Construction of interatomic potentials using large sets of DFT calculations and linear regression method

Akira Takahashi

2017

Contents

Chapter 1

General Introduction

Calculation of potential energy surface (PES) with respect to the atomic positions is crucially important to perform molecular dynamics (MD) simulation and estimate stability of systems. Those simulations are very useful and widely applied for material science and engineering. For example, MD simulation and estimation of stability is applied to study on deformation of materials. Dislocation processes [1] and deformation twinning [2] in nanocrystalline are studied by MD simulation. Estimation of generalized stacking fault energy (GSFE) is also common method to analyze deformation of materials using PES [3,4]. MD simulation can also estimate diffusion coefficients at limited temperature [5,6]. Apart from material science, MD simulation is also applied for biochemistry such as study on chemical property of protein [7].

To perform those calculations, fast and accurate method to estimate PES is desired. One of popular method is quantum mechanical method like density functional theory (DFT) calculations [8,9]. Even this method is reliable for many systems, however, the computational costs are large and usually limited to system including less than a few thousand atoms. For a large scale simulations, empirical potentials, which are simple functions to represent interatomic interaction energy and provide PES much faster than DFT, are widely used.

Conventionally, the functional form is determined with physical consideration and fitting parameters are determined with a set of experimental data or first principles results. For example, one of the simplest interatomic potential, Lennard-Jones (LJ) potential [10] is expressed as

$$E^{(i)} = \frac{1}{2} \sum_{j} 4\epsilon \left[\left(\frac{\sigma}{R_{ij}} \right)^{12} - \left(\frac{\sigma}{R_{ij}} \right)^{6} \right], \tag{1.1}$$

where $E^{(i)}$ is summation of interaction energy between atom i and neighbor atom j, R_{ij} is a distance between atom i and atom j, ϵ and σ are fitting parameters. First terms and second terms are repulsive interaction due to overlapping electron orbitals and dispersion force between two atoms, respectively. Parameters ϵ and σ can be determined to reproduce some experimental data for example equilibrium distance and bond-dissociation energy, or accurate DFT calculations.

Such method is very simple and can be applied to simulation of rare gas, however, it is known that LJ potential cannot provide accurate PES of bulk crystal. For example, because LJ potential is merely summation of pair interaction, elastic constants satisfy Cauchy relation [11] as

$$C_{12} = C_{14}, \tag{1.2}$$

which real cubic materials usually do not satisfy. To construct interatomic potentials of more accuracy and of more transferability, more interatomic potential models have been proposed. Embedded atom model (EAM) [12–14] includes many-body interactions from quantum mechanical consideration, and EAM is widely applied for simulations of bulk metals. Tersoff potentials [15–17] adopts bond angle among three atoms, and those are used for systems whose bonds are covalent like C, Si, and other semiconductors. Modified EAM [18, 19] is often used for transition metals.

To construct more accurate interatomic potential, machine learning interatomic potentials (MLIPs) are also proposed. In these procedures, the relationship between the energy of a crystal structure and a set of descriptors expressing the crystal structure is estimated. In beginning, application of MLIP had been limited to molecules containing few atoms [20, 21], however, recently MLIPs which can be applied to periodic systems are proposed [22–24]. In the literature, a few regression techniques such as artificial neural networks [22, 25–30], Gaussian process regression [23] and least absolute shrinkage and selection operator (LASSO) [24, 31] have been adopted to estimate MLIP for solids.

Once framework to apply machine learning technique and appropriate descriptor set are determined, MLIP can be automatically constructed to reproduce energies and forces of large set of DFT calculations. However, it is still difficult to think about how machine learning technique must be applied and what descriptors should be used. One solution of this problem is expressing interatomic potential model linearly (I call such potential linear potential) and using LASSO technique, which can select effective descriptors. Indeed, it is reported that such model can be applied to several elemental metals and optimal interatomic potential can reproduce several material properties like total energies of structures and phonon dispersion curve from DFT calculations [24, 31]. However, using LASSO technique, it is still difficult to choice candidate descriptors. In this thesis, firstly I show what descriptors should be used from mathematical consideration and reveal the relationship between conventional interatomic potentials and MLIPs. From approximate general potential model by polynomial, more general framework to construct linear potential can be obtained. Then, I applied the generalized framework for 31 elemental metals and estimated the accuracy of interatomic potentials.

In chapter 2, the general interatomic potential model is derived. The relationship between conventional interatomic potentials and MLIPs is also discussed. In chapter 3, I construct interatomic potential using such framework for 31 elemental metals by only radial descriptors and compare the accuracy of generalized framework with previous linear potential. The accuracy is also estimated by calculating several properties of metals. For Al and Cu, the properties obtained from linear potential are also compared with conventional interatomic potentials, such as EAM. In chapter 4, descriptors which include bonding angles are considered. I examine the effect of such descriptors for prediction of linear potentials. I also estimated some properties of metals by linear potentials. For Ti, interatomic potential is also compared with conventional EAM and modified EAM.

Chapter 2

General interatomic potential model

2.1 Introduction

Interatomic potential have been widely used for large scale molecular dynamics (MD) simulations. As shown in Chapter 1, many functional forms like embedded atom model (EAM), Tersoff, and modified EAM are developed to perform reliable MD simulation.

According to a general classification of the interatomic potential by milestone paper by Carlsson [32], interatomic potentials can be classified into four class, namely, pair potential, pair functional potential, cluster function potential, and cluster functional potential. Even present day, the most of interatomic potentials are classified into such classification. For example, EAM, which is used widely for simulation for bulk metal, is classified into pair functional potential. Modified EAM, which adopts bond angle and used for transition metals, is classified into cluster functional potential.

Incidentally, machine learning interatomic potentials (MLIP) are proposed and used recently. These potentials are constructed by estimating the relationship between the energy of a crystal structure and a set of descriptors expressing the crystal structure by machine learning techniques. In the literature, a few regression techniques such as artificial neural networks [22, 25–30], Gaussian process regression [23] and least absolute shrinkage and selection operator (LASSO) [24,31] have been adopted to estimate MLIP for solids. Also, several descriptors have been proposed [22–24, 33–39], while only a few of them have succeeded in obtaining accurate MLIP. Although MLIP model and descriptors set crucially determine the accuracy and computational costs of interatomic potential, method to apply machine learning and descriptor set are determined by trial and error rather than by clear theory based on physical and mathematical consideration. One way to select effective descriptors is expressing interatomic potential model linearly and the use of compressing sensing methods like LASSO technique. Indeed, it is reported that such model can be applied to several elemental metals and optimal interatomic potential can reproduce several material properties like total energies of structures and phonon dispersion curve from DFT calculations [24, 31]. Although LASSO potentials can select effective descriptors and construct accurate interatomic potentials, the preparation of candidate descriptors is still difficult. If we use unnecessary candidate descriptors, the computational costs of construction of interatomic potential becomes large and such extra descriptors may cause overfitting and decrease generalization ability of the interatomic potential model. Moreover, when the candidate set lacks necessary descriptors, we can never obtain the accurate model no matter how we select descriptors.

The difficulty of consideration of descriptor sets can be ascribed to the deficiency of mathematical interpretation of MLIP and comprehension of the relationship between MLIP and already existing interatomic potentials. It is not understood what type of Carlsson's classification MLIP is classified into. Machine learning has been used only as a tool for estimating a black-box function.

In this chapter, firstly, I enumerate the condition which interatomic potentials need to satisfy. Next, I derive the general functional form of interatomic potential from such conditions based on mathematical consideration. I also show the relationship among the general expression, MLIP, and conventional interatomic potential. Such consideration will make it easy to introduce existing physical insights into MLIP.

2.2 Derivation

In this section, the general functional form of interatomic potentials is derived and it is shown that when we consider sufficient terms, the general form can approximate any interatomic potential within any small error.

Interatomic potential determines a contribution to the total energy of the system of a target atom (in this study such energy is called atomic energy) from $3N$ coordinates $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N\}$ (the vectors of origin is the position of the target atom) of N atoms. For nonmagnetic unary system without external field, generally interatomic potential satisfies the following four conditions.

(A) Interaction is continuous:

$^\forall \epsilon > 0, ^\forall N, 0 <^\forall i < N, ^\exists \delta s.t.$ When $3N$ coordinates are displaced by small $\|\Delta \boldsymbol{r}_i\| < \delta$, or under the transformation $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_i, \ldots, \boldsymbol{r}_N\} \to \{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_i + \Delta \boldsymbol{r}_i, \ldots, \boldsymbol{r}_N\}$, the difference between the total energy before displacement and the total energy after displacement is smaller than ϵ.

(B) Interaction of long range can be ignored:

When the distance between target atom and the ith neighbor atom is larger than cutoff radius R_c, or $\|\boldsymbol{r}_i\| \geq R_c$, the existence and displacement of $\boldsymbol{r}_i$ has no effect on the atomic energy until the norm of $\boldsymbol{r}_i$ becomes shorter than R_c.

(C) The contribution energy has symmetry for permutation:

The atomic energy is invariable under any permutation like

$\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_i, \ldots, \boldsymbol{r}_j, \ldots, \boldsymbol{r}_N\} \to \{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_j, \ldots, \boldsymbol{r}_i, \ldots, \boldsymbol{r}_N\}$.

(D) The contribution energy has symmetry for rotation and mirroring:

The atomic energy is invariable under orthogonal group, $O(3)$ transformation. $O(3)$ is generated from rotation and mirroring transformation and can be represented as 3×3 matrix whose determinant is 1 or -1. Therefore, the condition (D) can be expressed that $^\forall R \in O(3)$, atomic energy is invariable under following the transformation where

$\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N\} \to \{R\boldsymbol{r}_1, R\boldsymbol{r}_2, \ldots, R\boldsymbol{r}_N\}$.

Here, first I define an atomic distribution function like

$$\rho(\boldsymbol{r}) = \sum_{n=1}^{N} p(\boldsymbol{r} - \boldsymbol{r}_n), \tag{2.1}$$

where $p(\boldsymbol{r})$ is a positive C^∞ function which has only one peak at $\boldsymbol{r} = \boldsymbol{0}$ and the range where the $p(\boldsymbol{r})$ has non-zero value is so small that any pair of functions $p(\boldsymbol{r} - \boldsymbol{r}_i)$ are not superposed. Clearly, the relationship between distribution functions and atomic environment of $3N$ coordinates is one-to-one because we can obtain one atomic distribution function from $3N$ coordinates and determine the set of $3N$ coordinates from positions of peak of the atomic distribution function. Then, interatomic potential can be interpreted as a mapping from atomic distribution function to one scalar. Using the notation of functional, we can express an interatomic potential function as $E\left[\rho(\boldsymbol{r})\right]$. Because atomic distribution function has symmetry for permutation of atoms, if the model which approximates $E\left[\rho(\boldsymbol{r})\right]$ is derived then condition (C) is automatically satisfied.

Next, continuity of $E\left[\rho(\boldsymbol{r})\right]$ has to be considered. The atomic distribution function which is displaced can be expressed as,

$$\rho_{\text{disp}}(\boldsymbol{r}) = \sum_{n=1}^{i-1} p(\boldsymbol{r} - \boldsymbol{r}_n) + p(\boldsymbol{r} - (\boldsymbol{r}_i + \Delta\boldsymbol{r}_i)) + \sum_{n=i+1}^{N} p(\boldsymbol{r} - \boldsymbol{r}_n) \tag{2.2}$$

Then, following condition is satisfied.

$${}^{\forall}\epsilon > 0,\ {}^{\exists}\delta\ s.t.\ \|\Delta\boldsymbol{r}_i\| < \delta \Rightarrow \|\rho - \rho_{\text{disp}}\|^2 < \epsilon, \tag{2.3}$$

where $\|\rho - \rho_{\text{disp}}\|^2 = \int \left[\rho(\boldsymbol{r}) - \rho_{\text{disp}}(\boldsymbol{r})\right]^2 d\boldsymbol{r}$. Therefore, if functional $E\left[\rho(\boldsymbol{r})\right]$ is constructed under the condition as

$${}^{\forall}\epsilon > 0,\ {}^{\exists}\delta\ s.t.\ \|\Delta\rho(\boldsymbol{r})\|^2 = \int \Delta\rho(\boldsymbol{r})^2 d\boldsymbol{r} < \delta \Rightarrow \|E\left[\rho(\boldsymbol{r}) + \Delta\rho(\boldsymbol{r})\right] - E\left[\rho(\boldsymbol{r})\right]\|^2 < \epsilon, \tag{2.4}$$

the functional satisfies the condition (A).

Incidentally, for any $\epsilon_1 > 0$, using sufficiently large number M and linearly independent function set $\{\rho_m(\boldsymbol{r})\}_{1\leq m\leq M}$, $p(\boldsymbol{r})$ can be expanded as

$$\rho(\boldsymbol{r}) = \sum_{m=1}^{M} c_m \rho_m(\boldsymbol{r}) + \epsilon_1(\boldsymbol{r}), \tag{2.5}$$

where c_m is weight parameters and $\epsilon_1(\boldsymbol{r})$ satisfies the relationship $\|\epsilon_1(\boldsymbol{r})\|^2 \leq \epsilon_1$. From Eqn.(2.4), $E\left[\rho(\boldsymbol{r})\right]$ can be accurately approximated by $E\left[\sum_{m=1}^{M} c_m \rho_m(\boldsymbol{r})\right]$ if sufficient basis functions are used. Because basis functions $\{\rho_m(\boldsymbol{r})\}_{1\leq m\leq M}$ are fixed, weight parameters $\{c_m\}_{1\leq m\leq M}$ completely specify $\sum_{m=1}^{M} c_m \rho_m(\boldsymbol{r})$ and functional $E\left[\sum_{m=1}^{M} c_m \rho_m(\boldsymbol{r})\right]$ can be rewrited by multivariable function as

$$E\left[\sum_{m=1}^{M} c_m \rho_m(\boldsymbol{r})\right] = E(c_1, c_2, \ldots,). \tag{2.6}$$

Weight parameters c_m are calculated from inner product of functions $\langle f, g\rangle$, expressed as

$$\langle f, g\rangle = \int f(\boldsymbol{r}) g(\boldsymbol{r}) d\boldsymbol{r}. \tag{2.7}$$

If $\{\rho_m(\boldsymbol{r})\}_{1\leq m\leq M}$ are normalized orthogonal functions, in other words, for any n and m functions have the relationship as

$$\langle f_n, f_m\rangle = \delta_{nm}, \tag{2.8}$$

where δ_{nm} is Kronecker's delta, then c_m can be expressed as

$$c_m = \langle f_m, \rho\rangle. \tag{2.9}$$

Then, Eqn.(2.6) can be expressed as

$$E\left[\rho(\boldsymbol{r})\right] = E(\langle f_1, \rho\rangle, \langle f_2, \rho\rangle, \ldots, \langle f_M, \rho\rangle). \tag{2.10}$$

Inner products are expressed as

$$\langle f_m, \rho\rangle = \sum_i \int p(\boldsymbol{r} - \boldsymbol{r}_i) f_m(\boldsymbol{r}) d\boldsymbol{r} = \sum_i F_m(\boldsymbol{r}_i), \tag{2.11}$$

where $F_m(\boldsymbol{r}_i) = \int p(\boldsymbol{r} - \boldsymbol{r}_i) f_m(\boldsymbol{r}) d\boldsymbol{r}$.

The calculation of $\langle f_m, \rho\rangle$ depends on what $p(\boldsymbol{r})$ is used in Eqn. (2.1) and what basis functions $f_m(\boldsymbol{r})$ are used. It is not always easy to calculate this integration. However, that is equivalent to prepare complete basis functions $\{g_m(\boldsymbol{r})\}_m$ and use descriptors expressed as $\sum_i g_m(\boldsymbol{r}_i)$ because set of $\{F_m(\boldsymbol{r})\}_m$ can be calculated by linear combination of $\{g_m(\boldsymbol{r})\}_m$ and it can be regarded E as composite function of function representing energies by $\{F_m(\boldsymbol{r})\}_m$ and such linear combination. For the same reason, linearly independent functions are sufficient even if functions are not orthogonal.

Next, condition (B) and condition (D) need to be considered. Then, using spherical coordinates, basis functions are set to $f_n(r)Y_{lm}(\theta,\phi)$, where $Y_{lm}(\theta,\phi)$ $(0{\leq}l, -l{\leq}m{\leq}l)$ is spherical harmonics function. When we don't consider long range interaction, $f_n(r)$ is zero when $r \geq 0$. Then condition (B) is satisfied.

Whereas radial parts $f_n(r)$ have obviously rotation symmetry and those descriptors $\sum_i f_n(r_i)$ can be used without translation, when angular basis functions are included, descriptors $c_{nlm} = \sum_i f_n(r_i)Y_{lm}(\theta_i,\phi_i)$ need to be translated to some functions which satisfy rotational symmetry.

When spherical harmonic functions Y_{lm} are translated into $R(\alpha,\beta,\gamma)Y_{lm}$ by rotation of Euler angle α, β and γ, following relationship is satisfied.

$$R(\alpha,\beta,\gamma)Y_{lm}(\theta,\phi) = \sum_{m'=-l}^{l} D^l_{mm'}(\alpha,\beta,\gamma)Y_{lm'}(\theta,\phi), \tag{2.12}$$

where $D^l_{mm'}(\alpha,\beta,\gamma)$ are elements of Wigner D-matrix. Therefore, descriptor of rotated distribution function $R(\alpha,\beta,\gamma)c_{nlm}$ also can be expressed as

$$R(\alpha,\beta,\gamma)c_{nlm} = \sum_{m'=-l}^{l} D^l_{mm'}(\alpha,\beta,\gamma)c_{nlm'}. \tag{2.13}$$

To satisfy rotational symmetry, descriptors need to converted by function G_p satisfying

$$f(c_{000},\ldots,c_{nlm},\ldots) = G_p(R(\alpha,\beta,\gamma)c_{000},\ldots R(\alpha,\beta,\gamma)c_{nlm},\ldots). \tag{2.14}$$

Because there are $2l+1$ coefficients of l and rotational operation in three dimensional space has two degree of freedom, there are $2l-1$ independent rotationally symmetrical functions. It is difficult to find such functions, however, because Wigner D-matrix is unitary matrix and norm of vector is invariant under unitary transformation, it can be easily found that one of those functions is

$$p_{nl} = \sum_{l=-m}^{l} \|c_{nlm}\|^2. \tag{2.15}$$

When radial function is not considered, such function is also known as bond order parameter, which represents local structures in liquid crystal and glass states [40]. Besides, by the combination of the third order bond order parameter and radial functions, other rotationally invariant can be created like

$$W_{nl} = \sum_{m_1,m_2,m_3=-l}^{l} \begin{pmatrix} l,l,l \\ m_1,m_2,m_3 \end{pmatrix} c_{nlm_1} c_{nlm_2} c_{nlm_3} \tag{2.16}$$

, where the coefficient written by the parentheses is the Wigner $3j$ symbol.

2.3 Relationship among general interatomic potentials, conventional potentials and MLIP

For the simplicity, firstly only radial descriptors are considered. Then, interatomic potential without angular descriptors are written as

$$E\left[\rho(\boldsymbol{r})\right] = E(\sum_i f_1(r_i), \sum_i f_2(r_i), \ldots,). \tag{2.17}$$

Incidentally, according to a general classification of the interatomic potential by milestone paper by Carlsson [32], interatomic potentials can be classified into four classes, namely,

(a) Pair potential

$$E = \frac{1}{2} \sum_i V_{\mathrm{pair}}\left(\boldsymbol{r}_i\right), \tag{2.18}$$

(b) Pair functional potential

$$E = \frac{1}{2} \sum_i V_{\mathrm{pair}}\left(\boldsymbol{r}_i\right) + U\left(\sum_i g_{\mathrm{pair}}(\boldsymbol{r}_i)\right), \tag{2.19}$$

(c) Cluster function potential

$$E = \frac{1}{2!} \sum_i V_{\mathrm{pair}}(\boldsymbol{r}_i) + \frac{1}{3!} \sum_{i,j} V_{\mathrm{three}}(\boldsymbol{r}_i, \boldsymbol{r}_j) + \ldots, \tag{2.20}$$

and (d) Cluster functional potential

$$E = \frac{1}{2} \sum_i V_{\mathrm{pair}}(\boldsymbol{r}_i) + U\left(\sum_i g_{\mathrm{pair}}(\boldsymbol{r}_i), \sum_{i,j} g_{\mathrm{three}}(\boldsymbol{r}_i, \boldsymbol{r}_j), \cdots\right). \tag{2.21}$$

When angular terms are not considered, conventional interatomic potentials are classified into pair potential expressed as Eqn.(2.18) or pair functional potential expressed as Eqn.(2.19). When the rotational symmetry is satisfied, the both potential is expressed by distance r_i. Then, these conventional potentials are also expressed by form of Eqn.(2.17). However, these conventional potentials can not express all potentials expressed by Eqn.(2.17). For example, potential like $[\sum_i f_1(r_i)]^2 + [\sum_i f_2(r_i)]^2$ can be expressed by Eqn.(2.17) but can not be expressed by neither Eqn.(2.18) nor Eqn.(2.19). It can be said that in conventional interatomic potential idea, the special case of Eqn.(2.17) are constructed on physical consideration and trial and error.

Next, MLIP is compared to Eqn.(2.17). A formula of neural network potential (NNP) proposed by Behler [22] and Gaussian approximation potentials (GAP) proposed by Bartók [23] using only radial descriptor can be interpreted that the model in which function E of Eqn.(2.17) is determined with artificial neural network method and with Gaussian process, respectively. LASSO potentials proposed by Seko [24, 31] is expressed as

$$E = \sum_{1\leq m\leq M, 1\leq p\leq P} w_{m,p} \left[\sum_i f_m(r_i)\right]^p + C, \tag{2.22}$$

where $w_{m,p}$ and C are fitting parameter determined by linear regression method. When functions $f_m(r_i)$ are linearly independent, this expression is less general than Eqn.(2.17) because some potentials like $[\sum_i f_1(r_i)]^2 + 3\left[\sum_i f_1(r_i)\right]\left[\sum_i f_2(r_i)\right] + [\sum_i f_2(r_i)]^2$ can not be expressed by Eqn.(2.22). The expression of interatomic potential by linear combination of descriptors should be considered by polynomial approximation of Eqn.(2.17) and expressed as

$$E = \sum_{1\leq s_1\leq\cdots\leq s_P\leq M+1} w_{s_1,s_2,\ldots,s_P} \left[\sum_i f_{s_1}(r_i) \sum_i f_{s_2}(r_i) \cdots \sum_i f_{s_P}(r_i)\right], \tag{2.23}$$

where P is order of polynomial approximation. The difference of this form and Eqn.(2.22) is only existence of cross terms, namely, product of summation of different basis functions. Evidently, this form is more general than previous linear potential. Construction of interatomic potentials by Eqn.(2.23) will be examined in Chapter 3.

Next, angular descriptors are considered. In this case, general form of interatomic potential is expressed as

$$E\left[\rho(\boldsymbol{r})\right] = E(\sum_i f_1(r_i), \sum_i f_2(r_i), \dots, p_{11}, \dots, p_{nl}, \dots, W_{11}, \dots, W_{nl}, \dots). \tag{2.24}$$

Firstly, the relationship between descriptor p_{nl} and bonding angle are discussed. For the simplicity, at first radial parts are not considered and only the case of unit vectors are shown. Then, let us represent descriptors when there are N atoms whose coordinates are $\boldsymbol{r}_i = (\theta_i, \phi_i)(1 \leq i \leq N)$, namely, $\sum_{m=-l}^{l} \|c_{lm}\|^2 = \sum_{m=-l}^{l} (\sum_i Y_{lm}(\theta_i, \phi_i) \sum_i Y_{lm}^*(\theta_i, \phi_i))$. Incidentally, the following relationship is satisfied.

$$P_l(\boldsymbol{x} \cdot \boldsymbol{y}) = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_{lm}(\theta', \phi') Y_{lm}^*(\theta, \phi), \tag{2.25}$$

where P_l is a Legendre polynomial of degree l and $\boldsymbol{x}$ and $\boldsymbol{y}$ are unit vectors which have spherical coordinates (θ, ϕ) and (θ', ϕ'). Then, by calculate the sum of Eqn.(2.25) over i, the following formula is obtained.

$$\begin{aligned} p_l = \sum_{m=-l}^{l} \left(\sum_i Y_{lm}(\theta_i, \phi_i) \sum_i Y_{lm}^*(\theta_i, \phi_i) \right) = \\ \sum_{m=-l}^{l} \left(\sum_{i,j} Y_{lm}(\theta_i, \phi_i) Y_{lm}^*(\theta_j, \phi_j) \right) = \\ \sum_{i,j} P_l(\boldsymbol{r}_i \cdot \boldsymbol{r}_j). \end{aligned} \tag{2.26}$$

Considering radial function,

$$\begin{aligned} p_{nl} = \sum_{m=-l}^{l} \left[\sum_i Y_{lm}(\theta_i, \phi_i) f_n(r_i) \right] \left[\sum_i Y_{lm}^*(\theta_i, \phi_i) f_n(r_i) \right] = \\ \sum_{m=-l}^{l} \left(\sum_{i,j} Y_{lm}(\theta_i, \phi_i) Y_{lm}^*(\theta_j, \phi_j) f_n(r_i) f_n(r_j) \right) = \\ \sum_{i,j} P_l(\frac{\boldsymbol{r}_i}{r_i} \cdot \frac{\boldsymbol{r}_j}{r_j}) f_n(r_i) f_n(r_j). \end{aligned} \tag{2.27}$$

Inner product of two vector is also expressed as $\frac{\boldsymbol{r}_i}{r_i} \cdot \frac{\boldsymbol{r}_j}{r_j} = \cos(\gamma_{ij})$, where γ_{ij} is bonding angle. Therefore, p_{nl} can be expressed by function of radial r_i, r_j and bonding angle γ_{ij}. The use of descriptor set p_{nl} is also equivalent to use exponential functions of $\cos^l(\gamma_{ij})$, namely $\sum_{i,j} f_n(r_i) f_n(r_j) \cos^l(\gamma_{ij})$. Then, using p_{nl} can be interpreted as expansion of three body potential. By the same idea, probably the set of W_{nl} expresses four body potential and other rotationally invariants express many body potential.

Conventional interatomic potentials including angular terms are classified into cluster function potential expressed as Eqn.(2.20) or cluster functional potential expressed as Eqn.(2.21). When the rotational symmetry is satisfied, the both potentials are expressed by distance r_i, r_j and γ_{ij}. As with the relationship among pair potential, pair functional potential and Eqn.(2.17), conventional cluster potential and cluster functional potential are less general than Eqn.(2.24). For example, potential like $\left[\sum_{i,j} f_1(r_i) f_1(r_j) \cos\gamma_{ij}\right]^2 + \left[\sum_{i,j} f_2(r_i) f_2(r_j) \cos\gamma_{ij}\right]^2$ can be expressed by Eqn.(2.24) but can not be expressed by neither Eqn.(2.20) nor Eqn.(2.21).

A formula of NNP and GAP using angular descriptor can be interpreted that the model in which function E of Eqn.(2.24) is determined with artificial neural network method and with Gaussian process, respectively. Interatomic potential model by linear combination of descriptors with angular terms can be considered by polynomial approximation of Eqn.(2.24. Construction of such interatomic potentials will be examined in Chapter 4.

2.4 Summary

In this study, I defined interatomic potential as continuous mapping from atomic distribution function to scalar which satisfies rotational symmetry. I also showed that such mapping can be represented by multivariable function of descriptors, which are rotationally symmetrical functions of coefficients of expansion of atomic distribution function by radial and spherical harmonics function. It was shown that this model is more general expression of interatomic potentials than conventional potentials. NNP and GAP can be interpreted the method in which such multivariable function is determine by NNP and GAP. LASSO potential previously supposed is less general than the formula proposed in this study. To construct interatomic potential model by linear combination of descriptors, polynomial approximation should be applied to the formula proposed in this study and cross terms of different descriptors should be added.

