Group-theoretical high-order rotational invariants for structural representations: Application to linearized machine learning interatomic potential

Atsuto Seko,1,2,3,4, * Atsushi Togo,2 and Isao Tanaka1,2,4,5
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 JST, PRESTO, Kawaguchi 332-0012, Japan
4 Center for Materials Research by Information Integration, National Institute for Materials Science, Tsukuba 305-0047, Japan
5 Nanostructures Research Laboratory, Japan Fine Ceramics Center, Nagoya 456-8587, Japan

(Received 15 January 2019; revised manuscript received 12 April 2019; published 26 June 2019)

Many rotational invariants for crystal structure representations have been used to describe the structure-property relationship by machine learning. The machine learning interatomic potential (MLIP) is one of the applications of rotational invariants, which provides the relationship between the energy and the crystal structure. Therefore, the enumeration of rotational invariants should be useful for constructing MLIPs with the desired accuracy. In this study, we introduce high-order linearly independent rotational invariants up to the sixth order based on spherical harmonics and apply them to linearized MLIPs for elemental aluminum. A set of rotational invariants is derived by the general process of reducing the Kronecker products of irreducible representations for the SO(3) group using a group-theoretical projector method. A high predictive power for a wide range of structures is accomplished by using high-order invariants with low-order invariants equivalent to pair and angular structural features.

DOI: 10.1103/PhysRevB.99.214108

I. INTRODUCTION

The machine-learning interatomic potential (MLIP) based on a large data set generated by density functional theory (DFT) calculations is beneficial for significantly improving the accuracy and transferability of interatomic potentials [1–22]. Applications of MLIP have been increasing for not only atomistic simulations in large systems but also global structure optimizations [23–25], which require a high predictive power over a wide range of configurations. Similarly to most of conventional interatomic potentials, the MLIP is based on the idea that the short-range part of the total energy of a structure may be divided into the energies of the constituent atoms of a system. In the formulation of MLIPs, the atomic energy originating from atomic interactions with neighboring atoms is formulated as a function of a set of numerous quantities depending on its neighboring environment called structural features or descriptors. A number of models have been employed to describe a function or a mapping from structural features to the atomic energy, including artificial neural network models [1–6], Gaussian process models [7–11], and linear models [12–18].

Recently, many studies on estimating the structure-property or compound-property relationship by machine learning have been reported [26–68]. In such a machine learning estimation, the invariant properties of a set of target systems such as translational and rotational invariances play a key role in generating structural or compound features. Regarding the structure-energy relationship of an MLIP, the total energy of a crystalline system has been modeled as a function of a wide range of invariant quantities, including order parameters depending on pairs and angles among three atoms [2,69–71], as used in conventional interatomic potentials, moments derived from the atomic distribution [18], the power spectrum, the bispectrum, and the smooth overlap of atomic positions (SOAPs) kernel based on spherical harmonics [70,72]. Simultaneously, group-theoretical methods have long been powerful tools for deriving invariants based on the symmetry of target systems, and some of the above invariants have been derived by group-theoretical methods. Such group-theoretical invariants have been widely used not only for estimating MLIPs but also for characterizing and analyzing local structures (e.g., bond-orientational order parameter (BOP) [73]) and deriving the Landau free energy based on supergroup-subgroup relationship [74], the potential energy surface for a molecule as a function of symmetry-adapted redundant coordinates [75], and the model Hamiltonian of a crystalline system based on its space group [76].

In this study, we introduce high-order linearly independent rotational invariants up to the sixth order based on spherical harmonics into an MLIP framework, i.e., a linearized MLIP. A set of rotational invariants is enumerated by the general process of reducing the Kronecker products of irreducible representations (Irreps) for the SO(3) group using a group-theoretical projector method. The enumeration of rotational invariants will be useful for constructing MLIPs with the desired accuracy. As an application of high-order invariants, we demonstrate a linearized MLIP for elemental aluminum, formulated by a linear polynomial function of invariants.
II. POTENTIAL ENERGY MODELS

In this section, we introduce an atomic energy model with high-order polynomial invariants of the SO(3) group derived from the neighboring atomic density. The SO(3) group is the group of all proper rotations in a three dimensional real vector space. We first give a general formulation of the relationship between the atomic energy and the neighboring atomic density including their definitions in Sec. II A. Here, we consider the neighboring atomic density in a structure composed of a single element for simplicity. Then, we introduce a linear polynomial model for the atomic energy with polynomial invariants in Sec. II B. In Sec. II C, a group-theoretical approach to construct linearly independent polynomial invariants of SO(3) is shown. In Sec. II D, a practical form of radial functions used to expand the neighboring atomic density is demonstrated. In Secs. II E and II F, we show two other models for the atomic energy used for comparison with the model with high-order polynomial invariants.

A. General formulation of interatomic potentials

To formulate the relationship between a structure and its total energy, we first define the neighboring atomic density for a given atom in the structure. The neighboring atomic density of atom \( i \) is defined as the density of atoms within a given cutoff radius \( r_c \) from atom \( i \) as schematically illustrated in Fig. 1. In a structure composed of a single element, the neighboring atomic density is written as

\[
\rho^{(i)}(\mathbf{r}) = \sum_{\text{neighboring } j} \delta(\mathbf{r} - \mathbf{r}_j),
\]

where \( \mathbf{r}_j \) denotes the position of neighboring atom \( j \) and the sum is taken over the neighboring atoms within the cutoff radius. Then, the short-range part of the total energy may be decomposed as \( E = \sum_i E^{(i)} \), where \( E^{(i)} \) denotes the contribution of atom \( i \) or the atomic energy. Finally, the atomic energy is assumed to be written in a functional form of the neighboring atomic density as

\[
E^{(i)} = \mathcal{F}[\rho^{(i)}].
\]  

The neighboring atomic density can be expanded in terms of a basis set \( \{b_n\} \) as

\[
\rho^{(i)}(\mathbf{r}) = \sum_n a_n^{(i)} b_n(\mathbf{r}).
\]

where \( a_n^{(i)} \) are the order parameters. This expansion replaces the functional form of the atomic energy to a function of order parameters. Then, let us consider a situation where an arbitrary rotation is applied to the neighboring atomic density. We define rotation operator \( \hat{R} \) acting on the basis functions as

\[
\hat{R} b_n(\mathbf{r}) = b_n(\hat{R}^{-1} \mathbf{r}),
\]

where \( \Gamma(\hat{R}) \) denotes the matrix representation of rotation \( \hat{R} \) for the basis set \( \{b_n\} \). This property holds for the atomic density expressed by a linear combination of the basis functions. Therefore, the atomic density is transformed by rotation \( \hat{R} \) to

\[
\hat{R} \rho^{(i)}(\mathbf{r}) = \sum_{n,n'} a_n^{(i)} \Gamma_{n,n'}(\hat{R}) b_{n'}(\mathbf{r}).
\]

This transformation of the atomic density can be also viewed as the change of the order parameters from \( a_n^{(i)} \) to \( \sum_n a_n^{(i)} \Gamma_{n,n'}(\hat{R}) \). Although an arbitrary rotation generally changes the neighboring atomic density, it does not change the atomic energy. This means all elements of the SO(3) group leave the atomic energy invariant, hence the atomic energy should be modeled by the invariants of the SO(3) group derived from \( \{a_n^{(i)}\}, \{d_n^{(i)}\} \), as

\[
E^{(i)} = F(a_1^{(i)}, a_2^{(i)}, \cdots).
\]

This formulation can describe both conventional iteratomic potentials and MLIPs, and \( \{d_n^{(i)}\} \) are called structural features in the context of the MLIP.

