

Machine learning potentials for multicomponent systems: The Ti-Al binary systemAtsuto Seko ^{*}*Department of Materials Science and Engineering, Kyoto University, Kyoto 606-8501, Japan*

(Received 24 August 2020; accepted 26 October 2020; published 9 November 2020)

Machine learning potentials (MLPs) are becoming powerful tools for performing accurate atomistic simulations and crystal structure optimizations. An approach to developing MLPs employs a systematic set of polynomial invariants including high-order ones to represent the neighboring atomic density. In this study, a formulation of the polynomial invariants is extended to the case of multicomponent systems. The extended formulation is more complex than the formulation for elemental systems. This study also shows its application to the Ti-Al binary system. As a result, an MLP with the lowest error and MLPs with high computational cost performance are selected from the many MLPs developed systematically. The predictive powers of the developed MLPs for many properties, such as the formation energy, elastic constants, thermodynamic properties, and mechanical properties, are examined. The MLPs exhibit high predictive power for the properties in a wide variety of ordered structures. The present scheme should be systematically applicable to other multicomponent systems.

DOI: [10.1103/PhysRevB.102.174104](https://doi.org/10.1103/PhysRevB.102.174104)**I. INTRODUCTION**

Machine learning potentials (MLPs) have been developed from extensive data sets generated by density functional theory (DFT) calculation and can significantly improve the accuracy and transferability of interatomic potentials. Therefore, MLPs are becoming useful tools for performing crystal structure optimizations and accurate large-scale atomistic simulations, which are prohibitively expensive by DFT calculation. Over the past decade, a number of methods that can be used to develop MLPs and their applications have been reported [1–23]. In these studies, the contribution of an atom to the potential energy is given as a function of quantities depending on its neighboring environment, called structural features. Also, several models are employed to describe a mapping from structural features to the atomic contribution, including artificial neural network models [1,2,4–7], Gaussian process models [3,8–11], and linear models [12–19].

Structural features play an essential role in controlling the accuracy and computational efficiency of MLPs, which are conflicting properties in general [24–26]. A systematic set of structural features is composed of polynomial invariants. The polynomial invariants include second- and third-order bond-orientational order parameters [27], angular Fourier series [28], the bispectrum [28,29], and moment tensors [19,30], which have been adopted to develop MLPs and machine learning models of physical properties in compounds. Recently, a group-theoretic procedure for enumerating the polynomial invariants derived from spherical harmonics, including high-order ones, was proposed [24]. The angular Fourier series, bond-orientational order parameters, and bispectrum can be

included in the enumeration by this procedure. Machine learning potentials developed with the polynomial invariants for a wide variety of elemental systems exhibit high predictive power for a wide range of structures. They are available in the Machine Learning Potential Repository [31,32]. However, the polynomial invariants and related potential energy models must be generalized for developing MLPs in multicomponent systems.

In this study, polynomial invariants for developing MLPs for a multicomponent system are formulated. The present formulation of the polynomial invariants should be helpful in developing MLPs for multicomponent systems even within other frameworks. Polynomial models combined with the polynomial invariants are also introduced to describe the potential energy. This study also shows an application of the polynomial models to the development of Pareto optimal MLPs in the Ti-Al binary alloy system. The predictive power of the Pareto optimal MLPs is examined for the cohesive energy, the formation energy, the elastic constants, the phonon density of states and dispersion curves, the thermal expansion, the energy profile along the Bain path, and the stacking fault properties.

Section II introduces potential energy models in multicomponent systems, including the polynomial invariants representing the neighboring atomic density and polynomial models for the potential energy. In Sec. III, data sets required to develop MLPs for the Ti-Al binary system are explained. Computational procedures for constructing data sets and estimating coefficients in the potential energy models are also shown. In Sec. IV, the development of Pareto optimal MLPs for the Ti-Al binary system is demonstrated. The predictive power of the MLPs for many properties, such as the formation energy, the elastic constants, the thermodynamic properties, and the mechanical properties, is also investigated. The study is summarized in Sec. V.

^{*}seko@cms.mtl.kyoto-u.ac.jp

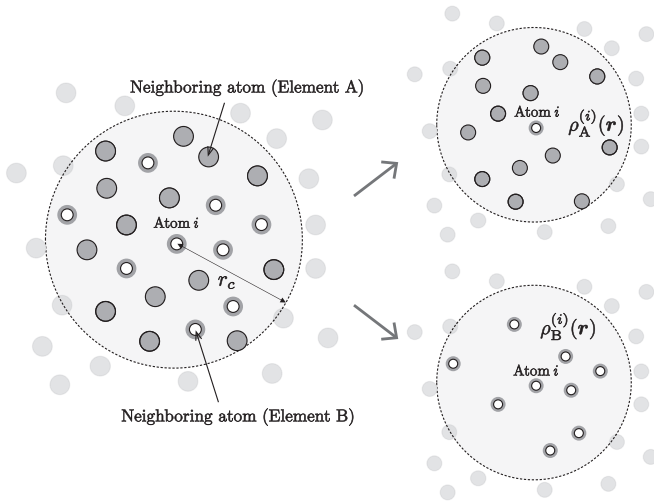


FIG. 1. Schematic illustration of the neighboring atomic density around atom i in a binary structure. Its decomposition into the neighboring atomic densities of elements A and B around atom i is also shown.

II. POTENTIAL ENERGY MODELS

This section shows an extension of polynomial models for the potential energy proposed to develop MLPs for elemental systems [24]. This section is composed of a general description of the potential energy that is useful for deriving potential energy models for a multicomponent system, a systematic set of structural features representing the neighboring atomic densities, and polynomial models for the potential energy used in this study.

A. General description of potential energy

Given a cutoff radius r_c , the short-range part of the potential energy for a structure E may be decomposed as

$$E = \sum_i E^{(i)}, \quad (1)$$

where $E^{(i)}$ denotes the contribution of atom i within cutoff radius r_c . The atomic contribution to the potential energy can be referred to as the atomic energy. The atomic energy is then assumed to be expressed in a functional form of the neighboring atomic densities. Figure 1 shows a schematic illustration of the neighboring atomic density around atom i within cutoff radius r_c and its decomposition into the neighboring atomic densities of the elements. In a multicomponent system composed of elements $\{A, B, \dots\}$, the atomic energy is written using functional \mathcal{F} dependent on the element of atom i as

$$E^{(i)} = \mathcal{F}_{s_i}[\rho_{(s_i,A)}^{(i)}, \rho_{(s_i,B)}^{(i)}, \dots], \quad (2)$$

where $\rho_{(s_i,s)}^{(i)}$ denotes the neighboring atomic density of element s ($s \in \{A, B, \dots\}$) around atom i of element s_i .

Subsequently, the neighboring atomic density of element s around atom i is expanded in terms of a basis set, because the expansion enables the functional form to be replaced with a function of its expansion coefficients. For a given basis set

$\{b_n\}$, the neighboring atomic densities can be expanded as

$$\begin{aligned} \rho_{(s_i,A)}^{(i)}(\mathbf{r}) &= \sum_n a_{n,(s_i,A)}^{(i)} b_n(\mathbf{r}), \\ \rho_{(s_i,B)}^{(i)}(\mathbf{r}) &= \sum_n a_{n,(s_i,B)}^{(i)} b_n(\mathbf{r}), \\ &\vdots \end{aligned} \quad (3)$$

where $a_{n,(s_i,s)}^{(i)}$ denotes an order parameter characterizing the neighboring atomic density of element s around atom i of element s_i . Using the order parameters, the atomic energy may be rewritten as

$$E^{(i)} = F'_{s_i}(a_{1,(s_i,A)}^{(i)}, a_{1,(s_i,B)}^{(i)}, \dots, a_{2,(s_i,A)}^{(i)}, a_{2,(s_i,B)}^{(i)}, \dots). \quad (4)$$

Although the function F' of Eq. (4) depends on the element of atom i , it is convenient to introduce a unified function that is independent of the element by combining functions F' for all elements. In this study, a unified function for the atomic energy is formulated using order parameters defined for unordered pairs of elements. This means that the order parameter $a_{n,(s_i,s_j)}^{(i)}$ of atom i and its swapped order parameter $a_{n,(s_j,s_i)}^{(j)}$ of atom j are considered as the same variable in the unified function. They are represented by $a_{n,\{s_i,s_j\}}^{(i)}$ and $a_{n,\{s_i,s_j\}}^{(j)}$, respectively. Defining the order parameter $a_{n,\{s_1,s_2\}}^{(i)}$ to be zero if s_i is not included in $\{s_1, s_2\}$ ($s_i \notin \{s_1, s_2\}$), the atomic energy is written in an independent form of the element of atom i as

$$\begin{aligned} E^{(i)} &= F''(a_{1,\{A,A\}}^{(i)}, a_{1,\{A,B\}}^{(i)}, a_{1,\{B,B\}}^{(i)}, \dots, a_{2,\{A,A\}}^{(i)}, \\ &a_{2,\{A,B\}}^{(i)}, a_{2,\{B,B\}}^{(i)}, \dots), \end{aligned} \quad (5)$$

where all combinations with the replacement of elements are enumerated for each n .

Moreover, an arbitrary rotation leaves the atomic energy invariant, although it generally changes the neighboring atomic densities and their order parameters. Therefore, the atomic energy is required to be a function of O(3) invariants $\{d_n^{(i)}\}$ as

$$E^{(i)} = F(d_1^{(i)}, d_2^{(i)}, \dots), \quad (6)$$

where the invariants are derived from the order parameters $\{a_{n,\{s_1,s_2\}}^{(i)}\}$. Hereafter, the invariants representing the neighboring atomic density are referred to as structural features. The present formulation is useful for deriving potential energy models, the accuracy and computational efficiency of which can be controlled by the selections of structural features and the function F .

B. Structural features

In this study, the neighboring atomic densities are expanded in terms of a basis set composed of radial functions or a basis set composed of products of radial functions and spherical harmonics [24,28], although it is also possible to use other basis sets in principle. When the neighboring atomic density is expanded in terms of radial functions $\{f_n\}$, the neighboring atomic density around atom i is expressed as

$$\rho_{(s_i,s)}^{(i)}(\mathbf{r}) = \sum_n a_{n,\{s_i,s\}}^{(i)} f_n(\mathbf{r}), \quad (7)$$

where r denotes the distance from the position of atom i . Since the order parameter $a_{n,\{s_i,s\}}^{(i)}$ is invariant for the O(3) group, it can be a pairwise structural feature defined as

$$d_{n0,t}^{(i)} = a_{n,t}^{(i)}, \quad (8)$$

where t identifies the unordered pair of elements, i.e., $t \in \{\{A, A\}, \{A, B\}, \dots\}$.