Chapter 3

Construction of interatomic potentials by generalized linear expression

3.1 Introduction

Machine learning interatomic potential (MLIP) has great potential for facilitating the discovery of new physics. Its frameworks applicable to periodic systems have recently been proposed [22–24], which derive a reliable interatomic potential by a combination of density functional theory (DFT) calculations for many different atomic configurations and a regression method. Unknown coefficients describing MLIP are estimated from the energies of the training data set using a regression method. Importantly, MLIP is described by some descriptors obtained from atomic positions, satisfying several invariances, such as translational and rotational invariance. General advantages of MLIP are as follows. Its accuracy should be much better than that of conventional interatomic potentials and very close to that of DFT calculation. On the other hand, the computational cost of MLIP is very close to that of conventional interatomic potentials. In addition, because the framework and descriptors are flexible, MLIP is applicable to a wide range of materials regardless of their nature of chemical bonding.

In the literature, only a few regression techniques such as artificial neural networks [22, 25–30], Gaussian process regression [23] and least absolute shrinkage and selection operator (LASSO) [24, 31] have been adopted to estimate MLIP for solids. Applications to solids have been actually limited to a small number of metallic [24, 25, 27, 31], covalent [22, 23, 28] and ionic materials [26]. Also, several descriptors have been proposed [22–24, 33–39], while only a few of them have succeeded in obtaining accurate MLIP.

MLIP provides a description for energy with high accuracy, however, it is still unclear how we should prepare the MLIP model and descriptor set. Although MLIP model and descriptors set crucially determine the accuracy and computational costs of interatomic potential, method to apply machine learning and descriptor set are determined by trial and error rather than by clear theory based on physical and mathematical consideration. One way to select effective descriptors is expressing interatomic potential model linearly and the use of compressing sensing methods like LASSO technique. Indeed, it is reported that such model can be applied to several elemental metals and optimal interatomic potential can reproduce several material properties like total energies of structures and phonon dispersion curve from DFT calculations [24, 31]. Although LASSO potentials can select effective descriptors and construct accurate interatomic potentials, the preparation of candidate descriptors is still difficult. If unnecessary candidate descriptors are used, the computational costs of construction of interatomic potential becomes large and such extra descriptors likely cause overfitting and decrease generalization ability of the interatomic potential model. Moreover, when the candidate set lacks necessary descriptors, we can never obtain the accurate model no matter how we select descriptors.

The difficulty of consideration of descriptor sets can be ascribed to the fact that mathematical interpretation of MLIP and the understanding the relationship between MLIP and already existing interatomic potentials are still lacking. Machine learning has been used only as a tool for estimating a black-box function. The main purpose of this study is to express general formula of interatomic potential and show the relationship between conventional interatomic potentials and MLIP.

In Chapter 2, however, I showed the general interatomic potential model and relationship among such model, conventional potential, and MLIP. I also showed that linear potential model should be constructed by polynomial approximation of general model and such model is more general than previous linear potential model. In this chapter, linear interatomic potentials of 31 elemental metals are constructed by the generalized framework and the previous framework [24, 31], and compare the accuracy of both models. At last, I estimated several properties of materials, namely energies of structures which is not used to fitting, elastic constants, and phonon dispersion curves. For Al and Cu, linear potential constructed by generalized framework is compared also with conventional EAMs and the applicability of the new framework proposed in this study is investigated.

3.2 Methodology

3.2.1 Linear expressions of interatomic potentials

In general interatomic potential idea, the total energy of a target system E^{total} is decomposed into contribution of each atoms included the system. Therefore, E^{total} can be written as

$$E^{\text{total}} = \sum_i E^{(i)}, \tag{3.1}$$

where $E^{(i)}$ is the contribution of atom i.

As discussed in Chapter 2, if target system is nonmagnetic and unary, a general interatomic potential model is written as

$$E^{(i)} = E\big(\sum_j f_1(\boldsymbol{r}_j - \boldsymbol{r}_i), \sum_j f_2(\boldsymbol{r}_j - \boldsymbol{r}_i), \ldots, \sum_j f_M(\boldsymbol{r}_j - \boldsymbol{r}_i)\big), \tag{3.2}$$

where functions $\{f_m(\boldsymbol{r})\}_{1 \leq m \leq M}$ are smooth basis functions and in the region where $r \geq R_c$, function is zero. R_c is a finite cutoff radius and if the distance between two atoms is larger than R_c, the interaction of such atoms is ignored.

When there is no external field, Eqn.(3.2) has to satisfy the rotational symmetry. The simplest way is preparing the basis functions which satisfy the rotational symmetry, however, it is not easy to prepare such basis functions. In this chapter, I only consider the function which is depends on only distance between atoms as the basis functions. For other functions which satisfy rotational symmetry will be written in the Chapter 4. When we consider only radial basis functions, Eqn.(3.2) can be rewrite as

$$E^{(i)} = E\big(\sum_j f_1(r_{ij}), \sum_j f_2(r_{ij}), \ldots, \sum_j f_M(r_{ij})\big), \tag{3.3}$$

where r_{ij} is a distance between atom i and atom j. For the simplicity, I denote the variable in the Eqn.(3.3) as

$$b_m^{(i)} = \sum_j f_m(r_{ij}) \tag{3.4}$$

and rewrite Eqn.(3.3) as

$$E^{(i)} = E\big(b_1^{(i)}, b_2^{(i)}, \dots\big). \tag{3.5}$$

A linear interatomic potential can be derived by expressing E of Eqn.(3.2) by a multivariable polynomial. Using radial descriptors, by expanding Eqn.(3.5), linear interatomic potential can be written as a polynomial of degree P, namely

$$E^{(i)} = \sum_{1 \le s_1 \le \dots \le s_P \le M+1} w_{s_1,s_2,\dots,s_P} b_{s_1}^{(i)} b_{s_2}^{(i)} \dots b_{s_P}^{(i)}, \tag{3.6}$$

where $b_{M+1} = 1$ and $w_{s_1,s_2,\dots,s_P}$ is fitting parameters. A framework to construct linear interatomic potential is proposed previously [24, 31] and in this framework, interatomic potential is expressed as

$$E^{(i)} = \sum_{1 \le m \le M, 1 \le p \le P} w_{m,p} \left[b_m^{(i)}\right]^p + C. \tag{3.7}$$

Comparing with a linear interatomic potential model previously, Eqn.(3.6) is more general form than previous model because Eqn.(3.6) includes cross terms of different descriptors whereas previous model uses only exponentials of same descriptor.

In this study, I used total energies of crystal structures and forces acting on atoms composing the crystal structures to construct interatomic potential model and estimate several material properties. The total energy of crystal structures of Eqn.(3.6) and Eqn.(3.7) can be expressed as

$$E^{\mathrm{Total}} = \sum_i \sum_{1 \le s_1 \le \dots \le s_P \le M+1} w_{s_1,s_2,\dots,s_P} b_{s_1}^{(i)} b_{s_2}^{(i)} \dots b_{s_P}^{(i)} \tag{3.8}$$

and

$$E^{\mathrm{Total}} = \sum_i \left[\sum_{1 \le m \le M, 1 \le p \le P} w_{m,p} \left[b_m^{(i)}\right]^p + C \right], \tag{3.9}$$

respectively. A force vector acting on atom k of the both models can be expressed as

$$-\frac{\delta E^{\mathrm{Total}}}{\delta \boldsymbol{x}_k} = -\sum_i \sum_{1\le s_1 \le \cdots \le s_P \le M+1} w_{s_1,s_2,\ldots,s_P} \left[\sum_{q=1}^{P} b_{s_1}^{(i)} b_{s_2}^{(i)} \ldots \frac{\delta b_{s_q}^{(i)}}{\delta \boldsymbol{x}_k} \ldots b_{s_P}^{(i)} \right] \tag{3.10}$$

and

$$-\frac{\delta E^{\mathrm{Total}}}{\delta \boldsymbol{x}_k} = -\sum_i \sum_{1\le m\le M, 1\le p\le P} w_{m,p} p \left[b_m^{(i)} \right]^{p-1} \frac{\delta b_m^{(i)}}{\delta \boldsymbol{x}_k}. \tag{3.11}$$

3.2.2 Regression method

After we prepare an interatomic potential model, the principal task of constructing an interatomic potential is to estimate regression coefficients from DFT calculation set with regression technique. In this study, I use total energies and forces acting on atoms computed with DFT calculations of many structures as samples in the regression process because all of them can be expressed by linear equations with the same expansions. As can be seen from Eqn.(3.8) and Eqn.(3.9), total energy of crystal structure n of both the interatomic potentials with cross terms and without cross terms can be denoted as

$$E^{(n,\mathrm{total})} = \sum_m w_m \left[\sum_i a_m^{(n,i)} \right], \tag{3.12}$$

where $a_m^{(n,i)}$ and w_m is the mth descriptor of atom i of structure n and its weight parameter, respectively. $a_m^{(n,i)}$ corresponds to $b_{s_1}^{(i)} b_{s_2}^{(i)} \ldots b_{s_P}^{(i)}$ of Eqn.(3.8) and $\left[b_m^{(i)} \right]^p$ of Eqn.(3.9). Force acting on atom k can be also expressed as

$$F^{(n,j)} = \frac{\delta E^{(n,\mathrm{total})}}{\delta \boldsymbol{r}_j} = -\sum_m w_m \left[\sum_i \frac{\delta a_m^{(n,i)}}{\delta \boldsymbol{r}_j} \right]. \tag{3.13}$$

Because the both of equations are linear connection of weight parameters w_m, w_m are determined with linear regression. Let us denote the explanation variables matrix X of M

descriptors of N structures as

$$X = \begin{pmatrix} \sum_i a_1^{(1,i)} & \dots & \sum_i a_m^{(1,i)} & \dots & \sum_i a_M^{(1,i)} \\ \vdots & & \vdots & & \vdots \\ \sum_i a_1^{(N,i)} & \dots & \sum_i a_m^{(N,i)} & \dots & \sum_i a_M^{(N,i)} \\ \sum_i \frac{\delta a_1^{(1,i)}}{\delta r_{1x}} & \dots & \sum_i \frac{\delta a_m^{(1,i)}}{\delta r_{1x}} & \dots & \sum_i \frac{\delta a_M^{(1,i)}}{\delta r_{1x}} \\ \sum_i \frac{\delta a_1^{(1,i)}}{\delta r_{1y}} & \dots & \sum_i \frac{\delta a_m^{(1,i)}}{\delta r_{1y}} & \dots & \sum_i \frac{\delta a_M^{(1,i)}}{\delta r_{1y}} \\ \sum_i \frac{\delta a_1^{(1,i)}}{\delta r_{1z}} & \dots & \sum_i \frac{\delta a_m^{(1,i)}}{\delta r_{1z}} & \dots & \sum_i \frac{\delta a_M^{(1,i)}}{\delta r_{1z}} \\ \sum_i \frac{\delta a_1^{(1,i)}}{\delta r_{2x}} & \dots & \sum_i \frac{\delta a_m^{(1,i)}}{\delta r_{2x}} & \dots & \sum_i \frac{\delta a_M^{(1,i)}}{\delta r_{2x}} \\ \vdots & & \vdots & & \vdots \\ \sum_i \frac{\delta a_1^{(N,i)}}{\delta r_{1x}} & \dots & \sum_i \frac{\delta a_m^{(N,i)}}{\delta r_{1x}} & \dots & \sum_i \frac{\delta a_M^{(N,i)}}{\delta r_{1x}} \\ \sum_i \frac{\delta a_1^{(N,i)}}{\delta r_{1y}} & \dots & \sum_i \frac{\delta a_m^{(N,i)}}{\delta r_{1y}} & \dots & \sum_i \frac{\delta a_M^{(N,i)}}{\delta r_{1y}} \\ \sum_i \frac{\delta a_1^{(N,i)}}{\delta r_{1z}} & \dots & \sum_i \frac{\delta a_m^{(N,i)}}{\delta r_{1z}} & \dots & \sum_i \frac{\delta a_M^{(N,i)}}{\delta r_{1z}} \\ \vdots & & \vdots & & \vdots \end{pmatrix}, \tag{3.14}$$

where r_{jx},r_{jy}, and r_{jz} are x,y, and z component of $\boldsymbol{r}_j$. The numbers of rows and columns are corresponds to the number of observed data and descriptors, respectively. For objective variables, DFT calculations are used. Let us denote the total energy of structure n and the component of force acting on atom k of structure n as $E_{\mathrm{DFT}}^{n,\mathrm{total}}$ and $F_{\mathrm{DFT}}^{n,k,l}(l = x, y, z)$. Objective variables vector $\boldsymbol{y}$ is expressed as

$$\boldsymbol{y} = \begin{pmatrix} E_{\mathrm{DFT}}^{1,\mathrm{total}} \\ \vdots \\ E_{\mathrm{DFT}}^{N,\mathrm{total}} \\ F_{\mathrm{DFT}}^{1,1,x} \\ \vdots \\ F_{\mathrm{DFT}}^{N,1,x} \\ \vdots \end{pmatrix}. \tag{3.15}$$

The dimension of $\boldsymbol{y}$ corresponds to the number of observed data set. The weight parameters vector w is expressed as

$$\boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{pmatrix}. \tag{3.16}$$

There are several linear regression methods like ordinary least-squares regression, partial least-squares regression, using pseudo-inverse matrix, and LASSO technique. In this study, for the simplicity, I used simple linear ridge regression to determine the regression coefficients to estimate and compare performance of the descriptor sets. In this method, the regression coefficients vector $\boldsymbol{w}$ is determined to minimize the minimization function expressed as

$$L(\boldsymbol{w}) = (\boldsymbol{y} - X\boldsymbol{w})^{\mathrm{T}}(\boldsymbol{y} - X\boldsymbol{w}) + \lambda \boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}, \tag{3.17}$$

where λ controls the magnitude of the penalty. The first terms of Eqn.(3.17) is the residual sum of squares error between the properties of samples and predicted properties with constructed model, and the second term is regularization terms, which is introduced to prevent overfitting due to excessively large regression coefficients.

$\boldsymbol{w}$ minimizing $L(\boldsymbol{w})$ can be easily determined by solving a normal equation. Because Eqn.(3.17) is a quadratic equation of fitting parameters of $\boldsymbol{w}$, $\boldsymbol{w}$ minimizing $L(\boldsymbol{w})$ satisfies the condition

$$\frac{\delta L(\boldsymbol{w})}{\delta \boldsymbol{w}} = \mathbf{0}. \tag{3.18}$$

From this equation, $\boldsymbol{w}$ can be denoted as

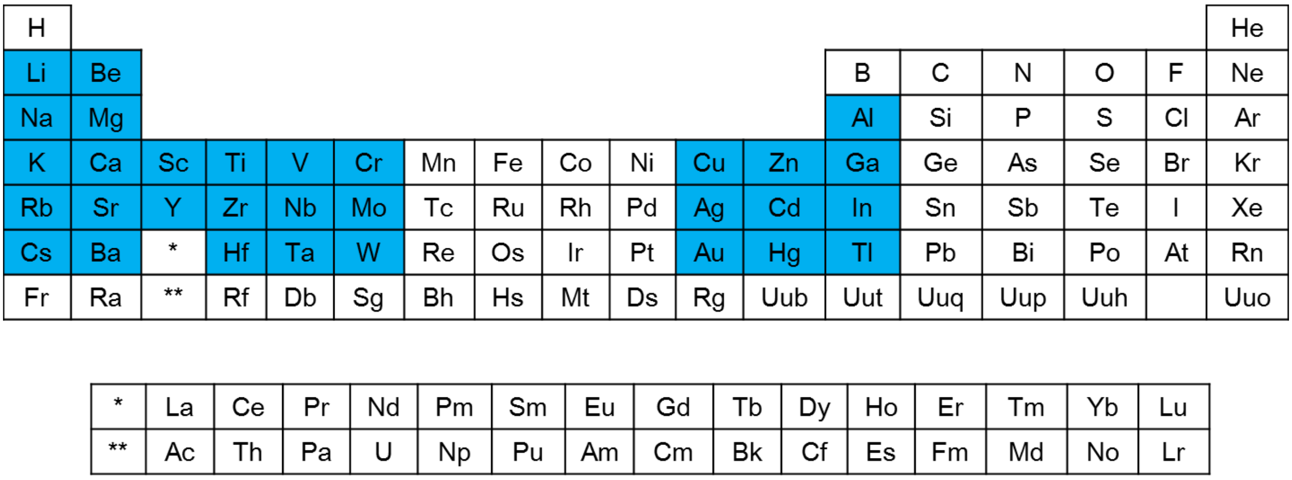
$$\boldsymbol{w} = (X^{\mathrm{T}}X + \lambda I)^{-1}X^{\mathrm{T}}y, \tag{3.19}$$

where I is an unit matrix whose dimension is the number of descriptors.

3.3 Target elements and structures

The training and test data sets are generated by the DFT calculation. To generate training and test data sets, I started to fully optimize the atomic positions and lattice constants of conventional FCC, BCC, HCP, simple cubic (SC), ω, and β-tin structures for each element. Next, from the six ideal structures, I constructed supercells, which are $2 \times 2 \times 2$ FCC (32 atoms), $3 \times 3 \times 3$ BCC (54 atoms), $3 \times 3 \times 3$ HCP (54 atoms), $4 \times 4 \times 4$ SC (64 atoms), $3 \times 3 \times 3$ ω (81 atoms), and $2 \times 2 \times 2$ β-tin (32 atoms) structures. Using the supercells, training and test set are generated by 30 isotropic expansions and 470 random expansions and distortions on the basis of the six ideal supercells. The volumes of 30 structures generated by isotropic expansions. Volume ratio of optimized structure and generated structures varies from 0.5 to 5.0. Because framework to construct MLIP has no physical background, such greatly compressed and expanded structures are necessary to make interatomic potential predicts high energy when distances between atoms are very close or when atoms isolate. Then, 500 deformed structures are generated for one structure and totally 3000 structures are generated for one elemental metal. Finally, the 3000 configurations are divided into 2700 training data sets and 300 test data sets of each element, respectively.

For all the configurations, energy and forces acting on each atom are calculated by the DFT calculation using the plane-wave basis projector augmented wave (PAW) method [41, 42] within the Perdew-Burke-Ernzerhof exchange-correlation functional [43] as implemented in the VASP code [44, 45]. The cutoff energy was set to 500 eV. The total energies converged to less than 10^{-3} meV/supercell. For optimization of lattice constants of ideal structures, the lattice constants were optimized until the residual forces became less than 10^{-3} eV/Å.



H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	**	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Uub	Uut	Uuq	Uup	Uuh		Uuo

*	La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
**	Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr

Figure 3.1: 31 nonmagnetic elements of which interatomic potentials are constructed.

3.4 Construction of interatomic potential

The principal task of constructing an interatomic potential is to estimate regression coefficients. To construct interatomic potentials, it is necessary to optimize some hyperparameters like cutoff radius, the order of expansion of polynomial approximation, and number of basis functions. In this section, how to optimize those parameters in this research and the dependence of accuracy on those input parameters are shown.

3.4.1 Cutoff radius

In this subsection, optimal cutoff radius for construction of interatomic potentials by framework with cross terms and that without cross terms is determined. I searched for an optimal cutoff radius for 31 elements by constructing interatomic potentials whose cutoff radius is ranging from 5 Åto 14 Åwith intervals of 1.0 Å. For all cutoff radius, 12 Gaussian functions is used as basis functions as expressed

$$f_n(r) = \exp\bigl(-a_n(r-b_n)^2\bigr). \tag{3.20}$$

$(a_n, b_n)_{1\leq n\leq 12}$, or 12 pairs of Gaussian parameters is determined by fixed a and arithmetic sequence from 0 to $b_{\max}$, or $(a_n, b_n)_{1\leq n\leq 12} = (a, (n-1)b_{\max}/11)_{1\leq n\leq 12}$. I constructed totally 81 interatomic potentials with $a = 0.1, 0.25, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5$ and $b_{\max}$ is varied from 5.0 Åto 13.0 Åfor each cutoff radii, and the interatomic potentials whose RMS error of test set is the lowest is selected as the optimal one of the cutoff radii. The order of polynomials was set to 3 for both the model with cross terms and the model without cross terms. The ridge penalty terms was set to $\lambda = 10^{-5}$.

Figure 3.2 and figure 3.3 shows the relationship between cutoff radius and test set root mean square (RMS) errors of total energies and the relationship between cutoff radius and test set RMS errors of forces acting on atoms, respectively. For all elements, converged cutoff radii of model with cross terms are not greatly different from that of model without cross terms. Table 3.1 shows cutoff radius which minimized RMS errors for energies. In this study, those cutoff radii were applied to construct interatomic potentials as optimal cutoff radii.

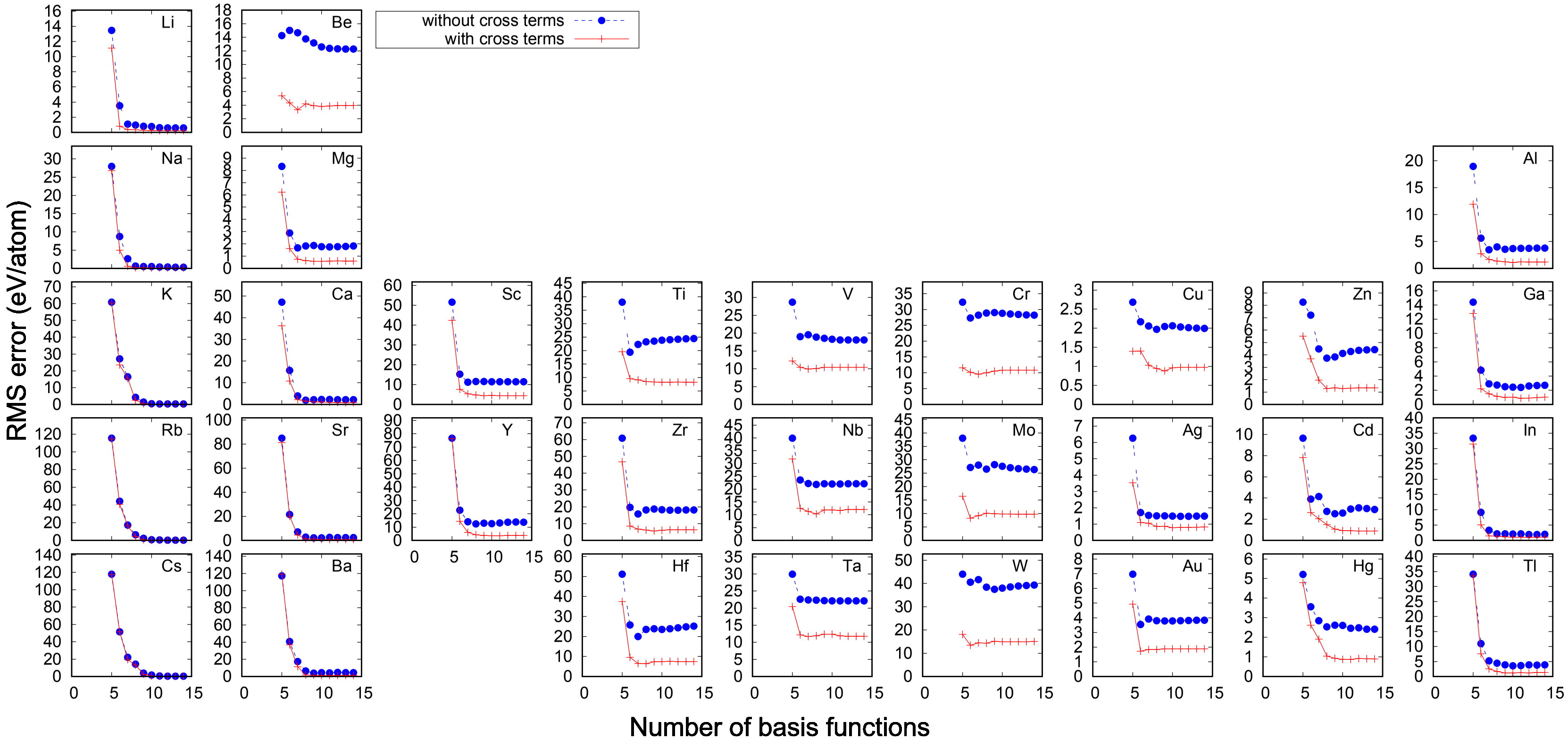


Figure 3.2: Dependence of RMS errors of total energies on cutoff radius.

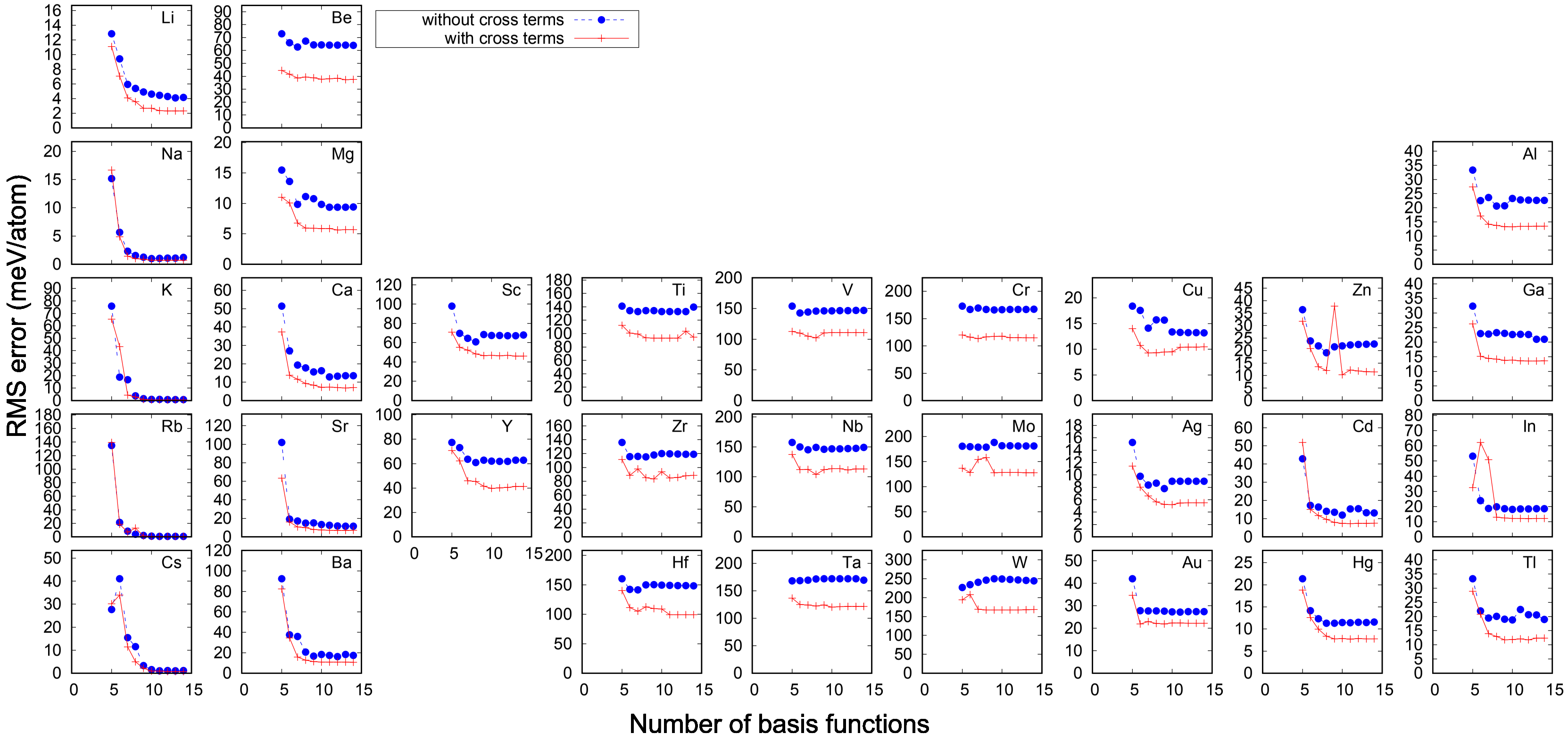


Figure 3.3: Dependence of RMS errors of forces acting on atoms on cutoff radius.

Table 3.1: The optimal cutoff radius for 31 elements.

Element	cutoff radius (Å) (with cross terms)	cutoff radius (Å) (without cross terms)
Li	12	13
Be	7	14
Na	12	14
Mg	10	7
Al	10	7
K	13	14
Ca	13	8
Sc	14	7
Ti	13	6
V	7	12
Cr	7	6
Cu	9	8
Zn	8	8
Ga	11	11
Rb	14	12
Sr	12	9
Y	11	8
Zr	9	7
Nb	8	8
Mo	6	14
Ag	10	12
Cd	13	9
In	12	12
Cs	14	14
Ba	10	9
Hf	8	7
Ta	7	14
W	6	9
Au	6	6
Hg	10	13
Tl	10	10

3.4.2 Basis functions and expansion order

Accuracy and the computational costs of interatomic potential depend on what and how many basis functions are used. In this subsection, I investigate the relationship and optimize species and number of basis functions.

