Original Article

Machine Learning Representation for Atomic Forces and Energies

Pham Tien Lam[1, 2,*], Nguyen Van Duy[1, 2], Nguyen Tien Cuong[3]

[1]Phenikaa Institute for Advanced Study, Phenikaa University, Ha Dong, Hanoi, Vietnam
[2]Faculty of Computer Science, Phenikaa University, Ha Dong, Hanoi, Vietnam
[3]VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam

Abstract: We present machine learning models for fast estimating atomic forces and energies. In our method, the total energy of a system is approximated as the summation of atomic energy which is the interaction energy with its surrounding chemical environment within a certain cutoff radius. Atomic energy is decomposed into two-body terms which are expressed as a linear combination of basis functions. For the force exerted on an atom, we employ a linear combination of a set of basis functions for representing pairwise force. We use least-square linear regression regularized by the l2-norm, known as Ridge regression, to estimate model parameters. We demonstrate that we can train linear model to accurately predict atomic forces and energies in comparison to density-functional-theory (DFT) calculations for crystalline and amorphous silicon. The machine learning force model is then applied to calculate the phonon dispersion of crystalline silicon. The result shows reasonable agreement with DFT calculations.

Keywords: Machine Learning, molecular dynamics, force field, materials informatics.

1. Introduction

The computation of the energies, especially the forces of a chemical system, play a central role in the computational design of matter and materials. The energies and forces can be obtained by performing electronic structure calculations such as those based on density functional theory (DFT) [1–3]. Although

_____

*Corresponding author.
 Email address: lam.phamtien@phenikaa-uni.edu.vn

DFT has become a standard method in the computer simulations of matter, it continues to remain computationally expensive. The energies and forces of a large system can also be approximated by using an empirical model with much lower computational cost. Typically, a potential energy surface (PES) is constructed as the sum of simple low-dimensional terms (structural elements) representing covalent bonds, angles, and dihedral angles [4]. These methods are efficient and are applied widely for simulating large biosystems, but they can hardly describe chemical reactions that involve the formation or dissociation of covalent bonds.

In recent years, alternative methods, which learn the PES from a set of materials structures and the corresponding DFT energies, have been developed [5–11]. These methods employ machine learning algorithms to determine the functional relationship between the structure and energy of a material. Normally, to build the machine learning models the total energy of a material is decomposed into the sum of all energies of the constituent atoms determined by interactions among these atoms and with atoms in the surrounding chemical environment within a cutoff radius.

Although atomic forces, which is a vector quantity, can be calculated by obtaining the gradient of the total energy, this approach would hardly be successful when the learned functional relationship between the material structure and energy is complicated, such as deep-learning models. Therefore, the machine learning models for representing the atomic forces are highly desired. In this work, we examine a linear representation for atomic energies on atomic forces. For the atomic forces, we develop linear models for the direct representation of atomic forces in terms of pairwise-force functions.

2. Computational Methods

2.1. Modeling Potential Energy Surface

To model the potential energy surface, we assume that the total energy of a chemical system is the summation of effective atomic energy of the constituent atoms [9,10,12–15]: $E = \sum_i E_i$, where E is the total energy and E_i is the effective energy measuring the contribution of atom i to the total energy. The effective energy of atom i can be represented based on its interactions with neighboring atoms within a cutoff radius, r_c. In this work, we only focus on the two-body interactions. In principle, we can extend the effective atomic energy to higher order.

$$E_i = \sum_j E_{ij} = \sum_j \sum_k c_k b_k(r_{ij}, v_i, v_j) = \sum_k c_k \sum_j b_k(r_{ij}, v_i, v_j) = \sum_k c_k x_k \tag{1}$$

where $x_k = \sum_j b_k(r_{ij}, v_i, v_j)$, $b_k(r_{ij}, v_i, v_j)$ are basis functions, c_k are expansion coefficients, and v_i and v_i are feature vectors encoding the information of atom i and j, respectively. To enhance the nonlinearity representation, we also include the polynomial terms of two-body term up to order p:

$$E_i = \sum_k c_k^{(1)} x_k + \sum_k c_k^{(2)} x_k^2 \dots \sum_k c_k^{(p)} x_k^p \tag{2}$$

