Conceptual and practical bases for the high accuracy of machine learning interatomic potentials: Application to elemental titanium

Akira Takahashi,[1,*] Atsuto Seko,[1,2,3,4,†] and Isao Tanaka[1,2,3,5]

[1]*Department of Materials Science and Engineering, Kyoto University, Kyoto 606-8501, Japan*
[2]*Center for Elements Strategy Initiative for Structure Materials (ESISM), Kyoto University, Kyoto 606-8501, Japan*
[3]*Center for Materials Research by Information Integration, National Institute for Materials Science, Tsukuba 305-0047, Japan*
[4]*JST, PRESTO, Kawaguchi 332-0012, Japan*
[5]*Nanostructures Research Laboratory, Japan Fine Ceramics Center, Nagoya 456-8587, Japan*

Machine learning interatomic potentials (MLIPs) based on a large data set obtained by density functional theory calculation have been developed recently. This study gives both conceptual and practical bases for the high accuracy of MLIPs, although MLIPs have been considered to be simply an accurate black-box description of atomic energy. We also construct the most accurate MLIP of elemental Ti ever reported using a linearized MLIP framework and many angular-dependent descriptors, which also corresponds to a generalization of the modified embedded atom method potential.

I. INTRODUCTION

Interatomic potentials (IPs) have played a central role in performing atomistic simulations, such as molecular dynamics simulation. A wide variety of conventional IPs have been developed by considering the nature of chemical bonding in specific systems of interest, such as Lennard-Jones [1], embedded atom method (EAM) [2–4], modified EAM (MEAM) [5,6], and Tersoff [7–9] potentials. However, the accuracy and transferability of conventional IPs are often lacking owing to the simplicity of their potential forms. As an example, the phonon dispersion relationships of hexagonal close-packed (hcp) Ti computed from several EAM and MEAM potentials are shown in Fig. 1, along with that computed on the basis of density functional theory (DFT). The overall phonon dispersions of EAM and MEAM potentials are scattered and markedly deviate from that obtained by DFT calculation.

On the other hand, the machine learning IP (MLIP) based on a large data set obtained by DFT calculation has great potential for improving its accuracy and transferability effectively. Once the MLIP is established, it does not increase the order of computational cost as compared with conventional IPs. The MLIP has also been increasingly applied to a wide range of materials regardless of their type of chemical bonding. Its frameworks and applications have recently been reported [21–31].

Although the MLIP can provide an accurate energy description, its physical interpretation or relationship with the existing IPs is still lacking. In this study, we introduce an interpretation of the MLIP on the basis of the framework of EAM and MEAM potentials. The interpretation provides a conceptual basis for the high accuracy of the MLIP. Second, we develop the most accurate MLIP of the elemental Ti ever reported using a linearized MLIP framework. As shown later, the high accuracy of the linearized MLIP implies that the high accuracy and transferability of MLIPs are based mainly on the use of

a large number of relevant descriptors, although it has been considered that the use of flexible black-box functions, such as neural network and Gaussian process models, is essential for modeling atomic energy.

II. INTERPRETATION OF MLIPs

A. Embedding atomic energy in EAM and MEAM potentials

The framework of EAM potentials is based on the concept of the embedding energy of an atom into a host described by electron density [32]. The embedding energy of atom i is defined as a functional of the host electron density $\rho(\boldsymbol{r})$ expressed as

$$E^{(i)} = \mathscr{F}^{(i)}[\rho(\boldsymbol{r})], \tag{1}$$

where $\mathscr{F}^{(i)}$ denotes the embedding energy functional for atom i. Although the application of this concept is not exclusive to metallic systems, the framework of EAM potentials is compatible only with metallic systems owing to the introduction of some approximations. A main approximation is the uniform density approximation (UDA), in which the embedding energy is assumed to be a function of the scalar local electron density, written as

$$E^{(i)} = F(\rho(\boldsymbol{r}_i)), \tag{2}$$

where $\boldsymbol{r}_i$ denotes the position of atom i. Another one is a pairwise approximation in which the local electron density is assumed to be equal to the sum of contributions from neighboring atoms expressed by a single pairwise function. Adding a short-range pairwise interaction, the EAM atomic energy is expressed as

$$E^{(i)} = F\left(\sum_j p(r_{ij})\right) + \frac{1}{2}\sum_j \phi(r_{ij}), \tag{3}$$

where $p(r_{ij})$ and $\phi(r_{ij})$ denote the pairwise contribution of the neighbor atom j to the local electron density and short-range pairwise interaction including repulsive energy,