When the neighboring atomic density is expanded in terms of products of radial functions $\{f_n\}$ and spherical harmonics $\{Y_{lm}\}$, the neighboring atomic density of element s at a position (r, θ, ϕ) in spherical coordinates centered at the position of atom i is expressed as

$$\rho_{(s_i,s)}^{(i)}(r, \theta, \phi) = \sum_{nlm} a_{nlm,\{s_i,s\}}^{(i)} f_n(r) Y_{lm}(\theta, \phi), \quad (9)$$

where the order parameter $a_{nlm,\{s_i,s\}}^{(i)}$ is the component nlm of the neighboring atomic density. Since the order parameters are not generally invariant for the O(3) group, polynomial invariants of the order parameters are adopted as structural features. A p th-order polynomial invariant for a radial index n and a set of pairs composed of the angular number and the element unordered pair $\{(l_1, t_1), (l_2, t_2), \dots, (l_p, t_p)\}$ is defined as a linear combination of products of p order parameters, expressed as

$$d_{n l_1 l_2 \dots l_p, t_1 t_2 \dots t_p, (\sigma)}^{(i)} = \sum_{m_1, m_2, \dots, m_p} c_{m_1 m_2 \dots m_p}^{l_1 l_2 \dots l_p, (\sigma)} a_{n l_1 m_1, t_1}^{(i)} a_{n l_2 m_2, t_2}^{(i)} \dots a_{n l_p m_p, t_p}^{(i)}. \quad (10)$$

Nonzero polynomial invariants are derived from the sets satisfying the condition that the intersection of t_p is not the empty set,

$$\prod_p t_p \neq \emptyset. \quad (11)$$

Therefore, possible sets of pairs composed of the angular number and the element unordered pair $\{(l_1, t_1), (l_2, t_2), \dots, (l_p, t_p)\}$ are enumerated for a given maximum angular number to obtain the entire set of polynomial invariants. An example where the intersection of element pairs becomes the empty set is the case that a polynomial invariant is composed of order parameters with $t_1 = \{A, A\}$ and $t_2 = \{B, B\}$. Such polynomial invariants are eliminated.

A coefficient set $\{c_{m_1 m_2 \dots m_p}^{l_1 l_2 \dots l_p, (\sigma)}\}$ is independent of the radial index n and the element unordered pair t . Therefore, linearly independent coefficient sets are obtained using the group-theoretic projector operation method for a given set $\{l_1, l_2, \dots, l_p\}$, as proposed in Ref. [24], ensuring that the linear combinations are invariant for arbitrary rotation. In terms of fourth- and higher-order polynomial invariants, multiple invariants are linearly independent for most of the set $\{l_1, l_2, \dots, l_p\}$, which are distinguished by index σ if necessary. Note that the second- and third-order invariants are equivalent to a multicomponent extension of the angular Fourier series and the bispectrum reported in the literature, respectively [28,29].

In this study, a finite set of Gaussian-type functions is adopted as radial functions in basis sets to expand the neigh-

boring atomic density, expressed as

$$f_n(r) = \exp[-\beta_n(r - r_n)^2] f_c(r), \quad (12)$$

where β_n and r_n are parameters. The cutoff function f_c is given by a cosine-based function as

$$f_c(r) = \begin{cases} \frac{1}{2} [\cos(\pi \frac{r}{r_c}) + 1] & (r \leq r_c) \\ 0 & (r > r_c). \end{cases} \quad (13)$$

The order parameters of atom i and element pair $\{s_i, s\}$ are approximately estimated from the neighboring atomic density of element s around atom i as

$$a_{nlm,\{s_i,s\}}^{(i)} = \sum_{\{j|r_{ij} \leq r_c, s_j=s\}} f_n(r_{ij}) Y_{lm}^*(\theta_{ij}, \phi_{ij}), \quad (14)$$

where $(r_{ij}, \theta_{ij}, \phi_{ij})$ denotes the spherical coordinates of neighboring atom j centered at the position of atom i . Although the Gaussian-type radial functions are not orthonormal, such an approximation of the order parameters is acceptable in the present polynomial-based framework, as also discussed in Ref. [24].

C. Polynomial models for atomic energy

Polynomial models are here employed to represent the atomic energy as a function of structural features. Given a set of structural features $D = \{d_1, d_2, \dots\}$, polynomial functions are written as

$$f_1(D) = \sum_i w_i d_i,$$

$$f_2(D) = \sum_{\{i',j'\}} w_{i'j'} d_{i'} d_{j'},$$

$$f_3(D) = \sum_{\{i',j',k'\}} w_{i'j'k'} d_{i'} d_{j'} d_{k'},$$

$$\vdots \quad (15)$$

where w denotes a regression coefficient. Although the polynomial functions are described by all combinations of structural features, only nonzero polynomial terms are retained, which is analogous to the enumeration of nonzero polynomial invariants. In a multicomponent system, a structural feature is composed of order parameters, each of which has an attribute on the element unordered pair t . Therefore, when the element pair of the p 'th-order parameter in the structure feature $d_{i'}$ of a polynomial term is denoted by $t_{i',p'}$, a nonzero polynomial term satisfies the condition that the intersection of $\{t_{i',p'}\}$ is not the empty set:

$$\left(\prod_{p'} t_{i',p'} \right) \cap \left(\prod_{p'} t_{j',p'} \right) \cap \dots \neq \emptyset. \quad (16)$$

For example, the intersection of element pairs becomes the empty set if a polynomial term is composed of structural features with $t_{i',p'_i} = \{A, A\}$ and $t_{j',p'_j} = \{B, B\}$. Such polynomial terms are eliminated from the polynomial functions.

The following polynomial models for the atomic energy are systematically applied to obtain Pareto optimal MLPs as described in Sec. IV A. The first model is a polynomial of pairwise structural features. When a set of pairwise structural

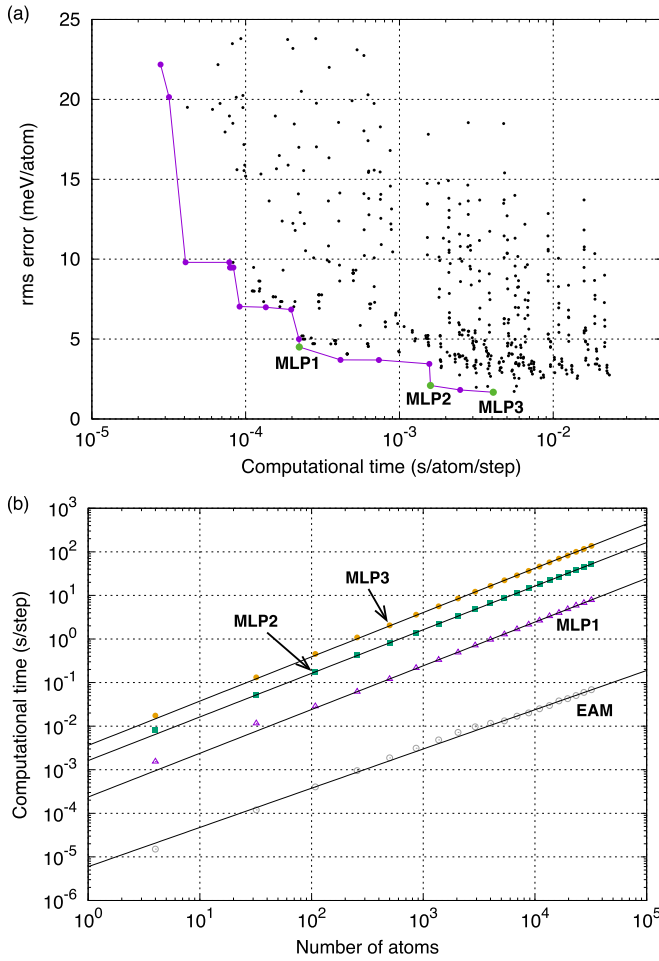


FIG. 2. (a) Distribution of MLPs for the Ti-Al binary system. The purple closed circles show Pareto optimal points of the distribution with different trade-offs between accuracy and computational efficiency. The green closed circles indicate the MLP with the lowest prediction error denoted by MLP3 and the Pareto optimal MLPs with high computational cost performance denoted by MLP1 and MLP2. The computational time indicates the elapsed time for a single-point calculation normalized by the number of atoms. The elapsed time is measured using a single core of Intel® Xeon® E5-2695 v4 (2.10 GHz). (b) Dependence of the computational time required for a single-point calculation on the number of atoms. The dependence of the computational time using the EAM potential [41] is also shown for comparison.

features is described as

$$D_{\text{pair}}^{(i)} = \{d_{n0,t}^{(i)}\}, \quad (17)$$

the first model is expressed as

$$E^{(i)} = f_1(D_{\text{pair}}^{(i)}) + f_2(D_{\text{pair}}^{(i)}) + f_3(D_{\text{pair}}^{(i)}) + \dots \quad (18)$$

The first model includes the special case that only powers of the pairwise structural features are considered, which was introduced for elemental systems in Refs. [12,13]. The first model can also be regarded as a straightforward extension of embedded-atom method (EAM) potentials [14].

