Learning Many-Body Wavefunction with Neural Network

Siyao Jia and Luning Zhao

August 28, 2018

1 Introduction

The electronic wave function Ψ is a fundamental object in quantum physics and chemistry that on its own, contains all the electronic information the systems have, which further decides the chemical properties of the system, such as an atom, a molecule or a solid. In principle, exponential amount of information is required to solve for the exact, many-body electronic wave function. However, in many cases only a small fraction of these exponential amount of information is needed to grasp the true physics due to less quantum entanglement effects in some systems. Therefore, dimensionality reduction is always needed to avoid explicit dealing with the exponentially large Hilbert space in which the wave function lives.

Among many different numerical approximations to the exact wave function. Restricted Hartree-Fock (RHF)[5] approximation is the simplest and most famous one. Within the formalism of RHF, the wave function is represented by a determinant of a specific matrix, in which the ijth entry represent the ith particle in jth orbital. The form of determinant is needed to enforce correct statistics of fermions. Since when two fermions are in the same orbital, indicting two columns or two rows are the same in the matrix, then the wave function should go to zero.

Due to the correct treatment of fermionic statistics and correlation between same-spin electrons, RHF has been proven to be a very sensible approximation in a lot of simple systems in which the correlation between electrons is not too large. For example, for hydrogen molecule (H_2), RHF is more than 95% of the exact wave function.[2]

In broader sense, finding good approximation to the many-body wave function resides in the realm of dimensionality reduction and feature extraction. This is essentially the same problem people are facing in data science. To tackle these problems, artificial neural network is the most commonly used technique nowadays.[3] The ANN has been shown to perform very well in image and speech recognition.

In general, ANNs are universal function approximators that has the ability to represent any function. But whether it is capable of representing many-body

wave function is still unexplored. In this study, we will use artificial neural net works to learn a RHF wave function for two simple molecules, H_2 and LiH. We will also try to decide the optimal values for hyperparameters of our model. In the end, with the well-trained neural network, we will use it to predict one of the most important properties in chemistry, the density of electrons

2 Restricted Hartree-Fock

Suppose we have M_α up-spin and M_β down-spin electrons in the molecule. In this study we only consider the case of the equal number of up-spin and down-spin electrons. We also have N orbitals that electrons can reside in. The restricted Hartree-Fock wave function is written as:

$$\Psi\left(r_1, \ldots r_M\right) = det(M) \tag{1}$$

in which $M_{ij} = \phi_i(r_j)$ and "det" stands for determinant. $\phi(r)$ is some functions that are obtained by minimizing the energy of this wave function.

3 Computational Details

The wave function samples for H_2 and LiH using restricted Hartree-Fock are obtained using QMCPACK code[1]. We use multi-layer perceptron (MLP) in Sklearn package[4] to fit our RHF data. For each molecule, 1.5 million samples are used to train the neural network and 0.5 million samples are then used to test the trained neural network. The overall training and testing error are plotted by averaging the training and testing error of 10 set of wave function samples. The bond distance used for H_2 and LiH is 1.398 and 4.535 atomic unit respectively.

4 Results and Discussions

In Figure 1 we plot the training and testing error of our neural network as a function of different layers and different number of neurons per layer for H_2. As we see in the plot, when the number of layers and the number of neurons per layer increases, both the training and testing error goes down. This is due to the increase of the model representability as the number of parameters increases. In all cases we test, we don't observe an increase of testing error in accompany with decreasing of training error, indicting that our neural network is correctly learning the true pattern of the wave function without hitting the overfitting problem. We therefore conclude that for H_2, our simple neural network is capable of learning the correct wave function.

Similarly, We show the training and testing error for LiH in Figure 2. As we see the error in this system in a lot more larger than that of H_2, especially when the number of neurons per layer are limited. Although LiH is only two

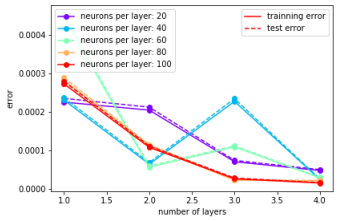

Figure 1: Training and testing error for H_2 with respect to number of layer and number of neurons per layer.