---
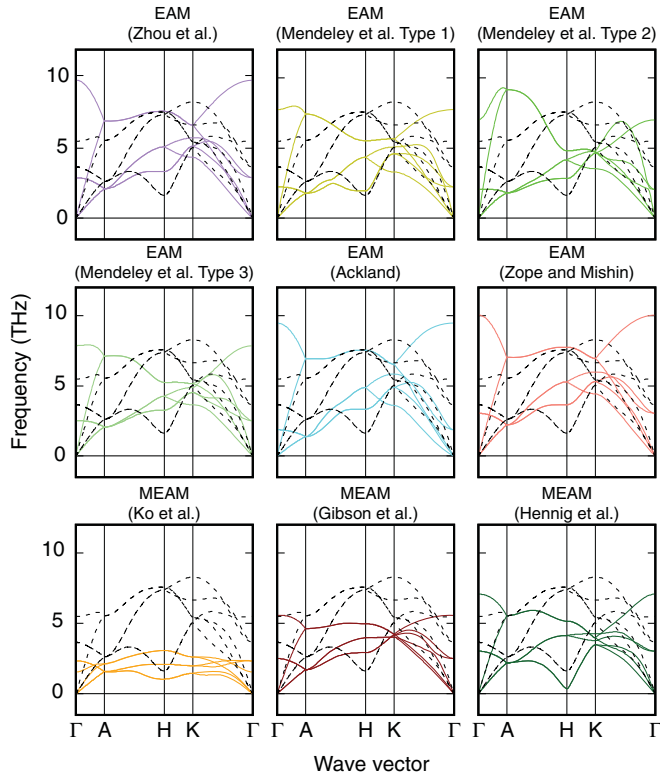*takahashi.akira.36m@gmail.com
†seko@cms.mtl.kyoto-u.ac.jp

FIG. 1. Phonon dispersion curves of elemental hcp Ti calculated using conventional EAM [10–13] and MEAM [14–17] potentials. Some of these curves are obtained from the interatomic potential repository project [18] and KIM project [19]. Black broken lines indicate the phonon dispersion curves obtained by DFT calculation. Force constants are calculated using the LAMMPS [20] code.

respectively [2]. In an extended manner, the MEAM atomic energy is given by

$$E^{(i)} = F(\rho(\boldsymbol{r}_i)) + \frac{1}{2}\sum_j \phi(r_{ij}), \qquad (4)$$

$$\rho(\boldsymbol{r}_i) = \sum_j p(r_{ij}) + \sum_{j,k} f(r_{ij})f(r_{ik})g(\cos\gamma_{jik}), \qquad (5)$$

where the local electron density is described by a three-body function g in addition to the pairwise contribution. Since the function forms of p, f, and g have not been established, a wide range of approximated forms have been proposed in the literature. In addition, polynomials and spline models have been simply used as a function F.

B. Atomic energy in MLIPs

On the other hand, all MLIPs with pairwise descriptors are formulated as

$$E^{(i)} = F\big(b_{10}^{(i)}, b_{20}^{(i)}, \dots, b_{n_{\max}0}^{(i)}\big), \qquad (6)$$

where $b_{n0}^{(i)}$ denotes a pairwise descriptor expressed as

$$b_{n0}^{(i)} = \sum_j f_n(r_{ij}). \qquad (7)$$

A large number of pairwise descriptors are generally used for formulating MLIPs, and neural network models, Gaussian process models, and polynomials have been used as functions F. This formulation is obviously a generalization of the EAM atomic energy. Similarly, most MLIPs with angular-dependent descriptors are formulated as

$$E^{(i)} = F\big(b_{10}^{(i)}, b_{20}^{(i)}, \dots, b_{11}^{(i)}, b_{21}^{(i)}, \dots, b_{n_{\max}l_{\max}}^{(i)}\big), \qquad (8)$$

where $b_{nl}^{(i)}$ denotes an angular-dependent descriptor. Most angular-dependent descriptors specified by number l belong to the class of angular Fourier series, which corresponds to a set of rotationally invariant descriptors derived from spherical harmonics [27]. The angular Fourier series is given by

$$b_{nl}^{(i)} = \sum_{j,k} f_n(r_{ij})f_n(r_{ik})\cos(l\gamma_{jik}) \quad (l \geqslant 1), \qquad (9)$$

where γ_{jik} denotes the bond angle between atoms j-i-k. From the comparison between Eqs. (4) and (8), the formulation of the MLIP with angular-dependent descriptors is clearly a generalization of the MEAM potential.

C. Derivation of MLIP atomic energy from embedding energy

We have demonstrated that the MLIP formulations can be regarded as the generalizations of the EAM and MEAM potentials by comparing their equations for atomic energy. We will show that the MLIP formulations can also be derived from the concept of embedding energy using a higher-order approximation beyond the UDA. This derivation interprets MLIPs. Using a higher-order approximation for the embedding energy functional [Eq. (1)], atomic energy may be described by a function of local electron density and its derivatives as

$$\begin{aligned} E^{(i)} &= \mathscr{F}^{(i)}[\rho(\boldsymbol{r})] \\ &= F\bigg(\rho(\boldsymbol{r}_i), \frac{\partial\rho}{\partial x}(\boldsymbol{r}_i), \frac{\partial\rho}{\partial y}(\boldsymbol{r}_i), \frac{\partial\rho}{\partial z}(\boldsymbol{r}_i), \dots\bigg). \end{aligned} \qquad (10)$$