The second model for the atomic energy is a linear polynomial of polynomial invariants given by Eq. (10). A set of

polynomial invariants is described as

$$D^{(i)} = D_{\text{pair}}^{(i)} \cup D_2^{(i)} \cup D_3^{(i)} \cup D_4^{(i)} \cup \dots, \quad (19)$$

where a set of p th-order polynomial invariants is given by

$$\begin{aligned} D_2^{(i)} &= \{d_{nl,t_1t_2}^{(i)}\}, \\ D_3^{(i)} &= \{d_{nl_1l_2l_3,t_1t_2t_3}^{(i)}\}, \\ D_4^{(i)} &= \{d_{nl_1l_2l_3l_4,t_1t_2t_3t_4,(\sigma)}^{(i)}\}. \end{aligned} \quad (20)$$

A second-order invariant is identified with a single l value because second-order linear combinations are invariant only when $l_1 = l_2$ [33,34]. The second model is then written as

$$E^{(i)} = f_1(D^{(i)}), \quad (21)$$

which was introduced in Ref. [24] for elemental systems. Note that a linear polynomial model with up to third-order invariants is equivalent to a spectral neighbor analysis potential (SNAP) [15], expressed as

$$E^{(i)} = f_1(D_{\text{pair}}^{(i)} \cup D_2^{(i)} \cup D_3^{(i)}). \quad (22)$$

An extension of the second model is a polynomial of polynomial invariants described as

$$E^{(i)} = f_1(D^{(i)}) + f_2(D^{(i)}) + f_3(D^{(i)}) + \dots \quad (23)$$

Note that a quadratic polynomial model of polynomial invariants up to the third order is equivalent to a quadratic SNAP [16]. Other extended models are also introduced, which are given by

$$\begin{aligned} E^{(i)} &= f_1(D^{(i)}) + f_2(D_{\text{pair}}^{(i)}) + f_3(D_{\text{pair}}^{(i)}), \\ E^{(i)} &= f_1(D^{(i)}) + f_2(D_{\text{pair}}^{(i)} \cup D_2^{(i)}), \\ E^{(i)} &= f_1(D^{(i)}) + f_2(D_{\text{pair}}^{(i)} \cup D_2^{(i)} \cup D_3^{(i)}). \end{aligned} \quad (24)$$

They are decomposed into a linear polynomial of structural features and a polynomial of a subset of the structural features.

The atomic energy in all the models is measured from the sum of the energies of isolated atoms. Moreover, the forces acting on atoms and the stress tensors in a multicomponent system can be derived in a manner similar to those in the elemental system derived in Ref. [24]. Note that the above polynomial function forms were also applied to develop MLPs for elemental systems included in the Machine Learning Potential Repository [31,32].

III. DATA SETS AND COMPUTATIONAL PROCEDURE

A straightforward development of training and test data sets for the Ti-Al binary system begins with a set of structure generators. In this study, prototype structures reported as

TABLE I. List of structure generators used for developing the training and test data sets.

ICSD CollCode	Structure type	ICSD CollCode	Structure type	ICSD CollCode	Structure type
239	Cu ₃ Se ₂	106786	Hg ₂ Pt	618295	MoC _{1-x}
5258	FeSi ₂	107998	MoNi ₄	618702	ScTe
16504	CrSi ₂	108707	HgMn	625334	Laves(2H)-MgZn ₂
16606	Nb ₃ Te ₄	108762	Hg ₄ Pt	626692	NiAs
30446	Fe ₂ B	150584	Fe ₁₃ Ge ₃	629380	Al ₃ Os ₂
42428	Fe ₃ Pt	155842	Co ₅ Fe ₁₁	629406	Cu ₄ Ti ₃
42472	CoO	161109	CoSn	633467	FeSe(<i>rP2</i>)
42773	IrGe ₄	161133	Fe ₂ Si(HT)	635060	FeSi
52294	GeTe	167735	Ru ₂ B ₃	635208	CoGa ₃
55492	BaPt	168897	LaI	635642	Hg ₅ Mn ₂
58471	CuZr ₂	169457	ZrH ₂	638227	CaF ₂
58607	Au ₂ Ti	181127	AuCu ₃	639037	HgIn
58745	Fe ₆ Ge ₆ Mg	181788	NaCl	639148	NiHg ₄
59508	AuCu	185626	Al ₃ Ni ₂	639227	Si ₂ U ₃
59586	Pd ₅ Th ₃	188260	Heusler-AlCu ₂ Mn	639879	In ₅ In ₄
69199	U ₃ Si	189695	CuHg ₂ Ti	640726	CuSmP ₂
69557	CdI ₂ (<i>hP9</i>)	189711	Heusler-AlLiSi	643301	Au ₃ Cd
73839	Ni ₃ S ₂	240119	AlLi	644708	WC
97006	InMg ₂	246555	Co ₂ Nd	648572	CuInPt ₂
99787	Fe ₃ Pt	248490	Pt ₂ Si	648748	Pd ₄ Se
100654	BiSe	260285	UCl ₃	649037	Ni ₃ Ti
102712	CoU	262070	AlLi(<i>hP8</i>)	650527	CsCl
103775	NaTl	409859	La ₂ Sb	652553	AlCr ₂ -MoSi ₂
103995	Ga ₃ Ti ₂	416747	Al ₃ Zr	655706	Cu ₂ Te(HT)
104506	Ni ₃ Sn	420250	LiPd ₂ Tl	659806	GeTe
105191	Al ₃ Ti	424636	MnGa ₄	659829	Al ₂ Li ₃
105521	Al ₅ W	609153	AlPt ₃	659856	LiPt
105636	PbU	610464	PbClF/Cu ₂ Sb		
105726	Pd ₅ Ti ₃	611176	Fe ₂ P		
105948	InNi ₂	611457	NbAs		
106325	BiIn	611618	TiAs		

binary alloy entries in the Inorganic Crystal Structure Database (ICSD) [35] are used as structure generators such that the data sets can cover a wide variety of structures. Moreover, the prototype structures are restricted to those represented by unit cells with up to eight atoms. A structure generator made by swapping elements in each prototype structure is also introduced; hence, the total number of structure generators is 150. The structure generators are listed in Table I.

Given the structure generators, the atomic positions and lattice constants of the structure generators are fully optimized by DFT calculation to obtain their equilibrium structures. A structure in the data sets is then generated by introducing random lattice expansion, random lattice distortion, and random atomic displacements into a supercell of each of the equilibrium structures. A mathematical description of the procedure can be found in Ref. [24]. By repeatedly applying the procedure to the structure generators, 27 394 binary structures are generated for the data sets. In addition to the binary structures and the equilibrium structures of the structure generators, existing data for elemental Ti and Al [31] are also included in the present data sets.

For a total of 41 508 structures, DFT calculations are performed using the plane-wave-basis projector augmented wave method [36] within the Perdew-Burke-Ernzerhof exchange-

correlation functional [37] as implemented in the VASP code [38–40]. The cutoff energy is set to 300 eV. The total energies converge to less than 10^{-3} meV/supercell. The atomic positions and lattice constants for the structure generators are optimized until the residual forces are less than 10^{-2} eV/Å.

The regression coefficients of a potential energy model are estimated by linear ridge regression. The DFT total energies, the DFT forces acting on atoms, and the DFT stress tensors of structures in the training data set are simultaneously used to estimate the regression coefficients, as adopted in Refs. [24,31]. Therefore, the total number of training data reaches 5 178 510.

IV. RESULTS AND DISCUSSION

A. Pareto optimal MLPs

The polynomial models given in Sec. II C are systematically applied to develop MLPs in the Ti-Al binary system. Because the accuracy and computational efficiency of the MLPs strongly depend on several input parameters, a systematic grid search is performed to find their optimal values. The parameters in the grid search are the cutoff radius, the type of structural features, the type of potential energy model, the number of radial functions, the polynomial order in the potential energy model, and the truncation of the polynomial

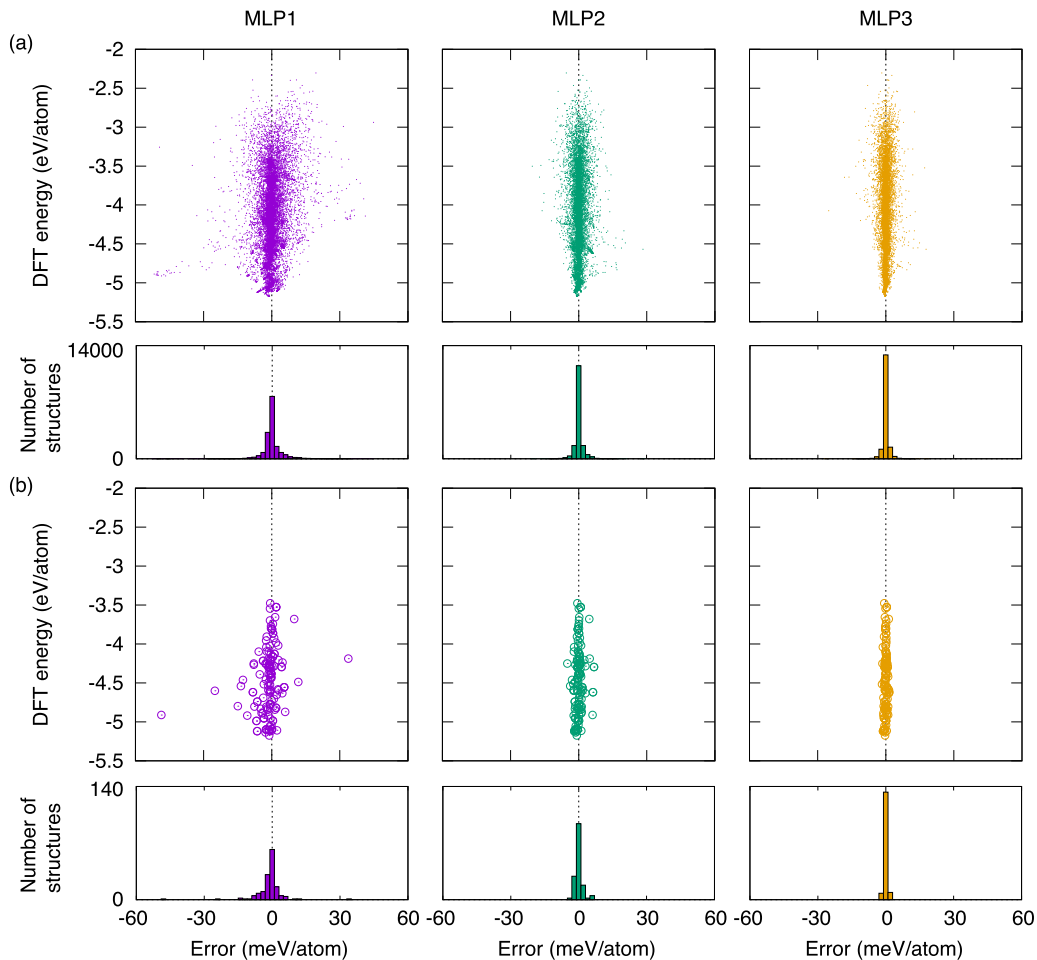


FIG. 3. (a) Distribution of the prediction errors for structures in the training and test data sets. Histograms with a bin size of 2 meV/atom are also shown. (b) Distribution of the prediction errors for the equilibrium structure generators listed in Table I.

invariants, i.e., the maximum angular numbers of spherical harmonics $\{I_{\max}^{(2)}, I_{\max}^{(3)}, \dots, I_{\max}^{(p_{\max})}\}$ and the polynomial order of the invariants p_{\max} .