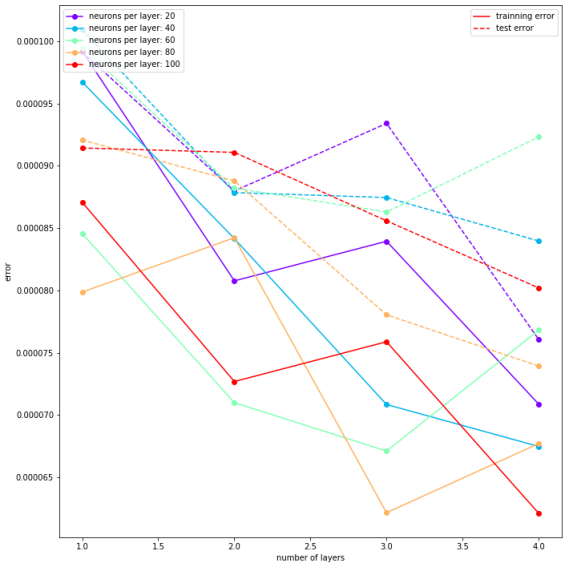

Figure 2: Training and testing error for LiH with respect to number of layer and number of neurons per layer.

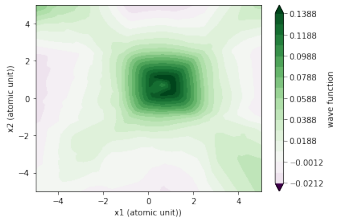

Figure 3: Predicted electronic probability density for H_2 using trained neural network.

times larger H_2 in terms of the number of electrons, the error has increased by an order of magnitude with the same number of layers and number of neurons per layer. Unlike H_2, in which both the training and testing error decrease monotonically with number of model parameters, in LiH we find that even though the training error is decreasing, the testing error grows as the number of model parameters. This behavior suggests that the model is overfitted when large number of parameters are used in our neural network.

Based on the testing error, We choose 3-layer and 100-neuron-per-layer as optimal hyperparameter since the testing error is at its minimal when such an architecture is applied. With neural network trained using such parameters for H_2. We compute $|\Psi|^2$, the probability density of electrons, on a 2-dimensional grid. The x and y axis of the 2D grid represents the position of first and second electron. We only plot $|\Psi|^2$ value along x-direction of electrons position since the hydrogen atoms are aligned on this axis.

The predicted electronic density is shown in Figure 3. As we can see in the plot, the electron density is peaked in the range where the positions of both electrons are between 0 and 1. This is as expected since first we know that these two electrons form chemical bond in H_2. In order to do so, they need to be in the same location to allow enough overlap of their wave function. As a result, we see the electron density is peaked at where the positions of two electrons are similar.

Secondly, since the two hydrogen atoms are located one at origin and another one around 1.5 atomic unit, we do see the electron density peaks in such area. Therefore our neural network successfully learn the physical feature that electrons need to be attracted by nuclei.

5 Conclusion

In this study, we use multi-layer perceptron neural network to learn one of the simplest approximation to many-electron wave function, the restricted Hartree-Fock wave function for H_2 and LiH molecule. We have tested the training and testing error of neural network as a function of number of layers and number of neurons per layer. We found that for H_2 both the training and testing error are

extremely small and no overfitting problem is observed as our model gets more and more complicated. However, for LiH we found the training and testing error are significantly larger and the overfitting problem starts to emerge.

With the well trained neural network for H_2, we use it to predict the density of electrons on the plane where two hydrogen atoms reside in. We found that the neural network successfully predicts the formation of chemical bonds by peaking the electron density at the location where two electrons are near to each other. Furthermore, it also predict that the electrons should be close to nuclei due to the coulomb attraction correctly.

However, although it looks promising in H_2, the large error we saw in LiH is concerning. It seems to imply that as the number of electrons increase, neural network will be less useful to model wave function, which limits its applicability in large systems. The reason for may be that when the number of electrons get larger. the wave function is highly oscillatory quantity due to the increase of complicatedness in nodal surface. Although neural network is known to be good at learning smooth functions, it is less useful in highly oscillatory functions. Therefore in the future, instead of using neural network to directly model wave function, one could use it as an addition to improve the current numerical approximation of the wave function. Research in these exciting directions is underway.

References

[1] Jeongnim Kim etc. Qmcpack : an open source ab initio quantum monte carlo package for the electronic structure of atoms, molecules and solids. *Journal of Physics: Condensed Matter*, 30(19):195901, 2018.

[2] T. Helgaker, P. Jøgensen, and J. Olsen. *Molecular Electronic Structure Theory*. John Wiley and Sons, Ltd, West Sussex, England, 2000.

[3] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.

[4] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[5] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, Mineola, N.Y., 1996.