Then, the local electron density is assumed to be described by direction-dependent contributions from neighbor atoms, $\rho(\boldsymbol{r}_i) = \sum_j p(\boldsymbol{r}_{ij})$. Equation (10) is rewritten as

$$\begin{aligned} E^{(i)} = F\bigg(&\sum_j p(\boldsymbol{r}_{ij}), \sum_j \frac{\partial}{\partial x}p(\boldsymbol{r}_{ij}), \\ &\sum_j \frac{\partial}{\partial y}p(\boldsymbol{r}_{ij}), \sum_j \frac{\partial}{\partial z}p(\boldsymbol{r}_{ij}), \dots\bigg). \end{aligned} \qquad (11)$$

Expanding the electron density contribution p using a basis set $\{f_n(\boldsymbol{r}_{ij})\}_{n=1,2,\dots,n_{\max}}$ as

$$p(\boldsymbol{r}_{ij}) = \sum_{n=1}^{n_{\max}} c_n f_n(\boldsymbol{r}_{ij}), \qquad (12)$$

the embedding atomic energy is written as

$$E^{(i)} = \tilde{F}\bigg(\sum_j f_1(\boldsymbol{r}_{ij}), \dots, \sum_j f_{n_{\max}}(\boldsymbol{r}_{ij})\bigg), \qquad (13)$$

where another symbol $\tilde{F}$ for the embedding energy function is derived from both function F and expansion coefficients $\{c_n\}_{n=1,2,\ldots,n_{\max}}$. Replacing the vector $\boldsymbol{r}_{ij}$ with the pair distance r_{ij}, Eq. (13) becomes the pairwise MLIP formulation. Generally, the basis set is not necessarily pairwise. When functions based on spherical harmonics are used as a basis set and function $\tilde{F}$ satisfying the rotational invariance, the angular-dependent MLIP [Eq. (8)] is derived. Thus, MLIP formulations are derived from the concept of embedding energy using an approximation beyond the UDA. This implies that the lack of accuracy and transferability of the EAM and MEAM potentials can be ascribed to their poor representation for embedding energy due to the limitation of the UDA [33].

III. METHODOLOGY FOR BUILDING MLIP

A. Linearized models for atomic energy

On the basis of the relationship between MLIPs and EAM potentials, we construct two MLIPs for the elemental Ti in this study. The first one is constructed by a third-order polynomial approximation of Eq. (6) expressed as

$$
E^{(i)} = w_0 + \sum_n w_{n0} b_{n0}^{(i)} + \sum_{n,n'} w_{n0,n'0} b_{n0}^{(i)} b_{n'0}^{(i)} \\
+ \sum_{n,n',n''} w_{n0,n'0,n''0} b_{n0}^{(i)} b_{n'0}^{(i)} b_{n''0}^{(i)}, \tag{14}
$$

where $w_0, w_{n0}, w_{n0,n'0}$, and $w_{n0,n'0,n''0}$ denote regression coefficients. The second one is constructed by a second-order polynomial approximation of Eq. (8) with angular Fourier series descriptors expressed as

$$
E^{(i)} = w_0 + \sum_{n,l} w_{nl} b_{nl}^{(i)} + \sum_{n,l,n',l'} w_{nl,n'l'} b_{nl}^{(i)} b_{n'l'}^{(i)}. \tag{15}
$$

Here, we fixed $l_{\max}$ to ten. Both of the models for the atomic energy are rewritten in the same form as

$$
E^{(i)} = \boldsymbol{w}^\top \boldsymbol{b}^{(i)}, \tag{16}
$$

using the vector $\boldsymbol{b}^{(i)} = [1, b_1^{(i)}, \ldots, b_M^{(i)}]^\top$ composed of the descriptors and their products, and the vector $\boldsymbol{w} = [w_0, w_1, \ldots, w_M]^\top$ composed of the regression coefficients, where M denotes the number of terms.

We used pairwise Gaussian-type functions as radial functions $f_n(r)$ expressed as

$$
f_n(r) = f_c(r) \exp[-p(r - q_n)^2], \tag{17}
$$

where $f_c(r)$ denotes a cosine-type cutoff function. p and q_n are given parameters, and we used a single p value and a set of q_n values given by an arithmetic sequence. Also in the EAM and MEAM potentials, Gaussian functions have sometimes been used for expressing the pairwise electron density contribution. In addition, a polynomial approximation for the embedding energy function F has been used for EAM and MEAM potentials. Therefore, the only difference between the MLIP and EAM (MEAM) potentials is in the number of descriptors being used in the formulation of atomic energy. Equations (14) and (15) are also a generalization of our previous linearized model where only the power of b_n is considered [23,34].

B. Total energy and forces acting on atoms

The total energy of a structure is expressed by the sum of the atomic energy. Therefore, the total energy E^{total} is given as

$$
E^{\text{total}} = \sum_i E^{(i)} = \boldsymbol{w}^\top \boldsymbol{x}, \tag{18}
$$

where $\boldsymbol{x} = \sum_i \boldsymbol{b}^{(i)}$. The forces acting on atoms are given by linear equations with respect to the regression coefficients as well as the total energy. The βth component of the force acting on atom a is expressed as

$$
F_{a,\beta} = -\frac{\partial E^{\text{total}}}{\partial r_{a,\beta}} \\
= \boldsymbol{w}^\top \boldsymbol{x}_{\text{force},a,\beta}, \tag{19}
$$

where $r_{a,\beta}$ denotes the βth component of the position of atom a provided in Cartesian coordinates (see the Appendix for details).