B. Interatomic potential with high-order polynomial invariants

In this study, we expand the atomic density in terms of a basis set corresponding to the Irreps of SO(3), i.e., spherical harmonics, using the procedure introduced by Bartók et al. [70]. Therefore, a rotation of the basis set is represented by the Irreps of SO(3) known as the Wigner D matrix. A rotation of products of radial functions \( \{f_n\} \) and spherical harmonics \( \{Y_{lm}\} \) is also represented by the Irreps of SO(3). When we expand the neighboring atomic density in terms of \( \{f_n Y_{lm}\} \), the neighboring atomic density at a position \( (r, \theta, \phi) \) in spherical coordinates centered at the position of atom \( i \) is expressed as

\[
\rho^{(i)}(r, \theta, \phi) = \sum_{nlm} d_{nlm}^{(i)} f_n(r) Y_{lm}(\theta, \phi),
\]

where \( a_{nlm}^{(i)} \) are order parameters.

Then, we adopt linearly independent polynomial invariants of SO(3) generated from order parameters as structural features for the atomic energy. A \( p \)-th order polynomial invariant...
for given $n$ and $\{l_1, l_2, \ldots, l_p\}$ is defined by a linear combination of products of $p$ order parameters, expressed as

$$d^{(l)}_{nl1, l_2, \ldots, l_p}(x) = \sum_{m_1, m_2, \ldots, m_p} c_{m_1 m_2 \ldots m_p}^{l_1, l_2, \ldots, l_p} \Gamma_{nl1} \Gamma_{nl2} \ldots \Gamma_{nlp}. \quad (7)$$

As will be shown in Sec. II C, a coefficient set $\{c_{m_1 m_2 \ldots m_p}^{l_1, l_2, \ldots, l_p}\}$ is constructed by solving the eigenvalue problem for a projector matrix, ensuring that the linear combination is invariant for arbitrary rotation. Since there are no $l$-positive first-order invariants, all first-order invariants are identical to radial order parameters, $d_{nl0}^{(l)} = d_{nl0}^{(0)}$. A second-order invariant is identified by a single value of $l$ because second-order linear combinations are invariant only when $l_1 = l_2$, which is a general feature of product groups [74,77]. In terms of fourth- and higher-order polynomial invariants, multiple invariants are linearly independent for most of the set $\{l_1, l_2, \ldots, l_p\}$, which are distinguished by index $i$ if necessary.

Finally, the atomic energy may be formulated as a function of the linearly independent polynomial invariants. We employ an atomic energy model based on a simple linear polynomial form of the invariants, written as

$$E^{(l)} = w_0 + \sum_n w_{nl} d^{(l)}_{nl0} + \sum_{n,l} w_{nl} d^{(l)}_{nl0} + \sum_{n,\{l_1, l_2, l_3\}} w_{nl1, l_2, l_3} d^{(l)}_{nl1, l_2, l_3} + \sum_{n,\{l_1, l_2, l_3, l_4\}} w_{nl1, l_2, l_3, l_4} d^{(l)}_{nl1, l_2, l_3, l_4} + \cdots, \quad (8)$$

where $w_0, w_{nl0}, w_{nl}, w_{nl1, l_2, l_3},$ and $w_{nl1, l_2, l_3, l_4}$ are regression coefficients.

Note that the second- and third-order invariants are equivalent to the angular Fourier series (AFS) and bispectrum reported in the literature, respectively [70,72]. Therefore, the formulation of Eq. (8) with up to third-order invariants is equivalent to spectral neighbor analysis potential (SNAP) [15]. When excluding radial parts, the second-order invariants are exactly the same as BOPs [73]. If we restrict the invariants to third-order symmetrized ones excluding radial parts, they are exactly the same as third-order BOPs.

C. Group-theoretical projector operation method

A set of polynomial invariants up to the sixth order is derived by the general process of reducing the Kronecker products of Irreps. The Kronecker products of Irreps have been widely used for many purposes in physics and chemistry such as the formulation of angular momentum coupling, the derivation of selection rules and the formulation of the Landau free energy for phase transitions. We employ the group-theoretical projection operator method [77] to derive polynomial invariants of SO(3).

Let us first consider the reduction of products of the same finite group $\mathcal{G}$. In the projection operator method, the projector matrix of the Kronecker product of Irreps $(\mu) \Gamma \otimes (\nu) \Gamma \otimes \cdots$ of the same finite group $\mathcal{G}$ to Irrep $\sigma$ is defined by

$$(\sigma) P = \sum_i (\sigma) P_{ii}$$

where $g$ and $d_{\mu \nu}$ denote the order of group $\mathcal{G}$ and the dimension of Irrep $\sigma$, respectively. Considering the reduction of the Kronecker product to the one-dimensional identity Irrep whose elements are all unity, the projection matrix is given by

$$P^{(l)} = \frac{1}{g} \sum_{R \in \mathcal{G}} (\mu) \Gamma(R) \otimes (\nu) \Gamma(R) \otimes \cdots. \quad (10)$$

By solving the eigenvalue problem for the projector matrix of each combination $\{\mu, \nu, \cdots\}$, expressed by

$$(\mu) P u = u, \quad (11)$$

eigenvector $u$ is obtained. Each eigenvector corresponds to a set of coefficients identifying a polynomial invariant of group $\mathcal{G}$ in the form of a linear combination of order parameters.

SO(3) is an infinite or continuous group, which has an infinite number of elements. In such a case, the average appearing in Eq. (10) is replaced with an integral over continuous parameters [78]. When a rotation is described by Euler angles $\{\alpha, \beta, \gamma\}$, the Wigner $D$-function $D^{(l)}(\alpha, \beta, \gamma)$ is an Irrep of SO(3). Therefore, the element of the projector matrix for the decomposition of $p$th-order Kronecker products into the identity Irrep is written as

$$(i = 0) P^{(l)}_{m_1 m_2 \cdots m_p} = \frac{1}{8\pi^2} \int_0^{2\pi} da \int_0^\pi d\beta \sin \beta \int_0^{2\pi} d\gamma D^{(l)}_{m_1 m_2} (\alpha, \beta, \gamma) \times D^{(l)}_{m_2 m_3} (\alpha, \beta, \gamma) \cdots D^{(l)}_{m_p m_1} (\alpha, \beta, \gamma), \quad (12)$$

where integrals are calculated only from Clebsch-Gordan coefficients for SO(3) or Wigner 3j symbols as shown in Appendix A. The projector matrix element has a nonzero value only when $m_1 + m_2 + \cdots + m_p = 0$ and $m_1' + m_2' + \cdots + m_p' = 0$. Then, a solution of the eigenvalue problem for a given set $\{l_1, l_2, \ldots, l_p\}$ corresponds to a polynomial invariant of SO(3).