The computational costs of prediction energy and force depend on not the number of total terms but on the number of basis functions. This is because exponential and cross terms can be calculated with only several multiple variables from sum of basis functions $\sum_j f_n(r_{ij})$ whereas sum of basis functions $\sum_j f_n(r_{ij})$ need calculation of function $f_n(r_{ij})$ as many as neighbor atoms (usually $10 \sim 100$ atoms), and because calculation of cross terms can be accelerated with recursive method. Therefore, in this study the number of basis function is used as criterion of the computational costs of prediction. In this section, I investigate the relationship between difference of prediction value from DFT calculations and number of basis functions for various species of basis functions and expansion order.

Firstly, I investigated the relationship between accuracy of interatomic potentials and the number of basis functions. I constructed using Bessel function $J_n(r_{ij})$, Neumann function $Y_n(r_{ij})$, spherical Bessel function $j_n(r_{ij})$, spherical Neumann function $y_n(r_{ij})$, cosine function $\cos(a_n r_{ij})$, modified Morlet wavelet (MMW) $\cos(a_n r_{ij})/\cosh(r_{ij})$, and Gaussian function like Eqn.(3.20) as f_n with interatomic potential model with cross terms and that without cross terms.

For Bessel, Neumann, spherical Bessel, and spherical Neumann functions, I set the minimum and interval of the n to zero and one, respectively. For cosine function and MMW function, a_n is given by arithmetic sequence. Minimum of a_n is fixed to zero and maximum of a_n is varied from 1.0 $\mathrm{\mathring{A}}^{-1}$ to 10.0 $\mathrm{\mathring{A}}^{-1}$ with intervals of 0.1 $\mathrm{\mathring{A}}^{-1}$, and the maximum of a_n which minimize the RMS error of energies of test set is adopted. For Gaussian function, a_n and b_n is optimized by the same manner as the previous Subsection 3.4.1.

The expansion power P is fixed to 3 for the both interatomic potential models. The same cutoff radii as previous Subsection 3.4.1 are used. The ridge penalty terms was set to $\lambda = 10^{-3}, 10^{-4}$, and 10^{-5} and penalty terms which minimize the RMS error of energies of test set is adopted.

Figure 3.4 and 3.5 show the convergence of the RMS errors for energies and forces with respect to the number of basis functions for various basis function types of interatomic potential model without cross terms. Similarly, figure 3.6 and 3.7 show the convergence of the RMS errors for energies and forces with respect to the number of basis functions for various basis function types of interatomic potential model with cross terms. For the both interatomic potential models, RMS errors of both potential models whose basis functions are Bessel and spherical Bessel functions were relatively large for the most of elements. Gaussian, cosine and MMW functions are good basis functions because both RMS errors for energies and forces are relatively low. RMS errors of Neumann and spherical Neumann are middle of all basis functions applied in this study. For all elements and both of two models, Gaussian functions construct the most accurate or close to the most accurate interatomic potentials. Therefore, I applied Gaussian functions as basis functions for both model with cross terms and model without cross terms.

Next, I investigated the relationships between expansion power P and RMS error. I used Gaussian function of Eqn.(3.20) as basis functions. Figure 3.8 and 3.9 show the dependency of RMS errors for energies and forces of both models on P. As can be seen, for all elements, both of the converged number of basis functions and RMS errors of model with cross terms are lower than those of model without cross terms. As mentioned above, because the number of basis functions crucially determines the computational costs, this result indicates the computational cost of model with cross terms is as much as that without cross terms. It can be seen that RMSEs of the most of 31 elements are converged at $P = 3$ for the both models.

Table 3.2 and 3.3 show the optimal parameters and RMS errors of energies and forces of the model without cross terms and the model with cross terms, respectively. For all 31 elements, using cross terms, RMS errors of energies decreased by more than 46% and RMS errors of forces decreased by more than 23%.

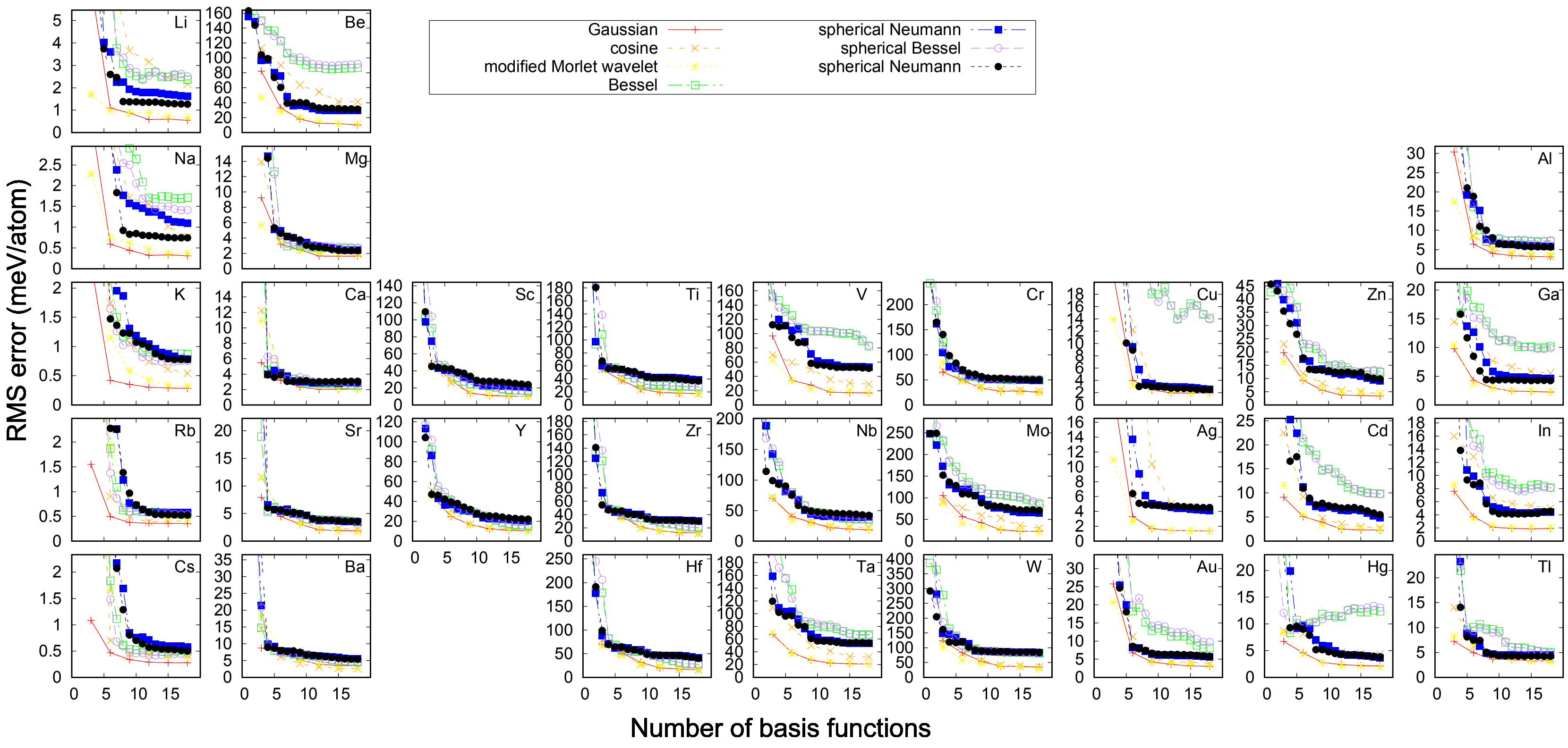


Figure 3.4: The dependency of RMS errors for energies on types of basis functions for model without cross terms.

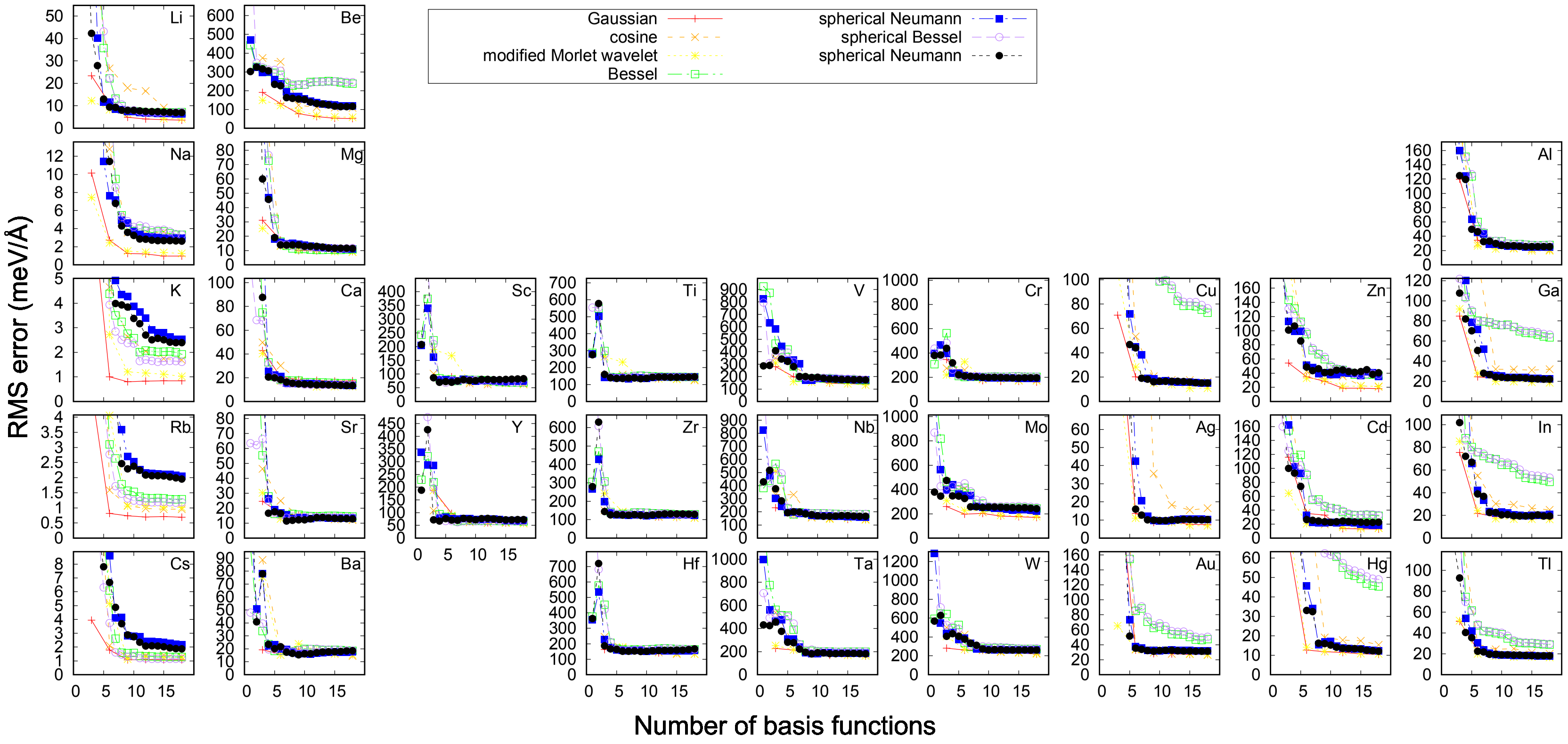


Figure 3.5: The dependency of RMS errors for forces on types of basis functions for model without cross terms.

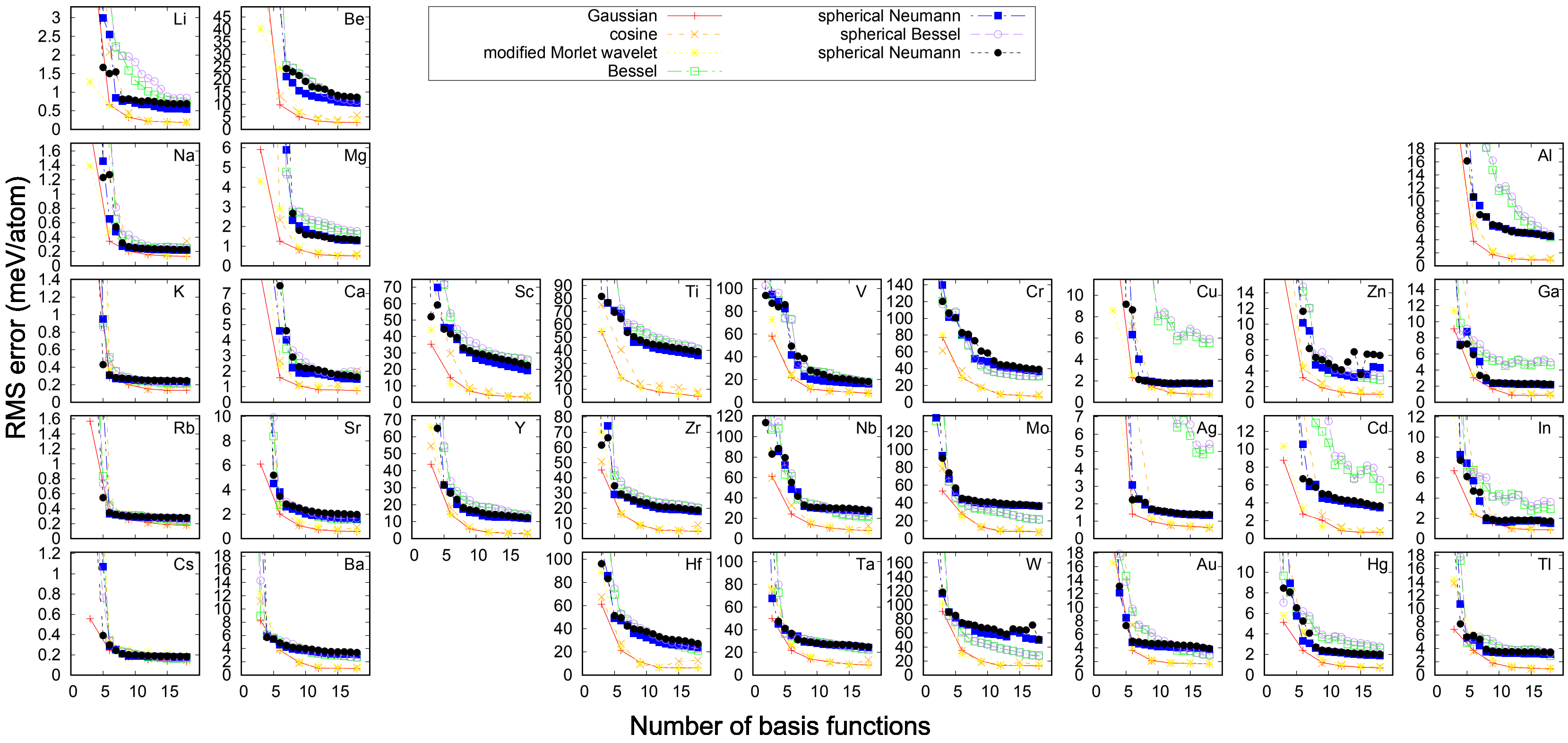


Figure 3.6: The dependency of RMS errors for energies on types of basis functions for model with cross terms.

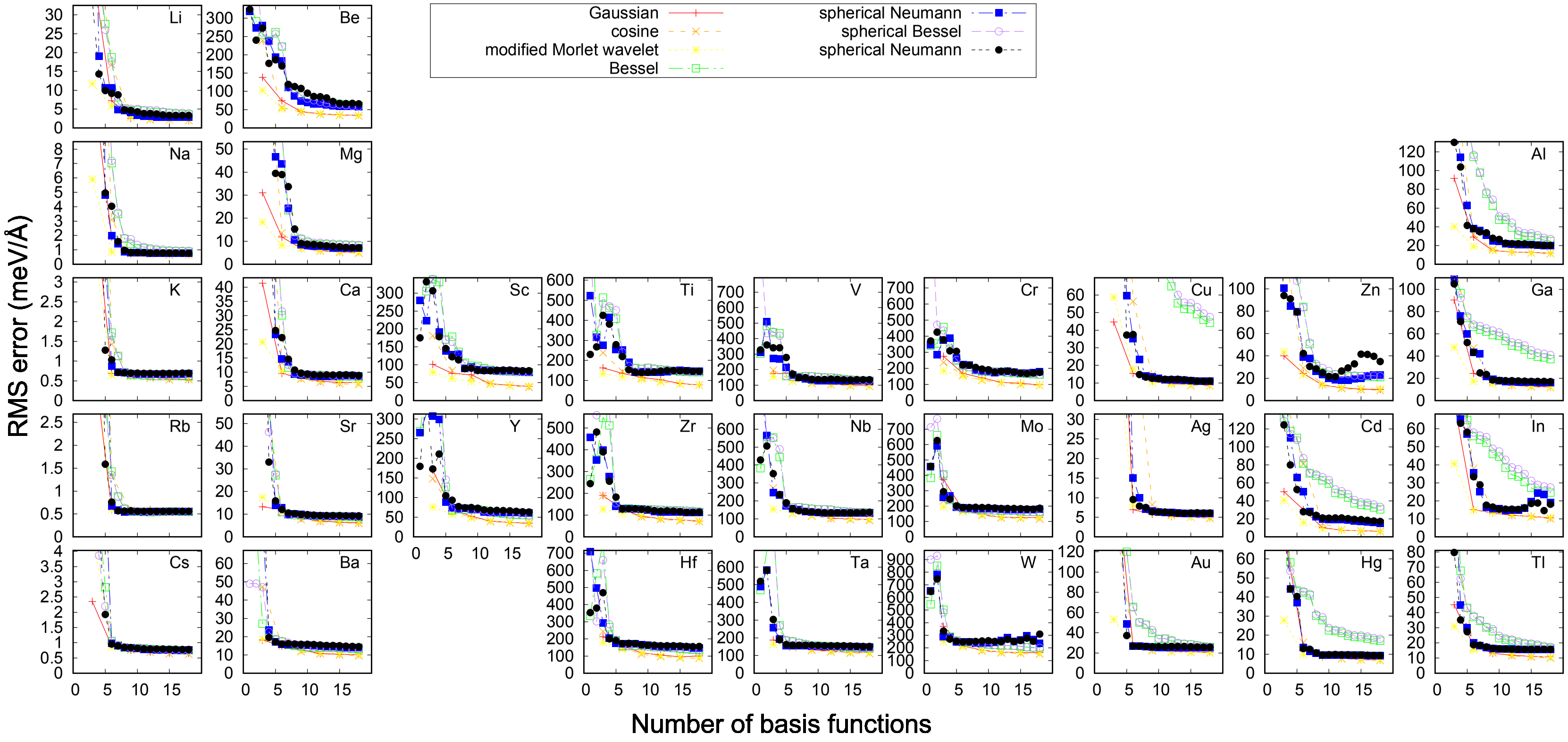


Figure 3.7: The dependency of RMS errors for forces on types of basis functions for model with cross terms.

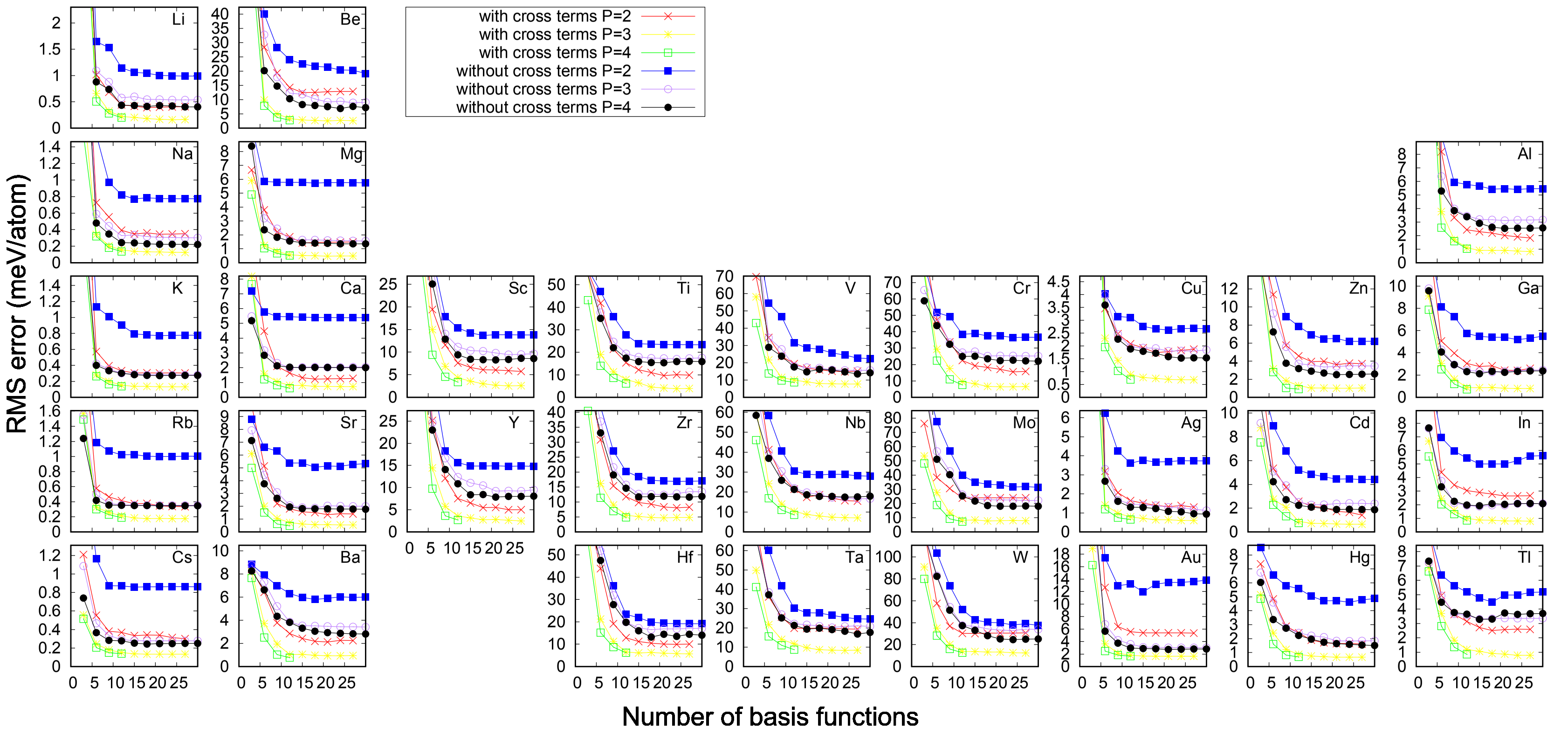


Figure 3.8: The dependency of RMS errors for energies on the expansion order P for both model with cross terms and model without cross terms.

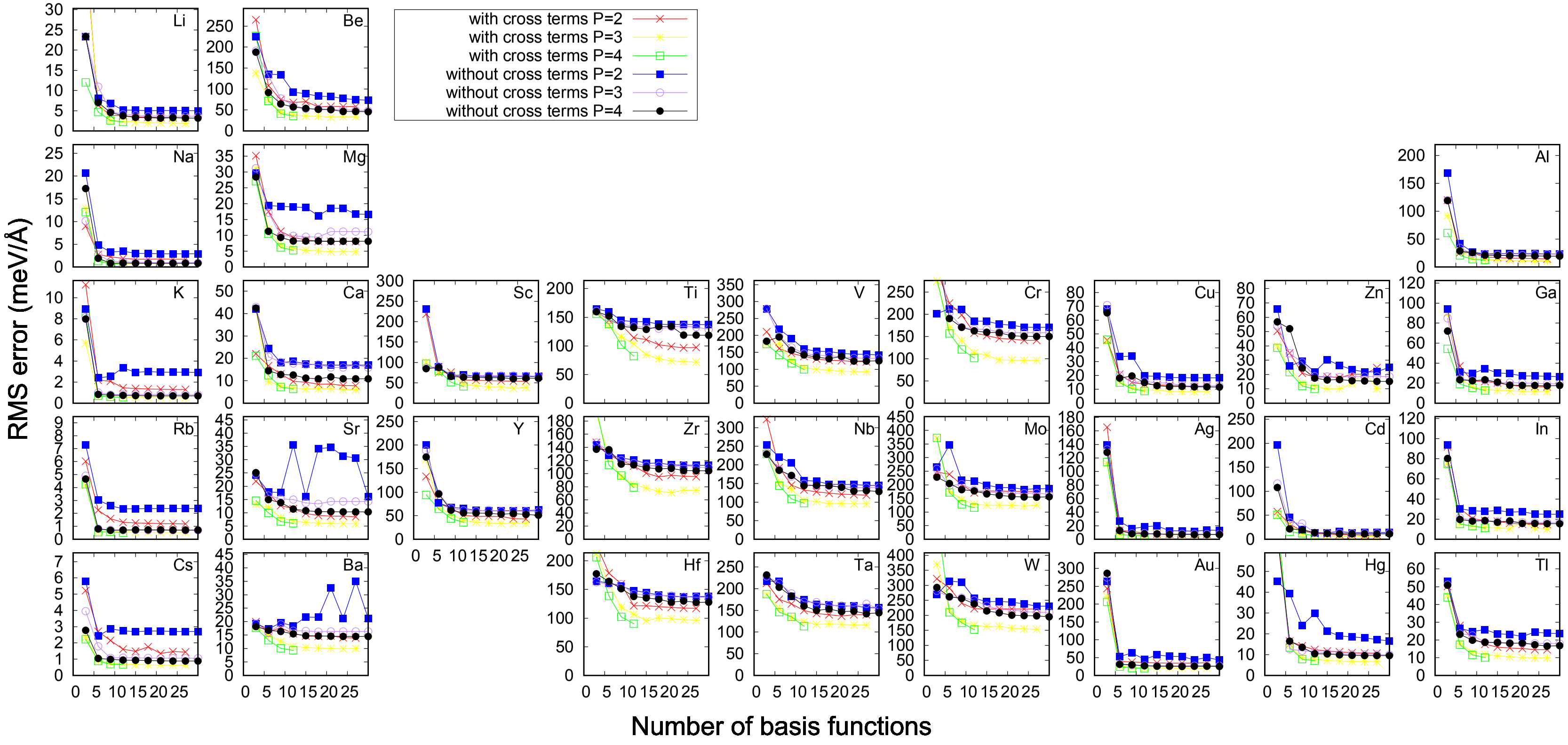


Figure 3.9: The dependency of RMS errors for forces on the expansion order P for both model with cross terms and model without cross terms.

Table 3.2: The optimal basis functions for 31 elements for framework without cross terms.

Element	Number of basis functions	a (Å^{-2})	b_{max} (Å)	RMS error (energy) (meV/atom)	RMS error (force) (eV/Å)
Li	27	0.5	10.0	0.5	0.004
Be	27	2.0	13.0	9.0	0.049
Na	30	0.25	13.0	0.3	0.001
Mg	30	1.0	13.0	1.6	0.011
Al	21	1.0	12.0	3.1	0.021
K	30	0.25	13.0	0.3	0.001
Ca	12	0.25	13.0	2.0	0.018
Sc	24	1.0	13.0	9.4	0.068
Ti	24	1.5	13.0	17.0	0.134
V	27	1.5	6.0	15.5	0.132
Cr	30	1.5	5.0	25.2	0.159
Cu	27	1.0	13.0	1.8	0.013
Zn	18	1.5	10.0	3.4	0.018
Ga	18	0.5	10.0	2.3	0.021
Rb	18	0.25	13.0	0.4	0.001
Sr	18	0.5	13.0	1.9	0.013
Y	27	1.0	13.0	9.2	0.060
Zr	21	0.5	13.0	12.4	0.111
Nb	30	1.5	11.0	16.6	0.140
Mo	27	1.5	11.0	22.1	0.170
Ag	30	1.0	11.0	1.1	0.008
Cd	15	1.0	8.0	2.3	0.013
In	15	0.5	13.0	1.9	0.018
Cs	21	0.1	13.0	0.3	0.001
Ba	30	0.25	13.0	3.4	0.016
Hf	18	1.0	9.0	16.6	0.139
Ta	21	2.0	8.0	20.2	0.161
W	24	1.0	8.0	32.9	0.212
Au	27	1.0	13.0	3.0	0.027
Hg	24	0.5	13.0	1.8	0.010
Tl	15	0.5	12.0	3.3	0.019

Table 3.3: The optimal basis functions for 31 elements for framework with cross terms. Values in brackets are ratio of RMSE of model with cross terms to RMSE of model without cross terms.