C. DFT data set

Training and test data sets were generated by DFT calculation for 2700 and 300 atomic configurations, respectively. We first optimized the atomic positions and lattice constants of face-centered cubic (fcc), body-centered cubic (bcc), hcp, simple cubic (sc), ω, and β-Sn structures, and supercells were then developed by the $2 \times 2 \times 2$, $3 \times 3 \times 3$, $3 \times 3 \times 3$, $4 \times 4 \times 4$, $3 \times 3 \times 3$, and $2 \times 2 \times 2$ expansions of their conventional unit cells, respectively. Atomic configurations were generated by isotropic expansion, random expansions, random distortions, and random displacements. DFT calculations were performed using the plane-wave basis projector augmented wave (PAW) method [35,36] within the Perdew-Burke-Ernzerhof exchange-correlation functional [37] as implemented in the VASP code [38,39]. The cutoff energy was set to 400 eV. The total energies converged to less than 10^{-3} meV/supercell. The lattice constants of the ideal structures were optimized until the residual forces became less than 10^{-3} eV/Å.

D. Estimation of models

The regression coefficients $\boldsymbol{w}$ are estimated by linear regression using a training data set. In this study, the energy and the forces acting on atoms computed by DFT calculations are used as observations for the training data. Therefore, the predictor matrix X and observation vector $\boldsymbol{y}$ are written as

$$
X = \begin{bmatrix} X_{\text{energy}} \\ X_{\text{force}} \end{bmatrix}, \qquad \boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_{\text{energy}} \\ \boldsymbol{y}_{\text{force}} \end{bmatrix}, \tag{20}
$$

where X_{energy} and X_{force} are composed of $\boldsymbol{x}$ and $\boldsymbol{x}_{\text{force},a,\beta}$ for all structures in the training data, respectively. Observation vectors of energy and forces, $\boldsymbol{y}_{\text{energy}}$ and $\boldsymbol{y}_{\text{force}}$, can be obtained by the energy and forces acting on atoms computed by DFT calculations, respectively. As a result, the total number of training data is 430 650.

To estimate MLIPs, we adopt linear ridge regression, which shrinks the regression coefficients to avoid overfitting by imposing the L_2 penalty. This method determines ridge coefficients to minimize the sum of the residual error of squares
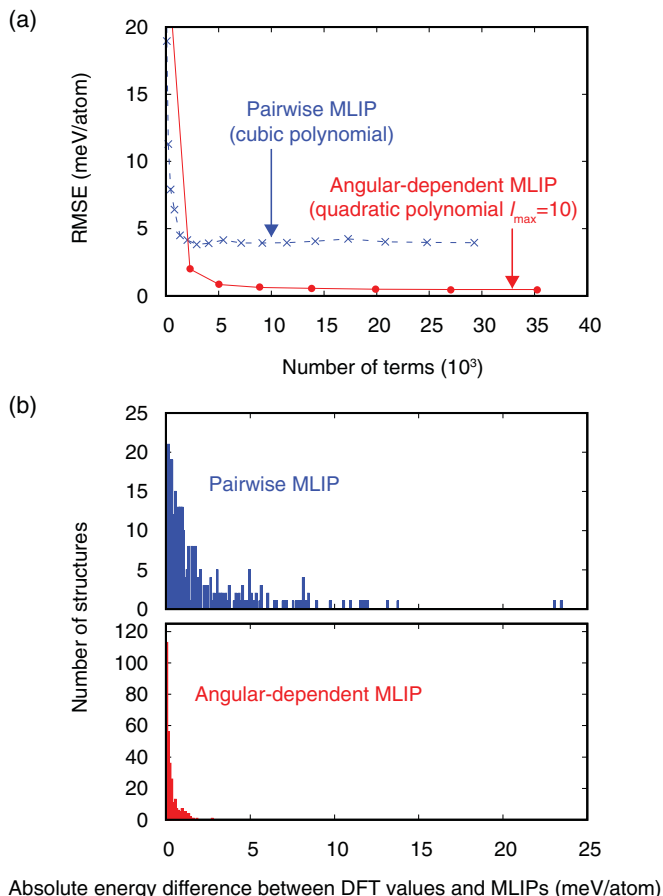
FIG. 2. (a) Dependence of RMSE of MLIPs on number of terms for elemental Ti. (b) Distribution of absolute energy difference between DFT values and MLIPs.

and L_2 penalty, that is,

$$L(\boldsymbol{w}) = ||\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}||_2^2 + \lambda||\boldsymbol{w}||_2^2, \qquad (21)$$

where λ controls the magnitude of the penalty. The solution is easily obtained only in terms of matrix operations as $\boldsymbol{w} = (\boldsymbol{X}^\top\boldsymbol{X} + \lambda\boldsymbol{I})^{-1}\boldsymbol{X}^\top\boldsymbol{y}$, where $\boldsymbol{I}$ denotes the unit matrix.