Figure 2(a) shows the distribution of MLPs obtained from the grid search. The root mean square (rms) error for the test data set is used as an estimator of the accuracy of MLPs. The computational time indicates the elapsed time for a single-point calculation normalized by the number of atoms. Figure 2(a) also shows the Pareto optimal MLPs when optimizing both the accuracy and computational efficiency simultaneously. As can be seen in Fig. 2(a), the accuracy and

computational efficiency of MLPs are conflicting properties; hence the Pareto optimal MLPs can be optimal ones with different trade-offs.

In performing an atomistic simulation, an appropriate MLP must be chosen from the Pareto optimal ones according to its target system and purpose. Therefore, a convenient score that can estimate the computational cost performance is required to find an MLP with high computational cost performance in a simplified manner. In this study, functions $t + \Delta E$ and $10t + \Delta E$ that should be minimized are introduced, where t and ΔE denote the computational time in units of ms/atom/step and the rms error in units of meV/atom, respectively.

Another approach to choosing an appropriate MLP is to examine the convergence behavior of the target property in terms of the computational time using the whole set of Pareto optimal MLPs. For example, the MLP that requires the lowest computational time among the MLPs showing convergence for the grain boundary energy can accurately predict the grain boundary energies of larger models [42]. Therefore, in Sec. IV G the convergence behavior of the stacking fault energy in terms of the computational time will be shown.

Figure 2(a) shows the MLP with the lowest rms error and two Pareto optimal MLPs showing high computational cost performance. The MLP with the lowest rms error is

TABLE II. Model parameters and prediction errors of MLP1, MLP2, and MLP3 for the Ti-Al binary system.

Parameter	MLP1	MLP2	MLP3
number of coefficients	7875	27,520	61,605
rms error (energy, meV/atom)	4.51	2.09	1.67
rms error (force, eV/Å)	0.138	0.074	0.066
cutoff radius (Å)	8.0	6.0	8.0
number of radial functions	10	10	15
polynomial order (function F)	2	2	2
polynomial order (invariants)	3	2	2
$\{I_{\max}^{(2)}, I_{\max}^{(3)}, \dots\}$	[0,0]	[4]	[4]

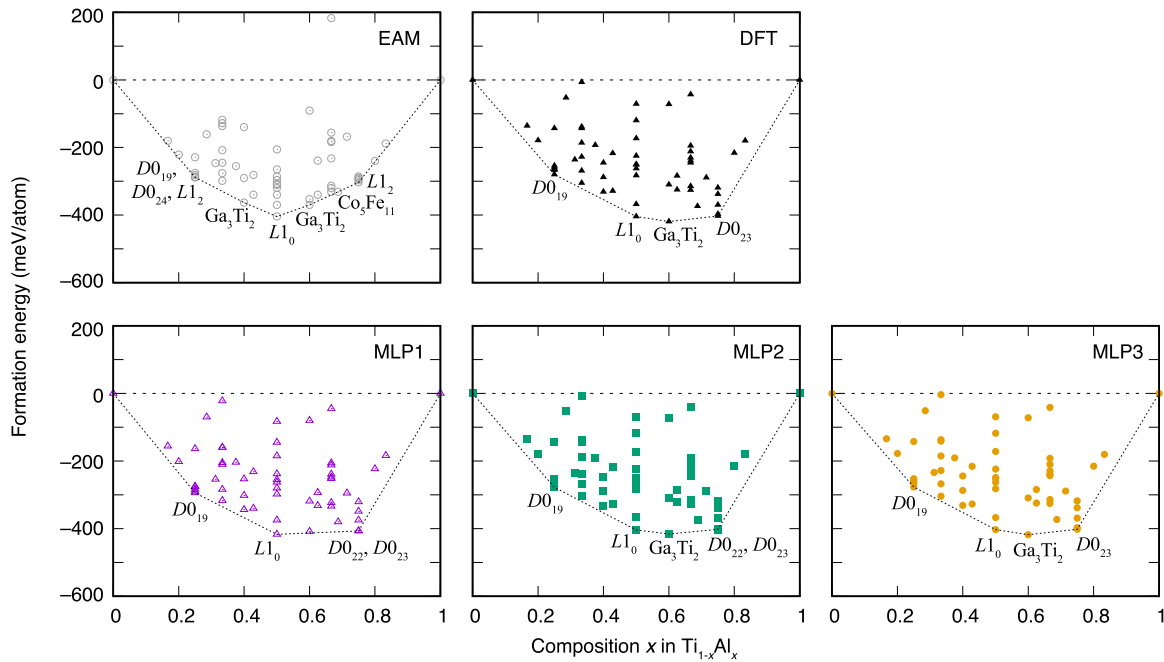


FIG. 4. Formation energies of selected ordered structures. The convex hull obtained from the formation energies of selected ordered structures is also shown.

denoted by MLP3, whereas the two MLPs showing high computational cost performance are denoted by MLP1 and MLP2; MLP1 and MLP2 are obtained by minimizing the scores $10t + \Delta E$ and $t + \Delta E$, respectively. As can be seen in Fig. 2(a), they exhibit high computational efficiency without significantly increasing the rms error. The model parameters and rms errors of MLP1, MLP2, and MLP3 are listed in Table II. The rms errors of MLP2 and MLP3 are close, 2.09 and 1.67 meV/atom, respectively. The rms error of MLP1 is 4.51 meV/atom, which is greater than those of MLP2 and MLP3, whereas MLP1 is ten times more computationally efficient than MLP2. As described in Table II, all the MLPs are derived from polynomial invariants. However, MLP1 is developed only using the $l = 0$ component of spherical harmonics. Therefore, MLP1 can be regarded as a polynomial model of pairwise structural features.

The present MLPs are derived from order parameters, given by Eq. (9), depending on both the central and neighboring element types. Therefore, they are compared with MLPs derived from order parameters depending only on the central element type, commonly used to develop MLPs in multicomponent systems. When developing MLPs using order parameters depending only on the central element type, the rms errors of MLP1, MLP2, and MLP3 models with 990, 3780, and 8370 regression coefficients are 48.79, 42.80, and 30.10 meV/atom, respectively. They are much greater than those of MLPs derived from order parameters depending on both the central and neighboring element types. This indicates that order parameters depending on the neighboring element type are useful to develop accurate MLPs in multicomponent systems. Note that the computational efficiency of an MLP with order parameters depending only on the central element type is almost the same as that of the corresponding present MLP.

Figure 2(b) shows the computational time required for a single-point calculation using the EAM potential [41], MLP1, MLP2, and MLP3. The computational time is evaluated for structures with up to 32 000 atoms constructed by the expansions of the four-atom unit cell of a AuCu-type ($L1_0$) structure with a lattice constant of 4 Å. The computational time is measured by implementing the present MLPs [43] in the LAMMPS code [44,45]. As can be seen in Fig. 2(b), the EAM potential and the MLPs show linear scaling of the computational time with respect to the number of atoms. Therefore, the computational time normalized by the number of atoms can be an estimator of the computational efficiency of MLPs and the computational time required for a simulation of n_{step} steps using a structure with n_{atom} atoms can be estimated as $t \times n_{\text{atom}} \times n_{\text{step}}$.

As described above, the accuracy and computational efficiency of the MLPs strongly depend on several input parameters. Therefore, the Pareto optimal MLPs also depend on the grid used for the grid search. The dependence of the Pareto optimal MLPs on the grid, i.e., the maximum value of the input parameter used for the grid search, is shown in Appendix A. Also, the accuracy of the MLPs depends on the size of the training data set. Hence, the learning curve or the dependence of the accuracy on the training data-set size is shown in Appendix B.

B. Cohesive energy

Figure 3(a) shows the distribution of the prediction errors for structures in the training and test data sets. The degree of scattering in the distribution decreases in the order of MLP1, MLP2, and MLP3, which coincides with the decreasing order of the rms errors shown in Table II. Figure 3(b) shows the distribution of the prediction errors for the equilibrium structure

generators listed in Table I. The distribution of the prediction errors for the structure generators indicates the predictive power for the cohesive energy. The prediction errors for some of the structure generators are significant in MLP1. On the other hand, the prediction errors are trivial for all the structure generators in MLP2 and MLP3, which indicates that MLP2 and MLP3 should have high predictive power for the cohesive energy in a wide variety of binary ordered structures.

C. Formation energy

The formation energy of a given ordered structure is more challenging to predict accurately than its cohesive energy because high predictive power is required not only for the ordered structure but also for the reference structures of hexagonal-close-packed Ti and face-centered-cubic Al. Furthermore, the local geometry relaxation for the ordered structure is crucial for evaluating the formation energy. Therefore, MLPs are needed to derive an accurate potential energy surface around the initial and equilibrium structures.

Figure 4 shows the formation energies of selected ordered structures predicted using the EAM potential, MLP1, MLP2, and MLP3, compared with those predicted by DFT calculation. The values of the formation energy are listed in Tables III and IV. Note that the ordered structures cover a wide range of compositions. The rms errors of the EAM potential, MLP1, MLP2, and MLP3 for the formation energy are 68.0, 13.8, 2.0, and 1.5 meV/atom, respectively. This reveals that MLP2 and MLP3 have high predictive power for the formation energy in a wide range of structures. The rms error of MLP1 for the formation energy is significant as a consequence of the systematic deviation of the formation energy for the overall ordered structures. As can be seen in Tables III and IV, the formation energies of most of the ordered structures predicted using MLP1 are approximately 10 meV/atom lower than those predicted by DFT calculation, which originates from the fact that the prediction error of MLP1 for hexagonal-close-packed Ti is significant (+27.4 meV/atom). Moreover, the prediction error of the EAM potential is much more significant and the EAM potential fails to reconstruct the hierarchy of the formation energies predicted by DFT calculation.