Element	Number of basis functions	a ($Å^{-2}$)	$b_{\max}$ (Å)	RMS error (energy) (meV/atom)	RMS error (force) (eV/Å)
Li	24	1.5	11.0	0.2 (0.30)	0.002 (0.53)
Be	21	3.0	8.0	2.6 (0.29)	0.033 (0.67)
Na	27	1.5	10.0	0.1 (0.41)	0.001 (0.70)
Mg	21	3.0	9.0	0.5 (0.30)	0.005 (0.44)
Al	27	2.5	9.0	0.8 (0.27)	0.011 (0.51)
K	24	1.0	12.0	0.1 (0.49)	0.001 (0.67)
Ca	27	2.0	10.0	0.7 (0.36)	0.006 (0.34)
Sc	24	2.0	10.0	2.5 (0.27)	0.036 (0.53)
Ti	24	2.0	9.0	3.8 (0.22)	0.073 (0.54)
V	27	2.0	7.0	7.7 (0.50)	0.093 (0.70)
Cr	24	3.0	6.0	6.4 (0.26)	0.096 (0.61)
Cu	27	3.5	7.0	0.7 (0.38)	0.008 (0.67)
Zn	21	3.0	7.0	1.0 (0.30)	0.014 (0.75)
Ga	27	2.0	8.0	0.8 (0.35)	0.012 (0.57)
Rb	24	1.0	13.0	0.2 (0.50)	0.001 (0.77)
Sr	27	2.0	11.0	0.5 (0.29)	0.006 (0.44)
Y	27	2.5	10.0	2.5 (0.27)	0.033 (0.56)
Zr	21	2.0	8.0	4.7 (0.38)	0.071 (0.64)
Nb	27	2.0	7.0	7.0 (0.42)	0.095 (0.68)
Mo	18	3.0	7.0	7.8 (0.35)	0.126 (0.74)
Ag	24	2.0	11.0	0.6 (0.53)	0.005 (0.62)
Cd	27	1.5	10.0	0.6 (0.28)	0.006 (0.44)
In	27	2.5	10.0	0.8 (0.42)	0.010 (0.58)
Cs	18	0.5	13.0	0.1 (0.48)	0.001 (0.61)
Ba	21	1.0	9.0	0.9 (0.27)	0.010 (0.61)
Hf	27	2.5	6.0	5.7 (0.34)	0.096 (0.69)
Ta	24	3.0	6.0	8.3 (0.41)	0.117 (0.72)
W	27	3.5	7.0	12.4 (0.38)	0.153 (0.72)
Au	18	2.0	6.0	1.6 (0.54)	0.021 (0.77)
Hg	24	1.5	9.0	0.7 (0.37)	0.007 (0.65)
Tl	27	2.0	9.0	0.8 (0.23)	0.010 (0.53)

3.5 Prediction of physical properties

3.5.1 Energies of test set structures

Here, I calculated energies of 300 structures of test set for each element by linear interatomic potential with cross terms and DFT calculations. Figure 3.10 shows the comparison of DFT and linear potential for all 31 elemental metals. As can be seen, for all elements, linear interatomic potentials very accurately reproduce the energies calculated by DFT calculation even energies of 300 structures vary very widely.

Figure 3.11 and 3.12 shows the comparison of DFT, linear potential, and conventional EAMs for Al and Cu, respectively Conventional EAM data set are obtained from potentials interatomic potential repository site [46, 47] and energies are calculated by LAMMPS code [48, 49]. For both of Al and Cu, all interatomic potentials reproduce energies of FCC and HCP structures well. Because parameters of conventional EAMs are constructed from properties of stable structures, even conventional EAMs can be applied to such stable structures. However, DFT energies of other structures such as BCC, SC, ω, and β-tin are reproduced accurately by only linear potentials. This can be ascribe to the fact that linear potential is constructed from such a large training data and the flexibility of linear interatomic potentials allows training data to include a few thousand structures. It is very difficult to construct conventional potentials with many parameters from such training data because usually coefficients are determined by algorithm require many steps like quasi-Newton method and calculation of gradient of error takes many computational costs, whereas linear potential can be constructed from a large training data by only solving normal equation.

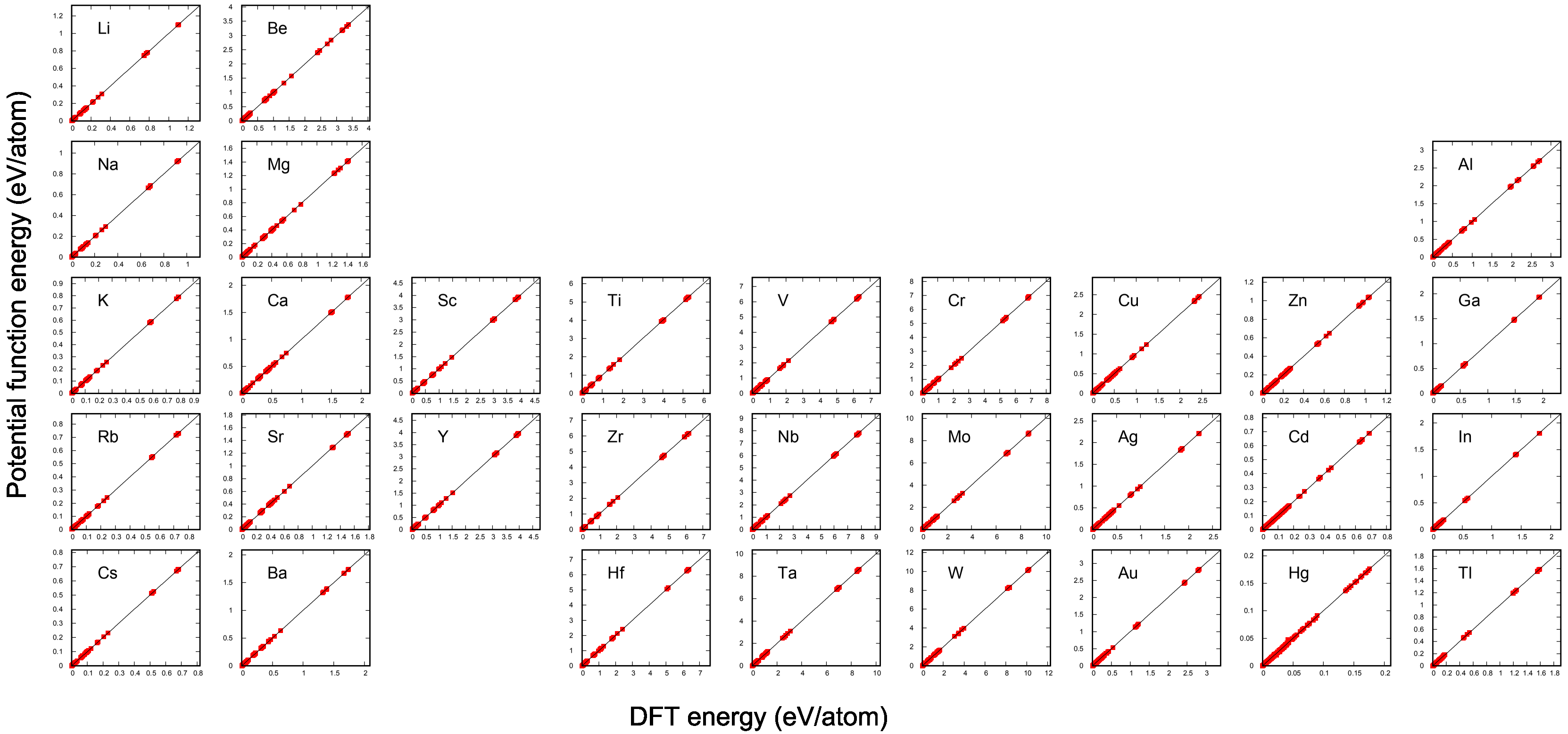


Figure 3.10: Energies of structures of test set predicted by the linear potential and DFT for 31 elemental metals measured from the lowest energy of 300 structures.

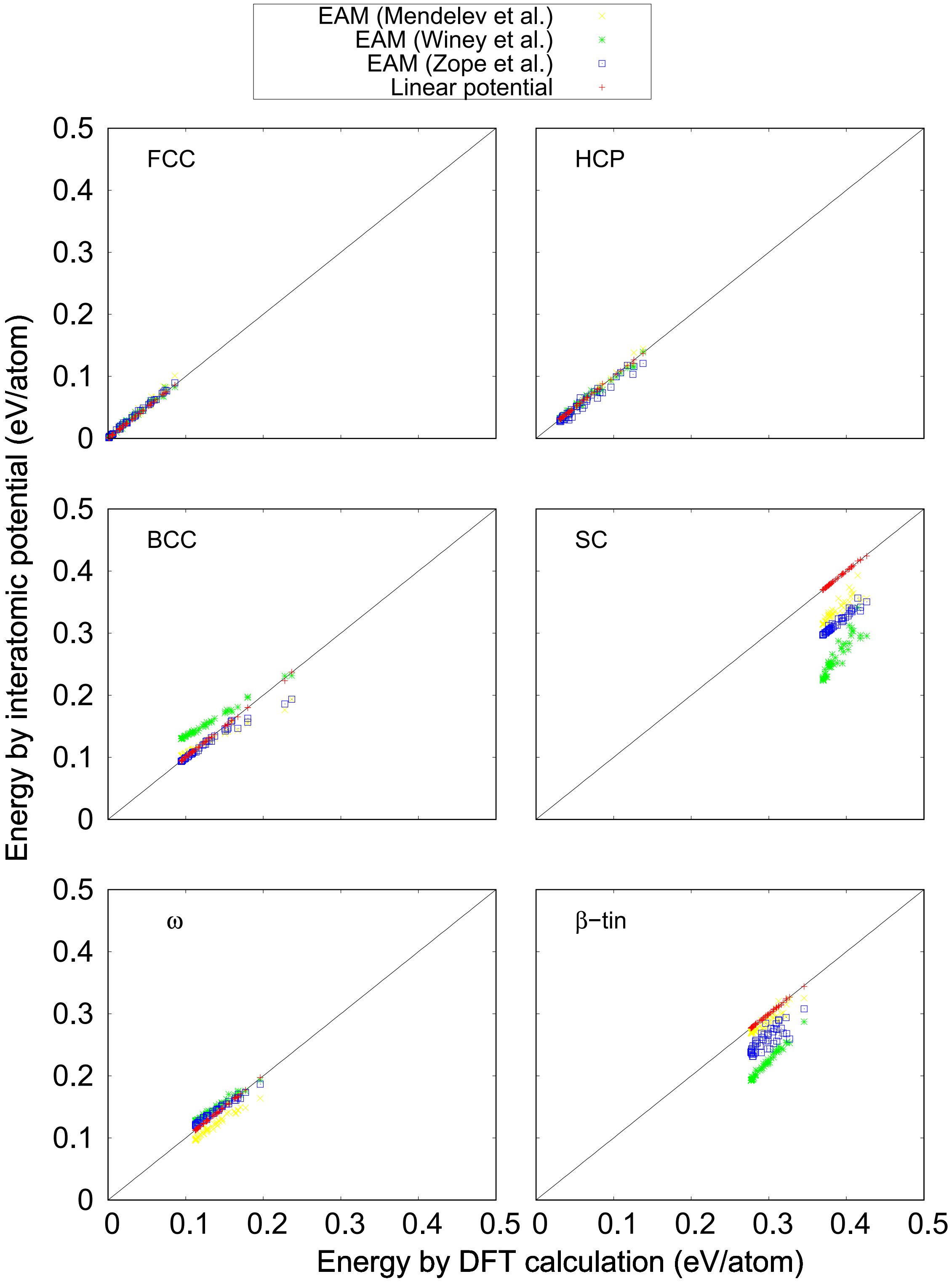


Figure 3.11: Energies of structures of test set predicted by the linear potential and DFT for Al elemental metals measured from the lowest FCC structures. Examined conventional EAMs are Mendelev EAM [50], Winey EAM [51, 52], and Zope EAM [53]. Extremely expanded or compressed structures are excluded.

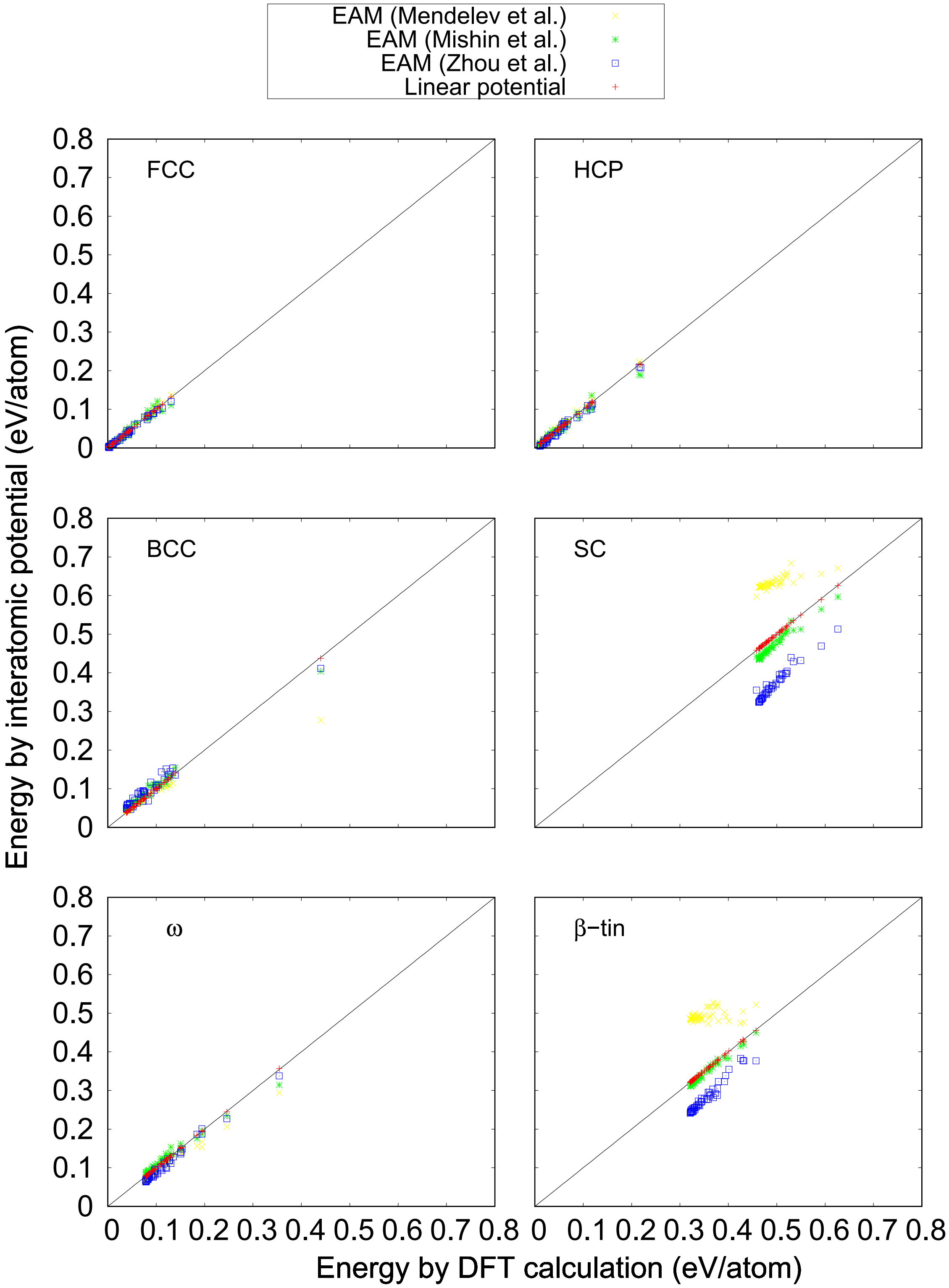


Figure 3.12: Energies of structures of test set predicted by the linear potential and DFT for Cu elemental metals measured from the lowest FCC structures. Examined conventional EAMs are Mendelev EAM [50], Mishin EAM [54], and Zhou EAM [55]. Extremely expanded or compressed structures are excluded.

3.5.2 Elastic constants and bulk modulus

Elastic constants are fundamental properties of bulk material and to evaluate them accurately is important when interatomic potentials are used to predict mechanical properties. In this study, I calculated the elastic constants of cubic crystals and hexagonal crystals with both interatomic potentials and DFT calculations. I estimated the elastic constants of cubic crystals by the same way as the book written by Finnis [11] and those of hexagonal crystals by the same way as literature of Fast *et al.* [56].

According to the literature of Finnis, Elastic constants can be estimated by energies of strained structures. Now I denote a homogeneous strain matrix as

$$\epsilon = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{12} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{13} & \epsilon_{23} & \epsilon_{33} \end{bmatrix}. \tag{3.21}$$

The elements of ϵ is defined as follows. If the strain carries the point $\boldsymbol{r} = (r_1, r_2, r_3)$ to a new position $\boldsymbol{r} + \boldsymbol{u}$, elements of ϵ are defined by

$$\epsilon_{ab} = \frac{1}{2}\left(\frac{\delta u_\alpha}{\delta x_\beta} + \frac{\delta u_\alpha}{\delta x_\beta}\right). \tag{3.22}$$

In practice, such transformation can be achieved by changing basis vectors from $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ to $\{\boldsymbol{a}'_1, \boldsymbol{a}'_2, \boldsymbol{a}'_3\}$ where

$$a'_{i\alpha} = a_{i\alpha} + \sum_\beta \epsilon_{\alpha\beta} a_{i\beta}, \tag{3.23}$$

in which $a_{i\alpha}$ is the Cartesian component along the x_α axis of basis vector $\boldsymbol{a}_i$.

Using Voigt notation, strain matrix can also be expressed as

$$\epsilon = \begin{bmatrix} \epsilon_1 & \epsilon_6 & \epsilon_5 \\ \epsilon_6 & \epsilon_2 & \epsilon_4 \\ \epsilon_5 & \epsilon_4 & \epsilon_3 \end{bmatrix}. \tag{3.24}$$

Then, stress-strain relationship is expressed as

$$\sigma_i = \sum_{j=1}^{6} C_{ij}\epsilon_j, \tag{3.25}$$

where σ_i and C_{ij} are elements of strain 6×6 matrix and elastic constants 6×6 matrix using Voigt notation, respectively. Then, elastic energy per unit volume stored homogeneously strained body is given by

$$E^{\mathrm{elas}} = \frac{1}{2}\sum_{ij} C_{ij}\epsilon_i\epsilon_j. \tag{3.26}$$

Therefore, considering appropriate symmetry transformation matrix T and scholar parameter γ and calculating $E(\gamma, T)$, which is energy when structure is strained by matrix γT, elastic constants matrix can be obtained from second derivative of the elastic energy with respective to strain, namely $\frac{\delta^2 E(\gamma,T)}{\delta\gamma^2}|_{\gamma=0}$.

For cubic system, there are three independent elastic constants C_{11}, C_{12}, and C_{44}. Elastic constant matrix can be expressed as

$$C = \begin{bmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix}. \tag{3.27}$$

Elastic energy per unit volume is

$$E^{\mathrm{elas}} = \frac{1}{2}C_{11}\left(\epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2\right) + C_{12}\left(\epsilon_1\epsilon_2 + \epsilon_2\epsilon_3 + \epsilon_3\epsilon_1\right) + C_{44}\left(\epsilon_4^2 + \epsilon_5^2 + \epsilon_6^2\right). \tag{3.28}$$

Elastic constants can be calculated by consideration of three matrices expressed as

$$T_e = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.29}$$

$$T_r = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix} \tag{3.30}$$

$$T_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.5 & 0 \\ 0 & 0 & -0.5 \end{bmatrix}. \tag{3.31}$$

From Eqn.(3.28), the relationships between elastic constants and second derivatives of energy respective to strain are

$$\left.\frac{\delta^2 E(\gamma, T_e)}{\delta\gamma^2}\right|_{\gamma=0} = 3V_0(C_{11} + 2C_{12}), \tag{3.32}$$

$$\left.\frac{\delta^2 E(\gamma, T_r)}{\delta\gamma^2}\right|_{\gamma=0} = \frac{3}{2}V_0(C_{11} - C_{12}), \tag{3.33}$$

$$\left.\frac{\delta^2 E(\gamma, T_t)}{\delta\gamma^2}\right|_{\gamma=0} = 3V_0 C_{44}, \tag{3.34}$$

where V_0 is volume per atoms of structure at $\gamma = 0$. Elastic constants are obtained by solving these equations.

For hexagonal system, there are five independent elastic constants C_{11}, C_{12}, C_{33}, C_{55}, and C_{13}. Elastic constants can be calculated by consideration of five matrices expressed as

$$T_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.35}$$

$$T_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.36}$$

$$T_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{3.37}$$

$$T_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.38}$$

$$T_5 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \tag{3.39}$$

The relationships between elastic constants and second derivatives of energy respective to strain are

$$\left.\frac{\delta^2 E(\gamma, T_1)}{\delta\gamma^2}\right|_{\gamma=0} = \frac{2}{9}V_0(2C_{11} + 2C_{12} + 4C_{13} + C_{33}), \tag{3.40}$$

$$\left.\frac{\delta^2 E(\gamma, T_2)}{\delta\gamma^2}\right|_{\gamma=0} = 2V_0(C_{11} + C_{12}), \tag{3.41}$$

$$\left.\frac{\delta^2 E(\gamma, T_3)}{\delta\gamma^2}\right|_{\gamma=0} = 2V_0(C_{11} - C_{12}), \tag{3.42}$$

$$\left.\frac{\delta^2 E(\gamma, T_4)}{\delta\gamma^2}\right|_{\gamma=0} = V_0 C_{33}, \tag{3.43}$$

$$\left.\frac{\delta^2 E(\gamma, T_5)}{\delta\gamma^2}\right|_{\gamma=0} = 4V_0 C_{55}. \tag{3.44}$$

Elastic constants are obtained by solving five equations.

In this study, to calculate second derivative, I calculated energies in the case γ ranging from -0.01 to 0.01 in steps of 0.001 for those strain matrices, fitted those energies by quartic polynomials, and obtained second derivative from coefficient of second-order term by DFT calculations, linear potentials without cross terms, linear potentials with cross terms, and conventional potentials. For all methods to calculate energies, lattice constants when $\gamma = 0$ are same as those obtained from DFT structure optimization.

Figure 3.13, 3.14, 3.15, 3.16, and 3.17 shows the second derivative of the energies for the strain matrices for 31 elements of FCC, BCC, SC, HCP, and ω structures, respectively. For alkali metals and noble metals, both of linear potentials with cross terms and linear potentials without cross terms reproduce elastic constants of DFT calculations well. For elements at the upper right of periodical table, such as Be, Mg, Al, Ga, In, and Tl, the prediction of potentials with cross terms is better than potential without cross terms. Because cross terms of linear potentials can be regarded as a parts of many-body interaction terms, it can be said that for those metals the detail description of many body interaction is more important than alkali and noble metals. However, even linear potentials with cross terms predict elastic second derivatives of those metals worse than those metals. It may be ascribed to the fact that electronic structures of those metals differs from free electron model and radial information may be not sufficient to describe PES of such elements. For transition metals of group 3, 4, 5, and 6, namely Sc, Y, Ti, Zr, Hf, V, Nb, Ta, Cr, Mo, and W, prediction of linear potential with cross terms is also improved, but there are relatively large discrepancies for some structures. Relatively large errors of some elements also may be ascribed to deficiency of angular information. Because transition metals have some d electrons, angular information is needed to construct interatomic potentials for those elements.

Figure 3.18, 3.19, 3.20 show the elastic constants and bulk modulus $B = \frac{1}{3}(C_{11} + 2C_{12})$ for 31 elements of FCC, BCC, and SC structures, respectively. Figure 3.21 and 3.22 also show the elastic constants and bulk modulus $B = \frac{2}{9}(2C_{11}+2C_{12}+4C_{13}+C_{33})$ for 31 elements of HCP and ω structures, respectively. For cubic structures of some elements, such as Mg-FCC structure, there are discrepancies between elastic constants predicted by DFT and linear potentials even there are a little difference between second derivatives predicted by DFT and linear potentials. This fact can be ascribed to linear regression optimize the error of all structures by same criterion. As can be seen fig. 3.13, 3.14, and 3.15, absolute value of errors of T_r and T_t are the same degree to errors of T_e. However, because degree of second derivatives of T_r and T_t is smaller than T_e, ratio of errors are large. When elastic constants are calculated by solving Eqn.(3.32), Eqn.(3.33), and Eqn.(3.34), such errors are enhanced due to multiplication of large coefficients. Therefore, to predict elastic constants well, it is necessary for potential to reproduce anisotropic transformation which shows little increase energies of DFT very accurately. To predict such elastic constants, maybe training data need more structures which are strained anisotropically. On the contrary, bulk modulus, which is equivalent to second derivative of energies respective to isotropic transformation of Eqn.(3.29), can be reproduced better than other elastic constants. Probably it is because such isotropic transformation largely increases energy and because training data include isotropically deformed structures. For hexagonal structures, accuracy of prediction of elastic constants has same tendency as that of second derivatives.

Figure 3.23 and 3.24 show the second derivatives and elastic constants predicted by DFT, linear potentials, and conventional EAMs for Al of cubic crystals and hexagonal crystals, respectively. Figure 3.25 and 3.26 show the second derivatives and elastic constants predicted by DFT, linear potentials, and conventional EAMs Cu of cubic crystals and hexagonal crystals, respectively. For both metals, linear potentials can predict elastic constants as well as conventional EAMs, whereas the training set of linear potential include very unstable structures.

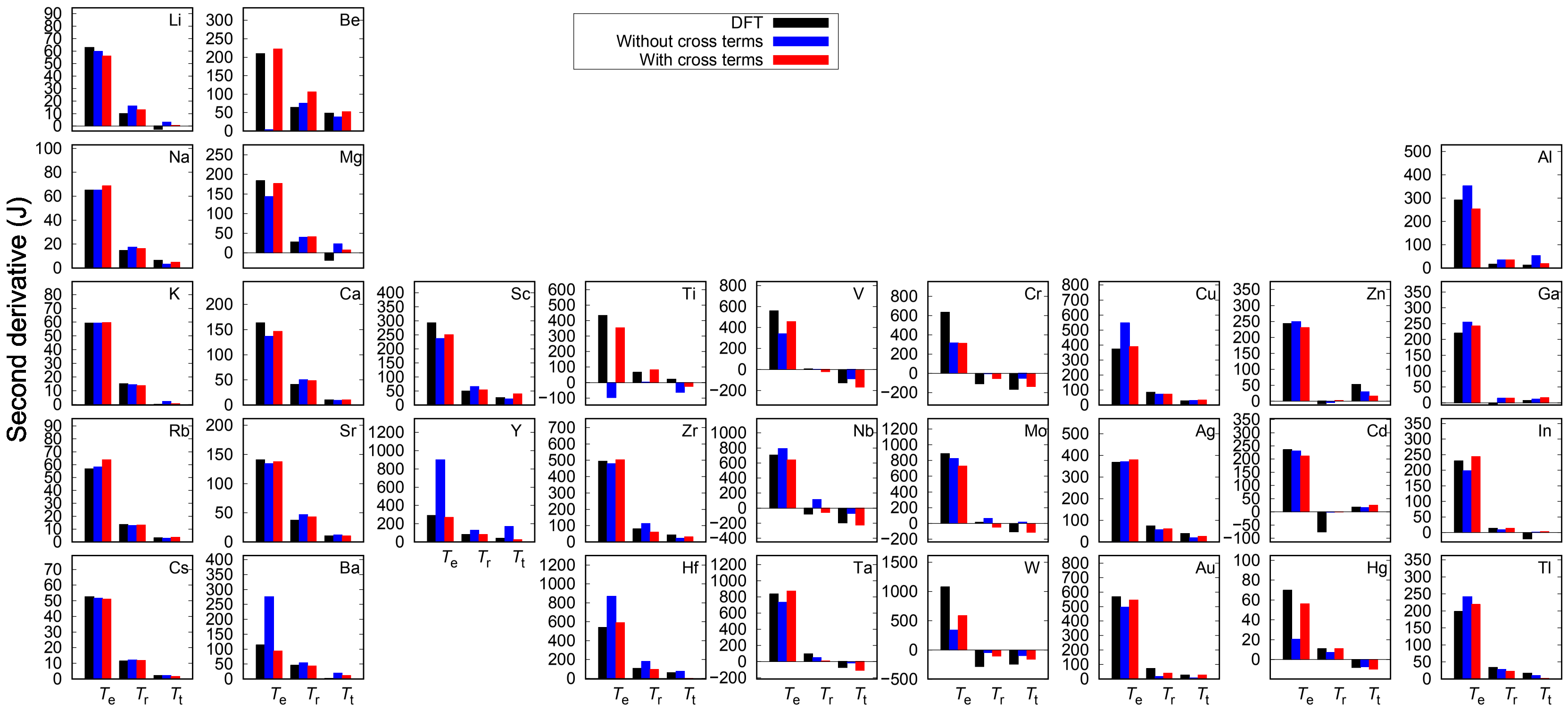


Figure 3.13: Second derivatives of energies respective to matrices expressed as Eqn.(3.29), (3.30), and (3.31) of FCC structures obtained from DFT calculations, linear potentials without cross terms, and linear potentials with cross terms.

