^{-k_2}(\frac{H}{2k_1}^2)) + O(H_1^3)
\end{aligned} \tag{16}$$

Now we have the altered Hamiltonian, H', and are ready to post probability on each state by the phase estimation algorithm (PEA). phase estimation algorithm is an algorithm with input a Hamiltonian and an eigenstate and outputs an eigenvalue, $U_{PEA}|\phi\rangle|0\rangle = |\phi\rangle|E\rangle$, where E is the corresponding eigenvalue of the input eigenstate $|\phi\rangle$. In this algorithm we have a system register (with $n+m$ qubits), energy register (with N qubits), scratchpad register (with k_1 qubits) and one final register. The steps of our algorithm are as below:

- Prepare the state:
$$|\Phi\rangle = \frac{1}{\sqrt{2^{m+n}}} \sum_{l=1}^{2^{m+n}} |\phi_l\rangle \otimes_{k=1}^{N} |0_k\rangle \otimes_{l=1}^{k_1} |0_l\rangle|0\rangle \tag{17}$$
where $|\phi_i\rangle = \otimes_{j=1}^{n}|\sigma_z^j\rangle \otimes_{k=1}^{m} |h_k\rangle$ is the combination with specific σ^z, h and N is number of digits we use to do PEA.

- Use our Hamiltonian to do the PEA and we get:
$$\frac{1}{\sqrt{2^{m+n}}} \sum_{l=1}^{2^{m+n}} |\phi_l\rangle|2arcsin(e^{E/2k_1}e^{-k_2})\rangle \otimes_{l}^{k_1} |0_l\rangle|0\rangle \tag{18}$$

- Post probability to all the ancillary l qubits by $R_y(e^{E/2k_1}e^{-k_2})$ as seen in **Figure 3 (b)** (of the main text), yielding:
$$\begin{aligned}
&\frac{1}{\sqrt{2^{m+n}}} \sum_{l=1}^{2^{m+n}} |\phi_l\rangle|2arcsin(e^{E/2k_1}e^{-k_2})\rangle \\
&\otimes_{l=1}^{k} e(\sqrt{1 - e^{E/k_1}e^{-2k_2}}|0_l\rangle + \sqrt{e^{E/k_1}e^{-2k_2}}|1_l\rangle)|0\rangle
\end{aligned} \tag{19}$$

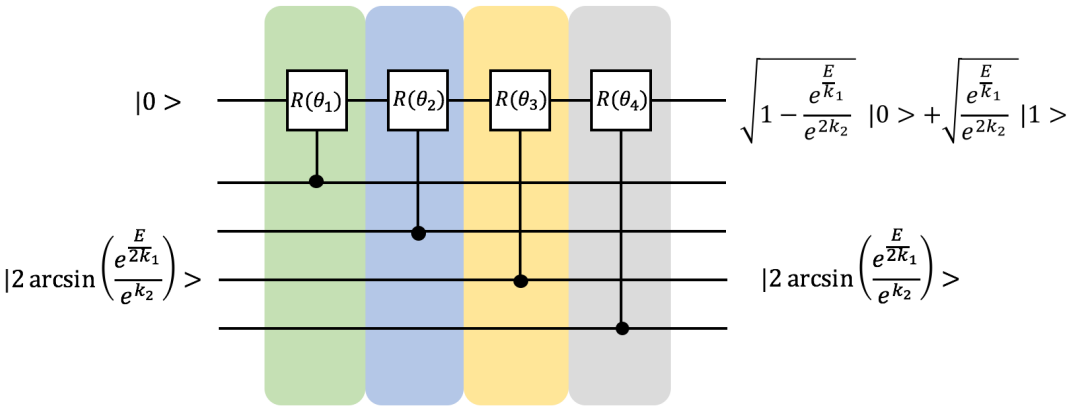


Figure 1: Post probability on ancillary qubit by Rotate-Y gate

The idea of our rotation gate is the control rotate of each digit of the total N number of digits. The controlled-rotation gate of the i_{th} qubit in the N is $CR_y(2\frac{2^{i-1}}{2^N})$.

- We can continue to do the CNOT gates controlled by all l qubits on the last qubit:

$$\frac{1}{\sqrt{2^{m+n}}}\sum_{l=1}^{2^{m+n}}|\phi_l\rangle|2arcsin(e^{E/2k_1}e^{-k_2})\rangle \\ (\sqrt{1-e^Ee^{-k_1k_2}}|\Psi\rangle|0\rangle + \sqrt{e^Ee^{-k_1k_2}}\otimes_{l=1}^{k_1}|1_l\rangle|1\rangle) \tag{20}$$

We perform a CNOT gate on the last qubits controlled by all k_1 qubits.