In this work, we focus on silicon systems and examine the Gaussian and cosine basis function for representing their potential energy surfaces. Since the atomic information is identical for all silicon atom, we can remove v_i and v_j in Eq. 2. Eq. 2 can be reduced to more simple form: $E_i = \sum_k c_k b(r_{ij})_k$. We use following forms: $b_k(r_{ij}) = e^{\eta_k (r-r_k)^2} f_c(r_{ij})$ and $b_k(r_{ij}) = \cos(\kappa_k r_{ij}) f_c(r_{ij})$ for Gaussian and cosine basis functions, respectively, where η_k, r_k, and κ_k are

parameters defined the basis functions, $f_c(r_{ij})$ is a smooth-cutoff function which ensures that the atomic energy varies smoothly at the cutoff radius:

$$f_c(r_{ij}) = \begin{cases} 0.5\left[cos\left(\dfrac{\pi r_{ij}}{r_c}\right) + 1\right] & if\ r_{ij} < r_c \\ 0 & if\ r_{ij} \geq r_c \end{cases} \tag{3}$$

We determine the linear coefficient, $c_k^{(p)}$ by a linear regression for DFT Energies. We use least-square linear regression regularized by the l2-norm, known as Ridge regression [9], to estimate model parameters.

2.2. Modeling Atomic Force

To model the force on an atom we only consider the interactions of its neighboring atoms within a cutoff radius. The force that acts on an atom is a superposition of all pairwise forces between the central atom and its neighboring atoms:

$$\vec{f_i} = \sum_j \vec{f_{ij}} = \sum_j F_{ij}(r_{ij}, v_i, v_j)\frac{\vec{r_{ij}}}{r_{ij}}f_c(r_{ij}) \tag{4}$$

where $\vec{f_i}$ is the force acting on atom i, $\vec{f_{ij}}$ is the force exerted on atom i by atom j, $F_{ij}(r_{ij}, v_i, v_j)$ is a pairwise-force function, and $f_c(r)$ is the smooth cutoff function. The pairwise-force function which depends the distance between atoms, r_{ij}, and the intrinsic properties of atoms represented by feature vectors, v_i and v_j. The pairwise-force function measures the strength of the force atom j exerts on atom i. Each component of the atomic force can be expressed by following equation:

$$f_{i\alpha} = \sum_j F_{ij}(r_{ij}, v_i, v_j)\frac{r_{i\alpha} - r_{j\alpha}}{r_{ij}}f_c(r_{ij}) \tag{5}$$

where $f_{i\alpha}$ is the α component of $\vec{f_i}$ where $\alpha \in \{x, y, z\}$. We represent the pairwise-function by a linear combination of basis functions.

$$F(r_{ij}, v_i, v_j) = \sum_k c_k b_k(r_{ij}, v_i, v_j) \tag{6}$$

Combining Eq. 2 and Eq. 3, we can express the Cartesian force components as follows:

$$f_{i\alpha} = \sum_k c_k \sum_j b_k(r_{ij}, v_i, v_j)\frac{r_{i\alpha} - r_{j\alpha}}{r_{ij}}f_c(r_{ij}) \tag{7}$$

Therefore, one can determine the linear coefficient, c_k by a linear regression for DFT forces, $f_{i\alpha}$. We also employ Ridge regression in order to estimate model parameters [9] and investigate the Gaussian and cosine basis functions, as described in section 1, for representation of the pairwise-force.

2.3. Data Preparation

We examine our method to silicon system. We used a 2×2×2 Si supercell that includes 64 Si atoms as training data and a 3×3×3 Si supercell with 216 atoms as the test data for the larger system. The systems were relaxed by performing a molecular dynamics (MD) calculations at 2,000 K in the NPT ensemble by the Car–Parrinello method [3]. Then, they were equilibrated at 1,000 K in the NPT ensemble. The pressure was standard ambient atmospheric pressure (1.01325 bar). We collected 3,000 structures in the equilibrated trajectories of the 2×2×2 cell training the data of the 3×3×3 cell as test

data. Additionally, we examined the applicability of our method to a large range of forces by adding uniform random displacement with a magnitude of 0.3Å to the coordinates of each atom in those structures equilibrated at 1,000 K. The energies and forces of the structures were calculated by using the PWSCF code [20, 21]. We employed the PBE exchange-correlation functional [22, 23] to process the exchange-correlation energy, plane wave basis set with the cutoff energy of 40 Ry, and ultrasoft-pseudopotential [24] for approximating the interaction between the valence electrons and core electrons. A 3×3×3 Monkhorst–Pack grid of k-points was employed for the Brillouin-zone integrations of 2×2×2 Si supercell, while a 2×2×2 grid was applied for 3×3×3 supercell.