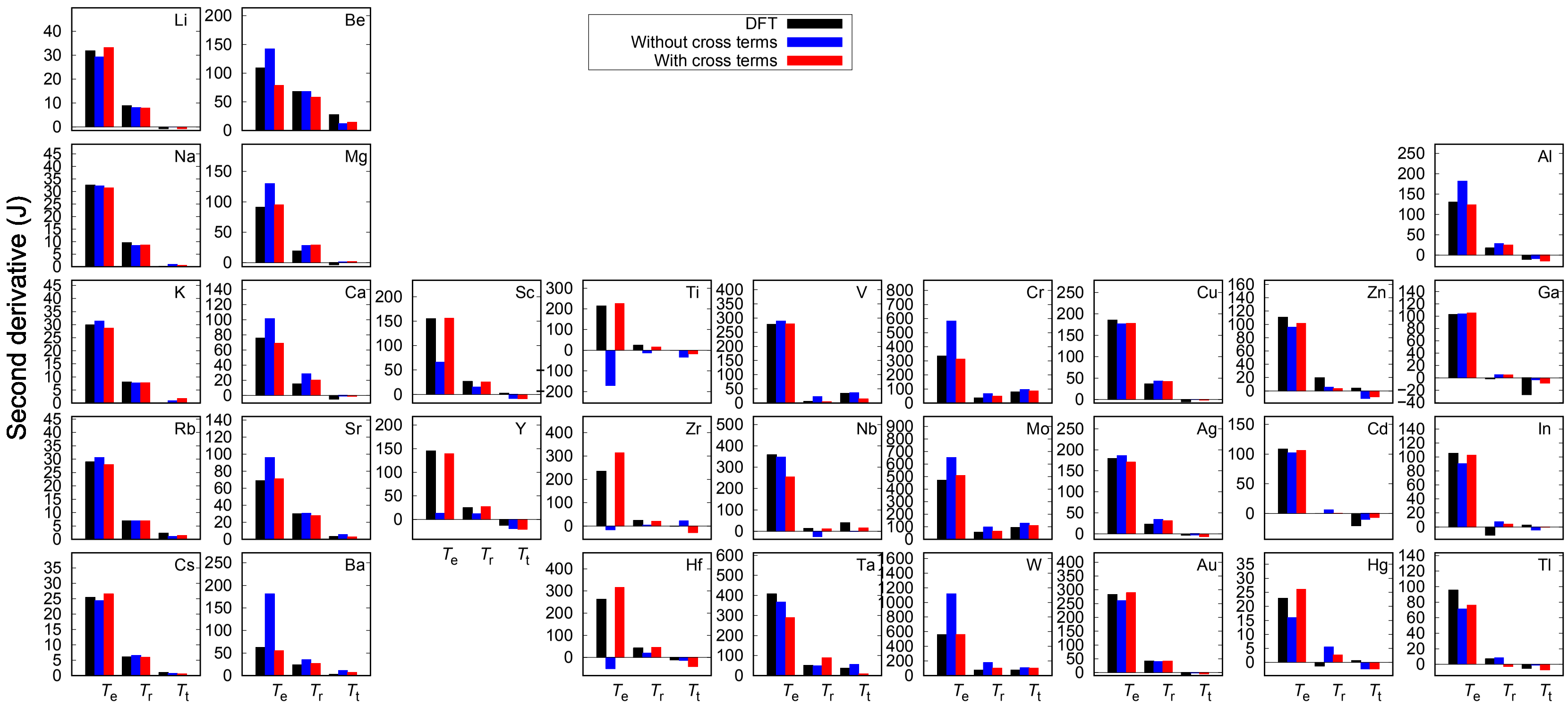


Figure 3.14: Second derivatives of energies respective to matrices expressed as Eqn.(3.29), (3.30), and (3.31) of BCC structures obtained from DFT calculations, linear potentials without cross terms, and linear potentials with cross terms.

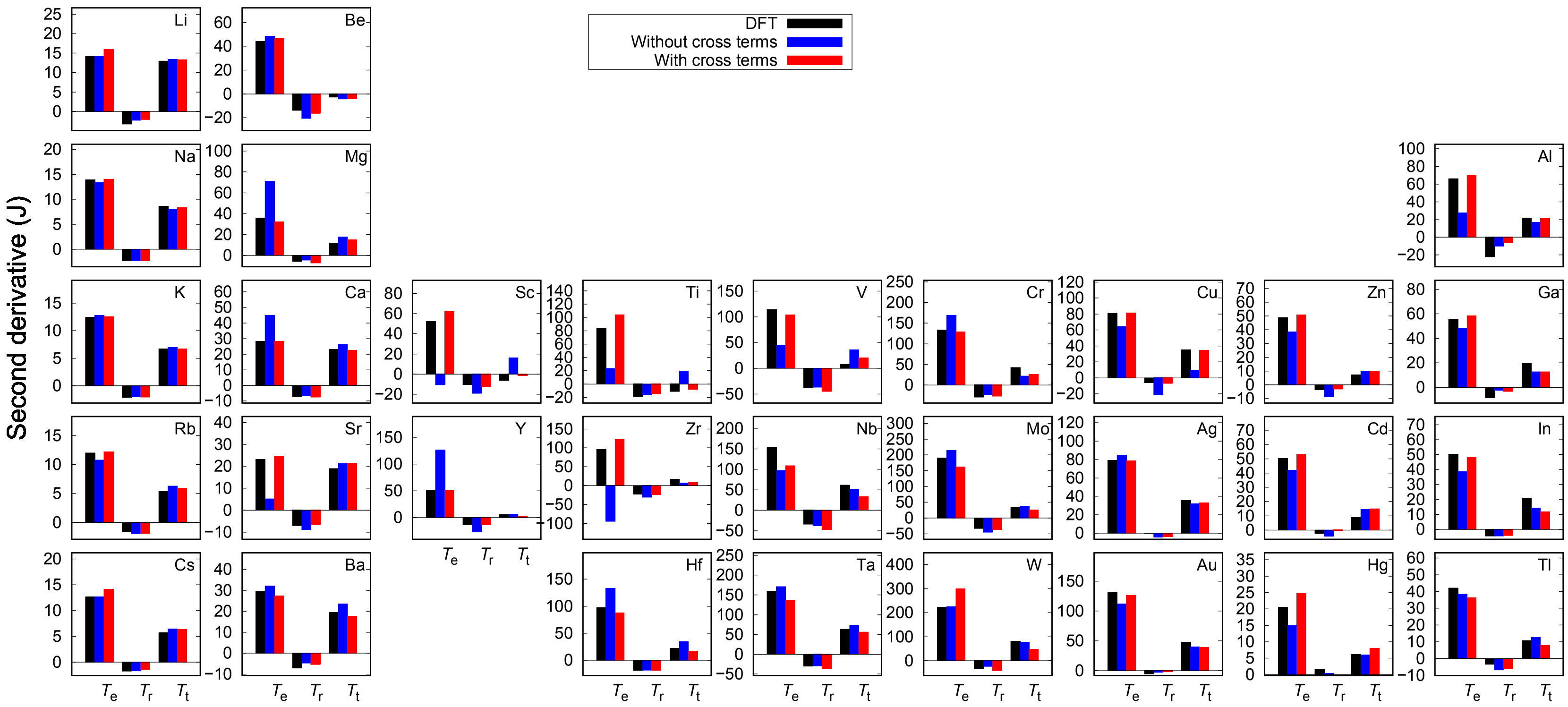


Figure 3.15: Second derivatives of energies respective to matrices expressed as Eqn.(3.29), (3.30), and (3.31) of simple cubic structures obtained from DFT calculations, linear potentials without cross terms, and linear potentials with cross terms.

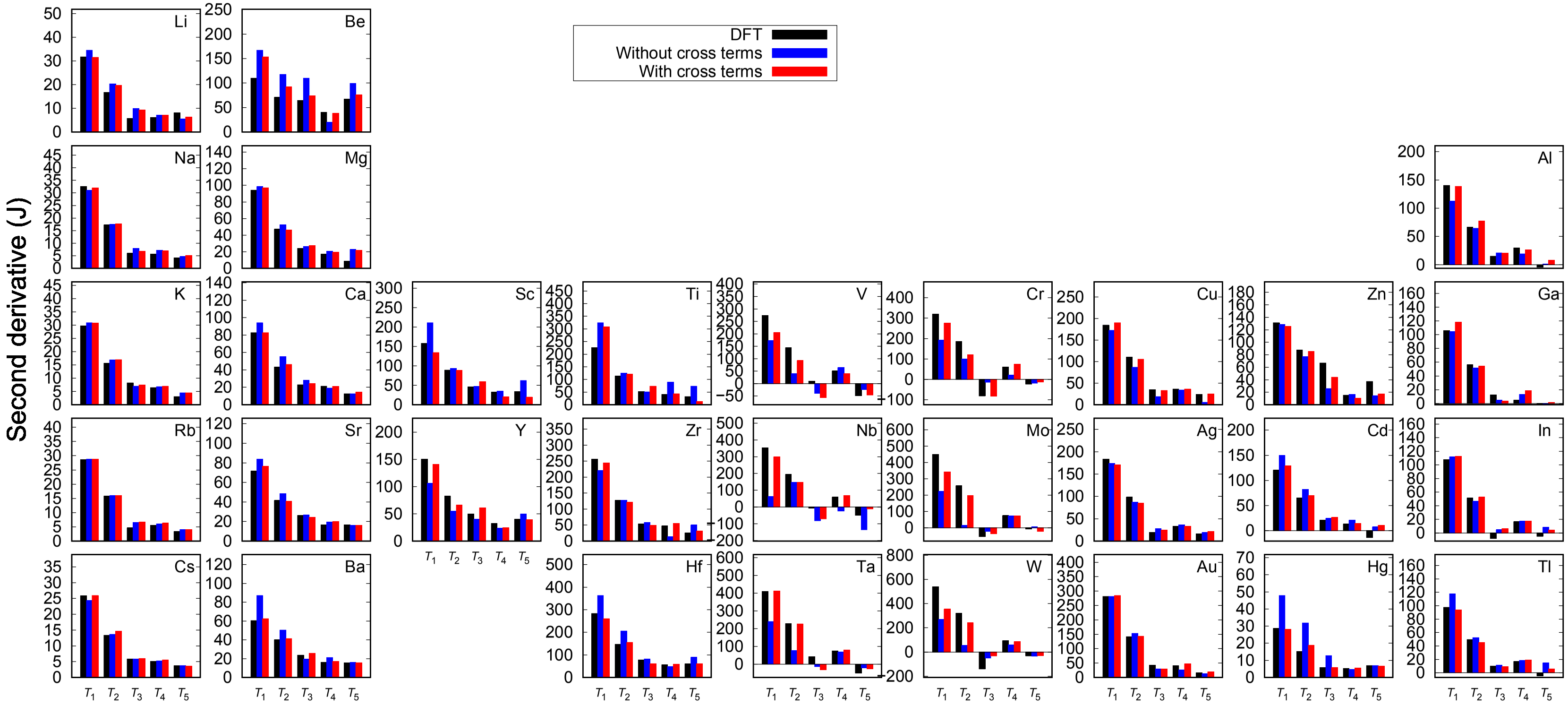


Figure 3.16: Second derivatives of energies respective to matrices expressed as Eqn.(3.35), (3.36), (3.37), (3.38), and (3.39) of HCP structures obtained from DFT calculations, linear potentials without cross terms, and linear potentials with cross terms.

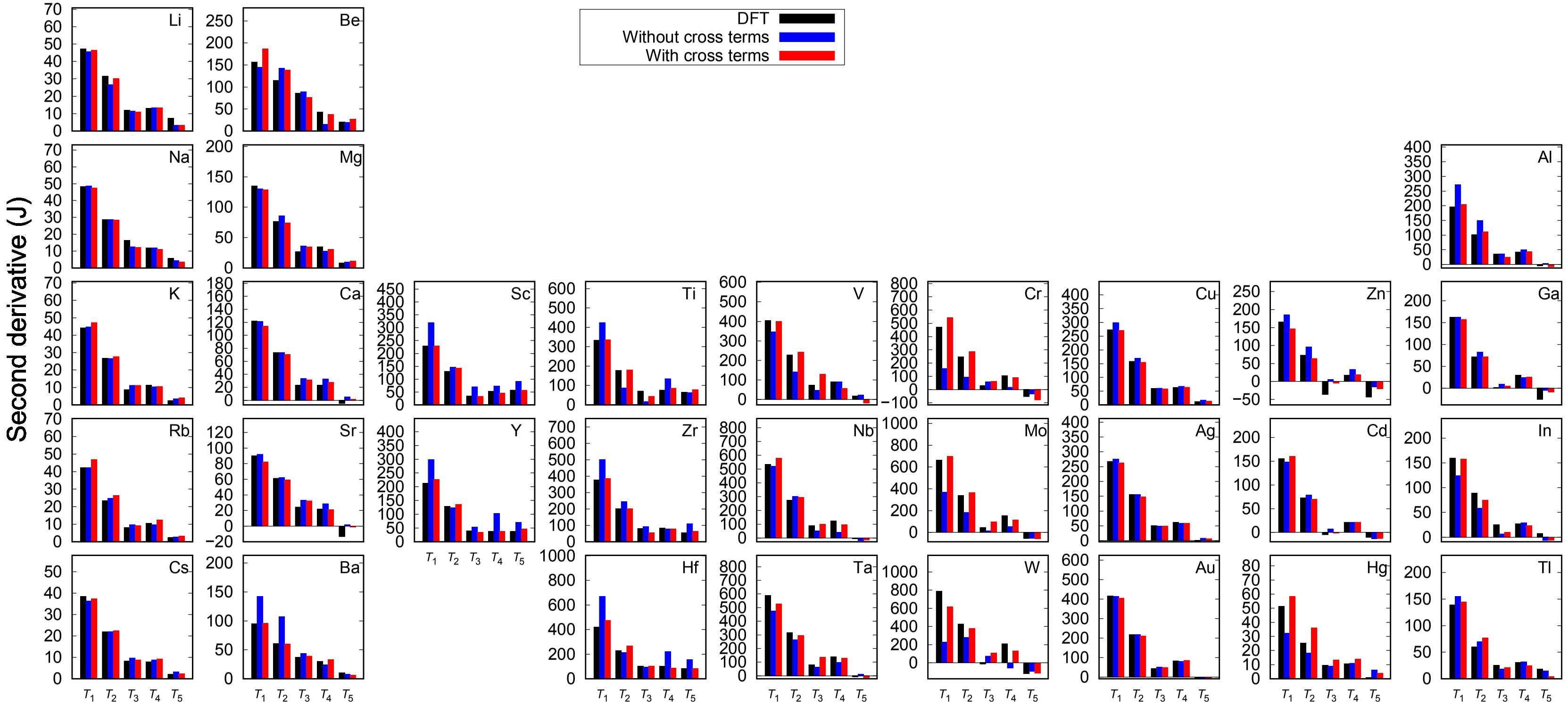


Figure 3.17: Second derivatives of energies respective to matrices expressed as Eqn.(3.35), (3.36), (3.37), (3.38), and (3.39) of ω structures obtained from DFT calculations, linear potentials without cross terms, and linear potentials with cross terms.

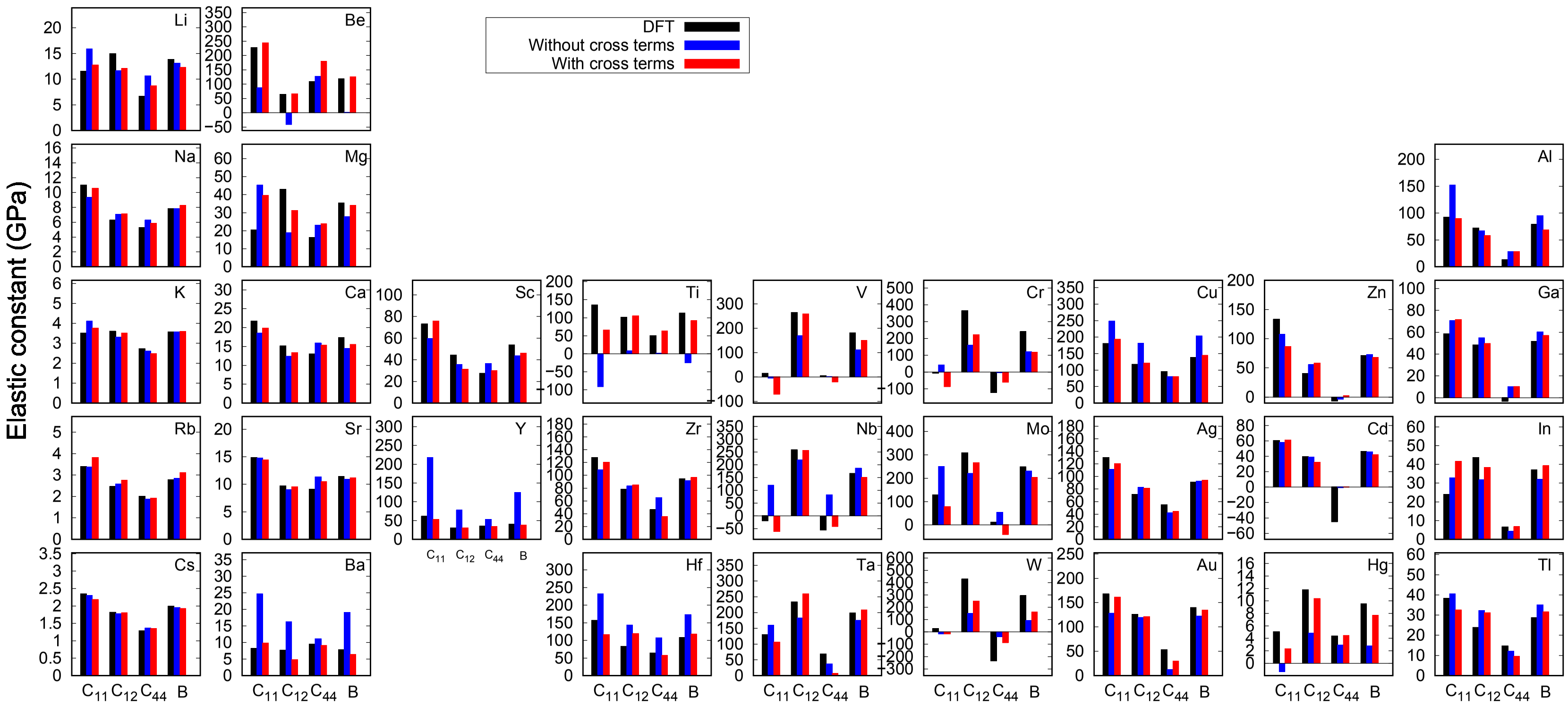


Figure 3.18: Elastic constants of FCC structures obtained from DFT calculations, linear potentials without cross terms, and linear potentials with cross terms.

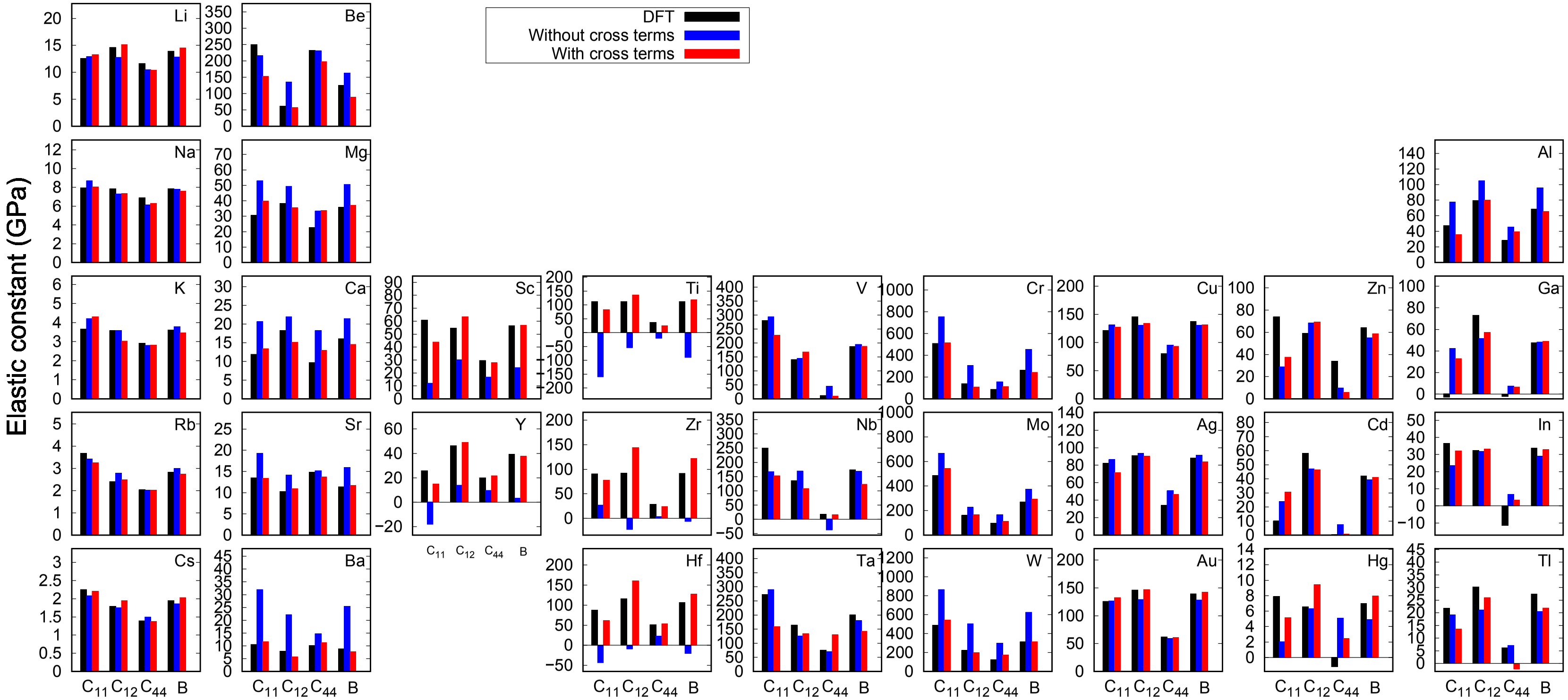


Figure 3.19: Elastic constants of BCC structures obtained from DFT calculations, linear potentials without cross terms, and linear potentials with cross terms.

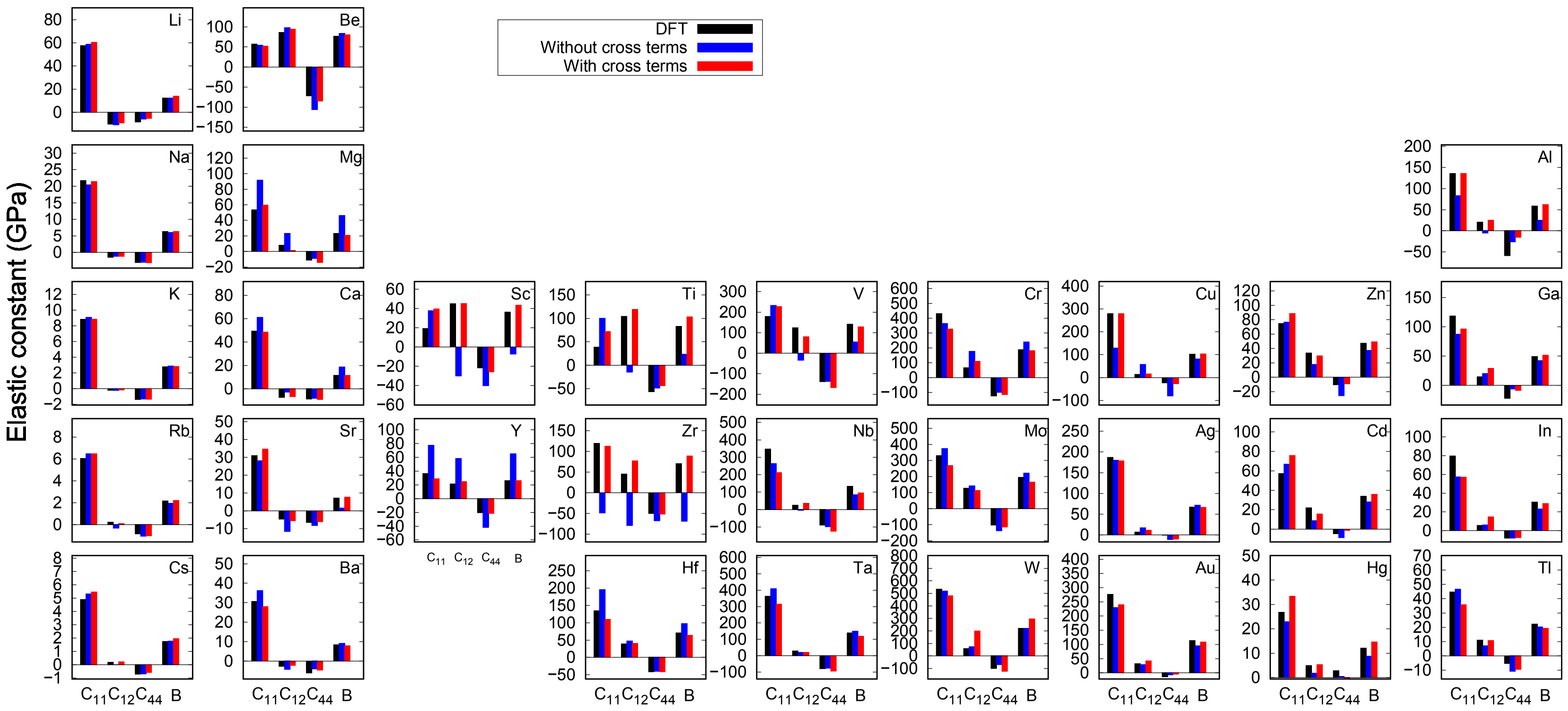


Figure 3.20: Elastic constants of simple cubic structures obtained from DFT calculations, linear potentials without cross terms, and linear potentials with cross terms.

