Restricted Boltzmann Machines for Quantum State Tomography

Andi Gu[1] Gerwin Koolstra[1] Irfan Siddiqi[1]

[1]Quantum Nanoelectronics Laboratory, University of California, Berkeley

Classical Machine Learning and Quantum Computing

Quantum computing has shown strong promise in simulating interesting and useful physical systems, especially in characterizing ground state energies. Methods towards this end, such as the variational quantum eigensolver (VQE) and quantum imaginary time evolution (QITE), have significant theoretical speedups compared to classical algorithms. However, one of the greatest challenges to practical quantum simulations on noisy immediate-scale quantum (NISQ) computers is the measurement overhead to achieve accurate estimates of relevant observables.

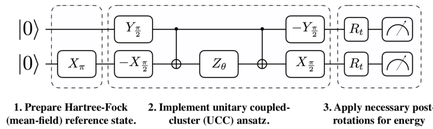

Figure 1. Circuit schematic for a 2-qubit VQE [1]

We aim to address this by employing classical machine learning techniques. We describe one approach to this problem using restricted Boltzmann machines, which have been shown to be promising tools for improving measurement precision [2].

Problem Formulation

Using quantum chemistry methods, we first calculate the Hamiltonian H for some molecule and write as a linear combination of Pauli matrices:

$$H = \sum_{k=1}^{K} c_k P_k, \quad P_k \in \{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}^{\otimes N}$$

With some circuit that parameterized a trial wavefunction $\Psi \in \mathbb{C}^n$, our goal is to **estimate the expectation of the Hamiltonian** $\langle\Psi|H|\Psi\rangle$. We take M projective measurements for each Pauli term in the Hamiltonian, giving us a collection of measurements $D_k = \{v_1^k, \ldots, v_M^k\}$ for each Pauli term. Typical methods estimate $\langle P_k\rangle$ by averaging:

$$\tilde{P}_k = \frac{1}{M}\sum_{m=1}^{M} \left\langle v_m^k \middle| P_k \middle| v_m^k \right\rangle$$

and then summing these estimates appropriately to estimate $\langle H\rangle$: $\tilde{H} \equiv \sum_{k=1}^{K} c_k \tilde{P}_k$. Due to quantum shot noise, $\tilde{P}_k$ (hence $\tilde{H}$) is a random variable with variance that decreases with $M^{-1/2}$. Using these averaging methods we require $M_{<\mathcal{E}} \sim \mathcal{O}(10^7)$ measurements to reach chemical accuracy on LiH – this is very costly. Our goal, then, is to **reduce the measurement requirement** $M_{<\mathcal{E}}$.

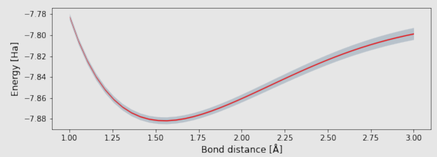

Figure 2. Ground state energies for LiH recovered via VQE simulations. The width of the red line indicates chemical accuracy $\mathcal{E} = 1.6 \times 10^{-3}$Ha and the blue is a 68% confidence interval of $\tilde{H}$ with $M = 1000$ measurements.

Applying the Restricted Boltzmann Machine

RBMs are networks designed to encode probability distributions over binary strings $P : \{0,1\}^n \to \mathbb{R}$. They do this by explicitly parameterizing a joint distribution over the 'visible' and 'hidden' units

$$P(v, h; a, b, W) = \frac{\exp\left(-a^T v - b^T h - v^T W h\right)}{Z}$$

where a, b, W are network parameters and Z is a normalizing factor. The distribution over binary strings can be found simply by marginalizing $P(v) = \sum_{h\in\{0,1\}^n} P(v,h)$.

Generalizing Complex Weights

We introduce a modification of the RBM that allows the weights a, b, W to be complex – making it a complex RBM (cRBM). The cRBM now encodes the wavefunction: $f : \mathbb{C}^n \to \mathbb{C}; v \mapsto \langle v|\Psi\rangle$. As usual, $P(v) = \|\langle v|\Psi\rangle\|^2 = \|f(v)\|^2$. We apply the cRBM to do **partial quantum state tomography** with our measurement dataset $\mathcal{D} = \{D_k \mid k = 1, \ldots, K\}$. We seek to find the network parameters that were most likely to have produced the measurements dataset $\mathcal{D}$. We therefore define the loss as the negative log likelihood:

$$\mathcal{L}(a, b, W) = -\frac{1}{MK}\sum_{k=1}^{K}\sum_{m=1}^{M} \log(P(v; a, b, W))$$

Using descent optimizers, we find $\text{argmin}_{a,b,W}\mathcal{L}$, and extract the encoded wavefunction Ψ' from the optimized cRBM, from which it is simple to directly calculate $\left\langle\Psi'\middle|H\middle|\Psi'\right\rangle$.