3. Results and Discussion

3.1. Prediction of Total Energy

We use the cutoff radius of 8 Å and investigate the dependence of the prediction of the potential energy surface on the number of basis functions. For the Gaussian functions, we use first use the functions with the center is the center atom, $r_k = 0.0$. The number of Gaussian basis functions is defined by the parameter, η_{max}. The parameters $\kappa_k \in \{0.05, 0.10, ..., \kappa_{max}\}$ are used for the cosine functions, and the number basis function is also determined by the parameter κ_{max}.

Table 1. The mean absolute error (MAE) in meV/atom for the test set derived from 2×2×2 supercell

Basis Function	#functions	p=1	p=2	p=3
Gaussian	20	8.9	8.5	8.0
	40	8.4	8.3	7.9
	60	8.1	7.9	8.2
Cosine	20	31.6	28.7	31.1
	40	10.4	10.8	10.6
	60	9.8	9.3	10.2

We first examine our model for the potentials energy surface of 2×2×2 system. We divide the data set of 3000 structures into the training set of 2400 structures and the test set of 600 structures. The performance of models derived from Gaussian and cosine basis functions is shown in Table 1. It can be seen that for this system, the polynomial terms, as described in Eq. 3, do not significantly improve the accuracy of the prediction. And the Gaussian basis functions seem to yield better performance in presenting the two-body terms for the potential energy surface of silicon systems.

We also extend the Gaussian basis functions with different centers: $r_k \in \{0, 1, 2, .., 7, 8\}$ and $\eta_k \in \{0.05, 0.10, ..., \eta_{max}\}$. By using η_{max} of 2.0, we can slightly improve the accuracy with MAE of 7.7 meV/atom. Using these basis functions and 3000 2×2×2-structures as training set, we can well reproduce the DFT energy for 3×3×3 system with similar accuracy.

3.2. Prediction of Atomic Forces

We use the same procedure to specify the basis functions for the pairwise forces between silicon atoms. We also first study the 2×2×2 systems with 2400 structures for the training set and 600 structures for the test set. Fig. 1 illustrates the dependence of the root mean square error (RMSE) on the number of basis functions. One can see that it requires smaller number of Gaussian function than that of cosine functions for the prediction of atomic forces.
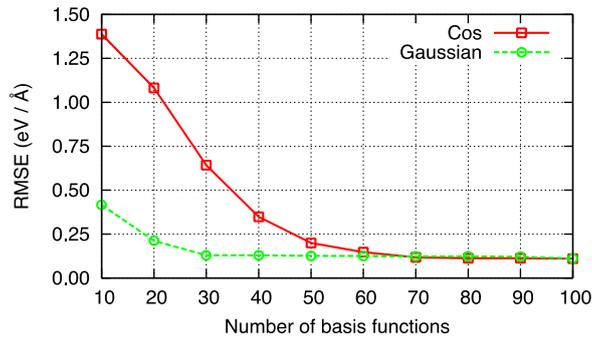
Fig. 1. The root mean square error (RMSE) estimated on the test set of the 2×2×2 supercell
as the function of the basis functions.

Fig. 2 shows the comparison between machine-learning forces and DFT forces for the training set (red) and the test set (blue). We can clearly see the good agreement between our model and the DFT calculations. For the model using 60 Gaussian basis function, we can obtain the accuracy with RMSE of 0.11 eV/ Å, MAE of 0.08 eV/ Å, and R^2 score of 0.99.
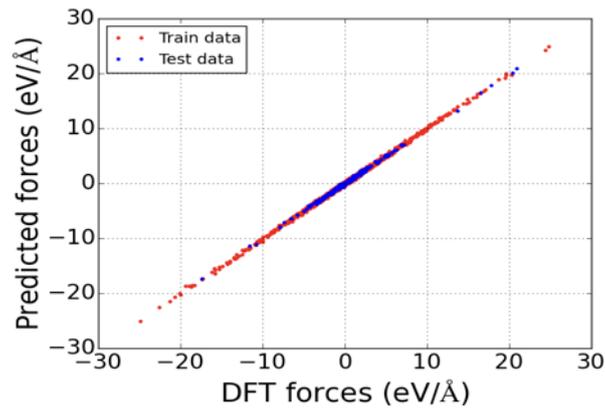