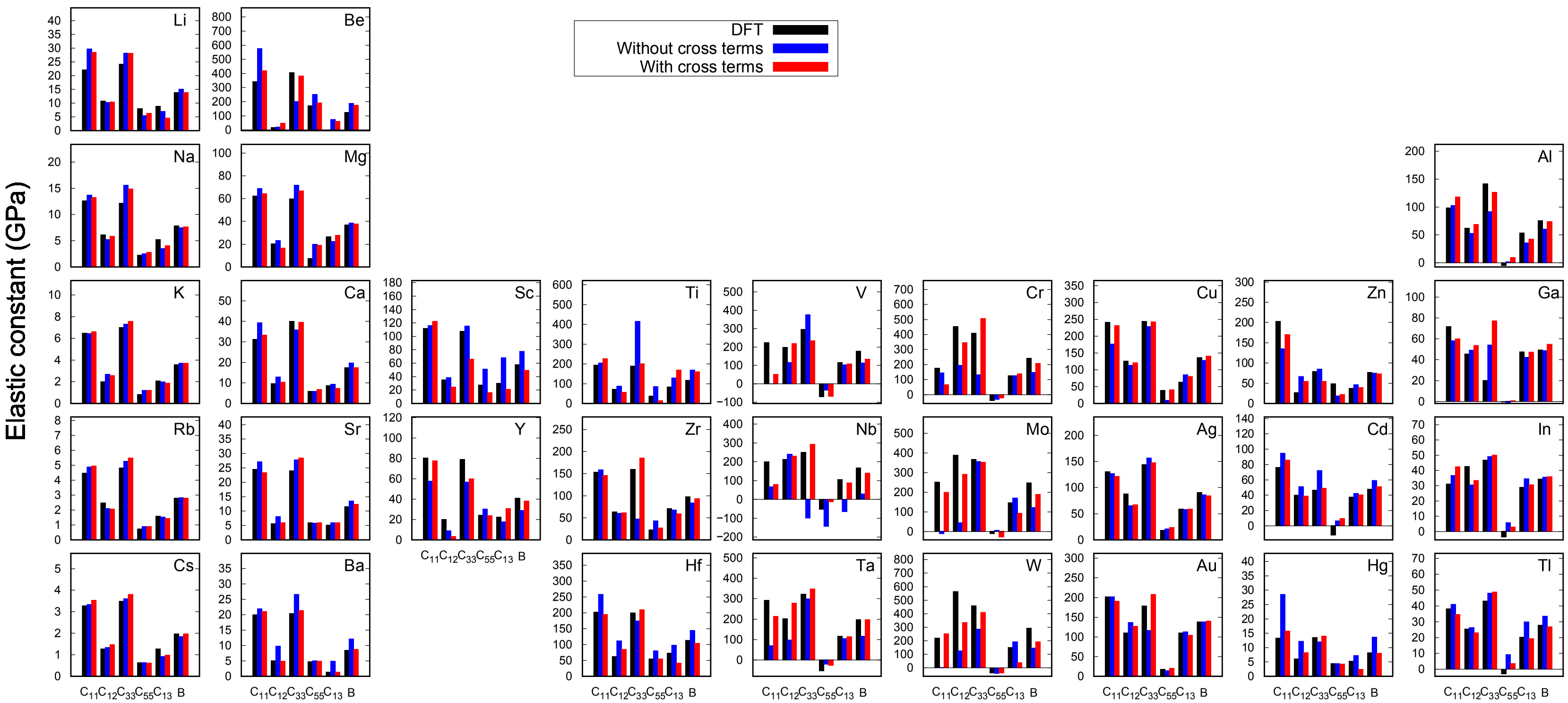


Figure 3.21: Elastic constants of HCP structures obtained from DFT calculations, linear potentials without cross terms, and linear potentials with cross terms.

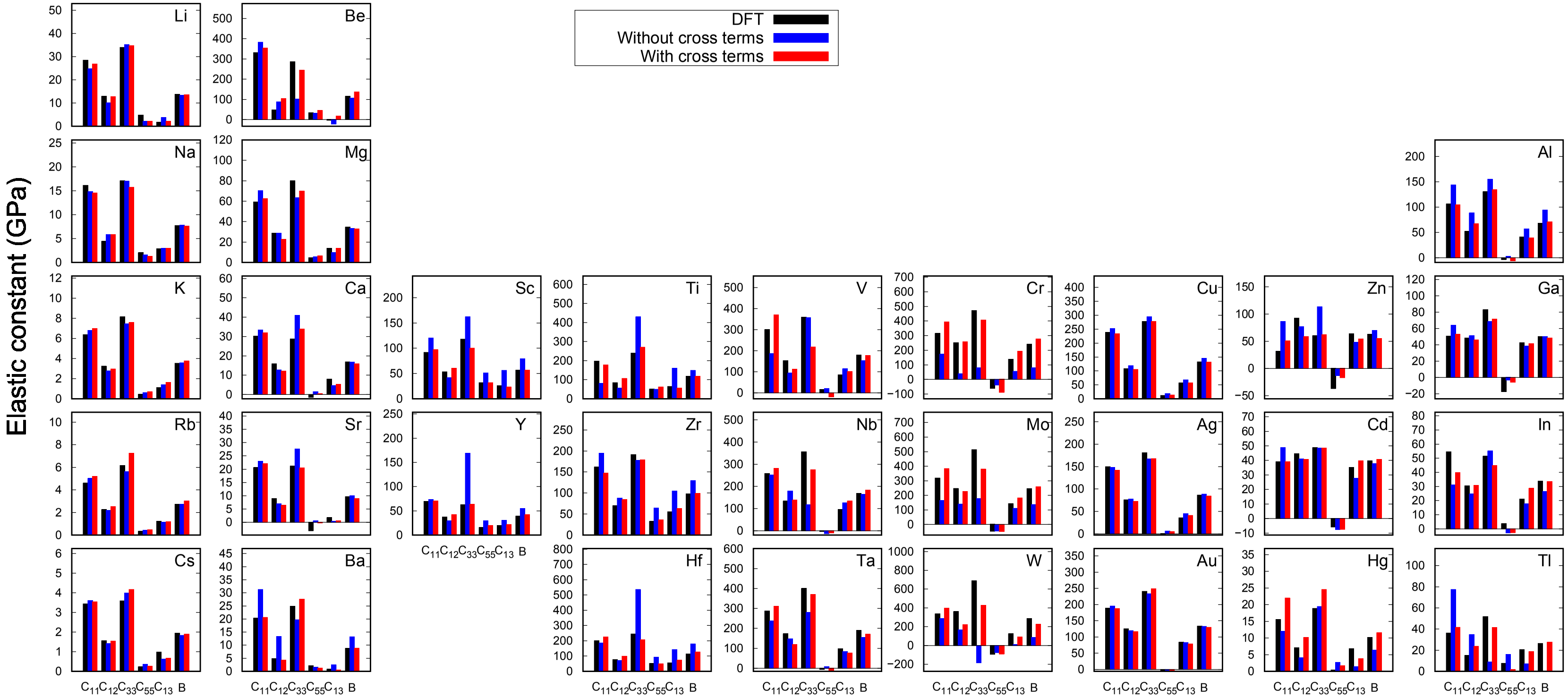


Figure 3.22: Elastic constants of ω structures obtained from DFT calculations, linear potentials without cross terms, and linear potentials with cross terms.

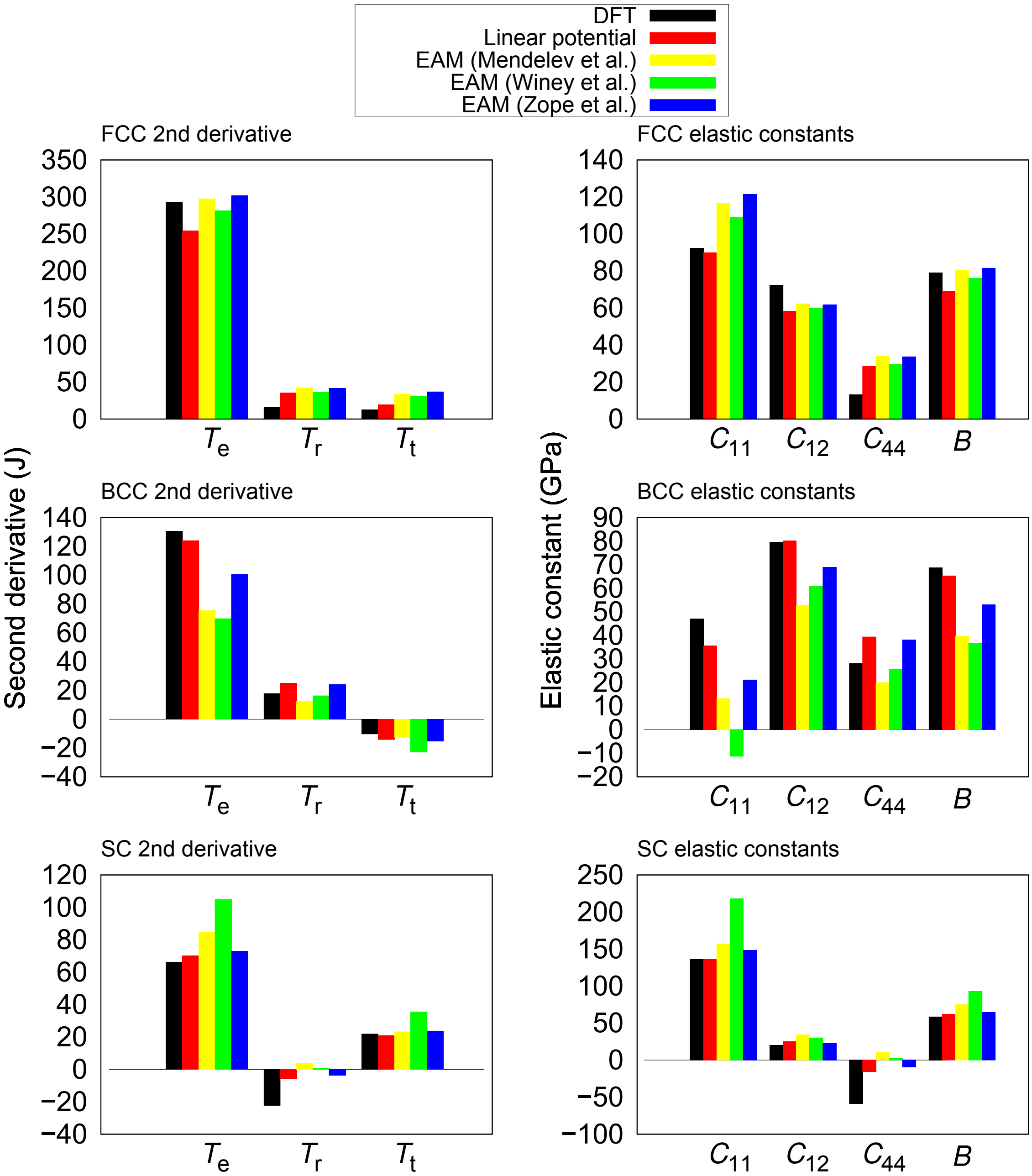


Figure 3.23: Second derivatives of energies respective to matrices expressed as Eqn.(3.29), (3.30), and (3.31) and elastic constants for Al BCC, FCC, and SC crystals obtained from DFT calculations, linear potentials, Mendelev EAM [50], Winey EAM [51, 52], and Zope EAM [53].

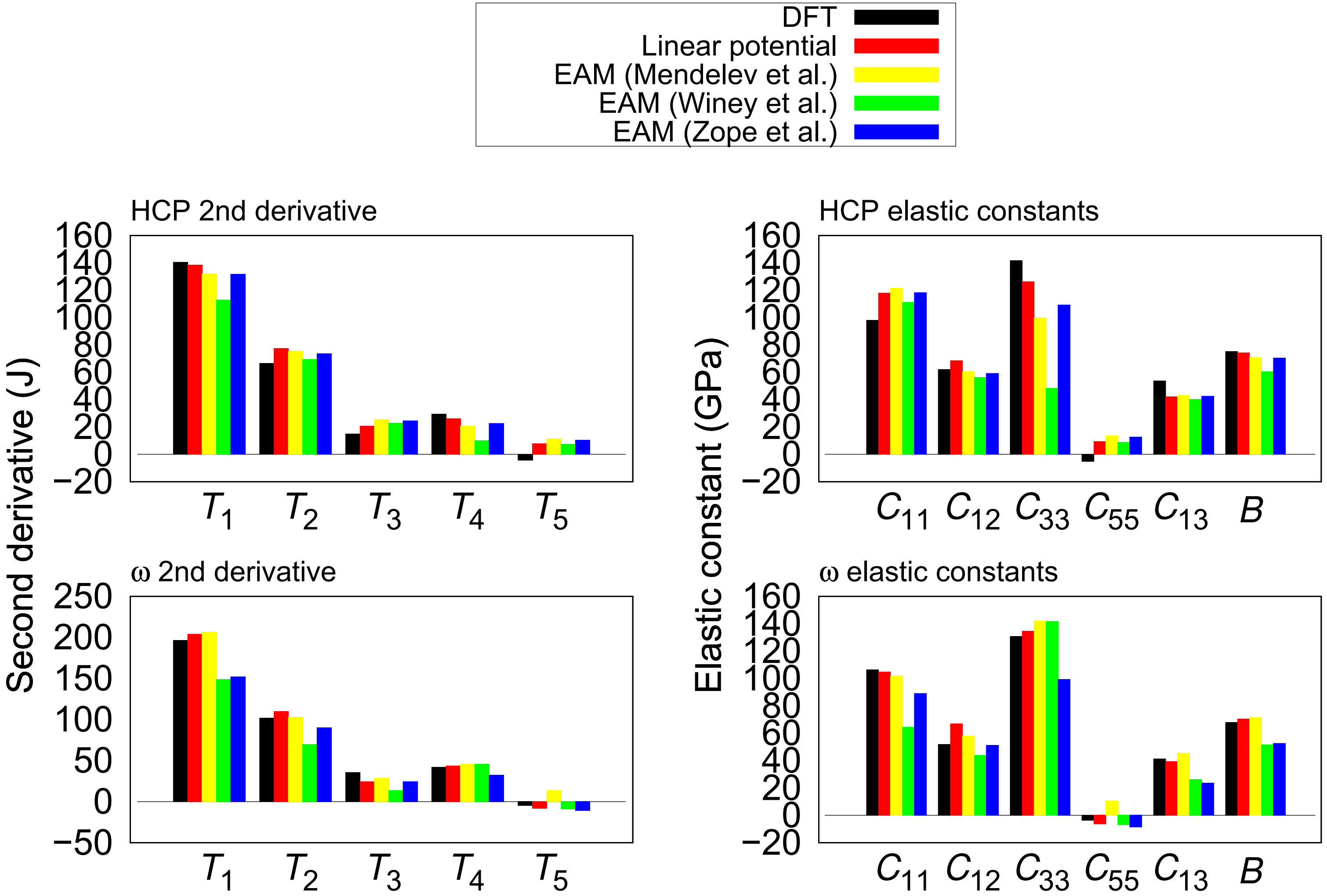


Figure 3.24: Second derivatives of energies respective to matrices expressed as Eqn.(3.35), (3.36), (3.37), (3.38) and (3.39), and elastic constants for Al HCP and ω crystals obtained from DFT calculations, linear potentials, Mendelev EAM [50], Winey EAM [51, 52], and Zope EAM [53].

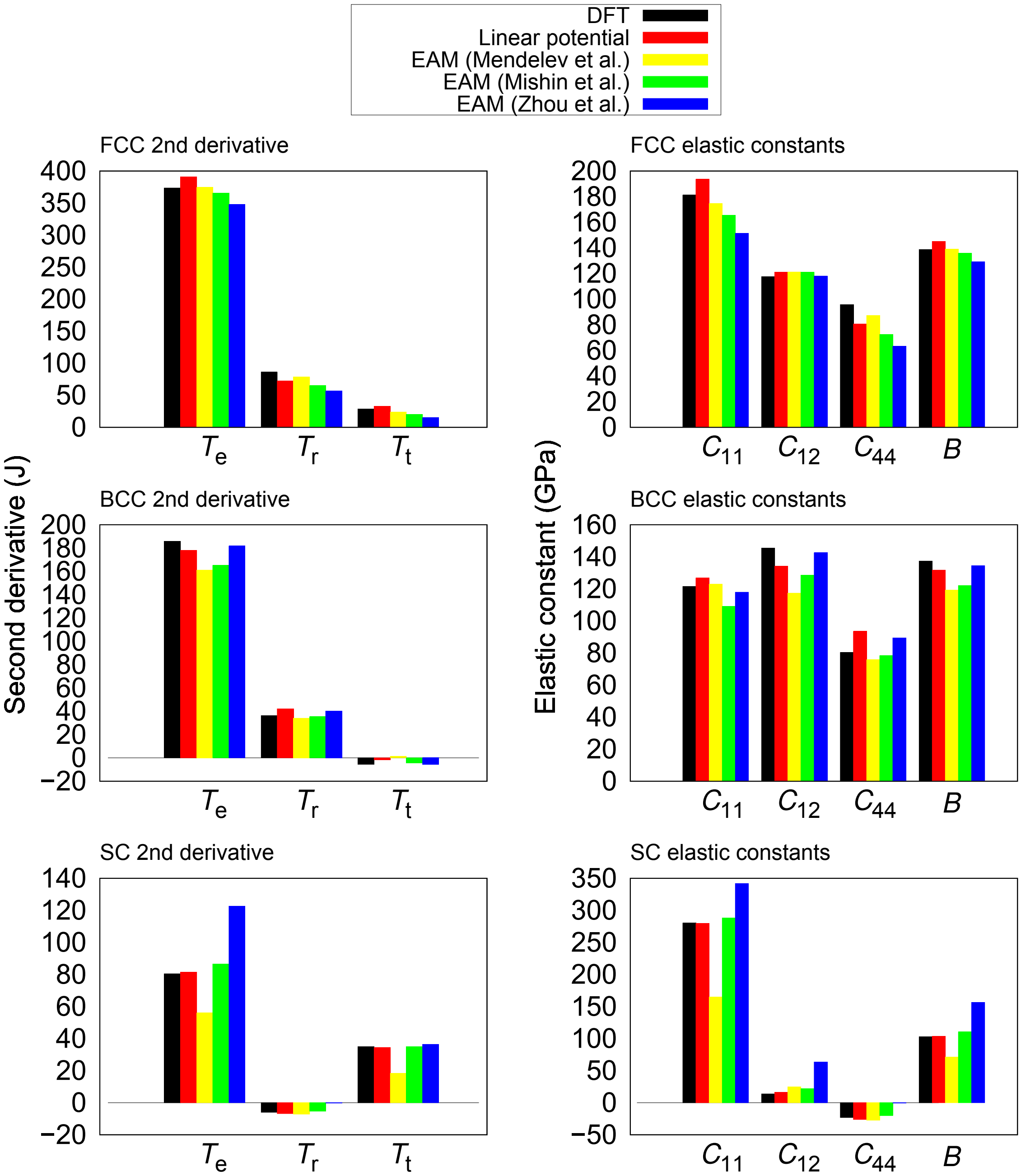


Figure 3.25: Second derivatives of energies respective to matrices expressed as Eqn.(3.29), (3.30), and (3.31) and elastic constants for Cu BCC, FCC, and SC crystals obtained from DFT calculations, linear potentials, Mendelev EAM [50], Mishin EAM [54], and Zhou EAM [55].

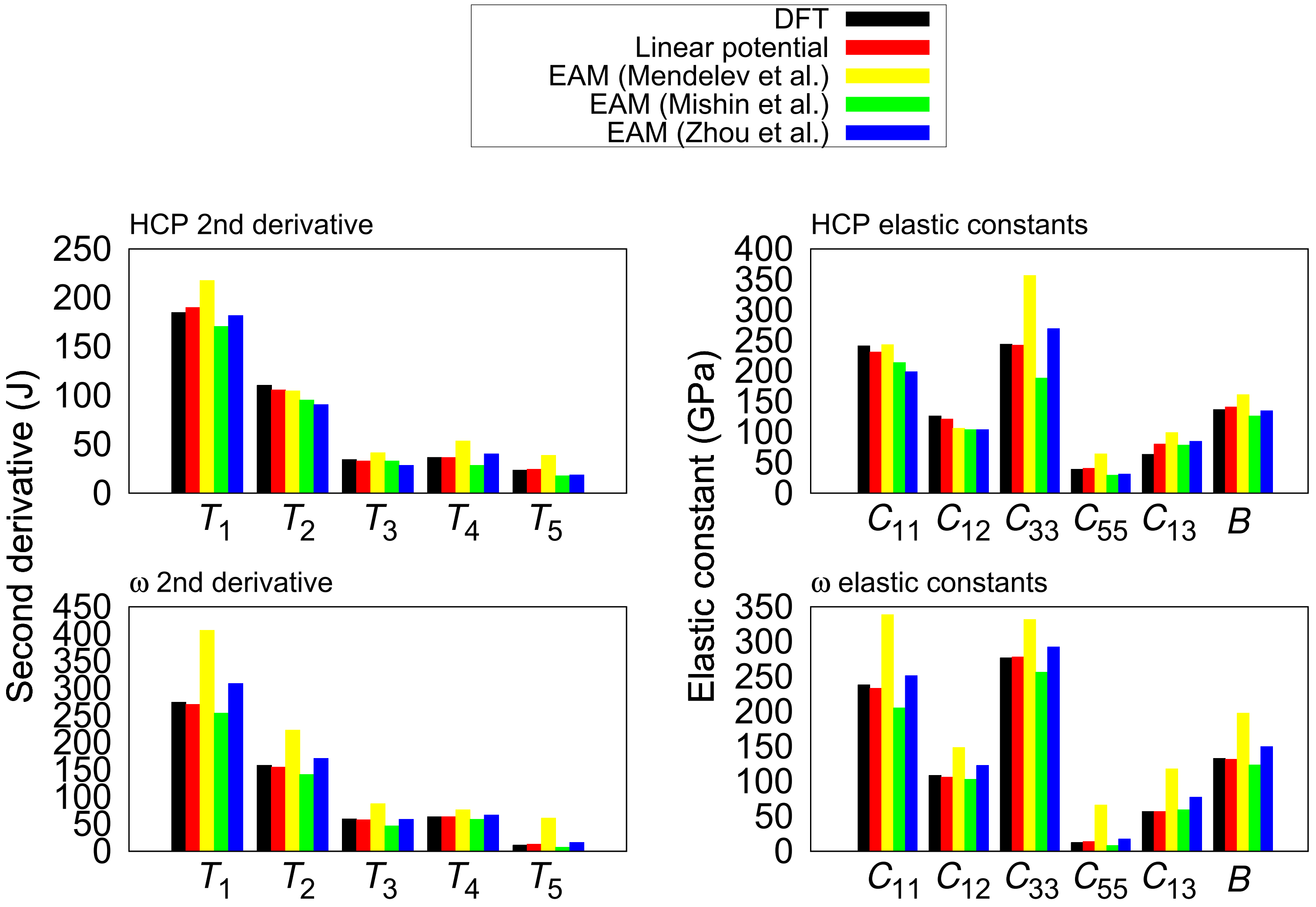


Figure 3.26: Second derivatives of energies respective to matrices expressed as Eqn.(3.35), (3.36), Eqn.(3.37), (3.38), and (3.39) and elastic constants for Cu HCP and ω crystals obtained from DFT calculations, linear potentials, Mendelev EAM [50], Mishin EAM [54], and Zhou EAM [55].

3.5.3 Phonon dispersion curves

For MD simulation, prediction of acting force on atoms and estimation of stability of structures play very important roles. Then, I examined the applicability of the interatomic potentials by comparing phonon dispersion curves computed with DFT calculation and interatomic potentials. The phonon dispersion curves were calculated with the supercell approach [57] for FCC, HCP, BCC, simple cubic, ω, and β-Sn structures with the equilibrium lattice constant. For all methods to calculate energies, lattice constants when $\gamma = 0$ are same as those obtained from DFT structure optimization. To evaluate the dynamical matrix, each symmetrically independent atomic position was displaced by 0.01 Å. The forces acting on atoms by the interatomic potential can then be analytically computed. Supercells were made by $4 \times 4 \times 4$ expansion of the conventional unit cells for all 6 structures. The phonon calculations were performed using the PHONOPY code [58].

Figure 3.27, 3.28, 3.29, 3.30, 3.31, and 3.32 show the dispersion curves with FCC, HCP, BCC, simple cubic, ω, and β-tin structures.

As can be seen, for the most metals linear potentials with cross terms predicted phonon dispersion curves better than linear potentials without cross terms. The degree of improvement shows the same tendency as the prediction of elastic constants. For alkali metals except Cs, namely Li, Na, K, and Rb, phonon dispersion curves with 6 structures predicted by interatomic potentials are very accurately reproduced those by DFT calculations. For Cs, phonon dispersion curve with β-tin structure is not very accurately reproduced. It could not reproduce imaginary modes of dispersion curves with HCP and BCC structures. For alkali earth metals like Be, Mg, Ca, Sr, and Ba, phonon dispersion curves are reproduced less accurately than alkali metals, however, still phonon dispersion curves by interatomic potentials are in good agreements with those predicted by DFT. For Be, phonon dispersion curves with FCC and ω structures are less accurately predicted than the other alkali earth metals. For transition metals of group 3, 4, 5, and 6, namely Sc, Y, Ti, Zr, Hf, V, Nb, Ta, Cr, Mo, and W, error between phonon dispersion curves predicted by interatomic potentials and DFT calculations are relatively large. Even noble metals like Cu, Ag, and Au are also transition metals, phonon dispersion curves by interatomic potentials for those noble metals are in good agreements with those predicted by DFT. For group 13 and 14 such as Zn, Cd, Hg, Al, Ga, In, and Tl, phonon dispersion curves by interatomic potentials are also in relatively good agreements with those predicted by DFT, however, for In, Hg, and Tl, there are large errors for some structures like ω structures.

Relatively large errors of some elements may be ascribed to deficiency of angular information. For example, because transition metals have some d electrons, angular information is needed to construct interatomic potentials for those elements.

Figure 3.33 and 3.34 show the phonon dispersion curves obtained from DFT, linear potentials, and conventional EAMs for Al and Cu, respectively. For both of elements, linear potentials can be predict phonon dispersion curves of six structures accurately, whereas discrepancy between DFT and EAMs can be found especially of unstable structures. To predict phonon dispersion curves accurately, it is necessary to forces of structures which is made by small displacement with preciseness of 1-10 meV/Å. It is notable that linear potentials can predict both energies of unstable structures as discussed in the above and phonon dispersion curves of stable structures.

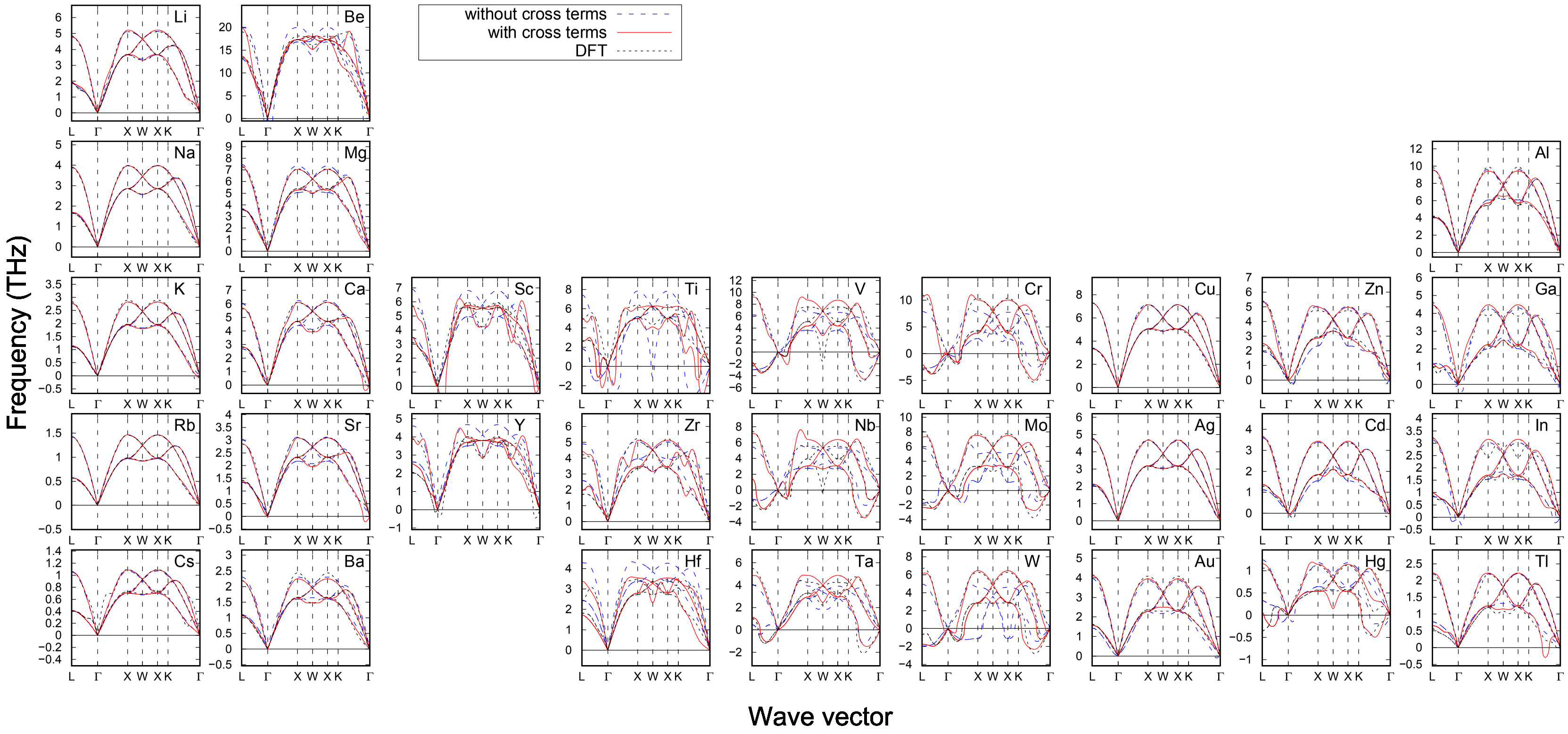


Figure 3.27: Phonon dispersion curves for 31 elemental metals with FCC structures by linear potentials without cross terms, linear potenitals with cross terms, and DFT calculations. Negative values indicate imaginary modes.