IV. RESULTS AND DISCUSSION

We will show the accuracy of MLIPs for elemental Ti. We regard the root mean square error (RMSE) for the energy of the test data set as a measure of prediction error. Figure 2(a) shows the dependence of prediction error on the number of regression coefficients. The number of regression coefficients was controlled using only the number of radial functions f_n for both pairwise and angular-dependent MLIPs. By examining the convergence of RMSE with respect to the number of regression coefficients, we obtained an optimized pairwise MLIP with a prediction error of 3.8 meV/atom (2925 coefficients). Similarly, we obtained an optimized angular-dependent MLIP with a prediction error of 0.5 meV/atom (35 245 coefficients), which means that it is very important to consider angular-dependent descriptors for expressing the interatomic interactions of the elemental Ti. Figure 2(b) also shows the distribution of the absolute energy difference
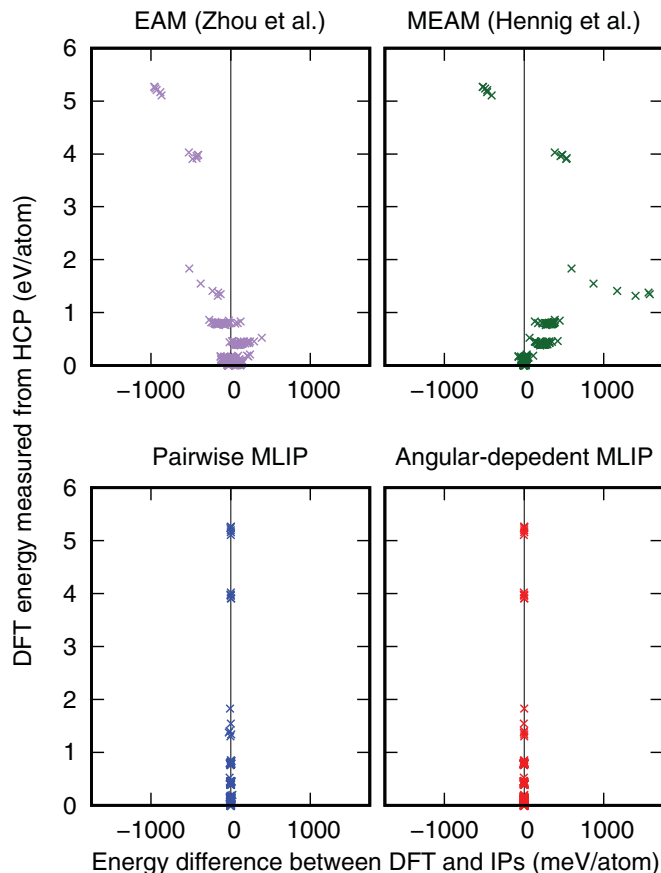

FIG. 3. Distribution of energy difference between DFT and IPs.

between DFT and MLIPs for the test data set. The distribution for the angular-dependent MLIP is much narrower than that for the pairwise MLIP, which is consistent with the degree of prediction error. For the angular-dependent MLIP, more than a hundred structures show the absolute energy difference within only 0.1 meV/atom. In addition, some outliers can be found in the distribution for the pairwise MLIP. A structure shows the maximum absolute energy difference of 23.0 meV/atom of the pairwise MLIP, whereas the absolute energy difference of the angular-dependent MLIP does not exceed 2.8 meV/atom.

We then compare the distribution of the energy difference between DFT and IPs for the test data, elastic constants, and phonon dispersion relationships obtained from EAM [10] and MEAM [17] potentials, the pairwise MLIP, and the angular-dependent MLIP along with a reference of the DFT calculation. Figure 3 shows the comparison of the distribution of energy difference between DFT and IPs for the test data set. EAM and MEAM potentials show very large energy differences for almost the entire test data set, while both the MLIPs show very small energy differences.

Figure 4 shows the elastic constants and bulk moduli of (a) hcp Ti and (b) bcc Ti obtained from EAM and MEAM potentials and the MLIPs. The elastic constants of EAM and MEAM potentials are close to those of DFT calculation, except for the C_{33} of hcp and the C_{44} of bcc obtained from the EAM potential. On the other hand, the pairwise MLIP is worst for predicting most of the elastic constants and bulk moduli of both hcp and bcc structures, despite its small prediction
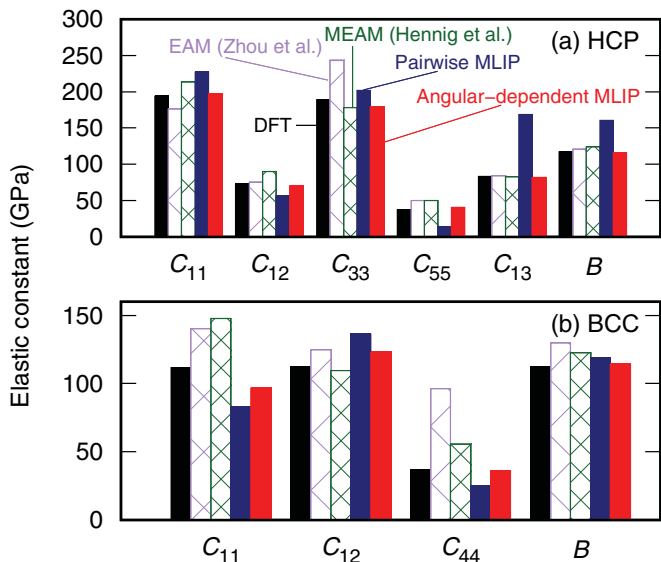
FIG. 4. Elastic constant C and bulk modulus B values of (a) hcp Ti and (b) bcc Ti calculated on the basis of DFT and IPs.

error. Including angular-dependent terms, the prediction of elastic constants and bulk moduli is much improved. This is consistent with the fact that the angular-dependent descriptors are essential for predicting the mechanical behavior of the elemental Ti.