- If we measure the last qubits and get $|1\rangle$ we then measure the first n qubits to obtain the probability distribution of σ^z.

Thus, the constants k_1 and k_2 are chosen to maintain $H/2k_1 \in (-1,1)$ and $e^{H/2k_1}e^{-k_2} \in (0,1)$. k_1 is also a integer. The error above after CNOT gate controlled by l qubits can be written as $O(k_1e^{-k_2}(\frac{H}{2k_1})^2) + O(k_1H_1^3) = O(k_1e^{-k_2}e^{H/2k_2}) = O(k_1e^{-k_2})$. We also employ Grover's Search Algorithm to increase the probability we get the last qubit as $|1\rangle$.

The complexity of our method is $O(Nmn + N^2 + k_1N + k_1) = O(Nmn + Nk_1)$ where $O(Nmn)$ comes from the recording of the Hamiltonian on eigenstates in PEA (because our Hamiltonian is diagonal, we can use $O(Nmn)$ gates to finish PEA). $O(N^2)$ comes from Inverse Quantum Fourier Transform, $O(k_1N)$ comes our posting probability on all l qubits and and $O(k_1)$ comes from final CNOT gate. We have $k_1 = O(mn/2)$ because k_1 is a upper bound of $|H|/2$ where $|H| = O(mn)$. The complexity comes to $O(Nmn)$; qubits requirement comes to $O(m+n+N+k_1) = O(mn/2)$ and error should be $O(k_1e^{-k_2})$. If we set restrictions on w in the parameters, we would have $k_1 = O(m+n)$ as now $|H| = O(m+n)$. The complexity comes to $O(Nmn)$, and the qubit requirement comes to $O(m+n+N+k_1) = O(m+n)$.

3.2 Directly Post Probability Approach

We can also directly write probabilities. In this algorithm we have a system register (with $n+m$ qubits), three different scratchpad registers (with n,m and nm qubits) and one final register

We first apply Hadamard gates to system register, which will give us a superposition of all possible σ^z and h with the same possibility.

Then the next step is to write $P(\sigma^z, h)$ to each possible σ^z and h. As we discussed in the classical part, we have:

$$P(\sigma^z, h) = e^{\sum_i a_i \sigma_z^i + \sum_j b_j h_j + \sum_{i=1}^{n} \sum_{j=1}^{m} w_{i,j} \sigma_z^i h_j}. \tag{21}$$

Our method is to apply the terms of $P(\sigma^z, h)$ one-by-one. First, we start at the application to a_i. For σ_z^i we add an ancillary qubit which has an initial state of $|0\rangle$. Then we add two controlled-rotation gates to σ_z^i and the ancillary qubits, denoted by $CR_{a_i,1} + CR_{a_i,2}$.

$$\begin{aligned} CR_{a_i,1} &= A_{\sigma_z^i} \otimes R_y(2arccos(\sqrt{e^{-a_i e^{-|a_i|}}})) + B_{\sigma_z^i} \otimes I \\ CR_{a_i,2} &= B_{\sigma_z^i} \otimes R_y(2arccos(\sqrt{e^{a_i e^{-|a_i|}}})) + A_{\sigma_z^i} \otimes I \end{aligned} \tag{22}$$

$$A_{\sigma_z^i} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$B_{\sigma_z^i} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

The operation is to rotate the ancillary qubit differently according to different input state of σ_z^i. If the input of σ_z^i is $|0\rangle$, $CR_{a_i,1}$ will rotate the ancillary qubit by $arccos(\sqrt{e^{-a_i e^{-|a_i|}}})$ and $CR_{a_i,2}$ will do nothing. If the input of σ_z^i is $|1\rangle$, $CR_{a_i,2}$ will rotate the ancillary qubit by $arccos(\sqrt{e^{a_i e^{-|a_i|}}})$ and $CR_{a_i,1}$ will do nothing.

