Interatomic force from neural network based variational quantum Monte Carlo

Yubing Qian[1,2,†], Weizhong Fu[1,2], Weiluo Ren[2], and Ji Chen[1]

1. School of Physics, Peking University, Beijing 100871, Peoples Republic of China
2. ByteDance Research, Zhonghang Plaza, No. 43, North 3rd Ring West Road, Haidian District, Beijing.

†phyqyb@pku.edu.cn

Introduction

Neural networks have been applied to ab initio calculations with VMC, such as FermiNet, enabling accurate energy calculation of materials. In addition to energy, interatomic force is an important quantity, but extra work is required due to the infinite variance problem. Several approaches are proposed in traditional QMC but are not guaranteed to work on neural networks. Here, we present the interatomic force obtained from neural networks with the modified traditional estimators.

Force Estimators

Bare Estimator

Direct evaluation of Coulomb force (F_{Bare}) from the Hellmann-Feynman theorem has infinite variance, making QMC evaluation impossible.

We need to find more efficient estimators with lower variance.

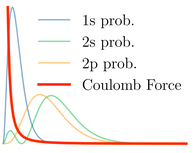

Fig. 1: Infinite variance contributed by the inner shell electron

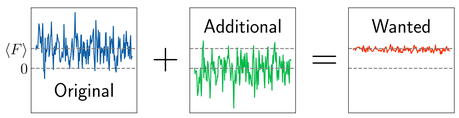

Fig. 2: Find estimators to reduce variance

Assaraf-Caffarel estimator

$$\widetilde{\mathbf{F}}_{A,\text{AC-ZVZB}} = \underbrace{\mathbf{F}_{\text{Bare}} - \frac{(H-E_L)\,\widetilde{\psi}_A}{\psi_T}}_{\widetilde{\mathbf{F}}_{A,\text{AC-ZV}}} + \underbrace{2\,(E_v-E_L)\,\frac{\widetilde{\psi}_A}{\psi_T}}_{\text{AC-Pulay term}}$$

(Variance lowered; Bias fixed)

It can be further simplified with an approximate derivative of the wavefunction:

$$\widetilde{\psi}_{A,\text{min}}(\mathbf{x}) = \mathbf{Q}_A\psi_T = Z_A \sum_{i=1}^{n_{\text{elec}}} \frac{\mathbf{r}_i - \mathbf{R}}{|\mathbf{r}_i - \mathbf{R}|}\psi_T$$

Space Warp Coordinate Transformation Estimator

Or we can just take the derivative of the local energy:

$$\widetilde{\mathbf{F}}_{A,\text{No-SWCT}} = -\frac{\partial}{\partial \mathbf{R}_A}E_L + 2(E_v-E_L)\frac{\partial}{\partial \mathbf{R}_A}\log\left(|\psi_T|\right)$$

...and space warp electrons with nuclei (Umrigar *et. al.*):

$$\bar{\mathbf{r}}_i = \mathbf{r}_i + \Delta\mathbf{R}_A\omega_A(\mathbf{r}_i), \qquad \omega_A(\mathbf{r}) = \frac{f(|\mathbf{r}-\mathbf{R}_A|)}{\sum_{j=1}^{n_{\text{atom}}} f(|\mathbf{r}-\mathbf{R}_j|)}$$

$$\widetilde{\mathbf{F}}_{A,\text{SWCT}} = \underbrace{-\frac{\mathrm{d}}{\mathrm{d}\mathbf{R}_A}E_L}_{\text{SWCT-HFM term}} + \underbrace{2(E_v-E_L)\frac{\mathrm{d}}{\mathrm{d}\mathbf{R}_A}\log\left(|J^{1/2}\psi_T|\right)}_{\text{SWCT-Pulay term}}$$

Results

The estimators are implemented based on the JAX version of FermiNet, and we calculated forces of H_2, Li_2, and N_2 along their potential energy curves.

Estimators work well in neural network based VMC

The forces from regularized estimators are all in **excellent agreement** with the reference and have small statistical errors, while the bare estimator has much larger systematic and statistical errors.

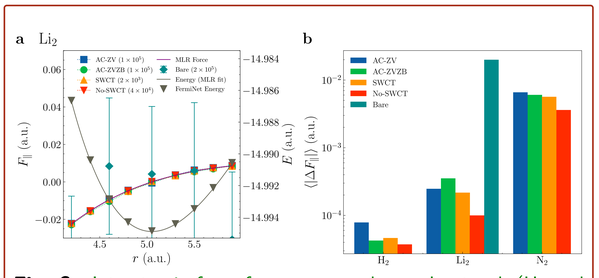

Fig. 3: Interatomic force from converged neural network (H_2 and N_2 not shown)

Force results are improved with better neural network quality

Different neural network checkpoints at different training steps are used as neural networks with different qualities. A better neural network gives better wavefunction. The error and variance of the force estimators can also be **improved with a better neural network**. Also, SWCT has lower variance and error among them.

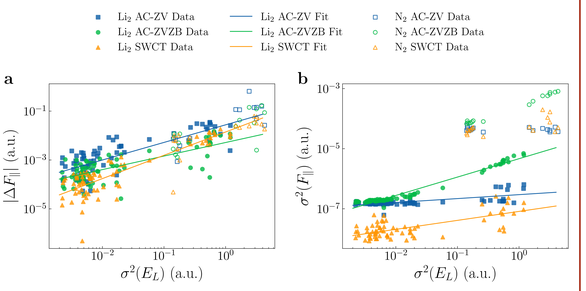

Fig. 4: Dependence on the quality of neural network

Scaling for estimators

Relative cost is the ratio between force estimator calculation time and training time per step. The results are consistent with the analysis of the asymptotic complexity.

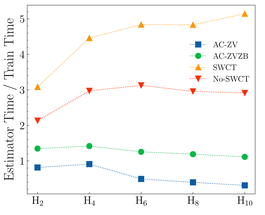

Fig. 5: Relative cost of estimators

Others

The contribution of different terms, the time consumption for each term, and the (crazy, $O(n^4)$) GPU memory consumption scaling of the SWCT estimator are also investigated.

Conclusions

- Wavefunction improved by the neural network ansatz also benefits the force, indicating a promising future for further application.
- Provide guidelines for future applications
  - SWCT estimator is the most accurate and achieves the best accuracy-cost balance when the system is not too large.
  - AC-ZV estimator is more efficient for large systems with low accuracy requirements.

Reference

Y. Qian, W. Fu, W. Ren, and J. Chen, Interatomic Force from Neural Network Based Variational Quantum Monte Carlo, The Journal of Chemical Physics 157, 164104 (2022).