Fig. 2. The comparison of machine-learning forces and DFT forces for the training set (red)
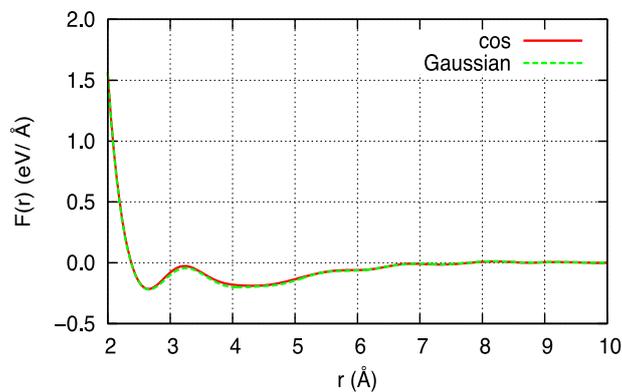and the test set (blue).

Fig. 3. The pairwise force function: by Gaussian (green) and cosine (red) basis functions.

We examine the pairwise functions represented by cosine and Gaussian functions, and the results are shown in Fig. 3. The cosine and Gaussian representations show a good agreement for approximating the pairwise function. It is observed that there might exist the attraction between Si atoms in the range of 2.3-7.0 Å. The interaction between Si atoms converges to zero at long distances. Interestingly, the interaction vanishes at 2.3 Å and 3.3 Å, which are corresponding to the first and second nearest neighbors in the crystalline silicon. The results imply that we can learn for the hidden knowledge of the pairwise interaction of Si atoms, which is not available in first-principles calculations.

We apply the our model to the calculation of phonon dispersion for Si with a 2×2×2 supercell. The phonon dispersions were calculated by using the supercell approach [16] implemented in Phonopy code [17]. Fig. 4, which shows the phonon dispersion for diamond Si using our machine learning model using 60 Gaussian and cosine basis functions, and DFT calculations, shows that the phonon dispersions obtained with these two methods are in good agreement. The slight difference indicates the limit of our method with respect to the prediction of small forces.
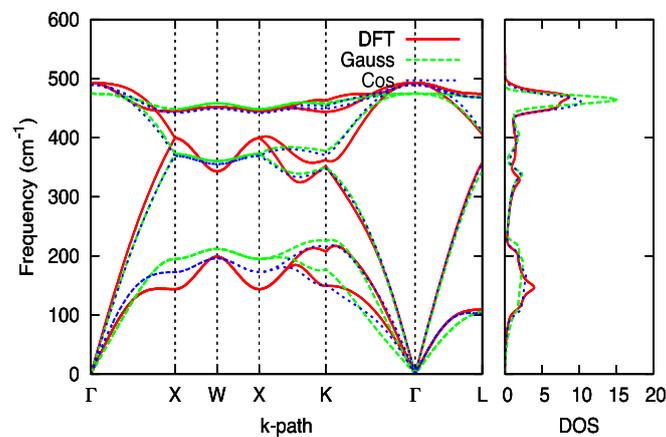

Fig. 4. The phonon band structure and density of state (DOS): by Gaussian (green)
and cosine (red) basis function.

We examined the transferability of our models trained by the 2×2×2 silicon system, by comparing the predicted forces and DFT forces of the 3×3×3 silicon system. A test set with 300 structures of the 3×3×3 supercell was used for the test. We obtained an the similar accuracy with RMSE of 0.11 eV/Å and MAE of 0.08 eV/Å for the 3×3×3 system. This result confirms that our model, which is trained with a small system, is transferable to large systems.

4. Conclusion

We report on a machine learning procedure for representing the potential energy surface and atomic forces from two-body terms. We can train linear models to accurately predict atomic forces and energies comparable to those obtained by density functional theory (DFT) calculations for crystalline and amorphous silicon. The machine learning force model is then applied to calculate the phonon dispersion of crystalline silicon. The result shows reasonable agreement with DFT calculations. We also show that the models learned from 2×2×2 data can extend to predict the potential energy surface and atomic force for 3×3×3 systems.