For b_j, we also have two controlled-rotation gates. For h_j, we add an ancillary qubit is initialized as $|0\rangle$. We then add two controlled-rotation gates to h_j and to the ancillary qubits, denoted as $CR_{b_i,1} + CR_{b_i,2}$.

$$\begin{aligned} CR_{b_j,1} &= A_{h_j} \otimes R_y(2arccos(\sqrt{e^{-b_j e^{-|b_j|}}})) + B_{h_j} \otimes I \\ CR_{b_j,2} &= B_{h_j} \otimes R_y(2arccos(\sqrt{e^{b_j e^{-|b_j|}}}) + A_{h_j} \otimes I) \end{aligned} \tag{23}$$

$$A_{h_j} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$B_{h_j} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

The operation is to rotate the ancillary qubit differently according to different input state of h_j. If the input of h_j is $|0\rangle$, $CR_{b_j,1}$ will rotate the ancillary qubit by $arccos(\sqrt{e^{-b_j e^{-|b_j|}}})$ and $CR_{b_j,2}$ will do nothing. If the input of h_j is $|1\rangle$, $CR_{b_j,2}$ will rotate the ancillary qubit by $arccos(\sqrt{e^{b_j e^{-|b_j|}}})$ and $CR_{b_j,1}$ will do nothing.

For w_{ij}, we have a single controlled-rotation gate with one ancillary qubit, denoted as $CR_{w_{ij},1} + CR_{w_{ij},2} + CR_{w_{ij},3} + CR_{w_{ij},4}$.

$$\begin{aligned} CR_{w_{ij},1} &= A_{\sigma_z^i,,h_j} \otimes R_y(2arccos(\sqrt{e^{w_{ij} e^{-|w_{ij}|}}})) + (B_{\sigma_z^i,,h_j} + C_{\sigma_z^i,,h_j} + D_{\sigma_z^i,,h_j}) \otimes I \\ CR_{w_{ij},2} &= B_{\sigma_z^i,,h_j} \otimes R_y(2arccos(\sqrt{e^{-w_{ij} e^{-|w_{ij}|}}})) + (A_{\sigma_z^i,,h_j} + C_{\sigma_z^i,,h_j} + D_{\sigma_z^i,,h_j}) \otimes I \\ CR_{w_{ij},2} &= C_{\sigma_z^i,,h_j} \otimes R_y(2arccos(\sqrt{e^{-w_{ij} e^{-|w_{ij}|}}})) + (A_{\sigma_z^i,,h_j} + B_{\sigma_z^i,,h_j} + D_{\sigma_z^i,,h_j}) \otimes I \\ CR_{w_{ij},2} &= D_{\sigma_z^i,,h_j} \otimes R_y(2arccos(\sqrt{e^{w_{ij} e^{-|w_{ij}|}}})) + (A_{\sigma_z^i,,h_j} + B_{\sigma_z^i,,h_j} + C_{\sigma_z^i,,h_j}) \otimes I \end{aligned} \tag{24}$$

$$A_{\sigma_z^i,,h_j} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$B_{\sigma_z^i,,h_j} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$C_{\sigma_z^i,,h_j} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$D_{\sigma_z^i,,h_j} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The operation is to rotate the ancillary qubit differently according to different input state of σ_z^i and h_j. If the input of σ_z^i and h_j is $|00\rangle$, $CR_{w_{ij},1}$ will rotate the ancillary qubit by $arccos(\sqrt{e^{w_{ij}}e^{-|w_{ij}|}})$ and left controlled-rotation gates will do nothing. If the input of σ_z^i and h_j is $|01\rangle$, $CR_{w_{ij},2}$ will rotate the ancillary qubit by $arccos(\sqrt{e^{-w_{ij}}e^{-|w_{ij}|}})$ and left controlled-rotation gates will do nothing. If the input of σ_z^i and h_j is $|10\rangle$, $CR_{w_{ij},3}$ will rotate the ancillary qubit by $arccos(\sqrt{e^{-w_{ij}}e^{-|w_{ij}|}})$ and left controlled-rotation gates will do nothing. If the input of σ_z^i and h_j is $|11\rangle$, $CR_{w_{ij},4}$ will rotate the ancillary qubit by $arccos(\sqrt{e^{w_{ij}}e^{-|w_{ij}|}})$ and left controlled-rotation gates will do nothing.

Finally, we have the states as $\frac{1}{\sqrt{2^{m+n}}}\sum_{\sigma^z,h}|\sigma^z\rangle|h\rangle(\sqrt{P(\sigma^z,h)/maxP}|1\rangle + \sqrt{1-P(\sigma^z,h)/maxP}|0\rangle)$, where $maxP = e^{m\times max|a_i|+n\times max|b_j|+mn\times max|w_{ij}|}$, where the ancillary qubit is the qubit used to store the sum of probabilities. The measurement step begins with measuring the last ancillary qubits, if it is $|1\rangle$, we can then measure $|reg_1\rangle$ which has the distribution $P(\sigma^z)$. The qubit requirement is $O(4mn)$, and the complexity is $O(4mn)$.