D. Elastic constants

Tables V–VII show the lattice constants and the elastic constants of nine structures, i.e., TiAl ($L1_0$), TiAl ($B2$), TiAl ($B8_1$), Ti₃Al ($D0_{19}$), Ti₃Al ($D0_{22}$), Ti₃Al ($L1_2$), TiAl₃ ($D0_{22}$), TiAl₃ ($D0_{19}$), and TiAl₃ ($L1_2$), which are predicted using the EAM potential, MLP1, MLP2, and MLP3, along with the lattice constants and the elastic constants obtained by DFT calculation. The lattice constants predicted using the EAM potential deviate from those obtained by DFT calculation in most of the structures. This deviation may arise from the fact that the EAM potential was developed by fitting to experimental lattice constants [41], excluding the descriptive power of the EAM potential. In contrast, the MLPs can compute the lattice constants accurately in all of the structures. Regarding the elastic constants, the elastic constants predicted using the EAM potential and MLP1 are close to those obtained by DFT calculation in many structures; however, the EAM potential and MLP1 fail to predict the elastic constants accurately in

TABLE III. Formation energies of selected ordered structures (units of meV/atom). The values of the formation energy closest to those obtained by the DFT are highlighted in boldface.

Structure type	EAM ^a	MLP1	MLP2	MLP3	DFT
		Ti ₅ Al			
Al ₅ W	-180	-156	-135	-134	-136
		Ti ₄ Al			
MoNi ₄ ($D1_a$)	-221	-202	-180	-178	-179
		Ti ₃ Al			
Ni ₃ Sn ($D0_{19}$)	-288	-294	-278	-277	-280
Ni ₃ Ti ($D0_{24}$)	-288	-290	-267	-265	-267
AuCu ₃ ($L1_2$)	-288	-283	-263	-262	-264
Al ₃ Zr ($D0_{23}$)	-282	-276	-258	-257	-259
Al ₃ Ti ($D0_{22}$)	-275	-275	-253	-252	-254
AlCu ₂ Mn ($L2_1$)	-229	-164	-143	-142	-143
		Ti ₅ Al ₂			
Hg ₅ Mn ₂	-161	-70	-53	-51	-53
		Ti ₁₁ Al ₅			
Co ₅ Fe ₁₁	-246	-254	-235	-234	-236
		Ti ₂ Al			
InNi ₂ ($B8_2$)	-208	-317	-305	-304	-305
CuZr ₂	-299	-284	-268	-267	-269
Fe ₂ P ($C22$)	-276	-205	-237	-228	-228
CrSi ₂ ($C40$)	-245	-210	-187	-185	-186
Fe ₂ B	-137	-160	-143	-141	-143
Cu ₂ Sb ($C38$)	-128	-160	-139	-137	-139
FeSi ₂	-118	-22	-7	-4	-7
		Ti ₅ Al ₃			
Pd ₅ Th ₃	-255	-204	-192	-191	-193
		Ti ₃ Al ₂			
Ga ₃ Ti ₂	-363	-344	-332	-331	-331
Al ₃ Os ₂	-290	-302	-288	-287	-288
Si ₂ U ₃ ($D5_a$)	-140	-253	-246	-244	-245
		Ti ₄ Al ₃			
Nb ₃ Te ₄	-340	-340	-327	-327	-329
Cu ₄ Ti ₃	-281	-232	-217	-216	-217
		TiAl			
AuCu ($L1_0$)	-404	-417	-404	-403	-404
PbU	-370	-375	-366	-367	-368
CoU (B_a)	-340	-298	-283	-282	-283
CsCl ($B2$)	-286	-280	-265	-263	-262
NiAs ($B8_1$)	-311	-263	-257	-251	-251
WC (B_h)	-296	-255	-250	-249	-250
ScTe	-307	-238	-223	-224	-225
BiSe	-265	-185	-172	-172	-173
FeSi ($B20$)	-206	-145	-118	-118	-120
NaTl ($B32$)	-319	-83	-70	-69	-71
		Ti ₂ Al ₃			
Ga ₃ Ti ₂	-370	-408	-416	-418	-419
Al ₃ Os ₂	-353	-319	-309	-309	-310
Si ₂ U ₃ ($D5_a$)	-91	-80	-73	-72	-72

^aReference [41].

a few structures. Also, the elastic constants predicted using MLP2 and MLP3 are almost the same as those obtained by DFT calculation in all structures.

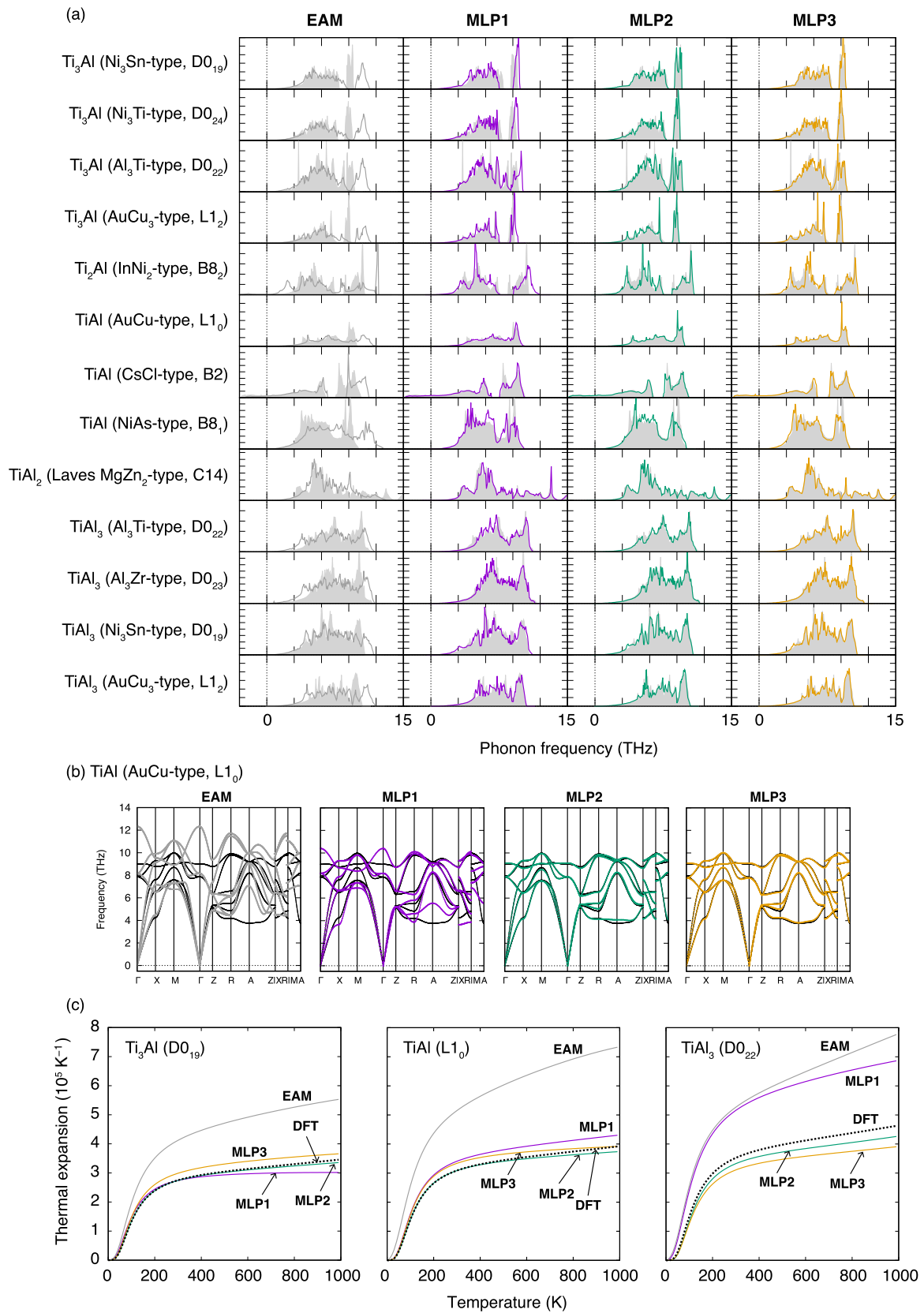


FIG. 5. (a) Phonon density of states for 13 selected compounds in the Ti-Al binary system, predicted using the EAM potential and the MLPs. The shaded region indicates the phonon density of states computed by DFT calculation. (b) Phonon dispersion curves for γ -TiAl predicted using the EAM potential and the MLPs. The solid black lines indicate the phonon dispersion curves predicted by the DFT calculation. (c) Temperature dependence of the thermal expansion predicted using the EAM potential and the MLPs. The dotted black lines show the thermal expansion computed by the DFT calculation.

TABLE IV. Formation energies of selected ordered structures (units of meV/atom).

Structure type	EAM ^a	MLP1	MLP2	MLP3	DFT
Ti ₃ Al ₅					
Pd ₅ Th ₃	-239	-332	-322	-324	-325
Pd ₅ Ti ₃	-340	-293	-285	-284	-284
TiAl ₂					
MgZn ₂ (<i>C14</i>)	-331	-334	-326	-325	-326
Co ₂ Nd	-313	-323	-315	-314	-315
CuZr ₂	-320	-252	-244	-243	-245
La ₂ Sb	-181	-248	-238	-238	-239
Cu ₂ Sb (<i>C38</i>)	-184	-239	-232	-229	-230
FeSi ₂	-156	-212	-210	-210	-212
Fe ₂ P (<i>C22</i>)	-232	-206	-191	-192	-195
Hg ₂ Pt	183	-45	-40	-41	-43
Ti ₅ Al ₁₁					
Co ₅ Fe ₁₁	-331	-380	-373	-373	-374
Ti ₂ Al ₅					
Hg ₅ Mn ₂	-168	-295	-289	-289	-289
TiAl ₃					
Al ₃ Zr (<i>D0₂₃</i>)	-297	-407	-402	-402	-403
Al ₃ Ti (<i>D0₂₂</i>)	-289	-407	-403	-397	-397
AuCu ₃ (<i>L1₂</i>)	-303	-375	-369	-369	-370
Ni ₃ Ti (<i>D0₂₄</i>)	-293	-349	-338	-338	-338
Ni ₃ Sn (<i>D0₁₉</i>)	-286	-321	-318	-318	-319
TiAl ₄					
MoNi ₄ (<i>D1_a</i>)	-240	-223	-216	-215	-216
TiAl ₅					
Al ₅ W	-188	-184	-180	-180	-180

^aReference [41].