The phonon dispersion curves were also calculated using the supercell approach [40] for hcp and bcc structures with the DFT equilibrium lattice constant. To evaluate a dynamical matrix, each symmetrically independent atomic position was displaced by 0.01 Å. The forces acting on atoms were then computed. Supercells were fabricated by the $4 \times 4 \times 4$ expansion of conventional unit cells for both hcp and bcc structures. Phonon calculations were performed using the PHONOPY code [41]. Figure 5 shows the phonon dispersion

curves of (a) hcp and (b) bcc structures computed from EAM and MEAM potentials, and the MLIPs. As shown in Fig. 5, the phonon dispersion curves from EAM and MEAM potentials differ largely from that obtained by DFT calculation. Imaginary phonon modes are observed in the DFT phonon dispersion for the bcc structure, but not in the EAM and MEAM phonon dispersions. Although the pairwise MLIP reproduces the DFT phonon dispersion better than the EAM and MEAM potentials, phonon frequencies tend to be overestimated. The angular-dependent MLIP significantly improves the inconsistency of phonon frequency.

V. SUMMARY

In summary, this study provides both conceptual and practical bases for the high accuracy of MLIPs. We have shown that MLIPs can be regarded as a description of embedding energy beyond the UDA, which is a fundamental approximation of both EAM and MEAM potentials. In other words, the high accuracy of MLIPs is based on the use of a higher-order approximation of embedding energy. We have then applied a linearized MLIP approach to the elemental Ti, which is also a generalization of the MEAM potential. An angular-dependent linearized MLIP predicts the energetics and phonon frequencies much more accurately than the existing MEAM potentials. The only difference between the MEAM potentials and linearized MLIP is in the number of descriptors being used. This indicates that the use of a systematic set of numerous descriptors is the most important practical feature for building MLIPs with high accuracy.
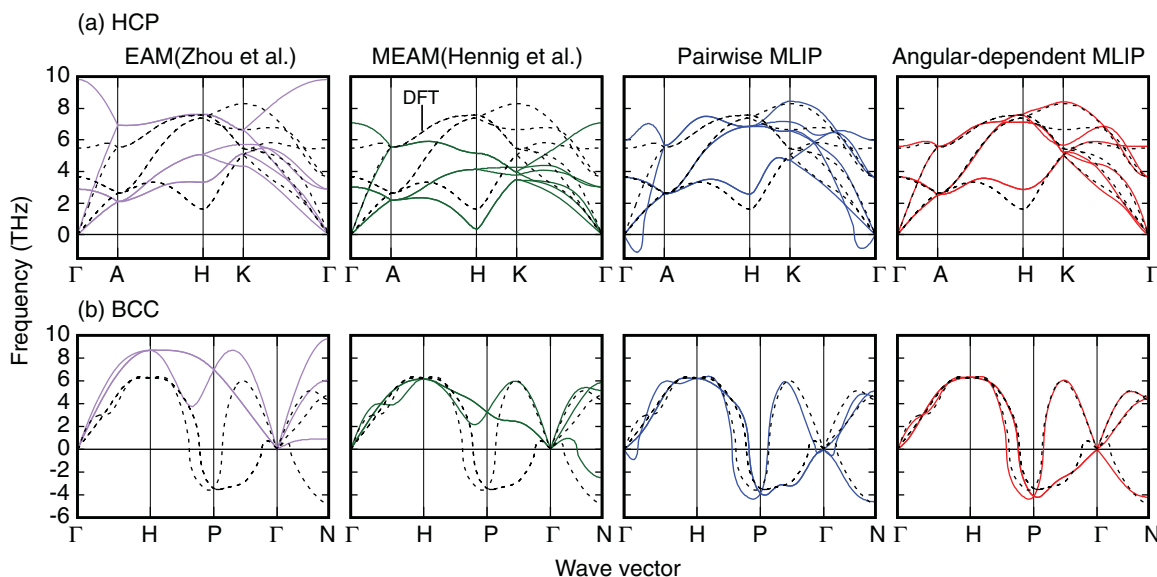
FIG. 5. Phonon dispersion curves of (a) hcp Ti and (b) bcc Ti calculated from IPs. Broken black lines indicate the phonon dispersion curves calculated on the basis of DFT.