4 Increase Amplitude

Our results for both algorithms are:

$$\frac{1}{\sqrt{2^{m+n}}}\sum_{\sigma^z,h}|\sigma_i^z\rangle|h_j\rangle(p_1|\psi_1\rangle|0\rangle + p_2|\psi_2\rangle|1\rangle), \tag{25}$$

where $|\psi_1\rangle$ $|\psi_2\rangle$ are ancillary qubits states and perpendicular to each other; and p_1 and p_2 are probabilities.

The Lemma from [40] states:

Let U and V be unitary matrices on $\mu + n$ qubits and n qubits, respectively, and let $\theta \in (0, \pi/2)$. Suppose that for any n-qubit state $|\psi\rangle$:

$$U|0^\mu\rangle|\psi\rangle = sin(\theta)|0^\mu\rangle V|\psi\rangle + cos(\theta)|\phi^\perp\rangle \tag{26}$$

where $|\phi^\perp\rangle$ is an $(\mu+n)$-qubit state that depends on $|\psi\rangle$ and satisfies $\prod|\phi^\perp\rangle = 0$, where $\prod = |0^\mu\rangle\langle 0^\mu| \otimes \mathbf{I}$. Let$R = 2\prod$ - $\mathbf{I}$ and $S = -URU^\dagger R$. Then for any $l \in Z$:

$$S^l U|0^\mu\rangle|\psi\rangle = sin((2l+1)\theta)|0^\mu\rangle|V\psi\rangle + cos((2L+1))|\phi^\perp\rangle \tag{27}$$

5 Electronic Structure Hamiltonian

Here we present the details of transforming original electronic structure Hamiltonian to a Hamiltonian consisting entirely of Pauli matrices. We use Hydrogen molecule, H_2, as an example.

We treat the Hydrogen molecule in a minimal STO-6G basis. By considering the spin functions, the four molecular spin orbitals in H_2 are:

$$|\chi_1\rangle = |\Psi_g\rangle |\alpha\rangle = \frac{|\Psi_{1s}\rangle_1 + |\Psi_{1s}\rangle_2}{\sqrt{2(1+S)}} |\alpha\rangle \tag{28}$$

$$|\chi_2\rangle = |\Psi_g\rangle |\beta\rangle = \frac{|\Psi_{1s}\rangle_1 + |\Psi_{1s}\rangle_2}{\sqrt{2(1+S)}} |\beta\rangle \tag{29}$$

$$|\chi_3\rangle = |\Psi_u\rangle |\alpha\rangle = \frac{|\Psi_{1s}\rangle_1 - |\Psi_{1s}\rangle_2}{\sqrt{2(1-S)}} |\alpha\rangle \tag{30}$$

$$|\chi_4\rangle = |\Psi_u\rangle |\beta\rangle = \frac{|\Psi_{1s}\rangle_1 - |\Psi_{1s}\rangle_2}{\sqrt{2(1-S)}} |\beta\rangle\,, \tag{31}$$

where $|\Psi_{1s}\rangle_1$ and $|\Psi_{1s}\rangle_2$ are the spatial-function for the two atoms respectively, $|\alpha\rangle$, $|\beta\rangle$ are spin up and spin down and $S = {}_1\langle\Psi_{1s}|\Psi_{1s}\rangle_2$ is the overlap integral. The one and two-electron integrals are giving by

$$h_{ij} = \int d\vec{r} \chi_i^*(\vec{r})(-\frac{1}{2}\nabla - \frac{Z}{r})\chi_j(\vec{r}) \tag{32}$$

$$h_{ijkl} = \int d\vec{r_1} d\vec{r_2} \chi_i^*(\vec{r_1})\chi_j^*(\vec{r_2})\frac{1}{r_{12}}\chi_k(\vec{r_2})\chi_l(\vec{r_1}) \tag{33}$$

Thus, we can write the second-quantization Hamiltonian of H_2:

$$\begin{aligned} H_{H_2} = {} & h_{00}a_0^\dagger a_0 + h_{11}a_1^\dagger a_1 + h_{22}a_2^\dagger a_2 + h_{33}a_3^\dagger a_3 + h_{0110}a_0^\dagger a_1^\dagger a_1 a_0 + h_{2332}a_2^\dagger a_3^\dagger a_3 a_2 + h_{0330}a_0^\dagger a_3^\dagger a_3 a_0 \\ & + h_{1221}a_1^\dagger a_2^\dagger a_2 a_1 + (h_{0220} - h_{0202})a_0^\dagger a_2^\dagger a_2 a_0 + (h_{1331} - h_{1313})a_1^\dagger a_3^\dagger a_3 a_1 \\ & + h_{0132}(a_0^\dagger a_1^\dagger a_3 a_2 + a_2^\dagger a_3^\dagger a_1 a_0) + h_{0312}(a_0^\dagger a_3^\dagger a_1 a_2 + a_2^\dagger a_1^\dagger a_3 a_0) \end{aligned} \tag{34}$$