E. Phonon properties

The predictive power of the MLPs for phonon properties and thermal expansion is examined. The phonon properties and thermal expansion are calculated using a finite dis-

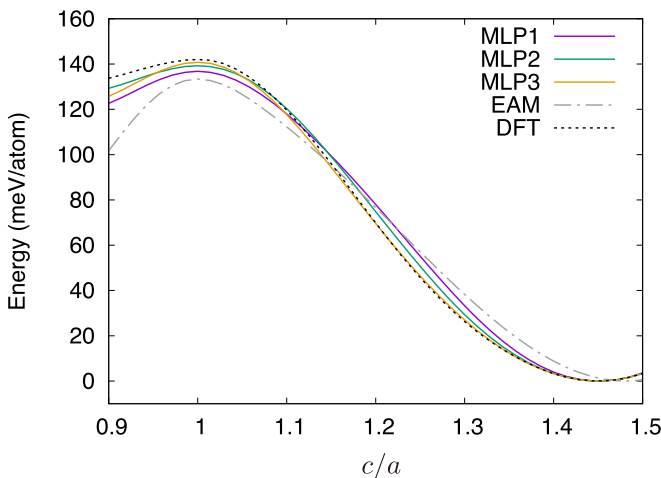


FIG. 6. Energy profiles along the Bain path between the *L1₀* structure and the *B2* structure.

TABLE V. Lattice constants and elastic constants of TiAl (*L1₀*, *B2*, and *B8₁*).

Constant	EAM ^a	MLP1	MLP2	MLP3	DFT
γ -TiAl (CuAu, <i>L1₀</i>)					
a_0 (Å)	2.827	2.812	2.812	2.812	2.813
c_0 (Å)	4.187	4.080	4.078	4.080	4.079
C_{11} (GPa)	237	219	195	190	195
C_{12} (GPa)	67	35	63	65	66
C_{13} (GPa)	114	87	90	90	89
C_{33} (GPa)	213	189	176	176	173
C_{44} (GPa)	92	112	114	114	113
C_{66} (GPa)	45	52	39	39	38
TiAl (CsCl, <i>B2</i>)					
a_0 (Å)	3.278	3.184	3.183	3.182	3.182
C_{11} (GPa)	80	67	82	69	74
C_{12} (GPa)	121	132	134	141	136
C_{44} (GPa)	95	80	87	82	66
TiAl (NiAs, <i>B8₁</i>)					
a_0 (Å)	2.853	2.880	2.878	2.877	2.879
c_0 (Å)	9.370	9.269	9.264	9.276	9.263
C_{11} (GPa)	157	149	126	132	136
C_{12} (GPa)	94	113	99	94	96
C_{13} (GPa)	93	88	65	72	74
C_{33} (GPa)	292	277	250	220	223
C_{44} (GPa)	67	73	73	71	75
C_{66} (GPa)	32	18	14	19	20

^aReference [41].

TABLE VI. Lattice constants and elastic constants of Ti₃Al (*D0₁₉*, *D0₂₂*, and *L1₂*).

Constant	EAM ^a	MLP1	MLP2	MLP3	DFT
Ti ₃ Al (Ni ₃ Sn, <i>D0₁₉</i>)					
a_0 (Å)	5.784	5.728	5.731	5.729	5.726
c_0 (Å)	4.750	4.646	4.643	4.644	4.646
C_{11} (GPa)	199	187	181	196	195
C_{12} (GPa)	89	113	104	99	90
C_{13} (GPa)	74	61	67	71	70
C_{33} (GPa)	224	237	235	231	232
C_{44} (GPa)	51	44	47	63	59
C_{66} (GPa)	55	37	39	48	53
Ti ₃ Al (TiAl ₃ , <i>D0₂₂</i>)					
a_0 (Å)	4.082	3.943	3.961	3.962	3.960
c_0 (Å)	8.252	8.479	8.423	8.433	8.444
C_{11} (GPa)	161	192	173	177	173
C_{12} (GPa)	100	145	105	92	97
C_{13} (GPa)	93	119	101	89	87
C_{33} (GPa)	146	260	168	151	171
C_{44} (GPa)	71	73	84	79	87
C_{66} (GPa)	78	98	94	98	89
Ti ₃ Al (Cu ₃ Au, <i>L1₂</i>)					
a_0 (Å)	4.089	4.036	4.036	4.036	4.035
C_{11} (GPa)	165	143	165	172	175
C_{12} (GPa)	97	102	91	89	89
C_{44} (GPa)	76	92	93	92	90

^aReference [41].

TABLE VII. Lattice constants and elastic constants of TiAl_3 ($D0_{22}$, $D0_{19}$, and $L1_2$).

Constant	EAM ^a	MLP1	MLP2	MLP3	DFT
TiAl_3 (TiAl_3 , $D0_{22}$)					
a_0 (Å)	4.049	3.838	3.842	3.842	3.841
c_0 (Å)	8.139	8.626	8.621	8.609	8.608
C_{11} (GPa)	170	190	213	188	196
C_{12} (GPa)	98	100	76	81	87
C_{13} (GPa)	89	2	26	38	47
C_{33} (GPa)	140	212	192	223	220
C_{44} (GPa)	62	70	83	83	95
C_{66} (GPa)	71	140	126	120	129
TiAl_3 (Ni_3Sn , $D0_{19}$)					
a_0 (Å)	5.704	5.565	5.565	5.564	5.563
c_0 (Å)	4.810	4.722	4.724	4.725	4.726
C_{11} (GPa)	205	219	203	206	209
C_{12} (GPa)	88	58	67	66	67
C_{13} (GPa)	62	53	63	59	60
C_{33} (GPa)	189	163	161	168	167
C_{44} (GPa)	34	46	53	62	65
C_{66} (GPa)	58	80	68	70	71
TiAl_3 (Cu_3Au , $L1_2$)					
a_0 (Å)	4.050	3.976	3.977	3.977	3.979
C_{11} (GPa)	179	183	181	190	191
C_{12} (GPa)	95	56	68	69	66
C_{44} (GPa)	73	84	77	76	77

^aReference [41].

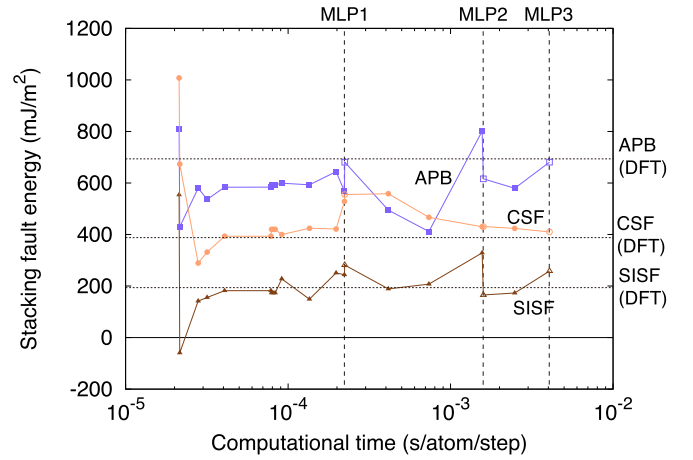
placement method implemented in the PHONOPY code [46]. Figure 5(a) shows the phonon density of states for 13 structures predicted using the EAM potential, MLP1, MLP2, and MLP3. The supercells required to compute the phonon properties are constructed by the expansions of the conventional unit cells of the structures. The number of atoms included in the supercells ranges from 54 to 162. The EAM potential predicts the phonon density of states well in the low-frequency region for many structures, while the deviation from the DFT phonon density of states is large in the high-frequency region. Conversely, the phonon density of states predicted using the MLPs and those predicted by DFT calculation overlap for all the structures, particularly those predicted using MLP2 and MLP3.

Figure 5(b) shows the phonon dispersion curves for TiAl (AuCu-type, $L1_0$) predicted using the EAM potential and the MLPs, compared with those predicted by the DFT calculation. The deviation of the EAM phonon dispersions from the DFT phonon dispersions is significant in the high-frequency region.

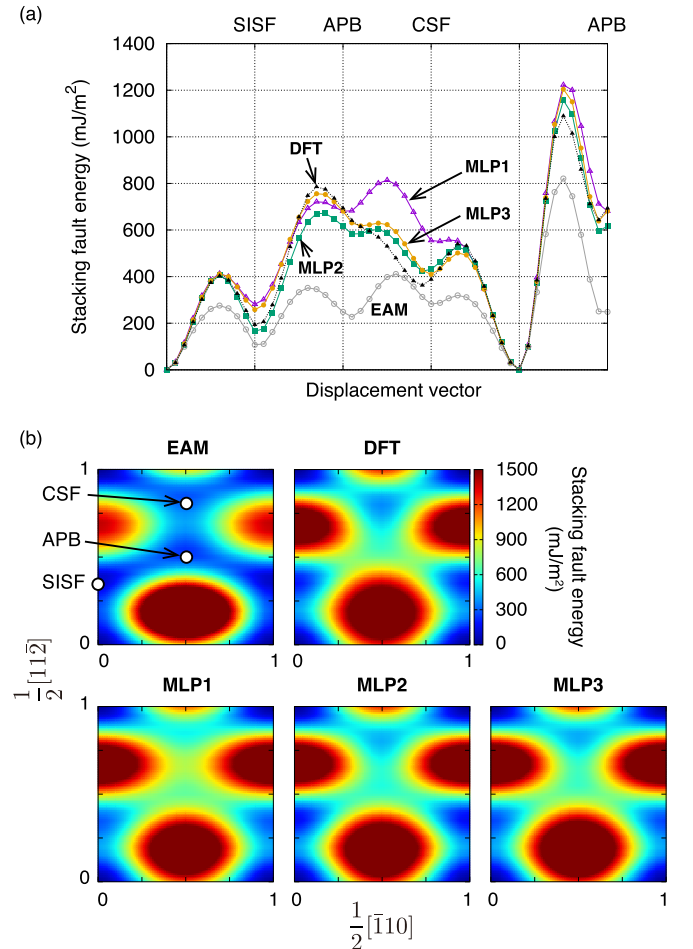
 TABLE VIII. Excessive energies of special stacking faults in γ -TiAl (units of mJ/m^2).

Stacking fault	EAM ^a	MLP1	MLP2	MLP3	DFT
SISF (111)	108	281	165	258	194
APB (111)	249	680	616	681	694
CSF (111)	282	555	431	410	388

^aReference [41].


 FIG. 7. Excessive energies of special stacking faults in γ -TiAl, predicted using the Pareto optimal MLPs.