Acknowledgment

References

[1] R. Car, M. Parrinello, Unified approach for molecular dynamics and density-functional theory, Phys. Rev. Lett. 55 (1985) 2471–2474. https://doi.org/https://doi.org/10.1103/PhysRevLett.55.2471.

[2] W. Kohn, L.J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, Phys. Rev. 140 (1965) 1133. https://doi.org/10.1103/PhysRev.140.A1133.

[3] P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, Phys. Rev. 136 (1964) 5188. https://doi.org/10.1103/PhysRev.136.B864.

[4] C.M. Handley, J. Behler, Next generation interatomic potentials for condensed systems, Eur. Phys. J. B. 87 (2014) 152. https://doi.org/https://doi.org/10.1140/epjb/e2014-50070-0.

[5] A.P. Bartók, G. Csányi, Gaussian approximation potentials: A brief tutorial introduction, Int. J. Quantum Chem. 115 (2015) 1051–1057. https://doi.org/https://doi.org/10.1002/qua.24927.

[6] S. De, A.P. Bartók, G. Csanyi, M. Ceriotti, Comparing molecules and solids across structural and alchemical space, Phys. Chem. Chem. Phys. 18 (2016) 13754–13769. https://doi.org/https://doi.org/10.1039/C6CP00415F.

[7] A. Seko, et al., A sparse representation for potential energy surface, Phys. Rev. B. 90 (2014) 24101. https://doi.org/https://doi.org/10.1103/PhysRevB.90.024101.

[8] M. Rupp, A. Tkatchenko, K.-R. Muller, O.A. von Lilienfeld, Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning, Phys. Rev. Lett. 108 (2012) 58301. https://doi.org/https://doi.org/10.1103/PhysRevLett.108.058301.

[9] T.L. Pham, H. Kino, K. Terakura, T. Miyake, H.C. Dam, Novel mixture model for the representation of potential energy surfaces, J. Chem. Phys. 145 (2016) 154103. https://doi.org/10.1063/1.4964318.

[10] J. Behler, M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, Phys. Rev. Lett. 98 (2007) 146401. https://doi.org/https://doi.org/10.1103/PhysRevLett.98.146401.

[11] N. Artrith, J. Behler, High-dimensional neural network potentials for metal surfaces: A prototype study for copper, Phys. Rev. B. 85 (2012) 45439. https://doi.org/https://doi.org/10.1103/PhysRevB.85.045439.

[12] J. Behler, Atom-centered symmetry functions for constructing high-dimensional neural network potentials, J. Phys. Chem. 134 (2011) 74106. https://doi.org/https://doi.org/10.1063/1.3553717.

[13] T. Lam Pham, H. Kino, K. Terakura, T. Miyake, K. Tsuda, I. Takigawa, H. Chi Dam, Machine learning reveals orbital interaction in materials, Sci Technol Adv Mater. 18 (2017) 756–765. https://doi.org/10.1080/14686996.2017.1378060

[14] T.L. Pham, N.D. Nguyen, V.D. Nguyen, H. Kino, T. Miyake, H.C. Dam, Learning structure-property relationship in crystalline materials: A study of lanthanide-transition metal alloys, J. Chem. Phys. 148 (2018). https://doi.org/10.1063/1.5021089.

[15] V.-D. Nguyen, T.-L. Pham, H.-C. Dam, Application of materials informatics on crystalline materials for two-body terms approximation, Comput. Mater. Sci. 166 (2019) 155–161. https://doi.org/10.1016/J.COMMATSCI.2019.04.030.

[16] K. Parlinski, Z.Q. Li, Y. Kawazoe, First-Principles Determination of the Soft Mode in Cubic ZrO_2, Phys. Rev. Lett. 78 (1997) 4063–4066. https://doi.org/10.1103/PhysRevLett.78.4063.

[17] A. Togo, I. Tanaka, First principles phonon calculations in materials science, Scr. Mater. 108 (2015) 1–5. https://doi.org/10.1016/J.SCRIPTAMAT.2015.07.021.