Note that some of the eigenvectors with nonzero eigenvalues correspond to invariants that are linearly dependent on invariants from the other eigenvectors. This originates from the fact that we consider Kronecker products of a single basis set. For the same reason, some of the eigenvectors correspond to invariants that are constantly zero. Therefore, they are removed from the set of invariants. In addition, we use only polynomials that are invariant for an arbitrary improper rotation. This means that we use polynomial invariants for a set of $\{l_1, l_2, \ldots, l_p\}$ whose sum is even.

D. Basis functions and order parameters

The order parameters $\{d_{nlm}^{(l)}\}$ for atom $i$ of a given structure are calculated from its neighboring atomic density. If one chooses an orthonormal set of radial functions, each of the order parameters is given by the inner product of the corresponding basis function and the neighboring atomic density as

$$d_{nlm}^{(l)} = \sum_{j \text{ neighbor}} f_{n}(r_{ij}) Y_{lm}^{*}(\theta_{ij}, \phi_{ij}), \quad (13)$$

where $(r_{ij}, \theta_{ij}, \phi_{ij})$ denotes the spherical coordinates of neighboring atom $j$ centered at the position of atom $i$. 

214108-3
We employ a finite basis set with Gaussian-type radial functions to expand the neighboring atomic density. The Gaussian-type radial functions are expressed as

\[ f_r(r) = \exp[-\beta r (r - r_n)^2] f_c(r), \]

where \( \beta_n \) and \( r_n \) are given parameters. Values of \( \{ \beta_n \} \) and \( \{ r_n \} \) are given by finite arithmetic progressions. Considering the cutoff function \( f_c \), we use the following cosine-based function:

\[ f_c(r) = \begin{cases} \frac{1}{2} \left[ \cos \left( \frac{\pi r}{r_c} \right) + 1 \right] & (r \leq r_c) \\ 0 & (r > r_c) \end{cases}, \]

where \( r_c \) denotes the cutoff radius. Since the Gaussian-type radial functions are not orthonormal, the order parameter \( a_{nlm}^{(i)} \) is given by the linear combination of the inner products with different \( n \). However, the inner products and their linear combinations are equivalent in the linear polynomial form for the atomic energy of Eq. (8). Therefore, we estimate the order parameters \( \{ a_{nlm}^{(i)} \} \) using Eq. (13) regardless of the orthonormality of the radial functions.

E. Pair functional model

For comparison with the atomic energy model with high-order polynomial invariants, we use the pair functional model introduced in Ref. [71]. This model was reported to be capable of predicting the energy within a root mean square (RMS) error of 2.7 meV/atom for data sets derived from single structure generators, which were averaged over 31 elemental metals including transition metals [71]. We use the second-order approximation of the pair functional model described as

\[ E = w_0 + \sum_n w_{nl} d_{nl}^{(i)} + \sum_{n,n'=1}^N w_{n,n'} d_{n,n'}^{(i)} + \sum_{n,n',l} w_{n,n',l} d_{n,n',l}^{(i)} + \sum_{n,n',l} w_{n,n',l} d_{n,n',l}^{(i)} + \sum_{n,n',l} w_{n,n',l} d_{n,n',l}^{(i)} \]

where Gaussian-type radial functions are adopted. This model was also applied to the 31 elemental metals using the data set mentioned in Sec. II E [71]. The RMS error averaged over the 31 elemental metals was reported to be 0.5 meV/atom [71]. We call this model a cluster functional MLIP because it is classified as a cluster functional potential.

III. ESTIMATION OF POTENTIAL ENERGY MODELS

A. Data sets

Training and test data sets are constructed from DFT calculations. The test data set is used to estimate the predictive power for structures that are not included in the training data set. To generate a wide range of structures for the training and test data sets, we adopt prototype structures reported in the Inorganic Crystal Structure Database (ICSD) [80] as structure generators. We restrict all ICSD entries to those with the ANX formula of “N” and eliminate duplicate entries that have the same ICSD structure type. In other words, we employ only unique ICSD prototype structures composed of single elements with zero oxidation state. The total number of structure generators is 86 and a list of structure generators is shown in Appendix B.

First, the atomic positions and lattice constants of the structure generators are fully optimized by DFT calculation to obtain their equilibrium structures. Then, a candidate structure used in each of the data sets is constructed by random lattice expansion, random lattice distortion, and random atomic displacements into a supercell of each of the structure generators. For a given parameter \( \epsilon \) controlling the degree of lattice expansion, lattice distortion, and atomic displacements, the lattice vectors of the candidate structure are expressed by

\[ A' = A + \epsilon R, \]

where \( A \) and \( A' \) denote the matrix representations of the lattice vectors of the supercell of the structure generator and the candidate structure, respectively. The \( (3 \times 3) \) matrix \( R \) is composed of uniform random numbers ranging from \(-1\) to 1. The displacement of an atom is described by the change of its fractional coordinates as

\[ f' = f + \epsilon A'^{-1} \eta, \]

where \( f \) and \( f' \) denote the fractional coordinates of the atom in the supercell of the structure generator and the candidate structure, respectively. The three-dimensional vector \( \eta \) consists of uniform random numbers ranging from \(-1\) to 1.

We generate a wide range of candidate structures using multiple values of \( \epsilon \). When we generate \( N_\text{atom} \) structures from a structure generator, the value of \( \epsilon \) for the \( N \)th structure, \( \epsilon_N \), is given by the finite arithmetic progression of length \( N_\text{atom} \) as \( \epsilon_N = 0.5N_\text{atom}^-N_\text{atom} / N_\text{atom} \). Therefore, it is worth emphasizing that the given displacements are much larger than the small displacements required to compute harmonic phonon force constants. Applying this procedure to all the structure generators, a total of 430,000 candidate structures are generated.

We sample 10,000 structures used for DFT calculations from the pool of candidate structures so that the variance or the uncertainty of the predicted energy becomes small in a similar manner to structure selection procedures used in the
cluster expansion for alloy systems [81–83]. We decrease the uncertainty more efficiently than a random structure selection using the following procedure. A DFT structure set composed of randomly selected 1000 structures is first prepared. Then 20 trial sets of additional ten structures are randomly selected from the pool of candidate structures. For each of the trial sets, the trace of the precision matrix, Tr[(X^T X)^{-1}], is computed, where X contains structural features for the combination of the DFT structure set and the trial set. In the process of the structure selection, we use a set of first- and second-order polynomial invariants as structure features. Among the 20 trials, the trial set showing the lowest trace of the precision matrix is added to the DFT structure set. They are repeated until the size of the DFT structure set reaches 10,000. We then split the DFT structure set into a set of 9,000 structures and the other set of 1,000 structures randomly for the training and test data sets, respectively.

For the total of 10,000 structures, DFT calculations were performed using the plane-wave-basis projector augmented wave (PAW) method [84] within the Perdew-Burke-Ernzerhof exchange-correlation functional [85] as implemented in the VASP code [86–88]. The cutoff energy was set to 400 eV. The total energies converged to less than 10^{-3} meV/supercell. The atomic positions and lattice constants were optimized for the structure generators until the residual forces were less than 10^{-2} eV/Å.