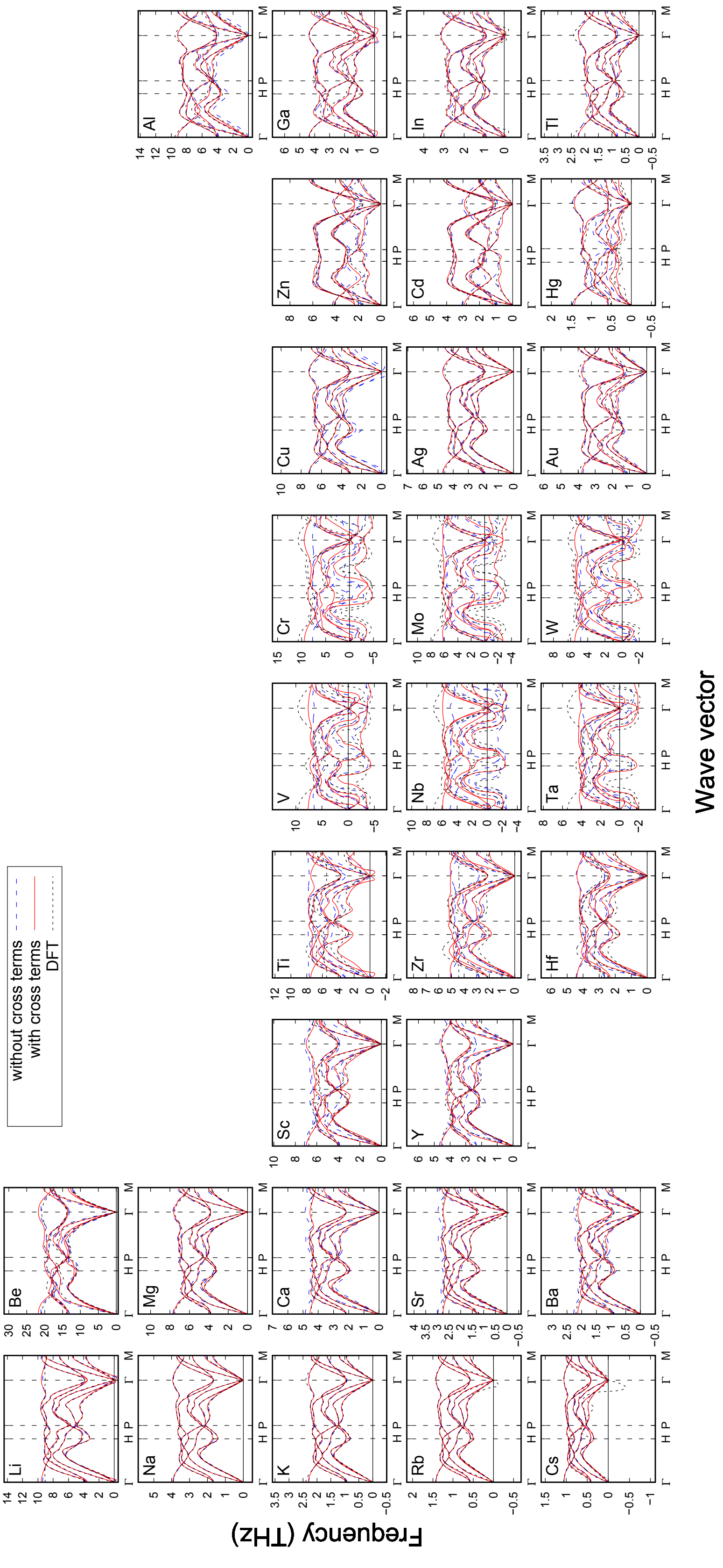


Figure 3.28: Phonon dispersion curves for 31 elemental metals with HCP structures by linear potentials without cross terms, linear potenitals with cross terms, and DFT calculations. Negative values indicate imaginary modes.

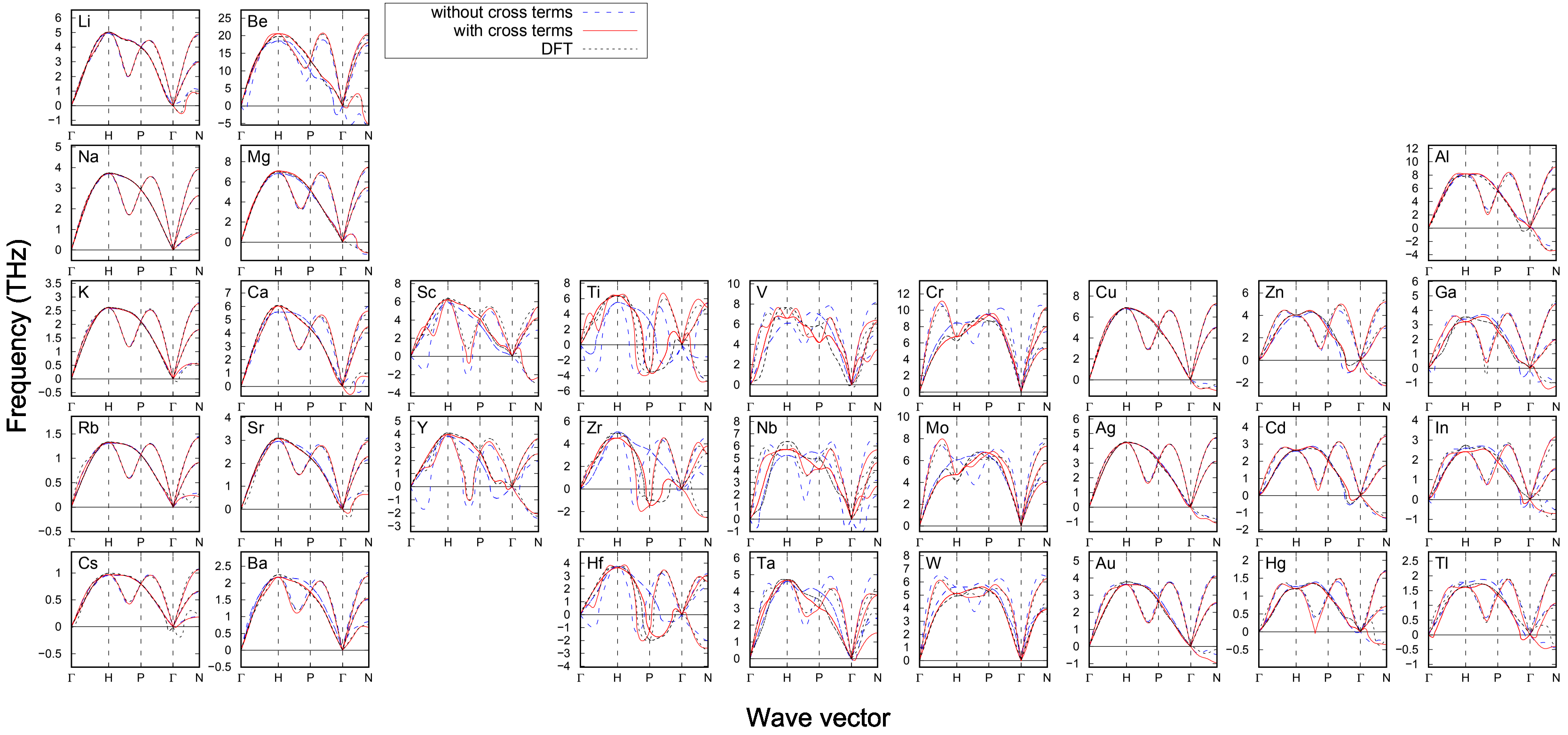


Figure 3.29: Phonon dispersion curves for 31 elemental metals with BCC structures by linear potentials without cross terms, linear potenitals with cross terms, and DFT calculations. Negative values indicate imaginary modes.

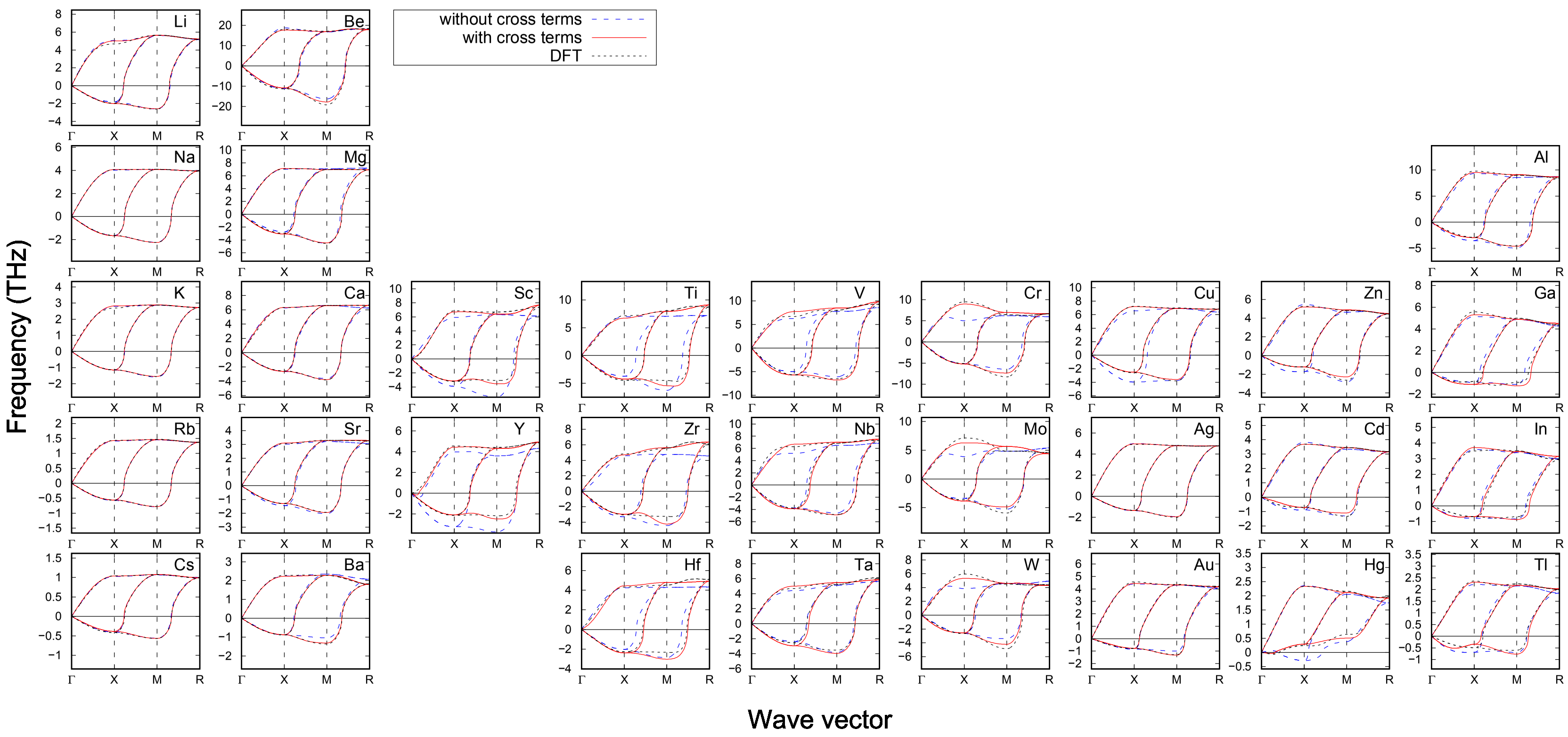


Figure 3.30: Phonon dispersion curves for 31 elemental metals with simple cubic structures by linear potentials without cross terms, linear potenitals with cross terms, and DFT calculations. Negative values indicate imaginary modes.

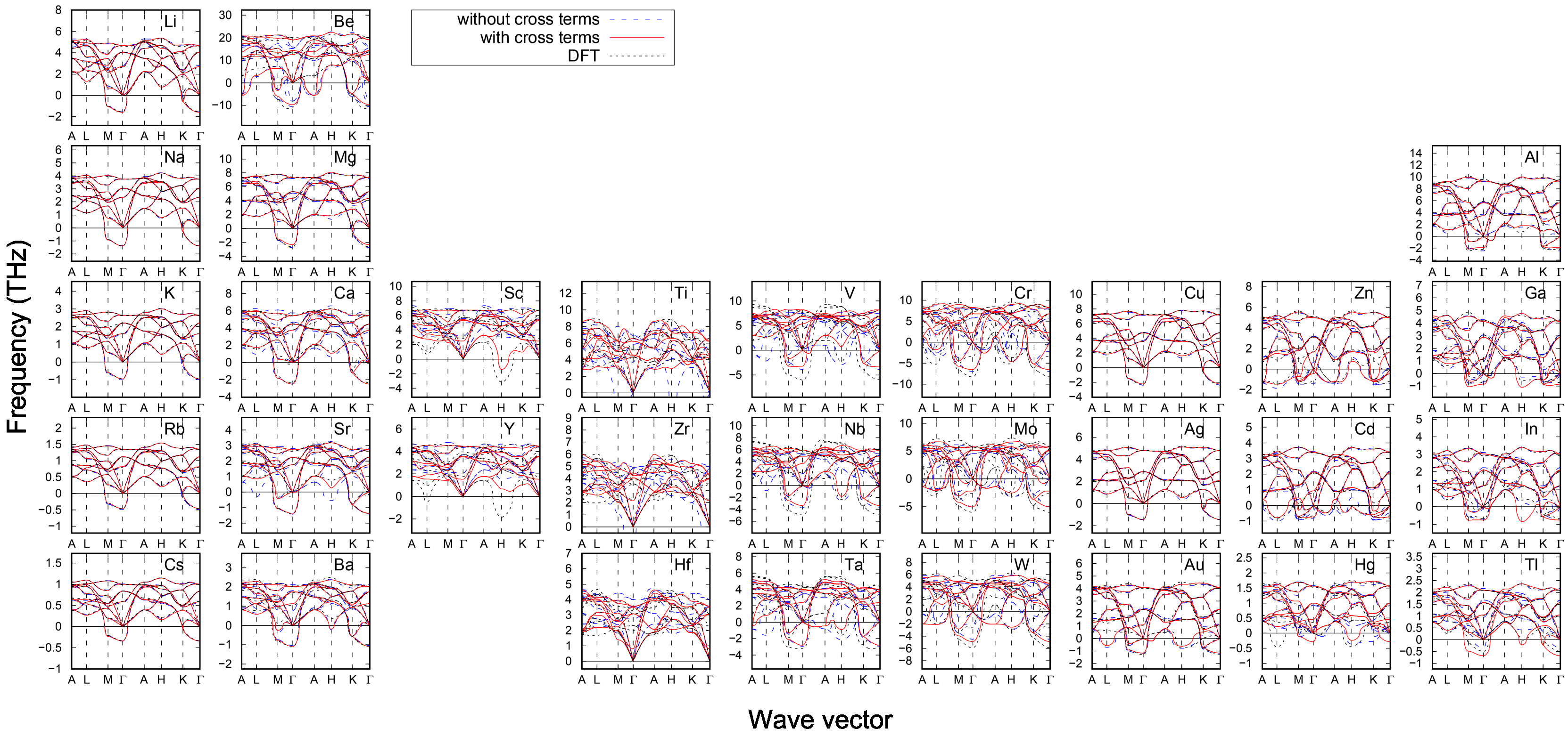


Figure 3.31: Phonon dispersion curves for 31 elemental metals with ω structures by linear potentials without cross terms, linear potenitals with cross terms, and DFT calculations. Negative values indicate imaginary modes.

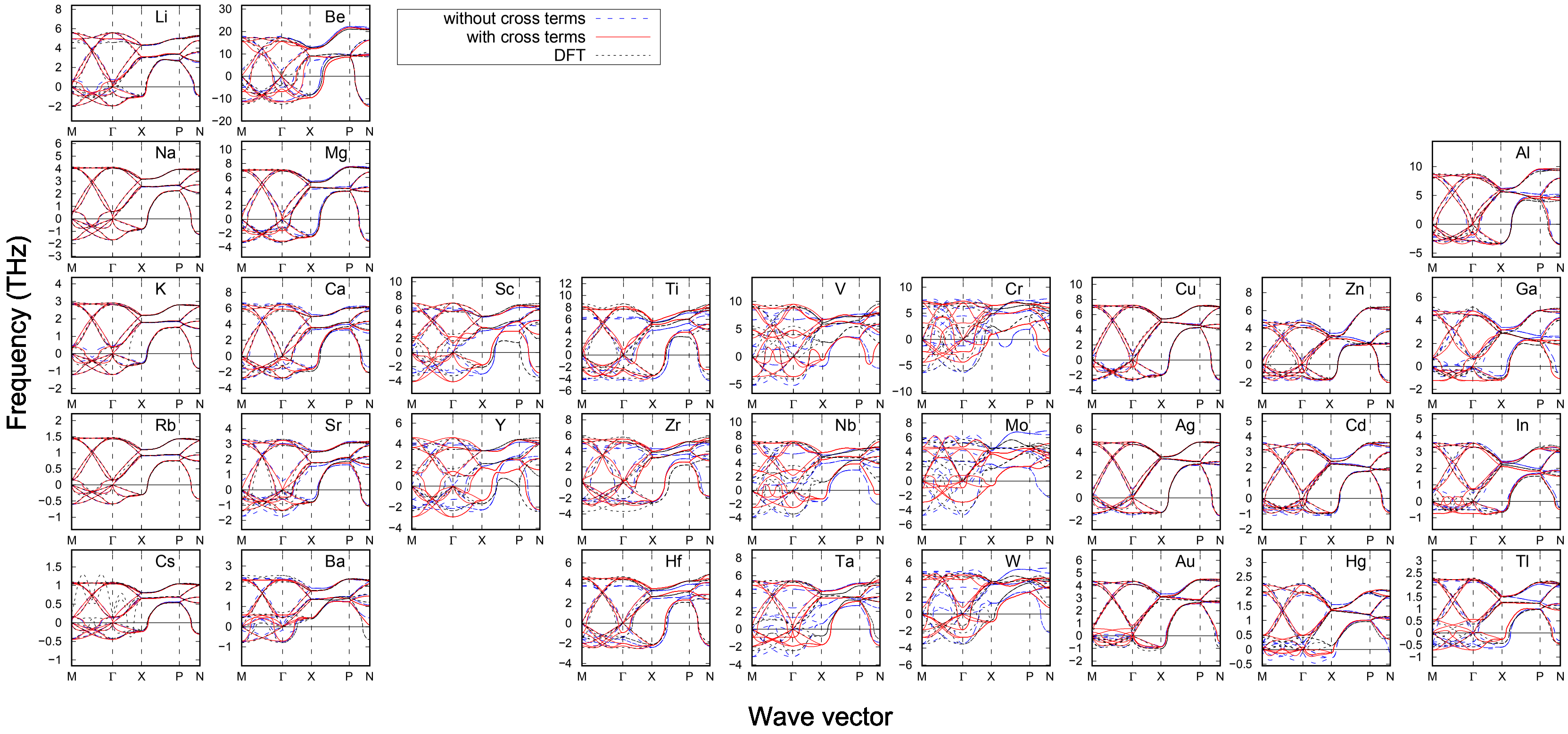


Figure 3.32: Phonon dispersion curves for 31 elemental metals with β-tin structures by linear potentials without cross terms, linear potenitals with cross terms, and DFT calculations. Negative values indicate imaginary modes.

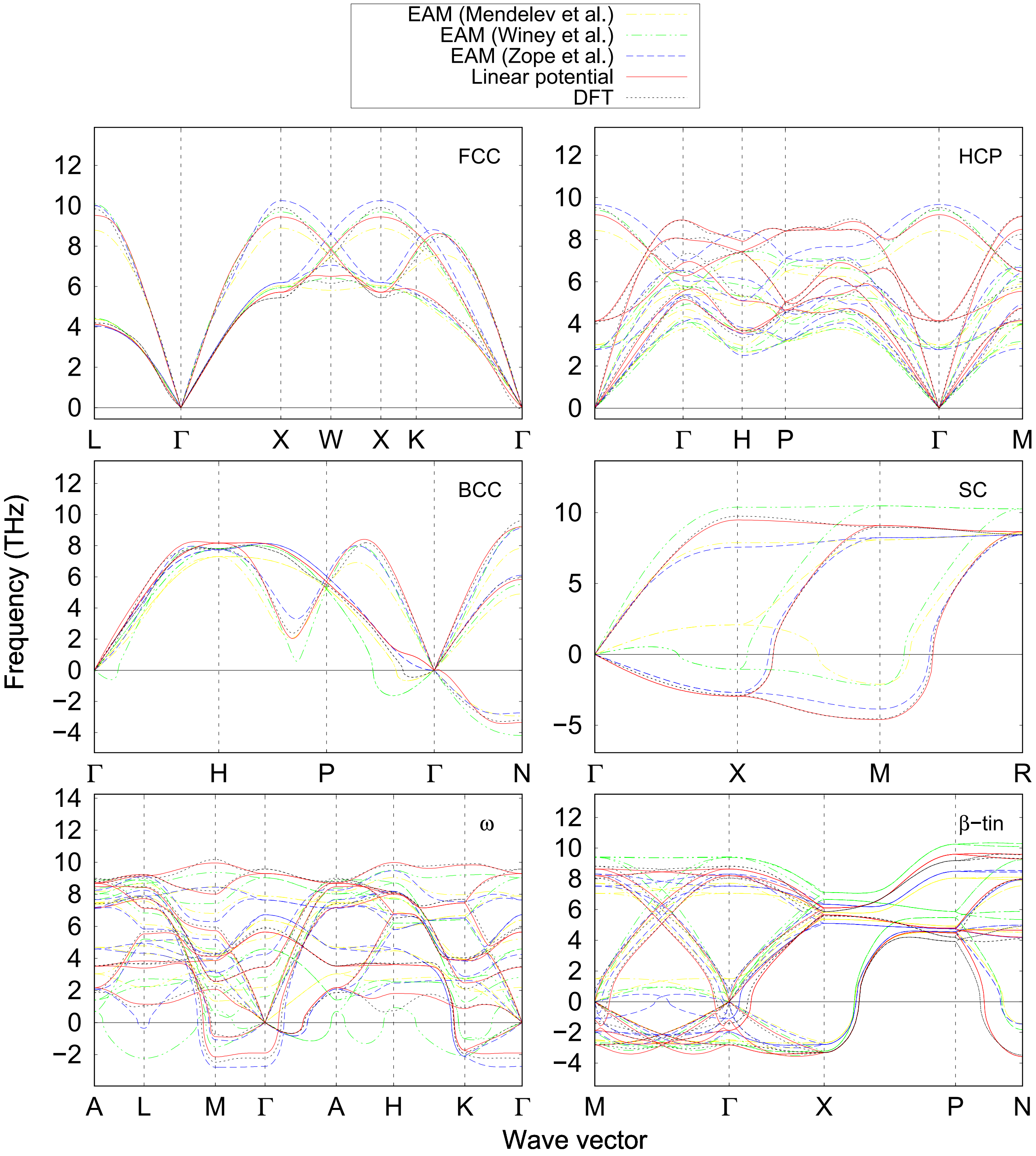


Figure 3.33: Phonon dispersion curves for Al metals of DFT, linear potential, Mendelev EAM [50], Winey EAM [51, 52], and Zope EAM [53].

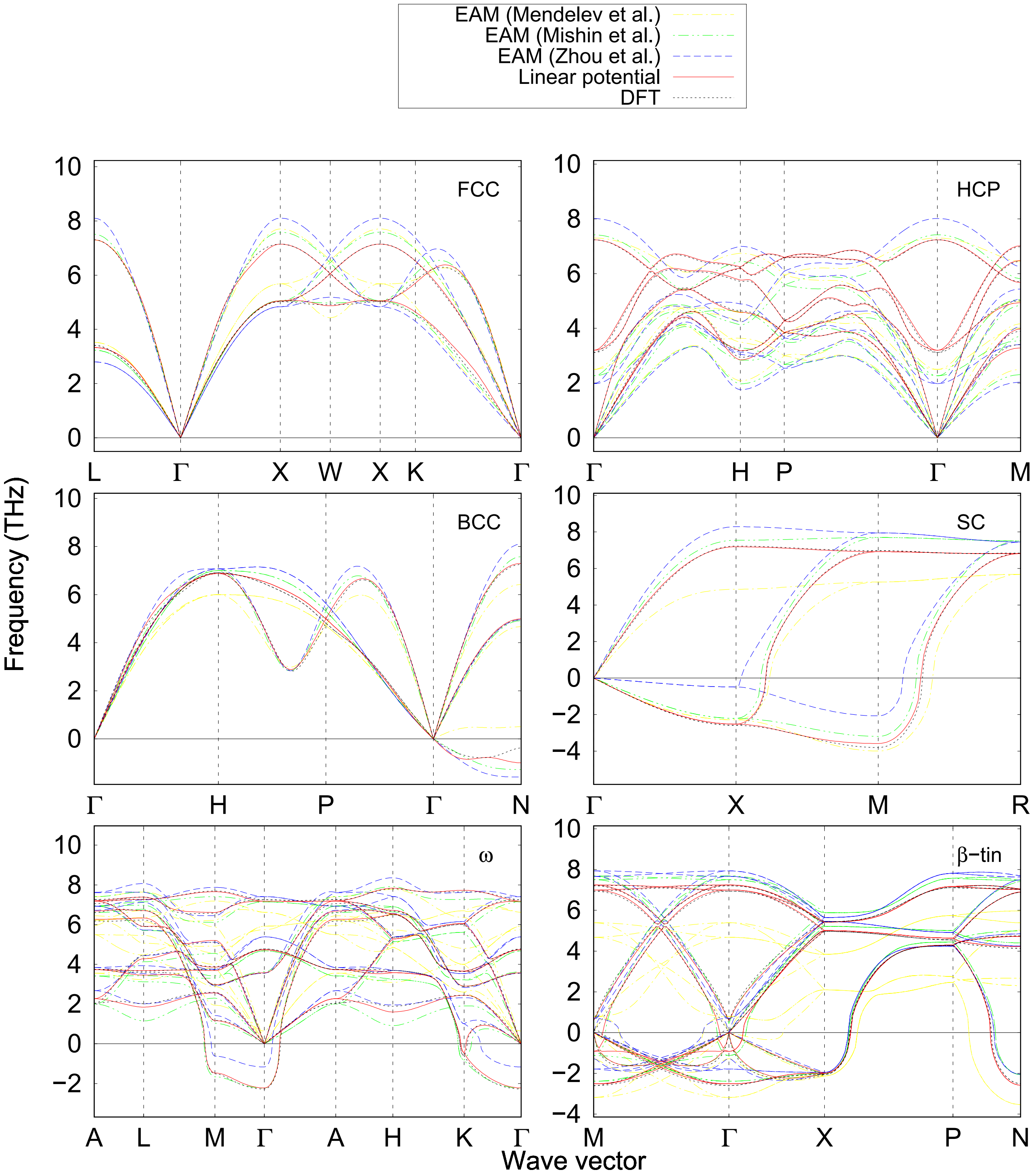


Figure 3.34: Phonon dispersion curves for Cu metals of DFT, linear potential, Mendelev EAM [50], Mishin EAM [54], and Zhou EAM [55].