APPENDIX: FORCES ACTING ON ATOMS

The force acting on atoms is obtained from the derivative of the total energy with respect to the atomic position given in Cartesian coordinates, expressed as

$$
F_{a,\beta} = -\frac{\partial E^{\text{total}}}{\partial r_{a,\beta}}
$$
$$
= -\sum_{m=0}^{M} w_m \frac{\partial x_m}{\partial r_{a,\beta}}. \qquad (A1)
$$

In the first model, x_m corresponds to each of the following terms, $\sum_i b_{n0}^{(i)}$, $\sum_i b_{n0}^{(i)} b_{n'0}^{(i)}$, and $\sum_i b_{n0}^{(i)} b_{n'0}^{(i)} b_{n''0}^{(i)}$. The derivative of $b_{n0}^{(i)}$ with respect to $r_{a,\beta}$ is simply given as

$$
\frac{\partial b_{n0}^{(i)}}{\partial r_{a,\beta}} = \sum_j f_n'(r_{ij}) \frac{\partial r_{ij}}{\partial r_{a,\beta}}. \qquad (A2)
$$

The derivatives of $b_{n0}^{(i)} b_{n'0}^{(i)}$ and $b_{n0}^{(i)} b_{n'0}^{(i)} b_{n''0}^{(i)}$ are calculated using the derivative of $b_{n0}^{(i)}$, expressed as

$$
\frac{\partial b_{n0}^{(i)} b_{n'0}^{(i)}}{\partial r_{a,\beta}} = \frac{\partial b_{n0}^{(i)}}{\partial r_{a,\beta}} b_{n'0}^{(i)} + b_{n0}^{(i)} \frac{\partial b_{n'0}^{(i)}}{\partial r_{a,\beta}} \qquad (A3)
$$

and

$$
\frac{\partial b_{n0}^{(i)} b_{n'0}^{(i)} b_{n''0}^{(i)}}{\partial r_{a,\beta}} = \frac{\partial b_{n0}^{(i)}}{\partial r_{a,\beta}} b_{n'0}^{(i)} b_{n''0}^{(i)} + b_{n0}^{(i)} \frac{\partial b_{n'0}^{(i)}}{\partial r_{a,\beta}} b_{n''0}^{(i)}
$$
$$
+ b_{n0}^{(i)} b_{n'0}^{(i)} \frac{\partial b_{n''0}^{(i)}}{\partial r_{a,\beta}}. \qquad (A4)
$$

In the second model, x_m corresponds to each of the following terms, $\sum_i b_{nl}^{(i)}$ and $\sum_i b_{nl}^{(i)} b_{n'l'}^{(i)}$. Although $b_{nl}^{(i)}$ is a three-body descriptor, we compute $b_{nl}^{(i)}$ using only two-body operations. Using the Chebyshev polynomial

$$
T_l(\cos\gamma) = \cos(l\gamma) = \sum_m T_{lm} \cos^m \gamma, \qquad (A5)
$$

b_{nl} is rewritten as

$$
b_{nl}^{(i)} = \sum_{j,k} f_n(r_{ij}) f_n(r_{ik}) \cos(l\gamma_{jik})
$$
$$
= \sum_m T_{lm} \sum_{j,k} f_n(r_{ij}) f_n(r_{ik}) \cos^m \gamma_{jik}. \qquad (A6)
$$

Then,

$$
\sum_{j,k} f_n(r_{ij}) f_n(r_{ik}) \cos^m \gamma_{jik}
$$
$$
= \sum_{j,k} f_n(r_{ij}) f_n(r_{ik}) \left(\frac{x_{ij}x_{ik} + y_{ij}y_{ik} + z_{ij}z_{ik}}{r_{ij}r_{ik}} \right)^m
$$
$$
= \sum_{p+q+r=m} \frac{m!}{p!q!r!} \sum_{j,k} f_n(r_{ij}) f_n(r_{ik}) \frac{x_{ij}^p y_{ij}^q z_{ij}^r x_{ik}^p y_{ik}^q z_{ik}^r}{r_{ij}^m r_{ik}^m}
$$
$$
= \sum_{p+q+r=m} \frac{m!}{p!q!r!} \left(\sum_j f_n(r_{ij}) \frac{x_{ij}^p y_{ij}^q z_{ij}^r}{r_{ij}^m} \right)^2, \qquad (A7)
$$

where x_{ij}, y_{ij}, and z_{ij} denote components of the vector between atoms i and j. Therefore, $b_{nl}^{(i)}$ is computed using the following equation,

$$
b_{nl}^{(i)} = \sum_m T_{lm} \sum_{p+q+r=m} \frac{m!}{p!q!r!} \left(\sum_j f_n(r_{ij}) \frac{x_{ij}^p y_{ij}^q z_{ij}^r}{r_{ij}^m} \right)^2. \qquad (A8)
$$