B. Regression

Regarding the training data, the total energy, the forces acting on atoms and the stress tensor computed by DFT calculations are available since they are all expressed by linear equations with the same regression coefficients. When we estimate the regression coefficients from all of the total energies, forces and stress tensors, the predictor matrix X is divided into three submatrices X_energy, X_force, and X_stress, which contain the structural features for the total energies, the forces acting on atoms and the stress tensors of structures in the training data set, respectively. The structural features for the forces are derived in Appendix C, and the structural features for the stress tensors can be easily derived in a similar manner to the derivation of the structural features for the forces. The observation vector also has three components of the total energy y_energy, forces y_force and stress tensor y_stress of structures in the training data set, obtained by DFT calculations. The predictor matrix and observation vector are simply written in a submatrix form as

\[ X = \begin{bmatrix} X_{\text{energy}} \\ X_{\text{force}} \\ X_{\text{stress}} \end{bmatrix}, \quad y = \begin{bmatrix} y_{\text{energy}} \\ y_{\text{force}} \\ y_{\text{stress}} \end{bmatrix}. \] (21)

Finally, the total number of training data reached 1,377,769.

The regression coefficients of a model for the atomic energy, which comprise the coefficient vector w, are estimated by linear ridge regression. The optimal ridge coefficients minimize the penalized residual sum of squares expressed as

\[ L(w) = ||Xw - y||_2^2 + \lambda||w||_2^2, \] (22)

where X and y denote the predictor matrix and observation vector, respectively. The regularization parameter \( \lambda \) controls the magnitude of the penalty. It is also beneficial to use sparse linear regressions such as the least absolute shrinkage and selection operator (Lasso) to decrease the cost of computing the energy and forces, while we adopt the linear ridge regression to estimate the regression coefficients rapidly and stably in this study.

IV. RESULTS AND DISCUSSION

A. Number of invariants

The number of polynomial invariants that occur in the decomposition of the Kronecker product of the irreps of the SO(3) group is obtained only from their characters. The number of pth-order polynomial invariants for the set \( \{l_1, l_2, \ldots, l_p\} \) is calculated using the following equation:

\[ n_{l_1, l_2, \ldots, l_p} = \frac{1}{8\pi^2} \int_0^{2\pi} d\alpha \int_0^{\pi} d\beta \sin \beta \times \int_0^{2\pi} d\gamma \sin [(2l_1 + 1)\omega/2] \sin [(2l_2 + 1)\omega/2] \cdots \sin [(2l_p + 1)\omega/2] \sin [\omega/2] \] (23)

where \( \chi^{(l)}(\alpha, \beta, \gamma) \) denotes the character of Irrep \( l \) for a rotation described by Euler angles \( \{\alpha, \beta, \gamma\} \). In a practical enumeration of invariants, the explicit forms of characters described by its rotation axis \( \phi \) and rotation angle \( \omega \) are more convenient than those described by Euler angles. When specifying the rotation in this manner, the character of Irrep \( l \) for rotation \( \hat{R} \) is simply expressed as

\[ \chi^{(l)}(\hat{R}) = \chi^{(l)}(\omega) = \frac{\sin [(2l + 1)\omega/2]}{\sin [\omega/2]} \] (24)

Therefore, the number of pth-order polynomial invariants is computed as

\[ n_{l_1, l_2, \ldots, l_p} = \frac{1}{4\pi^2} \int_0^{2\pi} d\omega \sin^2 \omega/2 \int_0^{\pi} d\theta \sin \theta \times \int_0^{2\pi} d\phi \sin [(2l_1 + 1)\omega] \cdots \sin [(2l_p + 1)\omega] \] (25)

where \( \theta \) and \( \phi \) identify the rotation axes.

Table I shows the number of pth-order invariants satisfying \( l_1 \leq l_{\text{max}}, l_2 \leq l_{\text{max}}, \ldots, l_p \leq l_{\text{max}} \) for a given \( l_{\text{max}} \). The integer sequences for second- and third-order invariants may correspond to the On-Line Encyclopedia of Integer Sequences (OEISs) A000027 and A002623, respectively [89]. Nonetheless, as described above, these numbers include linearly dependent and constantly zero invariants derived from a single basis set. Therefore, the number of available invariants shown in Table II is obtained by solving the eigenvalue problems for the projector matrix and removing such invariants.

Although we have considered invariants for all possible combinations of \( l \) thus far, we are also allowed to restrict them to symmetrized invariants, which are derived by reducing the pth power of Irrep \( l \) [77]. The symmetrized power of an Irrep
TABLE I. Cumulative number of SO(3) invariants for a given value of maximum \(l\).

<table>
<thead>
<tr>
<th>(l_{\text{max}})</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>5</td>
<td>9</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>14</td>
<td>22</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>35</td>
<td>52</td>
<td>38</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>59</td>
<td>80</td>
<td>59</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>87</td>
<td>117</td>
<td>84</td>
<td>51</td>
</tr>
</tbody>
</table>

TABLE II. Cumulative number of nonzero linearly independent SO(3) invariants of Eq. (7) for each \(n\). Cumulative number of invariants for an arbitrary improper rotation among them is also shown in parentheses.

<table>
<thead>
<tr>
<th>(l_{\text{max}})</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>6</td>
<td>10</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>13</td>
<td>20</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>30</td>
<td>48</td>
<td>35</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>51</td>
<td>75</td>
<td>59</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>73</td>
<td>104</td>
<td>84</td>
<td>51</td>
</tr>
</tbody>
</table>

TABLE III. Cumulative number of symmetrized invariants for a given value of maximum \(l\), \(l_{\text{max}}\).

<table>
<thead>
<tr>
<th>(l_{\text{max}})</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>5</td>
<td>9</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>14</td>
<td>22</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>35</td>
<td>52</td>
<td>38</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>59</td>
<td>80</td>
<td>59</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>87</td>
<td>117</td>
<td>84</td>
<td>51</td>
</tr>
</tbody>
</table>

has played an important role in many subjects such as the construction of symmetrized tensors and the Landau theory.

B. Application to elemental aluminum

We first demonstrate a procedure to construct an optimal MLIP with polynomial invariants for elemental Al. An MLIP is identified by a given maximum value of order \(p\), \(p_{\text{max}}\), and a given set of maximum values of \(l\) for order \(p\), \((l_{\text{max}}^{(2)}, l_{\text{max}}^{(3)}, \ldots, l_{\text{max}}^{(p_{\text{max}})})\). This means that all polynomial invariants satisfying \(p \leq p_{\text{max}}\) and \(l_{\text{max}} \leq l_{\text{max}}^{(p)}\) are included in the set of structural features for a given \(p_{\text{max}}\) and \((l_{\text{max}}^{(2)}, l_{\text{max}}^{(3)}, \ldots, l_{\text{max}}^{(p_{\text{max}})})\). We hereafter describe an MLIP as \((l_{\text{max}}^{(2)}, l_{\text{max}}^{(3)}, \ldots, l_{\text{max}}^{(p_{\text{max}})})\). To find an optimal MLIP,
we consider \( l_{\text{max}}^{(p)} \) up to nine for the second order, seven for the third order, three for the fourth order, and two for both the fifth and sixth orders. For symmetrized invariants, \( l_{\text{max}}^{(p)} \) up to nine for the second order, eight for the third order, eight for the fourth order, four for the fifth order, and three for the sixth order are taken into account. In the optimization of \( \{l_{\text{max}}^{(2)}, l_{\text{max}}^{(3)}, \ldots, l_{\text{max}}^{(p_{\text{max}})}\} \) and \( p_{\text{max}} \), we use the number of radial functions, the parameters in the radial functions and the cutoff radius that are optimized by trial and error. This means that we use 22 sets of parameters \( \beta_n \) and \( r_n \) corresponding to all combinations of finite arithmetic progressions of \( \beta_n = \{0.25, 0.5\} \) and \( r_n = \{0, 1, \ldots, 10\} \). The cutoff radius is set to 10 Å.