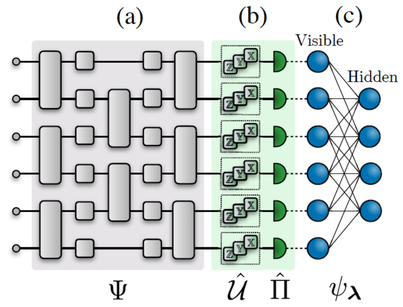

Figure 3. Measurement and recovery process: (a) is a series of unitary gates that parameterizes a ground truth state Ψ, (b) is a set of projective measurements $\hat{U}$. The RBM (c) takes the results of the measurements as input to the visible layer, and together with the hidden layer, parameterizes the state $\psi_{\boldsymbol{\lambda}}$. Adapted from [2].

cRBM Variance Reduction for LiH

The following are the results for the LiH molecule. We calculate the Hamiltonian H for LiH (which requires $N = 4$ qubits). We then train 100 RBMs on different randomly simulated datasets of fixed size MK with a ground truth state Ψ that represents the ground state of H.

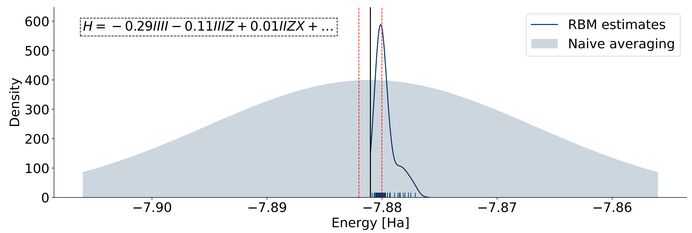

Figure 4. Distribution of recovered energy for LiH via naive averaging and RBM, with $M = 1000$. The solid black line marks the truth ground state energy and red lines mark threshold for chemical accuracy.

We observe a variance reduction of up to a factor of 100 compared to widely used 'naive averaging' methods. However, there is a nonzero bias in the recovered distribution due to the variational principle; recovered states Ψ' will always have a higher energy than the true ground state.

cRBM Efficiency for Various Molecules

This reduced variance far outweighs any increased bias – therefore, we find the RBM enables a significant reduction in measurement requirements $M_{<\mathcal{E}}$. We list the results for three molecules (H_2, LiH, BeH_2) with different numbers of qubits ($N = 2, 4, 6$ respectively).

Molecule	# of Pauli terms K	Variance reduction	Measurement reduction factor
H_2	5	30	10
LiH	100	150	60
BeH_2	95	50	20

Table 1. RBM improvements for three molecules

Conclusions

We find that the restricted Boltzmann machine is a very useful technique for quantum simulation. It has been demonstrated to be

- effective in reducing measurement requirements by **at least an order of magnitude**, and up to two orders of magnitude depending on the ground truth state Ψ and relative magnitude of coefficients $|c_k|^2$ in H
- **increasingly powerful on more complex Hamiltonians** (i.e. larger number of Pauli terms) in part due to increasing numbers of commuting Pauli groups in H
- useful for circumventing the prohibitively high memory costs of classical tomography techniques which require $\mathcal{O}(2^N)$ memory (the RBM requires $\mathcal{O}(N^2)$ memory)

Further Work

This work is a promising indication of the efficacy of classical machine learning techniques in enhancing quantum computation. We propose a few areas of possible further investigation.

- The theoretical demonstration of RBMs efficacy remains to be tested on experimental data. We aim to verify the theoretical results by applying the RBM to a 2 qubit transverse field Ising model simulated on QNL's quantum computer.
- Extend the RBM to encode full density matrices by adding a hidden layer to implicitly purify the state
- Integrate RBMs with groupings of complete sets of commuting operators to further reduce variance while decreasing number of necessary different measurement configurations

References

[1] Yudong Cao et al.
Quantum chemistry in the age of quantum computing.
Chemical reviews, 2019.

[2] Giacomo Torlai, Guglielmo Mazzola, Giuseppe Carleo, and Antonio Mezzacapo.
Precise measurement of quantum observables with neural-network estimators.
Physical Review Research, 2(2), Jun 2020.