On the other hand, the phonon dispersions of MLP2 and MLP3 are consistent with those of the DFT calculation.


 FIG. 8. (a) Profile of the stacking fault energy in γ -TiAl along the path γ -TiAl \rightarrow SISF \rightarrow APB \rightarrow CSF \rightarrow γ -TiAl \rightarrow APB. (b) GSFE surface in γ -TiAl obtained using the EAM, the MLPs, and the DFT calculation.

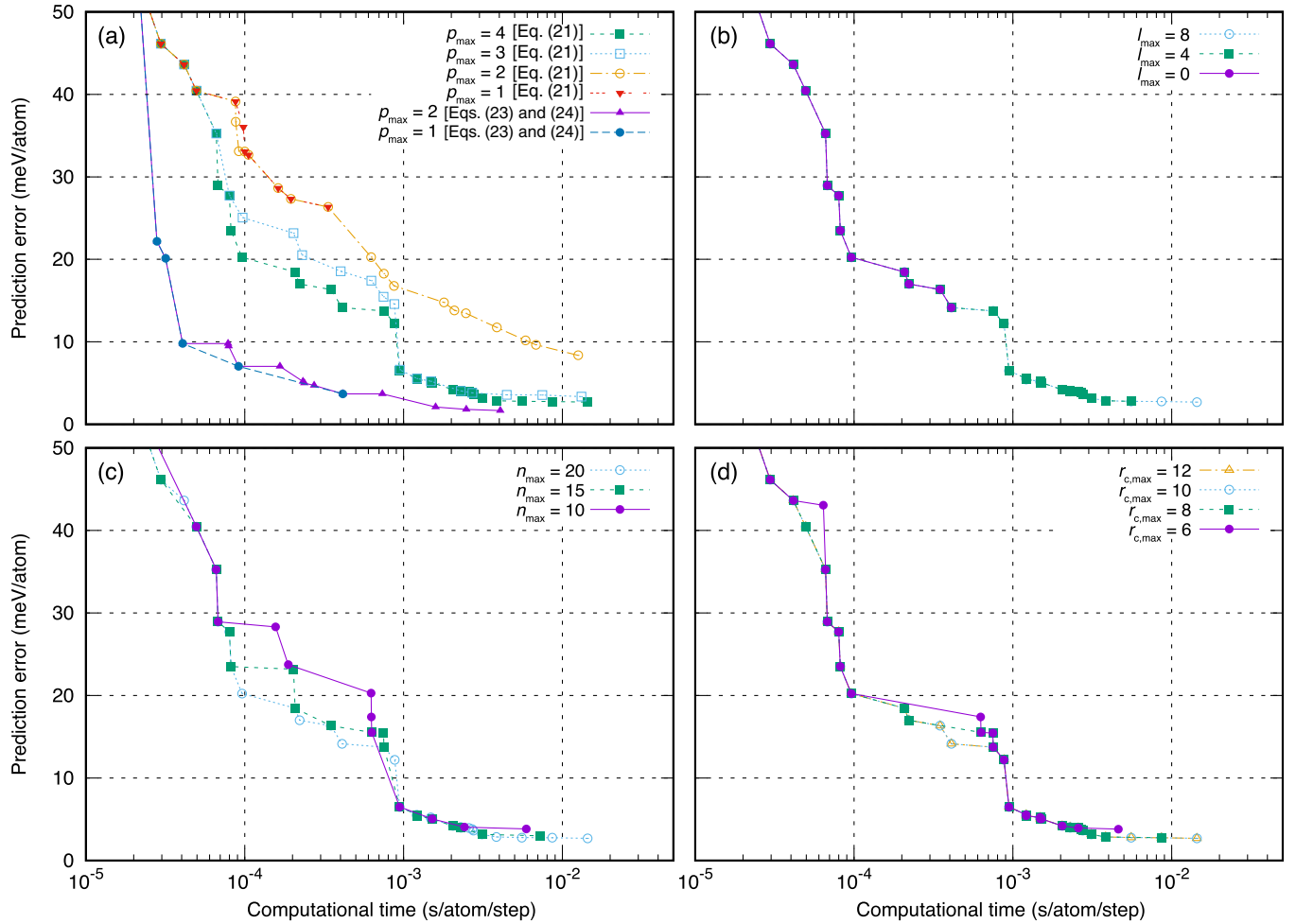


FIG. 9. Dependences of the Pareto optimal MLPs on (a) the maximum order of polynomial invariants p_{\max} , (b) the maximum angular number of spherical harmonics l_{\max} , (c) the maximum number of radial functions n_{\max} , and (d) the maximum cutoff radius $r_{c,\max}$. The Pareto optimal MLPs in (b)–(d) are obtained only from linear polynomial models given by Eq. (21).

Figure 5(c) shows the temperature dependence of the thermal expansion, calculated using a quasiharmonic approximation, in Ti_3Al ($D0_{19}$), TiAl ($L1_0$), and TiAl_3 ($D0_{22}$), which are experimentally observed in the Ti–Al binary system. As can be seen in Fig. 5(c), the thermal expansion of the EAM potential differs from that of the DFT calculation in all the structures. On the other hand, MLP2 and MLP3 derive the temperature dependence of the thermal expansion accurately in all the structures, even though the thermal expansion is more challenging to predict accurately than the phonon density of states and the phonon dispersion curves. The accurate prediction of the thermal expansion indicates that MLP2 and MLP3 can accurately evaluate the volume dependence of the whole range of phonon frequencies.

F. Bain path between γ -TiAl and $B2$

The energy profile along the Bain path between the γ -TiAl ($L1_0$) structure and the $B2$ structure is calculated using the EAM potential and the MLPs and compared with that obtained by DFT calculation. The structures required to compute the energy profile are obtained by transforming the c/a

ratio while keeping their volumes fixed to that of the equilibrium $L1_0$ structure. The energy profile is then evaluated by single-point calculations without geometry optimization for the structures. Figure 6 shows the energy profiles along the Bain path predicted using the EAM potential, the MLPs, and DFT calculation. The energy profiles along the Bain path of the MLPs, particularly MLP2 and MLP3, are consistent with that obtained by DFT calculation. The EAM profile is also close to the DFT profile for a c/a ratio of 1.0–1.5, although the c/a ratio of the equilibrium $L1_0$ structure of the EAM potential is slightly different from that obtained by DFT calculation. The behavior of the energy profile along the Bain path predicted by the EAM and the deviation of the c/a ratio are the same as those discussed by Zope and Mishin in Ref. [41], who developed the EAM potential.

G. Stacking faults in γ -TiAl

Computational models for the superlattice intrinsic stacking fault (SISF), the antiphase boundary (APB), and the complex stacking fault (CSF) are constructed using the procedure shown in Ref. [47]. First, the supercell is con-

structed by the expansion along the $\langle 111 \rangle$ direction of γ -TiAl with ideal cubic lattice parameters and the resultant supercell is composed of 24 atoms. The equilibrium structure of the supercell is then obtained by local geometry optimization. A computational model with a stacking fault is generated by introducing a tilt into the equilibrium structure. A displacement vector defining a tilt is given by a linear combination of two vectors corresponding to the $\langle \bar{1}10 \rangle/2$ and $\langle 11\bar{2} \rangle/2$ directions in ideal cubic γ -TiAl. The displacement vectors for the SISF, the APB, and the CSF are expressed as

$$\begin{aligned} \mathbf{b}_{\text{SISF}} &= \frac{1}{3} \left[\frac{1}{2} \langle 11\bar{2} \rangle \right], \\ \mathbf{b}_{\text{APB}} &= \frac{1}{2} \left[\frac{1}{2} \langle \bar{1}10 \rangle \right] + \frac{1}{2} \left[\frac{1}{2} \langle 11\bar{2} \rangle \right], \\ \mathbf{b}_{\text{CSF}} &= \frac{1}{2} \left[\frac{1}{2} \langle \bar{1}10 \rangle \right] + \frac{5}{6} \left[\frac{1}{2} \langle 11\bar{2} \rangle \right], \end{aligned} \quad (25)$$

respectively. Finally, single-point calculations are performed for the tilted structures.

Table VIII summarizes the stacking fault energies computed using the EAM potential, the MLPs, and DFT calculation. The cohesive energy of the equilibrium structure of γ -TiAl is used as a reference to compute the stacking fault energy. As can be seen in Table VIII, the MLPs predict the stacking fault energies of the SISF, the APB, and the CSF accurately. On the other hand, the EAM potential lacks the predictive power for the stacking fault energy of the APB.

Figure 7 shows the convergence behavior of the stacking fault energies in terms of the computational time obtained using the whole set of Pareto optimal MLPs. Figure 7 indicates the slow convergence of the stacking fault energy, which is relevant to the slow convergence of the prediction error. Therefore, it is reliable to use MLP2 or MLP3 for predicting the stacking fault energy accurately. Simultaneously, several Pareto optimal MLPs with lower computational cost also predict the stacking fault energy accurately. Therefore, an accurate large-scale atomistic simulation related to stacking faults in γ -TiAl may be possible in their careful use.

The stacking fault energy can be defined not only for these special stacking faults but also for other general stacking faults defined by displacement vectors. A collection of the excessive energies of the general stacking faults comprises a generalized stacking fault energy (GSFE) surface. The displacement vector identifying a tilt of the supercell for a general stacking fault is given by

$$\mathbf{b} = u \left[\frac{1}{2} \langle \bar{1}10 \rangle \right] + v \left[\frac{1}{2} \langle 11\bar{2} \rangle \right], \quad (26)$$

where u and v denote the fractional coordinates defined by the vectors $\langle \bar{1}10 \rangle/2$ and $\langle 11\bar{2} \rangle/2$, respectively.