Figure 2(a) shows the dependence of the RMS error for the test data on the number of invariants. Each data point corresponds to the RMS error of an MLIP \( \{l_{\text{max}}^{(2)}, l_{\text{max}}^{(3)}, \ldots, l_{\text{max}}^{(p_{\text{max}})}\} \). As can be seen in Fig. 2(a), the RMS error decreases as the number of invariants increases. The MLIP with the lowest RMS error of 2.27 meV/atom is composed of 2399 invariants (9-7-3-2-2). The fitting error for the training data of 2.23 meV/atom is almost the same as the RMS error for the test data.

On the other hand, the pair functional MLIP has a large RMS error of 19.88 meV/atom although it was reported to have an RMS error of 0.89 meV/atom for the test data constructed from simple structure generators such as fcc-, bcc-, and hcp-type structures [71]. The pair functional MLIP is composed of 276 regression coefficients (22 invariants), and its RMS error is a converged one with respect to the number of invariants. Clearly, the difference in RMS error originates from the different structures used to estimate the prediction error, because the present test data is obtained from a wide range of structure generators as shown in Sec. IIIA. Therefore, the pair functional MLIP provides an accurate description of the atomic interactions in simple structures, whereas it has no power to describe the atomic interactions for a wide range of structures. This limitation of the transferability has also long been recognized in conventional pair functional interatomic potentials such as embedded atom method (EAM) potentials, which are regarded as reductions of the pair functional MLIP. Although the use of many pairwise structure features in the pair functional MLIP improves the descriptive power of the pair functional model for simple structures, this result indicates that systematically increasing the number of pairwise structural features is not a useful way of increasing the transferability to a wide range of structures.

Figure 2(b) shows the dependence of the RMS error on the maximum order of the invariants. Only the MLIP with the lowest RMS error among all MLIPs identified by a value of \( p_{\text{max}} \) is shown in Fig. 2(b). The best MLIP composed of only second-order invariants, which is equivalent to an MLIP with AFS structural features, has an RMS error of 7.72 meV/atom. Using both second- and third-order invariants, the RMS error significantly decreases to 2.83 meV/atom. Upon adding higher-order invariants, the RMS error gradually decreases and MLIP (9-7-3-2-2) exhibits the lowest RMS error of 2.10 meV/atom. Even when restricting the invariants to symmetrized ones, the RMS error decreases with increasing maximum order of the invariants. MLIP (9-8-8-4-3) composed of only symmetrized invariants has an RMS error of 3.77 meV/atom, which is the lowest among the MLIPs with only symmetrized invariants.

However, the RMS error almost converges at \( p_{\text{max}} = 4 \). To further increase the accuracy of the MLIP, both a large number of additional high-order invariants and a large number of additional training data may be required, because the RMS error decreases very slowly as the number of invariants increases as shown in Fig. 2(a). In addition, the potential performance of higher-order invariants may be evaluated more accurately simply by significantly increasing the number of training data and high-order invariants.

We measure in more detail the accuracy for a structure group, which is defined as a set of structures derived from a structure generator. Figure 3 shows the dependence of the RMS error on the structure group of MLIP (9-7-3-2-2) in comparison with that of the pair functional MLIP. The RMS error for a structure group is averaged over the structures in the group. The pair functional MLIP exhibits large errors for some structure groups, such as high-pressure and covalent structures, and shows relatively small errors for
The structures significantly different from those of typical metallic and high-pressure structures originate from their electronic systematic deviations and related huge errors for covalent structures are distributed around the diagonal line. The energy is equal to the DFT energy, although most of the whole data set combining the training and test data sets.

In the comparison of the pair functional MLIP and [Fig.4(a)] MLIP (9-7-3-2-2) for the other structure groups. On the other hand, the use of high-order polynomial invariants significantly improves the accuracy for the whole range of structure groups. Figure 2(b) also shows the standard deviation of the errors for the structure groups. Figure 2(b) indicates that high-order polynomial invariants decrease not only the error averaged over all structures but also the structure group dependence of the error. In addition to the energy, the other physical properties depending on the structure group are predicted with a higher accuracy by MLIP (9-7-3-2-2) as shown in Figs. 2(c) and 2(d).

Figure 4 shows a comparison of the energy computed by DFT calculation with those predicted by [Fig. 4(a)] the pair functional MLIP and [Fig. 4(c)] MLIP (9-7-3-2-2) for the whole data set combining the training and test data sets. The energy is measured from the DFT energy of the fcc equilibrium structure. In the comparison of the pair functional MLIP with the DFT calculation, a series of structures belonging to structure groups with large errors in Fig. 3 shows a systematic deviation from the diagonal line indicating that the MLIP energy is equal to the DFT energy, although most of the structures are distributed around the diagonal line. The systematic deviations and related huge errors for covalent and high-pressure structures originate from their electronic structures significantly different from those of typical metallic structures such as fcc structure. On the other hand, no such systematic deviation is seen in the comparison of MLIP (9-7-3-2-2) with the DFT calculation. The DFT and MLIP energies are almost the same for all structures in the energy scale of Fig. 4(c).

Figure 4 also shows the energy dependence of the error for [Fig. 4(b)] the pair functional MLIP and [Fig. 4(d)] MLIP (9-7-3-2-2). It is clearly found that the pair functional MLIP has a wider error distribution than MLIP (9-7-3-2-2). Although a clear energy dependence of the error is not observed in Figs. 4(b) and 4(d), the error of the structures with a low DFT energy is much lower than those of the other structures. Moreover, there are several other systematic deviations in the error distribution of the pair functional MLIP in addition to a clear systematic deviation that can also be recognized in Fig. 4(a).

Undoubtedly, a useful MLIP requires not only a low prediction error but also predictive power for properties related to the energies of structures that are not included in the data sets, i.e., transferability. Here, we evaluate the prediction error for the phonon frequencies, vacancy formation energy, and grain boundary energy to estimate the transferability of MLIPs. Table IV shows the RMS error for the phonon frequencies, vacancy formation energy, and symmetric tilt grain boundary (STGB) energy. The RMS error for the phonon frequency is
evaluated as

\[
(RMS_{\text{phonon}})^2 = \frac{1}{N_{\text{kpoints}}} \sum_{k,n} \left( \omega_{k,n}^{\text{MLIP}} - \omega_{k,n}^{\text{DFT}} \right)^2 ,
\]

(27)

where \(\omega_{k,n}^{\text{MLIP}}\) and \(\omega_{k,n}^{\text{DFT}}\) denote the phonon frequencies of band \(n\) at \(k\) point \(k\) predicted by the MLIP and DFT calculation, respectively. Table IV indicates that the phonon frequencies predicted by MLIP (9-7-3-2-2) are very close to those predicted by DFT calculation. The pair functional MLIP also has good predictive power for the phonon properties of fcc, bcc, and hcp structures, although it has large errors ranging from 5–8 meV/atom for fcc, bcc, and hcp structure groups as shown in Fig. 3. The high predictive power for the phonon properties of the pair functional MLIP is ascribed to the fact that structures with a low DFT energy have very small prediction errors in the pair functional MLIP as shown in Fig. 4(b).

