Machine Learning Potential Function Generation for *ab initio* Electronic Structure Calculations

Duong Hoang*
Paul Rulis
dtht2d@umsystem.edu
rulisp@umkc.edu
University of Missouri - Kansas City
Kansas City, Missouri, USA

ABSTRACT

First-principles electronic structure calculations based on density functional theory (DFT) are well-known to have a high computational cost that scales algorithmically as $O(N^3)$, where N is the number of electrons. Reducing that cost is a key goal of the computational materials physics community and machine learning (ML) is viewed as an essential tool for that task. However, ML model training requires an appropriate match between the input descriptors and the target property as well as copious quantities of training data.

Therefore, we present a computer program that is designed to automate the generation of local atomic environment descriptors for single element systems that may be used for training neural networks to predict electronic potential function coefficients, $\{A_i\}$, which are used within the DFT based orthogonalized linear combination of atomic orbitals (OLCAO) method [2].

In our approach, the total electronic potential function of a periodic crystal, $V_{cry}(\vec{r})$, is expressed as a sum of localized atom-centered Gaussian functions.

$$V_{cry}(\vec{r}) = \sum_B V_B(\vec{r} - \vec{t_B}), \quad V_B(\vec{r}) = \sum_{i=1}^{N_A} A_i e^{-\alpha_i r^2} \quad (1)$$

Each Gaussian function, i, in the set of all Gaussian functions has a fixed α_i coefficient. The set of $\{A_i\}$ coefficients are updated in each cycle of the self-consistent field (SCF) iterations in accordance with the charge density $\rho(\vec{r})$ that was computed in the previous SCF step. However, if the choice of coefficients $\{A_i\}$ can be accurately predicted for a given system, then the SCF process can be skipped entirely, satisfying an important requirement of our goal to reduce the computational cost. The prediction method uses suitable neural networks (NNs) where the input values are a set of local atomic environment descriptors and the output values are the $\{A_i\}$ coefficients for a targeted system. The descriptors we opted to use are the bispectrum components but other additional descriptors may be incorporated. Bispectrum components are geometric calculations that smoothly capture subtle variations in the local atomic

environment and that are invariant under translation, rotation, and permutation of neighborhood atoms. The bispectrum components can also easily incorporate different types and numbers of elements, and they have been used by others for a similar purpose [3, 4]. Those requirements are difficult to achieve using other methods such as a list of bond angles and bond lengths toward nearest neighbor atoms while maintaining a fixed number of NN input features.

$$B(j_1, j_2, j) = \sum_{m_1, m_1' = -j_1}^{j_1} \sum_{m_2, m_2' = -j_2}^{j_2} \sum_{m, m' = -j}^{j} \left(u_{mm'}^j \right)^* \\ H_{j_1 m_1 m_1', j_2 m_2 m_2'}^{jmm'} u_{m_1 m_1'}^{j_1} u_{m_2 m_2'}^{j_2} \quad (2)$$

where $u_{mm'}^j$, $u_{m_1 m_1'}^{j_1}$, $u_{m_2 m_2'}^{j_2}$ are expansion coefficients,

$$H_{j_1 m_1 m_1', j_2 m_2 m_2'}^{jmm'} \equiv C_{j_1 m_1 j_2 m_2}^{jm} C_{j_1 m_1' j_2 m_2'}^{jm'}$$

is the coupling coefficient for four-dimensional spherical harmonics, analogous to the Clebsch-Gordan coefficients for rotations in three dimensional space.

One challenge in this research is defining a suitable cut-off radius for evaluation of the bispectrum component to avoid neglecting the interaction between a targeted atom and its neighbors. The cut-off radius is weighted as a function of the elements involved to accommodate different types of bonding (e.g., ionic, covalent, metallic). Additionally, for properly defining and training a neural network (see below), it is vital that we provide a clear correlation between the physical (geometric) features of the bispectrum components and the electronic features that may simultaneously be present to avoid too much redundancy in the input data. This lack of understanding can limit the development of methods to predict the electronic structure properties based on the bispectrum components, underscoring the need for further research in this area.

A supervised training framework for a proposed neural network is demonstrated using a data set of pure Si models that includes amorphous Si, crystalline Si, Si with a passive defect, and Si with self-interstitials. Other models will be implemented to compare efficiency. For each model, the input/target output training pairs consist of local environment descriptors - bispectrum components (input) that encode the structure of neighboring atoms relative to the central atom i at a specific point in real space, along with the converged potential functions obtained by the SCF process (target output). The data set must be partitioned into training, test, and validation sets for use in subsequent iterations of training and

validation to evaluate and optimize the model's performance during the training process.

In OLCAO, the total electronic potential function of a crystal is expressed as a sum of atom-centered potential functions. Each atom-centered potential function is represented as a sum of Gaussian functions. However, it is vital to recognize that although the potential function is an assembly of site-centered functions it cannot be said that the potential function from a given site is the potential function "of" the atom at that site. Rather, the potential function at a given site is determined by the influence of all nearby atoms. Therefore, it is intuitive to seek a ML model that follows a similar structure. In this case, it is important to find a way for the input data structure to incorporate that feature of the potential function, which consists of a mixture of influences derived from the neighboring atoms. Each component of this mixture represents a cluster or subpopulation within the local region. To capture this structure, we propose a neural network framework based on Mixture Density Network (MDN) [1] for the training process. This approach involves encoding the local, medium-range, and long-range (global) influences for each atom. In many cases, electron interactions are considered 'short-sighted,' meaning that they are mainly affected by nearby atoms only. However, our proposed method overcomes this limitation and effectively addresses novel long-range electronic structure properties such as those found in metallic or certain magnetic materials.

Results regarding the optimization of the run time for calculating the bispectrum component is discussed, including a comparison with key function program code that uses third-party libraries such as SymPy. A computer program is developed to automatically generate bispectrum components for a single-element system in a periodic unit cell. We investigated the symmetric properties of the bispectrum components, which align with the proof established in [4]. However, further development and testing of the program are necessary before it can be applied to multiple-element systems.

Overall, this research contributes to the ongoing effort to develop new and improved neural network frameworks for predicting the electronic structure properties of materials with desirable features. When combined with other unique aspects of the OLCAO method it is expected that this approach will enable us to overcome the $O(N^3)$ algorithmic complexity scaling problem and thereby address multi-scale physics problems that require both direct access to the electronic wave function and a large number of atoms to realistically model.

ACM Reference Format:
Duong Hoang and Paul Rulis. 2023. Machine Learning Potential Function Generation for *ab initio* Electronic Structure Calculations. In *Proceedings of Computing for the Common Good (PEARC'23)*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3569951.3597581

REFERENCES

[1] Christopher M. Bishop. 1994. *Mixture Density Networks*. Technical Report. Aston University.

[2] Wai-Yim Ching and Paul Rulis. 2012. *Electronic Structure Methods for Complex Materials: The Orthogonalized Linear Combination of Atomic Orbitals*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199575800.001.0001

[3] Lenz Fiedler, Normand A. Modine, Steve Schmerler, Dayton J. Vogel, Gabriel A. Popoola, Aidan P. Thompson, Sivasankaran Rajamanickam, and Attila Cangi. 2022. Predicting Electronic Structures at Any Length Scale with Machine Learning. http://arxiv.org/abs/2210.11343

[4] Aidan P. Thompson, Laura P. Swiler, Christian R. Trott, Stephen M. Foiles, and Garritt J. Tucker. 2015. A Spectral Analysis Method for Automated Generation of Quantum Accurate Interatomic Potentials. *J. Comput. Phys.* (2015). http://arxiv.org/abs/1409.3880