Figure 8(a) shows the profiles of the stacking fault energy along the path γ -TiAl \rightarrow SISF \rightarrow APB \rightarrow CSF \rightarrow γ -TiAl \rightarrow APB, predicted using the EAM potential, the MLPs, and DFT calculation. The displacement vector changes continuously along the path. The stacking fault energy profiles of MLP2 and MLP3 are close to that obtained by the DFT calculation. The MLP1 profile agrees with the DFT profile along a major part of the path, while it deviates from the DFT profile along the path between the APB and the CSF. The EAM profile also deviates from the DFT profile along the

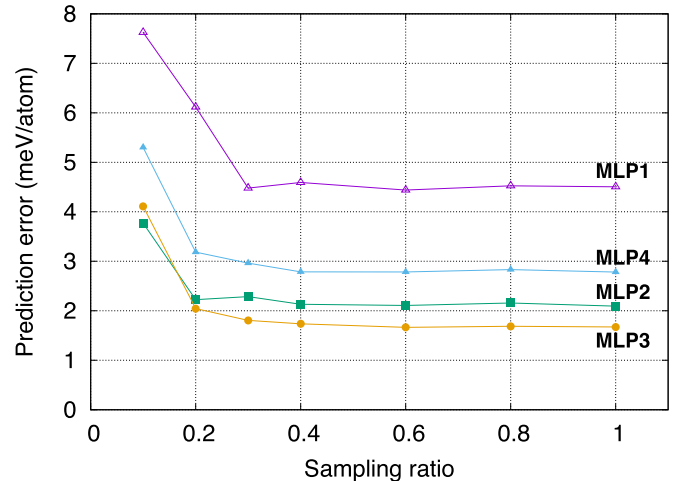


FIG. 10. Training data-set size dependence of the prediction error for MLP1, MLP2, MLP3, and an MLP with up to fourth-order polynomial invariants denoted by MLP4. The cutoff radius, the number of radial functions, and the maximum angular numbers of spherical harmonics used to develop MLP4 are 10 Å, 20, and [4,4,2], respectively. MLP4 is a linear polynomial model of the polynomial invariants represented by Eq. (21); hence it consists of 6940 regression coefficients.

path between the SISF and the CSF. Figure 8(b) shows the GSFE surface in γ -TiAl predicted using the EAM potential, the MLPs, and DFT calculation. The GSFE surfaces predicted using the MLPs and DFT calculation are similar, while that predicted using the EAM potential is different from that predicted by the DFT calculation. Thus, the present MLPs should have high predictive power for the stacking faults and related properties.

V. CONCLUSION

In this study, $O(3)$ polynomial invariants representing the neighboring atomic density in a multicomponent system have been formulated. Polynomial models have also been introduced to describe the relationship between the atomic energy and the polynomial invariants. Although the present formulation to develop MLPs in a multicomponent system is more complex than the formulation for elemental systems, it enables the accuracy and computational efficiency of MLPs to be controlled systematically.

This study also showed an application of the formulation of MLPs to the Ti-Al binary alloy system. Pareto optimal MLPs have been developed by applying the polynomial models combined with the polynomial invariants. These MLPs are available in the Machine Learning Potential Repository [31,32]. Moreover, a user package that combines the MLPs with atomistic simulations using the LAMMPS code and input script files used in this study are also available on a website [43]. The predictive power of the Pareto optimal MLPs has been examined for the cohesive energy, the formation energy, the elastic constants, the phonon density of states and dispersion curves, the thermal expansion, the energy profile along the Bain path, and the stacking fault properties. The MLP with the lowest prediction error (MLP3) and that with high

computational cost performance (MLP2) have high predictive power for all the properties, whereas an MLP showing higher computational cost performance (MLP1) than MLP2 fails to predict some of the properties with high accuracy. This study has revealed that the present framework provides a systematic way to develop MLPs with high computational cost performance in multicomponent systems.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Scientific Research (B) (Grant No. 19H02419) and a Grant-in-Aid for Scientific Research on Innovative Areas (Grant No. 19H05787) from the Japan Society for the Promotion of Science.

APPENDIX A: HYPERPARAMETER DEPENDENCE OF PARETO OPTIMAL MLPs

Figure 9 shows the dependence of the Pareto optimal MLPs on the parameter range used for the grid search, i.e., the maximum order of polynomial invariants [Fig. 9(a)], the maximum angular number of spherical harmonics [Fig. 9(b)], the maximum number of radial functions [Fig. 9(c)], and the maximum cutoff radius [Fig. 9(d)]. The dependence of

the Pareto optimal MLPs on the polynomial model is also shown in Fig. 9(a). Figure 9(a) indicates that the polynomial models of Eqs. (23) and (24) are significantly useful to increase the accuracy and computational efficiency of the MLPs. Moreover, second-, third-, and fourth-order polynomial invariants can increase the accuracy of the MLPs, although high-order polynomial invariants only with small angular numbers are considered in this study. As shown in Fig. 9(b), the spherical harmonics are essential for developing accurate MLPs in the Ti-Al system. The number of radial functions and cutoff radius are also essential factors for developing MLPs, while their importance is not recognized in the current grid search, as shown in Figs. 9(c) and 9(d).

APPENDIX B: LEARNING CURVE

Figure 10 shows the training data-set size dependence of the prediction error for MLP1, MLP2, MLP3, and an MLP with up to fourth-order polynomial invariants and 6940 regression coefficients. A set of structures for a sampling ratio is randomly selected from the whole training data set. As can be seen in Fig. 10, the prediction error converges well at the number of structures in the training data set of approximately 8000–12 000.

-
- [1] S. Lorenz, A. Groß, and M. Scheffler, *Chem. Phys. Lett.* **395**, 210 (2004).
- [2] J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [3] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Phys. Rev. Lett.* **104**, 136403 (2010).
- [4] J. Behler, *J. Chem. Phys.* **134**, 074106 (2011).
- [5] J. Han, L. Zhang, R. Car, and E. Weinan, *Commun. Comput. Phys.* **23**, 629 (2018).
- [6] N. Artrith and A. Urban, *Comput. Mater. Sci.* **114**, 135 (2016).
- [7] N. Artrith, A. Urban, and G. Ceder, *Phys. Rev. B* **96**, 014112 (2017).
- [8] W. J. Szlachta, A. P. Bartók, and G. Csányi, *Phys. Rev. B* **90**, 104108 (2014).
- [9] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, *Phys. Rev. X* **8**, 041048 (2018).
- [10] Z. Li, J. R. Kermode, and A. De Vita, *Phys. Rev. Lett.* **114**, 096405 (2015).
- [11] A. Glielmo, P. Sollich, and A. De Vita, *Phys. Rev. B* **95**, 214302 (2017).
- [12] A. Seko, A. Takahashi, and I. Tanaka, *Phys. Rev. B* **90**, 024101 (2014).
- [13] A. Seko, A. Takahashi, and I. Tanaka, *Phys. Rev. B* **92**, 054113 (2015).
- [14] A. Takahashi, A. Seko, and I. Tanaka, *Phys. Rev. Materials* **1**, 063801 (2017).
- [15] A. Thompson, L. Swiler, C. Trott, S. Foiles, and G. Tucker, *J. Comput. Phys.* **285**, 316 (2015).
- [16] M. A. Wood and A. P. Thompson, *J. Chem. Phys.* **148**, 241721 (2018).
- [17] C. Chen, Z. Deng, R. Tran, H. Tang, I.-H. Chu, and S. P. Ong, *Phys. Rev. Materials* **1**, 043603 (2017).
- [18] X.-G. Li, C. Hu, C. Chen, Z. Deng, J. Luo, and S. P. Ong, *Phys. Rev. B* **98**, 094104 (2018).
- [19] A. V. Shapeev, *Multiscale Model. Simul.* **14**, 1153 (2016).
- [20] V. L. Deringer, C. J. Pickard, and G. Csányi, *Phys. Rev. Lett.* **120**, 156001 (2018).
- [21] E. V. Podryabinkin, E. V. Tikhonov, A. V. Shapeev, and A. R. Oganov, *Phys. Rev. B* **99**, 064114 (2019).
- [22] K. Gubaev, E. V. Podryabinkin, G. L. Hart, and A. V. Shapeev, *Comput. Mater. Sci.* **156**, 148 (2019).
- [23] T. Mueller, A. Hernandez, and C. Wang, *J. Chem. Phys.* **152**, 050902 (2020).
- [24] A. Seko, A. Togo, and I. Tanaka, *Phys. Rev. B* **99**, 214108 (2019).
- [25] A. Hernandez, A. Balasubramanian, F. Yuan, S. A. Mason, and T. Mueller, *npj Comput. Mater.* **5**, 112 (2019).
- [26] Y. Zuo, C. Chen, X. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A. V. Shapeev, A. P. Thompson, M. A. Wood, and S. P. Ong, *J. Phys. Chem. A* **124**, 731 (2020).
- [27] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, *Phys. Rev. B* **28**, 784 (1983).
- [28] A. P. Bartók, R. Kondor, and G. Csányi, *Phys. Rev. B* **87**, 184115 (2013).
- [29] R. Kondor, Ph.D. thesis, Columbia University, 2008.
- [30] J. Flusser, B. Zitova, and T. Suk, *Moments and Moment Invariants in Pattern Recognition* (Wiley, New York, 2009).
- [31] A. Seko, [arXiv:2007.14206](https://arxiv.org/abs/2007.14206).
- [32] A. Seko, Machine Learning Potential Repository at Kyoto University, <https://sekocha.github.io/repository/index-e.html>.
- [33] M. El-Batanouny and F. Wooten, *Symmetry and Condensed Matter Physics: A Computational Approach* (Cambridge University Press, Cambridge, 2008).

- [34] J.-C. Tolédano and P. Tolédano, *The Landau Theory of Phase Transitions: Application to Structural, Incommensurate, Magnetic and Liquid Crystal Systems* (World Scientific, Singapore, 1987).
- [35] G. Bergerhoff and I. D. Brown, in *Crystallographic Databases*, edited by F. H. Allen, G. Bergerhoff, and R. Sievers (International Union of Crystallography, Chester, 1987), pp. 77–95.
- [36] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- [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] G. Kresse and D. Joubert, *Phys. Rev. B* **59**, 1758 (1999).
- [41] R. R. Zope and Y. Mishin, *Phys. Rev. B* **68**, 024102 (2003).
- [42] T. Nishiyama, A. Seko, and I. Tanaka, [arXiv:2007.15944](https://arxiv.org/abs/2007.15944).
- [43] A. Seko, lammps-mlip-package, <https://github.com/sekocha/lammps-mlip-package>.
- [44] LAMMPS code, <http://lammps.sandia.gov>.
- [45] S. Plimpton, *J. Comput. Phys.* **117**, 1 (1995).
- [46] A. Togo and I. Tanaka, *Scr. Mater.* **108**, 1 (2015).
- [47] P. Dumitraschkewitz, H. Clemens, S. Mayer, and D. Holec, *Appl. Sci.* **7**, 1193 (2017).