We next compare the vacancy formation energies for fcc and hcp structures predicted by MLIPs with those predicted by DFT calculation. Supercells are formed by the \(3 \times 3 \times 3\) expansions of the conventional unit cells of the fcc and hcp structures. As shown in Table IV, all MLIPs show similar predictive powers for the vacancy formation energy. We also examine the prediction error for grain boundary models not included in the training and test data sets. Here, we introduce STGB models with a \((110)\) tilt direction. They have \([113], [112], [111],\) and \([221]\) grain boundary planes and misorientation angles of \(50.479, 70.529, 109.471,\) and \(141.058^\circ\), respectively. They are respectively composed of 880, 284, 384, and 512 atoms including two boundary planes. Starting from their initial structures taken from Ref. [90], we optimize the STGB models by using MLIPs and by DFT calculation. Table IV shows the grain boundary energies of the STGB models predicted by DFT calculation and MLIPs. As can be seen in Table IV, all MLIPs predict the grain boundary energy very accurately. These results are associated with the above discussion on the high predictive power for the phonon properties, i.e., structures with a low DFT energy can be predicted with very small prediction errors in the pair functional MLIP. These results indicate that the pair functional and cluster functional MLIPs with AFS structure features can be highly useful for a limited range of applications, although an MLIP with high-order polynomial invariants shows the highest predictive power in general.

Finally, we discuss the computational efficiency of MLIP. Figure 5 shows the distribution of MLIPs in the plane of the RMS error and the normalized elapsed time to compute the energy, the forces and the stress tensors of a configuration of

**TABLE IV.** Transferability of pair functional MLIP, cluster functional MLIP with AFS and MLIP (9-7-3-2-2) in elemental Al.

<table>
<thead>
<tr>
<th></th>
<th>Pair func.</th>
<th>Cluster func. (AFS)</th>
<th>MLIP (9-7-3-2-2)</th>
<th>DFT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RMS error in phonon frequency (THz)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCC</td>
<td>0.118</td>
<td>0.273</td>
<td>0.052</td>
<td>–</td>
</tr>
<tr>
<td>BCC</td>
<td>0.282</td>
<td>0.631</td>
<td>0.289</td>
<td>–</td>
</tr>
<tr>
<td>HCP</td>
<td>0.185</td>
<td>0.491</td>
<td>0.053</td>
<td>–</td>
</tr>
<tr>
<td><strong>Vacancy formation energy (eV)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCC</td>
<td>0.727</td>
<td>0.663</td>
<td>0.660</td>
<td>0.614</td>
</tr>
<tr>
<td>HCP</td>
<td>0.652</td>
<td>0.681</td>
<td>0.681</td>
<td>0.585</td>
</tr>
<tr>
<td><strong>Grain boundary energy (mJ/m²), STGB(110)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[113] (50.479°)</td>
<td>187.5</td>
<td>197.1</td>
<td>187.4</td>
<td>158.4</td>
</tr>
<tr>
<td>[112] (70.529°)</td>
<td>376.6</td>
<td>388.1</td>
<td>379.8</td>
<td>348.3</td>
</tr>
<tr>
<td>[111] (109.471°)</td>
<td>88.8</td>
<td>64.3</td>
<td>57.5</td>
<td>39.2</td>
</tr>
<tr>
<td>[221] (141.058°)</td>
<td>397.0</td>
<td>430.1</td>
<td>421.2</td>
<td>386.0</td>
</tr>
</tbody>
</table>
features, which represent effectively many-body interactions. The present result implies that one can obtain more Pareto optimal MLIPs by introducing polynomials of the higher-order invariants.

V. CONCLUSION

In this study, SO(3) polynomial invariants up to the sixth order representing atomic distributions have been enumerated. We have applied them to construct accurate MLIPs for elemental Al by formulating the atomic energy as a linear polynomial form of the polynomial invariants. The high-order invariants play an essential role in constructing an MLIP with a high predictive power for a wide range of crystal structures. The list of invariants and the group-theoretical procedure to derive the invariants should be useful for constructing not only MLIPs but also prediction models for the other physical properties that depend on the crystal structure.

ACKNOWLEDGMENT

This study was supported by PRESTO from JST.

APPENDIX A: PROJECTOR MATRIX ELEMENTS

In this Appendix, we show the formulation used to compute the projector matrix elements for the decomposition of the Kronecker product of Irreps into the identity Irrep. The projector matrix elements for SO(3) invariants are calculated from integrals involving the Wigner D functions. The projector matrix elements for the second- and third-order invariants are expressed [93] as

\begin{equation}
\begin{aligned}
(1) Pl_{1, l_1, m_1, m_1'}^{l_1; m_1, m_1; m_1'} &= \frac{1}{8\pi^2} \int_0^{2\pi} d\alpha \int_0^\pi d\beta \sin \beta \int_0^{2\pi} d\gamma D_{m_1, m_1'}^{l_1}(\alpha, \beta, \gamma) \\
& \times (\alpha, \beta, \gamma)D_{m_2, m_2'}^{l_2}(\alpha, \beta, \gamma) \\
& = (-1)^{m_1-m_1'} \frac{1}{2l_2+1} C_{l_1, m_1, m_1'} D_{m_1, m_1'}^{l_1}(\alpha, \beta, \gamma) \\
&D_{m_2, m_2'}^{l_2}(\alpha, \beta, \gamma) \\
& = (-1)^{m_1-m_1'} \frac{1}{2l_2+1} C_{l_1, m_1, m_1'} C_{l_2, m_2, m_2'}. \tag{A1}
\end{aligned}
\end{equation}

where \(C_{l_1, m_1, l_2, m_2}^{m_1, m_2}\) denotes the Clebsch-Gordan coefficient for the SO(3) group. The projector matrix elements for the fourth-, fifth-, and sixth-order invariants are reduced to integrals involving products of two or three Wigner D functions using the Clebsch-Gordan expansion given as

\begin{equation}
\begin{aligned}
D_{m_1, m_1'}^{l_1}(\alpha, \beta, \gamma)D_{m_2, m_2'}^{l_2}(\alpha, \beta, \gamma) \\
& = \sum_{l_1+l_2} \sum_{m_{11}, m_{12}} C_{l_1, m_{11}, m_1}^{l_{11}} C_{l_2, m_{12}, m_2}^{l_{12}} D_{m_1, m_1'}^{l_1}(\alpha, \beta, \gamma)C_{l_2, m_{21}, m_2'}^{l_{21}}. \tag{A3}
\end{aligned}
\end{equation}
Therefore, the projector matrix elements for the fourth-, fifth-, and sixth-order invariants are derived as

\[
(1) P_{m_{1}m_{2}l_{1}l_{2}}^{l_{3}l_{4}}_{m_{1}m_{2}m_{3}m_{4}n} = \frac{1}{8\pi^{2}} \int_{0}^{2\pi} \int_{0}^{2\pi} d\beta \sin \beta \int_{0}^{2\pi} d\gamma D_{m_{1}m_{2}l_{1}l_{2}}^{(l_{3}l_{4})} D_{m_{3}m_{4}l_{3}l_{4}}^{(l_{1}l_{2})} D_{m_{1}m_{2}l_{1}l_{2}}^{(l_{3}l_{4})} D_{m_{3}m_{4}l_{3}l_{4}}^{(l_{1}l_{2})}