By using Bravyi-Kitaev transformation [41], we have:

$$\begin{aligned} &a_0^\dagger = \frac{1}{2}\sigma_x^3\sigma_x^1(\sigma_x^0 - i\sigma_y^0) \quad a_0 = \frac{1}{2}\sigma_x^3\sigma_x^1(\sigma_x^0 + i\sigma_y^0) \quad a_1^\dagger = \frac{1}{2}(\sigma_x^3\sigma_x^1\sigma_z^0 - i\sigma_x^3\sigma_y^1) \quad a_1 = \frac{1}{2}(\sigma_x^3\sigma_x^1\sigma_z^0 + i\sigma_x^3\sigma_y^1) \\ &a_2^\dagger = \frac{1}{2}\sigma_x^3(\sigma_x^2 - i\sigma_y^2)\sigma_z^1 \quad a_2 = \frac{1}{2}\sigma_x^3(\sigma_x^2 + i\sigma_y^2)\sigma_z^1 \quad a_3^\dagger = \frac{1}{2}(\sigma_x^3\sigma_z^2\sigma_z^1 - i\sigma_y^3) \quad a_3 = \frac{1}{2}(\sigma_x^3\sigma_z^2\sigma_z^1 + i\sigma_y^3). \end{aligned} \tag{35}$$

Thus, the Hamiltonian of H_2 takes the following form:

$$\begin{aligned} H_{H_2} = {} & f_0\mathbf{1} + f_1\sigma_z^0 + f_2\sigma_z^1 + f_3\sigma_z^2 + f_1\sigma_z^0\sigma_z^1 + f_4\sigma_z^0\sigma_z^2 + f_5\sigma_z^1\sigma_z^3 + f_6\sigma_x^0\sigma_z^1\sigma_x^2 + f_6\sigma_y^0\sigma_z^1\sigma_y^2 \\ & + f_7\sigma_z^0\sigma_z^1\sigma_z^2 + f_4\sigma_z^0\sigma_z^2\sigma_z^3 + f_3\sigma_z^1\sigma_z^2\sigma_z^3 + f_6\sigma_x^0\sigma_z^1\sigma_x^2\sigma_z^3 + f_6\sigma_y^0\sigma_z^1\sigma_y^2\sigma_z^3 + f_7\sigma_z^0\sigma_z^1\sigma_z^2\sigma_z^3. \end{aligned} \tag{36}$$

We can utilize the symmetry that qubits 1 and 3 never flip to reduce the Hamiltonian to the following form which just acts on only two qubits:

$$H_{H_2} = g_0\mathbf{1} + g_1\sigma_z^0 + g_2\sigma_z^1 + g_3\sigma_z^0\sigma_z^1 + g_4\sigma_x^0\sigma_x^1 + g_4\sigma_y^0\sigma_y^1 = g_0\mathbf{1} + H_0 \tag{37}$$

$$g_0 = f_0 \; g_1 = 2f_1 \; g_2 = 2f_3 \; g_3 = 2(f_4 + f_7) \; g_4 = 2f_6 \tag{38}$$

$$\begin{aligned}
g_0 &= 1.0h_{00} + 0.5h_{0000} - 0.5h_{0022} + 1.0h_{0220} + 1.0h_{22} + 0.5h_{2222} + 1.0/R \\
g_1 &= -1.0h_{00} - 0.5h_{0000} + 0.5h_{0022} - 1.0h_{0220} \\
g_2 &= 0.5h_{0022} - 1.0h_{0220} - 1.0h_{22} - 0.5h_{2222} \\
g_3 &= -1.0h_{00} - 0.5h_{0000} + 0.5h_{0022} - 1.0h_{0220} \\
g_4 &= 0.5h_{0022}
\end{aligned} \tag{39}$$

Where $\{g_i\}$ depends on the fixed bond length of the molecule.

Similar to H_2 and other molecules, next we treat LiH molecule with 4-electrons in a minimal basis STO-6G and use the Jordan-Wigner transformation. Using the technique defined above [42] we can reduce the locality to a Hamiltonian with 558 terms on 8 qubits. And same as H_2O, we also took the technique in [43] to reduce the number of qubits, finally we have 444 terms on 8 qubits [44].