The derivative of $b_{nl}^{(i)}$ is also computed from Eq. (A8) as

$$
\frac{\partial b_{nl}^{(i)}}{\partial r_{a,\beta}} = \sum_m T_{lm} \sum_{p+q+r=m} \frac{m!}{p!q!r!} \frac{\partial \left(\sum_j f_n(r_{ij}) \frac{x_{ij}^p y_{ij}^q z_{ij}^r}{r_{ij}^m} \right)^2}{\partial r_{a,\beta}}
$$
$$
= \sum_m T_{lm} \sum_{p+q+r=m} \frac{m!}{p!q!r!} 2 \left(\sum_j f_n(r_{ij}) \frac{x_{ij}^p y_{ij}^q z_{ij}^r}{r_{ij}^m} \right)
$$
$$
\times \sum_j \left[\frac{\partial r_{ij}}{\partial r_{a,\beta}} \frac{1}{r_{ij}^{m+1}} [f_n'(r_{ij})r_{ij} - f_n(r_{ij})m] x_{ij}^p y_{ij}^q z_{ij}^r \right.
$$
$$
\left. + \frac{(\partial x_{ij}^p y_{ij}^q z_{ij}^r)}{\partial r_{a,\beta}} \frac{f_n(r_{ij})}{r_{ij}^m} \right]. \qquad (A9)
$$

The derivative of $b_{nl}^{(i)} b_{n'l'}^{(i)}$ is calculated using the derivative of $b_{nl}^{(i)}$ in a similar way to Eq. (A3).

[1] J. E. Lennard-Jones, Proc. R. Soc. London, Ser. A **106**, 463 (1924).

[2] M. S. Daw and M. I. Baskes, Phys. Rev. Lett. **50**, 1285 (1983).

[3] M. S. Daw and M. I. Baskes, Phys. Rev. B **29**, 6443 (1984).

[4] M. S. Daw, S. M. Foiles, and M. I. Baskes, Mater. Sci. Rep. **9**, 251 (1993).

[5] M. I. Baskes, Phys. Rev. Lett. **59**, 2666 (1987).

[6] M. I. Baskes, Phys. Rev. B **46**, 2727 (1992).

[7] J. Tersoff, Phys. Rev. Lett. **56**, 632 (1986).

[8] J. Tersoff, Phys. Rev. B **38**, 9902 (1988).

[9] J. Tersoff, Phys. Rev. Lett. **61**, 2879 (1988).

[10] X. W. Zhou, R. A. Johnson, and H. N. G. Wadley, Phys. Rev. B **69**, 144113 (2004).

[11] M. I. Mendelev, T. L. Underwood, and G. J. Ackland, J. Chem. Phys. **145**, 154102 (2016).

[12] G. J. Ackland, Philos. Mag. A **66**, 917 (1992).

[13] R. R. Zope and Y. Mishin, Phys. Rev. B **68**, 024102 (2003).

[14] W.-S. Ko, B. Grabowski, and J. Neugebauer, Phys. Rev. B **92**, 134107 (2015).

[15] P. Zhang and D. R. Trinkle, Comput. Mater. Sci. **124**, 204 (2016).

[16] J. Gibson, Titanium model for multi-state modified embedded atom method, https://openkim.org/cite/MO_309653492217_000.

[17] R. G. Hennig, T. J. Lenosky, D. R. Trinkle, S. P. Rudin, and J. W. Wilkins, Phys. Rev. B **78**, 054121 (2008).

[18] NIST Interatomic Potentials Repository, http://www.ctcms.nist.gov/potentials.

[19] E. Tadmor, R. Elliott, J. Sethna, R. Miller, and C. Becker, Knowledgebase of interatomic models (KIM), https://openkim.org.

[20] S. Plimpton, J. Comput. Phys. **117**, 1 (1995).

[21] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).

[22] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010).

[23] A. Seko, A. Takahashi, and I. Tanaka, Phys. Rev. B **90**, 024101 (2014).

[24] J. Behler, J. Chem. Phys. **134**, 074106 (2011).

[25] N. Artrith, T. Morawietz, and J. Behler, Phys. Rev. B **83**, 153101 (2011).

[26] K. J. Jose, N. Artrith, and J. Behler, J. Chem. Phys. **136**, 194111 (2012).

[27] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013).

[28] J. Behler, J. Phys.: Condens. Matter **26**, 183001 (2014).

[29] W. J. Szlachta, A. P. Bartók, and G. Csányi, Phys. Rev. B **90**, 104108 (2014).

[30] S. A. Ghasemi, A. Hofstetter, S. Saha, and S. Goedecker, Phys. Rev. B **92**, 045131 (2015).

[31] V. Botu and R. Ramprasad, Int. J. Quantum Chem. **115**, 1074 (2015).

[32] M. J. Stott and E. Zaremba, Phys. Rev. B **22**, 1564 (1980).

[33] Even if $p(r_{ij})$ is expressed by the linear combination of two functions $f_1(r_{ij})$ and $f_2(r_{ij})$, Eq. (3) is not enough to express some functions such as $[\sum_j f_1(r_{ij})]^2 + [\sum_j f_2(r_{ij})]^2$.

[34] A. Seko, A. Takahashi, and I. Tanaka, Phys. Rev. B **92**, 054113 (2015).

[35] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[36] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).

[37] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[38] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).

[39] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[40] K. Parlinski, Z.-Q. Li, and Y. Kawazoe, Phys. Rev. Lett. **78**, 4063 (1997).

[41] A. Togo and I. Tanaka, Scr. Mater. **108**, 1 (2015).