= (-1)^{m_{4}-m_{4}} \frac{1}{2l_{1}+1} \sum_{l=|l_{1}-l_{2}|}^{l_{1}+l_{2}} C_{l_{1}l_{2}m_{1}m_{2}}^{(l_{1}l_{2})} C_{l_{2}l_{1}m_{1}m_{2}}^{(l_{1}l_{2})} C_{l_{1}l_{2}m_{3}m_{4}}^{(l_{3}l_{4})} C_{l_{2}l_{1}m_{3}m_{4}}^{(l_{3}l_{4})}

(2) P_{m_{1}m_{2}l_{1}l_{2}l_{3}l_{4}}^{l_{5}l_{6}}_{m_{1}m_{2}m_{3}m_{4}m_{5}m_{6}} = \frac{1}{8\pi^{2}} \int_{0}^{2\pi} \int_{0}^{2\pi} d\beta \sin \beta \int_{0}^{2\pi} d\gamma D_{m_{1}m_{2}l_{1}l_{2}l_{3}l_{4}}^{(l_{5}l_{6})} D_{m_{3}m_{4}l_{3}l_{4}l_{5}l_{6}}^{(l_{1}l_{2})} D_{m_{1}m_{2}l_{1}l_{2}l_{3}l_{4}}^{(l_{5}l_{6})} D_{m_{3}m_{4}l_{3}l_{4}l_{5}l_{6}}^{(l_{1}l_{2})}

= (-1)^{m_{4}-m_{4}} \frac{1}{2l_{1}+1} \sum_{l=|l_{1}-l_{2}|}^{l_{1}+l_{2}} \sum_{l=|l_{1}-l_{2}|}^{l_{1}+l_{2}} C_{l_{1}l_{2}m_{1}m_{2}}^{(l_{1}l_{2})} C_{l_{2}l_{1}m_{1}m_{2}}^{(l_{1}l_{2})} C_{l_{1}l_{2}m_{3}m_{4}}^{(l_{3}l_{4})} C_{l_{2}l_{1}m_{3}m_{4}}^{(l_{3}l_{4})}
\times C_{l_{3}l_{4}m_{5}m_{6}}^{(l_{5}l_{6})} C_{l_{4}l_{3}m_{5}m_{6}}^{(l_{5}l_{6})} C_{l_{3}l_{4}m_{5}m_{6}}^{(l_{5}l_{6})} C_{l_{4}l_{3}m_{5}m_{6}}^{(l_{5}l_{6})}

(3) P_{m_{1}m_{2}l_{1}l_{2}l_{3}l_{4}l_{5}l_{6}}^{l_{7}l_{8}}_{m_{1}m_{2}m_{3}m_{4}m_{5}m_{6}m_{7}m_{8}} = \frac{1}{8\pi^{2}} \int_{0}^{2\pi} \int_{0}^{2\pi} d\beta \sin \beta \int_{0}^{2\pi} d\gamma D_{m_{1}m_{2}l_{1}l_{2}l_{3}l_{4}l_{5}l_{6}}^{(l_{7}l_{8})} D_{m_{3}m_{4}l_{3}l_{4}l_{5}l_{6}l_{7}l_{8}}^{(l_{1}l_{2})} D_{m_{1}m_{2}l_{1}l_{2}l_{3}l_{4}l_{5}l_{6}}^{(l_{7}l_{8})} D_{m_{3}m_{4}l_{3}l_{4}l_{5}l_{6}l_{7}l_{8}}^{(l_{1}l_{2})}

= (-1)^{m_{4}-m_{4}} \frac{1}{2l_{1}+1} \sum_{l=|l_{1}-l_{2}|}^{l_{1}+l_{2}} \sum_{l=|l_{1}-l_{2}|}^{l_{1}+l_{2}} \sum_{l=|l_{1}-l_{2}|}^{l_{1}+l_{2}} C_{l_{1}l_{2}m_{1}m_{2}}^{(l_{1}l_{2})} C_{l_{2}l_{1}m_{1}m_{2}}^{(l_{1}l_{2})} C_{l_{1}l_{2}m_{3}m_{4}}^{(l_{3}l_{4})} C_{l_{2}l_{1}m_{3}m_{4}}^{(l_{3}l_{4})}
\times C_{l_{3}l_{4}m_{5}m_{6}}^{(l_{5}l_{6})} C_{l_{4}l_{3}m_{5}m_{6}}^{(l_{5}l_{6})} C_{l_{3}l_{4}m_{5}m_{6}}^{(l_{5}l_{6})} C_{l_{4}l_{3}m_{5}m_{6}}^{(l_{5}l_{6})}
\times C_{l_{5}l_{6}m_{7}m_{8}}^{(l_{7}l_{8})} C_{l_{6}l_{5}m_{7}m_{8}}^{(l_{7}l_{8})} C_{l_{5}l_{6}m_{7}m_{8}}^{(l_{7}l_{8})} C_{l_{6}l_{5}m_{7}m_{8}}^{(l_{7}l_{8})}

A\text{\.}

\text{APPENDIX B: LIST OF STRUCTURE GENERATORS}

Table V shows the structure generators used for the training and test data sets in this study.
APPENDIX C: FORCES ACTING ON ATOMS

In this Appendix, we derive the formulation for the forces acting on atoms from the atomic energy model with high-order polynomial invariants of Eq. (8) used in this study. The forces acting on atoms are evaluated from the derivatives of the total energy $E$ with respect to the Cartesian coordinates of the atoms. The Cartesian component $\alpha$ of the force acting on atom $k, F_{k,\alpha}$, is written as

$$F_{k,\alpha} = -\frac{\partial E}{\partial x_{k,\alpha}},$$

where $x_{k,\alpha}$ denotes the $\alpha$ component of the Cartesian coordinates of atom $k$. Because the total energy is expressed as the sum of the atomic energies written as Eq. (8), the force component is derived as

$$F_{k,\alpha} = -\sum_{i} \left[ \sum_{n} w_{nl} \frac{\partial a^{(i)}_{nl, \mu \alpha}}{\partial x_{k,\alpha}} + \sum_{n} w_{nl, l_1 l_2} \frac{\partial a^{(i)}_{nl, l_1 l_2, \mu \alpha}}{\partial x_{k,\alpha}} + \cdots \right].$$

Therefore, the derivatives of polynomial invariants are required to compute the forces. At the same time, the minus sign of the derivatives of the polynomial invariants corresponds to the structural features for the forces, which are used to estimate regression coefficients as shown in Sec. III B. The derivative on the $p$th-order polynomial invariant is expressed as

$$\frac{\partial a^{(i)}_{nl, l_1 \cdots l_p}}{\partial x_{k,\alpha}} = \sum_{m_{l_1} m_{l_2} \cdots m_{l_p}} \frac{\partial a^{(i)}_{nl, m_{l_1} m_{l_2}}}{x_{k,\alpha}} \cdot a^{(i)}_{nl, m_{l_2}, \mu} + \frac{\partial a^{(i)}_{nl, m_{l_1} m_{l_2}}}{\partial x_{k,\alpha}} \cdot a^{(i)}_{nl, m_{l_2} m_{l_3}} + \cdots + \frac{\partial a^{(i)}_{nl, m_{l_1} m_{l_2}}}{\partial x_{k,\alpha}} \cdot a^{(i)}_{nl, m_{l_2}, \mu},$$

where the derivative of $a^{(i)}_{nl, m}$ is given by

$$\frac{\partial a^{(i)}_{nl, m}}{\partial x_{k,\alpha}} = \sum_{j \text{neighbor}} \frac{\partial}{\partial x_{k,\alpha}} [f_{n}(r_{ij}) Y_{lm}^{*}(\theta_{ij}, \phi_{ij})].$$

The derivative of the right side has a nonzero value when $k = j$ or $k = i$. When $k = j$, the derivative is computed using the following set of equations:

$$\frac{\partial}{\partial r_{ij}} [f_{n}(r) Y_{lm}(\theta, \phi)] = \frac{y_{j} - y_{i}}{r} f'_{n}(r) Y_{lm}(\theta, \phi) + f_{n}(r) \left[ \frac{\cos \theta \cos \phi}{r} \frac{\partial Y_{lm}(\theta, \phi)}{\partial \theta} - \frac{\sin \theta}{r \sin \phi} \frac{\partial Y_{lm}(\theta, \phi)}{\partial \phi} \right]$$

$$\frac{\partial}{\partial r_{ij}} [f_{n}(r) Y_{lm}(\theta, \phi)] = \frac{y_{j} - y_{i}}{r} f'_{n}(r) Y_{lm}(\theta, \phi) + f_{n}(r) \left[ \frac{\cos \theta \sin \phi}{r} \frac{\partial Y_{lm}(\theta, \phi)}{\partial \theta} + \frac{\sin \phi}{r \sin \theta} \frac{\partial Y_{lm}(\theta, \phi)}{\partial \phi} \right]$$

$$\frac{\partial}{\partial r_{ij}} [f_{n}(r) Y_{lm}(\theta, \phi)] = \frac{z_{j} - z_{i}}{r} f'_{n}(r) Y_{lm}(\theta, \phi) + f_{n}(r) \left[ -\frac{\sin \theta}{r} \frac{\partial Y_{lm}(\theta, \phi)}{\partial \theta} \right].$$

where the derivative of spherical harmonics with respect to $\theta$ and $\phi$ is given by

$$\frac{\partial Y_{lm}(\theta, \phi)}{\partial \theta} = m \cot \theta Y_{lm}(\theta, \phi) + \sqrt{(l-m)(l+m+1)} e^{-i \phi} Y_{l(m+1)}(\theta, \phi)$$

$$\frac{\partial Y_{lm}(\theta, \phi)}{\partial \phi} = i m Y_{lm}(\theta, \phi).$$

When $k = i$, the derivatives are given as the minus signs of Eqs. (C5)–(C7).

APPENDIX D: CHARACTER OF SYMMETRIZED $p$th POWER OF AN IRREP

The $p$th tensor product of an Irrep is reducible into symmetrized and antisymmetrized representations. The symmetrized $p$th power of Irrep $^{(\mu)} \Gamma$ and its character are denoted as $^{(\mu)} \Gamma^p(\hat{R})$ and $^{[\mu]} \Gamma^p(\hat{R})$, respectively, for an arbitrary operation $\hat{R}$ of a group. The character of an operator $\hat{R}$ in the symmetrized $p$th power of Irrep $\mu$ is expressed as

$$^{[\mu]} \Gamma_{\mu}^{2}(\hat{R}) = \frac{1}{2}^{(\mu)} \chi^{2}(\hat{R}) + \frac{1}{2}^{(\mu)} \chi^{2}(\hat{R})$$

$$^{[\mu]} \Gamma_{\mu}^{3}(\hat{R}) = \frac{1}{2}^{(\mu)} \chi^{3}(\hat{R}) + \frac{1}{2}^{(\mu)} \chi^{3}(\hat{R}) + \frac{1}{4}^{(\mu)} \chi^{3}(\hat{R}) + \frac{1}{4}^{(\mu)} \chi^{3}(\hat{R})$$

$$^{[\mu]} \Gamma_{\mu}^{4}(\hat{R}) = \frac{1}{24}^{(\mu)} \chi^{4}(\hat{R}) + \frac{1}{8}^{(\mu)} \chi^{2}(\hat{R})^{2} \chi^{2}(\hat{R}) + \frac{1}{8}^{(\mu)} \chi^{2}(\hat{R})^{2} \chi^{2}(\hat{R}) + \frac{1}{8}^{(\mu)} \chi^{2}(\hat{R})^{2} + \frac{1}{4}^{(\mu)} \chi^{4}(\hat{R})$$

$$^{[\mu]} \Gamma_{\mu}^{5}(\hat{R}) = \frac{1}{120}^{(\mu)} \chi^{5}(\hat{R}) + \frac{1}{12}^{(\mu)} \chi^{3}(\hat{R}) \chi^{2}(\hat{R}) + \frac{1}{8}^{(\mu)} \chi^{3}(\hat{R}) \chi^{2}(\hat{R})$$

$$+ \frac{1}{8}^{(\mu)} \chi^{3}(\hat{R}) \chi^{2}(\hat{R}) + \frac{1}{4}^{(\mu)} \chi^{2}(\hat{R})^{3} \chi^{2}(\hat{R}) + \frac{1}{4}^{(\mu)} \chi^{2}(\hat{R})^{3} \chi^{2}(\hat{R}) + \frac{1}{4}^{(\mu)} \chi^{2}(\hat{R})^{3} \chi^{2}(\hat{R}) + \frac{1}{4}^{(\mu)} \chi^{2}(\hat{R})^{3} \chi^{2}(\hat{R})$$

214108-12
\[
X^{(\mu)} \Gamma^0 (\hat{R}) = \frac{1}{720} X^{(\mu)}(\hat{R}) + \frac{1}{45} X^{(4)(\mu)}(\hat{R})^2 + \frac{1}{16} X^{(2)(\mu)}(\hat{R})^2 X(\hat{R})^2 + \frac{1}{18} X^{(3)(\mu)}(\hat{R})^2 X(\hat{R})^2 + \frac{1}{18} X^{(3)(\mu)}(\hat{R})^2 X(\hat{R})^2 \\
+ \frac{1}{6} X(\hat{R})^2 \Gamma^0 (\hat{R})^2 X(\hat{R})^2 + \frac{1}{16} X^{(2)(\mu)}(\hat{R})^2 X(\hat{R})^4 + \frac{1}{8} X^{(2)(\mu)}(\hat{R})^2 X(\hat{R})^4 + \frac{1}{8} X(\hat{R})^2 \Gamma^0 (\hat{R})^2 X(\hat{R})^4 + \frac{1}{5} X(\hat{R})^2 \Gamma^0 (\hat{R})^5.
\]