3.6 Conclusion

In this chapter, I showed generalized formula of linear potential and construction of interatomic potentials using the generalized formula. Firstly I constructed by both previous and generalized model for 31 elemental metals using only radial descriptor. For all elements, compared with the previous model, generalized model constructed more accurate interatomic potentials for the same computational costs. Next, I predicted several properties of elements, namely, energies of structures which are not used for construction of interatomic potentials, elastic constants of cubic crystals, and phonon dispersion curves. For typical elements, interatomic potentials constructed in this study reproduced energies, bulk modulus, and phonon dispersion curves obtained from DFT calculations accurately. The predictions of elastic constants excepting bulk modulus were relatively poor. It may be ascribed to training data set. For transition metals, the predictions of elastic constants and phonon dispersion curves were relatively worse than that for typical elements. In terms of predictions of elastic constants and phonon dispersion curves of the most of metals, linear interatomic potentials generalized formula are better than linear potentials obtained by previous model. It may be ascribed to deficiency of angular information. For Al and Cu, these properties are also compared to conventional EAMs. The predictions of this work were more accurate than conventional EAMs, especially for unstable structures. This fact implies that generalized potential model can construct interatomic potentials of wide transferability and good accuracy.

Chapter 4

Descriptors of angular information

4.1 Introduction

In Chapter 3, only radial descriptors are considered to construct linear potentials. Even angular information was not considered, for typical elements predictions were accurate.

However, for transition metals and some elements in upper right of the periodical tables, some properties were not reproduced well. To construct interatomic potential models for such materials, usually bonding angle are considered. For example, for semiconductor or insulator, widely Tersoff potentials [15–17] are used and simulations for transition metals are often performed using modified EAMs [18, 19]. For construction of MLIP, some descriptors adopting bonding angle [22, 35] are proposed.

In this chapter, I constructed linear potentials including angular terms and predicted some properties of metals by those potentials.

4.2 Methodology

In general interatomic potential idea, the total energy of a target system E^{total} is decomposed into contribution of each atoms included the system. Therefore, E^{total} can be written as

$$E^{\mathrm{total}} = \sum_i E^{(i)}, \tag{4.1}$$

where $E^{(i)}$ is the contribution of atom i.

In the previous chapter, I defined radial descriptor $b_m^{(i)}$ as

$$b_m^{(i)} = \sum_j f_m(r_{ij}) \tag{4.2}$$

and expressed interatomic potential model as

$$E^{(i)} = \sum_{1 \le s_1 \le \cdots \le s_P \le M+1} w_{s_1,s_2,\ldots,s_P} b_{s_1}^{(i)} b_{s_2}^{(i)} \ldots b_{s_P}^{(i)}, \tag{4.3}$$

where $b_{M+1} = 1$ and $w_{s_1,s_2,\ldots,s_P}$ is fitting parameters.

Let us add the angular terms to Eqn.(4.3). As discussed in Chapter 2, interatomic potentials with angular terms can be expressed as

$$E^{(i)} = E(\sum_j f_1(r_{ij}), \sum_j f_2(r_{ij}), \ldots, p_{11}^{(i)}, \ldots, p_{nl}^{(i)}, \ldots, W_{11}^{(i)}, \ldots, W_{nl}^{(i)}, \ldots), \tag{4.4}$$

where r_{ij}, $f_n(r_{ij})$, $p_{nl}^{(i)}$ and $W_{nl}^{(i)}$ are distance between atom i and atom j, radial basis function, second order bond order parameters and third order bond order parameters with radial functions, respectively. Because I construct for unary system, long range interaction can be ignored and I use radial basis functions which are zero if distance is larger than cutoff radius R_c. In this study, only p_{nl} are used as angular descriptors. p_{nl} is expressed as

$$p_{nl}^{(i)} = \sum_{j,k} f_n(r_{ij}) f_n(r_{ik}) P_l(\cos \gamma_{ijk}), \tag{4.5}$$

where γ_{ijk} is bonding angle among atom i, j and k, and P_l is a Legendre polynomial function.

The interatomic potential model of linear combination of descriptors can be obtained from polynomial approximation of Eqn.(4.4). However, if polynomial expansion is naively applied, too many terms are generated. Such many terms causes overfitting and make difficult regression computationally due to demanding of large memory. In this study, only angular Fourier series (AFS), expressed as

$$\mathrm{AFS}_{n,n',l}^{(i)} = \sum_{j,k} f_n(r_{ij}) f_{n'}(r_{ik}) \cos^l(\gamma_{ijk}) \tag{4.6}$$

are added to Eqn.(4.3). Therefore, the model with angular terms is expressed as

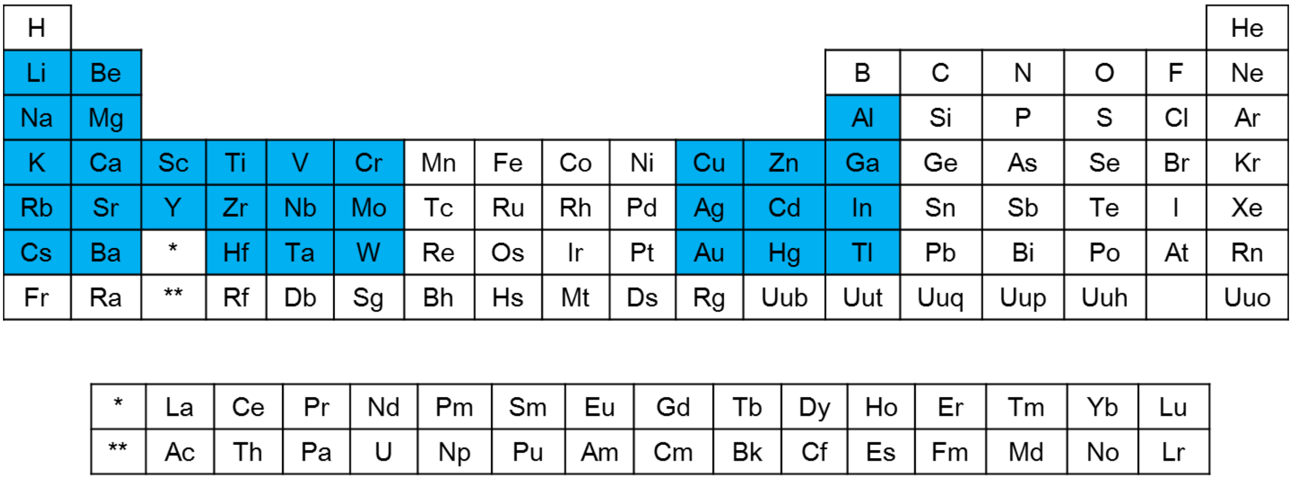
$$E^{(i)} = \sum_{1 \le s_1 \le \cdots \le s_P \le M+1} w_{s_1, s_2, \ldots, s_P} b_{s_1}^{(i)} b_{s_2}^{(i)} \ldots b_{s_P}^{(i)} + \sum_{n, n', l} w_{n,n',l}^{\mathrm{AFS}} \mathrm{AFS}_{n,n',l}^{(i)}, \tag{4.7}$$

where $w_{n,n',l}^{\mathrm{AFS}}$ is fitting parameters for angular terms.

4.3 Target elements and structures

The training and test data sets are generated by the DFT calculation. To generate training and test data sets, I started to fully optimize the atomic positions and lattice constants of conventional FCC, BCC, HCP, simple cubic (SC), ω, and β-tin structures for each element. Next, from the six ideal structures, I constructed supercells, which are $2 \times 2 \times 2$ FCC (32 atoms), $3 \times 3 \times 3$ BCC (54 atoms), $3 \times 3 \times 3$ HCP (54 atoms), $4 \times 4 \times 4$ SC (64 atoms), $3 \times 3 \times 3$ ω (81 atoms), and $2 \times 2 \times 2$ β-tin (32 atoms) structures. Using the supercells, training and test set are generated by 30 isotropic expansions and 470 random expansions and distortions on the basis of the six ideal supercells. The volumes of 30 structures generated by isotropic expansions. Volume ratio of optimized structure and generated structures varies from 0.5 to 5.0. Because framework to construct MLIP has no physical background, such greatly compressed and expanded structures are necessary to make interatomic potential predicts high energy when distances between atoms are very close or when atoms isolate. Then, 500 deformed structures are generated for one structure and totally 3000 structures are generated for one elemental metal. Finally, the 3000 configurations are divided into 2700 training data sets and 300 test data sets of each element, respectively.

For all the configurations, energy and forces acting on each atom are calculated by the DFT calculation using the plane-wave basis projector augmented wave (PAW) method [41, 42] within the Perdew-Burke-Ernzerhof exchange-correlation functional [43] as implemented in the VASP code [44, 45]. The cutoff energy was set to 500 eV. The total energies converged to less than 10^{-3} meV/supercell. For optimization of lattice constants of ideal structures, the lattice constants were optimized until the residual forces became less than 10^{-3} eV/Å.



H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	**	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Uub	Uut	Uuq	Uup	Uuh		Uuo

*	La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
**	Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr

Figure 4.1: 31 nonmagnetic elements of which interatomic potentials are constructed.

4.4 Construction of interatomic potential

In this section, interatomic potentials including AFS expressed as Eqn.(4.7) are constructed for 31 elements. To estimate the effect of AFS for regression error and construct optimal potential, interatomic potentials are constructed varying $l_{\max}$ from 0 to 3. When $l_{\max} = 0$, the potential model coincide with Eqn.(4.3), namely potential model used in previous chapter 3. The potential is constructed by linear ridge regression, in which the coefficients w is determined by solving

$$\boldsymbol{w} = (X^{\mathrm{T}}X + \lambda I)^{-1}X^{\mathrm{T}}y, \tag{4.8}$$

where X, y and I are descriptors matrix, DFT data and an unit matrix whose dimension is the number of descriptors. For training data, energy and forces acting on atom are included. The detail about regression method is written in subsection 3.2.2.

The expansion power P is fixed to 3. Gaussian functions expressed as

$$f_n(r) = \exp\left(-a_n(r - b_n)^2\right) \tag{4.9}$$

are used as radial basis set. (a_n, b_n), namely pairs of Gaussian parameters is determined by fixed a and arithmetic sequence from 0 to $b_{\max}$, or $(a_n, b_n)_{1\leq n\leq N} = (a, (n-1)b_{\max}/(N-1)_{1\leq n\leq N}$. I constructed totally 81 interatomic potentials with $a = 0.1, 0.25, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5$ and $b_{\max}$ is varied from 5.0 Å to 13.0 Å for each cutoff radii, and the interatomic potentials whose

RMS error of test set is the lowest is selected as the optimal one of the cutoff radii. The same cutoff radii as subsection 3.4.1 are used. The ridge penalty terms was set to $\lambda = 10^{-3}, 10^{-4}$, and 10^{-5} and penalty terms which minimize the root mean square (RMS) error of energies of test set is adopted.

Figure 4.2 and 4.3 show the dependency of RMS errors for energies and forces of both models on l_{max}. As can be seen, RMS errors of both energies and forces for 31 elements are decreased by adopting AFS.

Table 4.4 shows the optimal parameters and RMS errors of energies and forces of the model with AFS. For all 31 elements, using AFS, RMS errors of energies and forces are less than those of interatomic potential without angular terms shown in table 3.3.

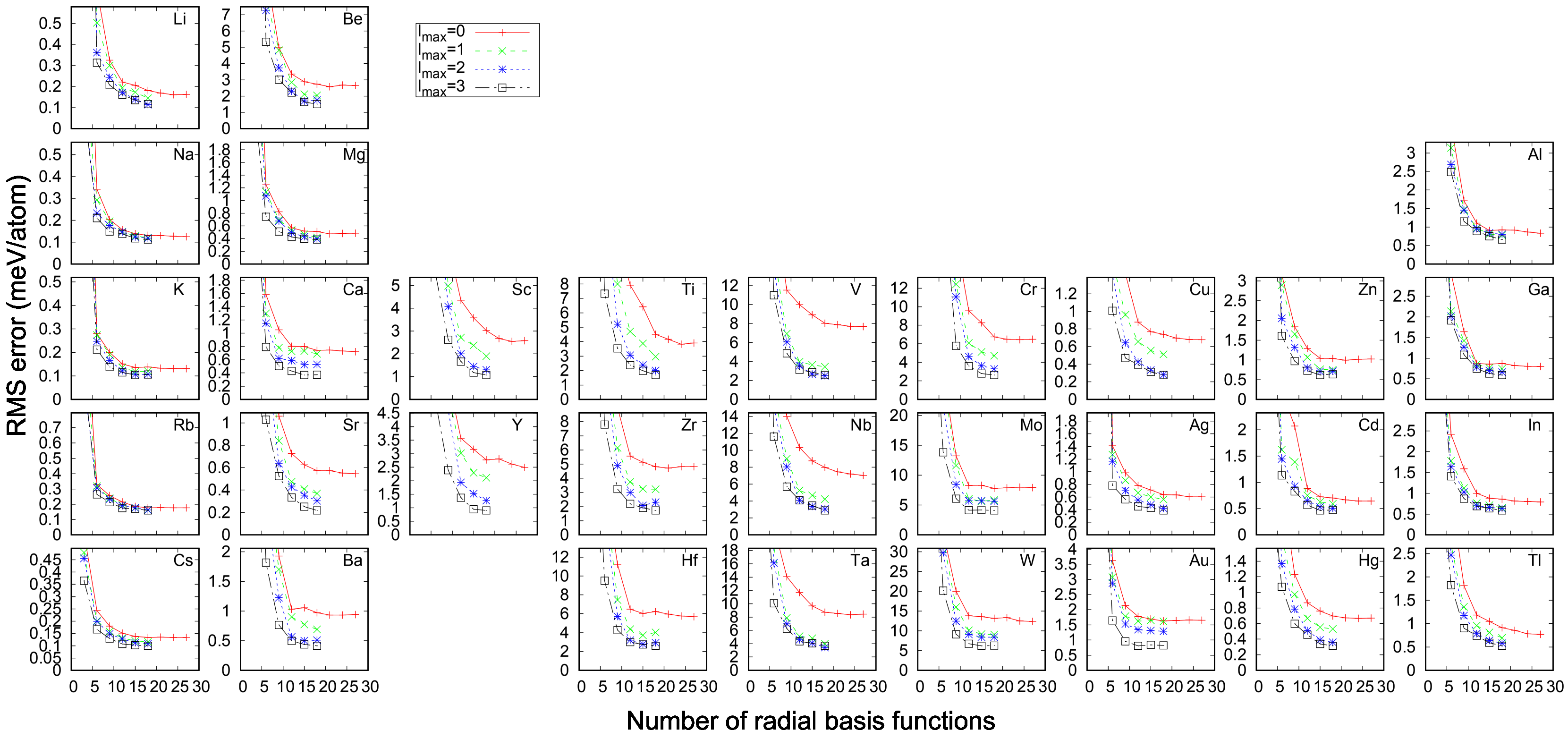


Figure 4.2: The dependency of RMS errors for energies on $l_{\max}$ and number of basis radial functions.

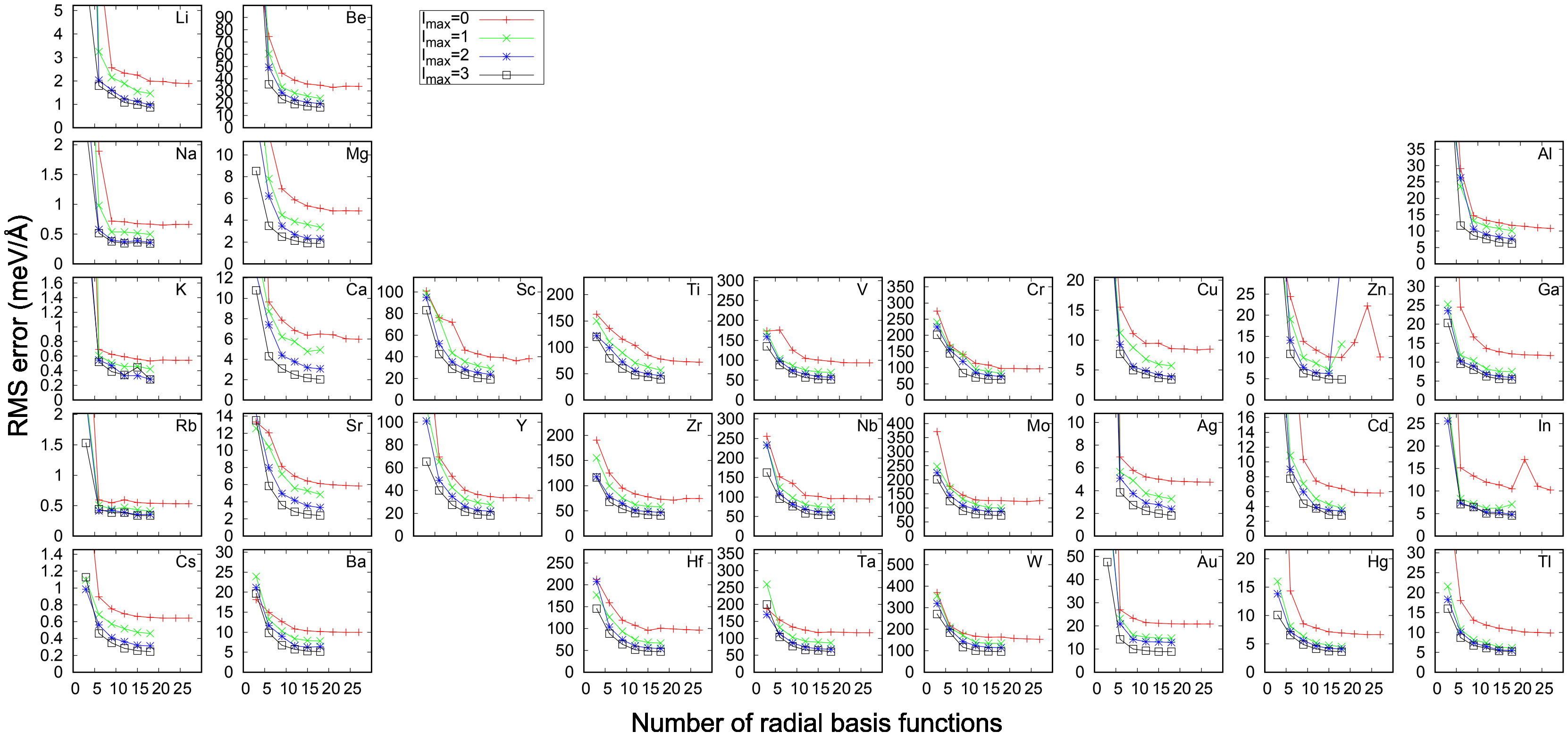


Figure 4.3: The dependency of RMS errors for forces on $l_{\max}$ and number of basis radial functions.

Table 4.1: The optimal basis functions for 31 elements for framework with angular terms.

Element	Number of basis functions	a (Å^{-2})	b_{max} (Å)	l_{max}	RMS error (energy) (meV/atom)	RMS error (force) (eV/Å)
Li	18	1.0	12.0	2	0.1	0.001
Be	18	2.5	6.0	3	1.5	0.017
Na	18	1.5	11.0	3	0.1	0.000
Mg	18	1.0	11.0	3	0.4	0.002
Al	18	1.5	10.0	3	0.7	0.006
K	15	1.5	11.0	3	0.1	0.000
Ca	15	1.0	12.0	3	0.4	0.002
Sc	18	1.5	9.0	3	1.1	0.019
Ti	18	1.5	10.0	3	1.7	0.039
V	18	3.0	6.0	2	2.5	0.057
Cr	18	2.0	6.0	3	2.6	0.063
Cu	18	2.0	9.0	3	0.3	0.003
Zn	15	1.5	8.0	3	0.6	0.005
Ga	18	1.0	10.0	3	0.6	0.005
Rb	18	1.0	12.0	3	0.2	0.000
Sr	18	1.0	11.0	3	0.2	0.002
Y	18	1.0	10.0	3	0.9	0.018
Zr	18	2.0	8.0	3	1.7	0.041
Nb	18	2.0	7.0	3	2.9	0.052
Mo	18	1.5	8.0	3	4.1	0.073
Ag	18	1.5	10.0	3	0.4	0.002
Cd	15	1.0	12.0	3	0.5	0.003
In	18	1.5	10.0	3	0.6	0.005
Cs	18	0.5	13.0	3	0.1	0.000
Ba	18	1.0	9.0	3	0.4	0.005
Hf	18	1.5	7.0	3	2.6	0.047
Ta	18	2.0	7.0	2	3.4	0.068
W	15	2.0	6.0	3	6.2	0.097
Au	12	2.0	7.0	3	0.8	0.009
Hg	18	1.0	9.0	3	0.3	0.004
Tl	18	1.5	10.0	3	0.5	0.005

4.5 Prediction of physical properties

4.5.1 Elastic constants and bulk modulus

Elastic constants are fundamental properties of bulk material and to evaluate them accurately is important when interatomic potentials are used to predict mechanical properties. In this study, I calculated the elastic constants of cubic crystals and hexagonal crystals with both interatomic potentials and DFT calculations. I estimated the elastic constants of cubic crystals by the same way as the book written by Finnis [11] and those of hexagonal crystals by the same way as literature of Fast *et al.* [56].

According to the literature of Finnis, Elastic constants can be estimated by energies of strained structures. Now I denote a homogeneous strain matrix as

$$\epsilon = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{12} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{13} & \epsilon_{23} & \epsilon_{33} \end{bmatrix} . \tag{4.10}$$

The elements of ϵ is defined as follows. If the strain carries the point $\boldsymbol{r} = (r_1, r_2, r_3)$ to a new position $\boldsymbol{r} + \boldsymbol{u}$, elements of ϵ are defined by

$$\epsilon_{ab} = \frac{1}{2}\left(\frac{\delta u_\alpha}{\delta x_\beta} + \frac{\delta u_\alpha}{\delta x_\beta}\right) . \tag{4.11}$$

In practice, such transformation can be achieved by changing basis vectors from $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ to $\{\boldsymbol{a}'_1, \boldsymbol{a}'_2, \boldsymbol{a}'_3\}$ where

$$a'_{i\alpha} = a_{i\alpha} + \sum_\beta \epsilon_{\alpha\beta} a_{i\beta}, \tag{4.12}$$

in which $a_{i\alpha}$ is the Cartesian component along the x_α axis of basis vector $\boldsymbol{a}_i$.

Using Voigt notation, strain matrix can also be expressed as

$$\epsilon = \begin{bmatrix} \epsilon_1 & \epsilon_6 & \epsilon_5 \\ \epsilon_6 & \epsilon_2 & \epsilon_4 \\ \epsilon_5 & \epsilon_4 & \epsilon_3 \end{bmatrix} . \tag{4.13}$$

Then, stress-strain relationship is expressed as

$$\sigma_i = \sum_{j=1}^{6} C_{ij}\epsilon_j, \tag{4.14}$$

where σ_i and C_{ij} are elements of strain 6×6 matrix and elastic constants 6×6 matrix using Voigt notation, respectively. Then, elastic energy per unit volume stored homogeneously strained body is given by

$$E^{\text{elas}} = \frac{1}{2}\sum_{ij} C_{ij}\epsilon_i\epsilon_j. \tag{4.15}$$

Therefore, considering appropriate symmetry transformation matrix T and scholar parameter γ and calculating $E(\gamma, T)$, which is energy when structure is strained by matrix γT, elastic constants matrix can be obtained from second derivative of the elastic energy with respective to strain, namely $\frac{\delta^2 E(\gamma,T)}{\delta\gamma^2}|_{\gamma=0}$.

For cubic system, there are three independent elastic constants C_{11}, C_{12}, and C_{44}. Elastic constant matrix can be expressed as

$$C = \begin{bmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix}. \tag{4.16}$$

Elastic energy per unit volume is

$$E^{\text{elas}} = \frac{1}{2}C_{11}\left(\epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2\right) + C_{12}\left(\epsilon_1\epsilon_2 + \epsilon_2\epsilon_3 + \epsilon_3\epsilon_1\right) + C_{44}\left(\epsilon_4^2 + \epsilon_5^2 + \epsilon_6^2\right). \tag{4.17}$$

Elastic constants can be calculated by consideration of three matrices expressed as

$$T_e = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.18}$$

$$T_r = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix} \tag{4.19}$$

$$T_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.5 & 0 \\ 0 & 0 & -0.5 \end{bmatrix}. \tag{4.20}$$

From Eqn.(4.17), the relationships between elastic constants and second derivatives of energy respective to strain are

$$\left.\frac{\delta^2 E(\gamma, T_e)}{\delta\gamma^2}\right|_{\gamma=0} = 3V_0(C_{11} + 2C_{12}), \tag{4.21}$$

$$\left.\frac{\delta^2 E(\gamma, T_r)}{\delta\gamma^2}\right|_{\gamma=0} = \frac{3}{2}V_0(C_{11} - C_{12}), \tag{4.22}$$

$$\left.\frac{\delta^2 E(\gamma, T_t)}{\delta\gamma^2}\right|_{\gamma=0} = 3V_0 C_{44}, \tag{4.23}$$

where V_0 is volume per atoms of structure at $\gamma = 0$. Elastic constants are obtained by solving these equations.

For hexagonal system, there are five independent elastic constants C_{11}, C_{12}, C_{33}, C_{55}, and C_{13}. Elastic constants can be calculated by consideration of five matrices expressed as

$$T_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.24}$$

$$T_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{4.25}$$

$$T_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{4.26}$$

$$T_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.27}$$

$$T_5 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \tag{4.28}$$

The relationships between elastic constants and second derivatives of energy respective to strain are

$$\left.\frac{\delta^2 E(\gamma, T_1)}{\delta\gamma^2}\right|_{\gamma=0} = \frac{2}{9}V_0(2C_{11} + 2C_{12} + 4C_{13} + C_{33}), \tag{4.29}$$

$$\left.\frac{\delta^2 E(\gamma, T_2)}{\delta\gamma^2}\right|_{\gamma=0} = 2V_0(C_{11} + C_{12}), \tag{4.30}$$

$$\left.\frac{\delta^2 E(\gamma, T_3)}{\delta\gamma^2}\right|_{\gamma=0} = 2V_0(C_{11} - C_{12}), \tag{4.31}$$

$$\left.\frac{\delta^2 E(\gamma, T_4)}{\delta\gamma^2}\right|_{\gamma=0} = V_0 C_{33}, \tag{4.32}$$

$$\left.\frac{\delta^2 E(\gamma, T_5)}{\delta\gamma^2}\right|_{\gamma=0} = 4V_0 C_{55}. \tag{4.33}$$

Elastic constants are obtained by solving five equations.

In this study, to calculate second derivative, I calculated energies in the case γ ranging from -0.01 to 0.01 in steps of 0.001 for those strain matrices, fitted those energies by quartic polynomials, and obtained second derivative from coefficient of second-order term by DFT calculations, linear potentials, and conventional potentials. For all methods to calculate energies, lattice constants when $\gamma = 0$ are same as those obtained from DFT structure optimization. Elastic constants of conventional potentials are calculated by LAMMPS code [48, 49].

Figure 4.4, 4.5, 4.6, 4.7, and 4.8 show the second derivatives of the energies for the strain matrices for 31 elements of FCC, BCC, SC, HCP, and ω structures, respectively. Figure 4.9, 4.10, 4.11 4.12, and 4.13 show the elastic constants and bulk modulus 31 elements of FCC, BCC, SC, HCP, and ω structures, respectively. For most of typical metals, there is little difference between second derivatives obtained from linear potentials without angular terms and that with angular terms. Angular terms improve the prediction of the most of all elastic constants. Especially, for some transition metals such as Sc, Y, Ti, V, Nb, Ta, Cr, Mo, and W, improvement is remarkable. For some of elements at the upper right of periodical table, such as Be, Mg, Zn, Cd, Hg, Al, Ga, In, and Tl, there are still rather large discrepancy between DFT and linear potentials for some structures. As discussed in Chapter 3, it can be ascribed to training set and/or as discussed in Chapter 2 it can be ascribed to lack of other descriptors.

Figure 4.14 and 4.15 show the second derivatives and elastic constants predicted by DFT, linear potentials, and conventional EAM and conventional modified EAM for Ti of cubic crystals and hexagonal crystals, respectively. For stable structures such as HCP and FCC, the linear potential with angular terms reproduces elastic constants obtained from DFT as well as conventional potentials. For unstable structures such as SC, the linear potential with angular terms reproduces elastic constants obtained from DFT better than conventional potentials.

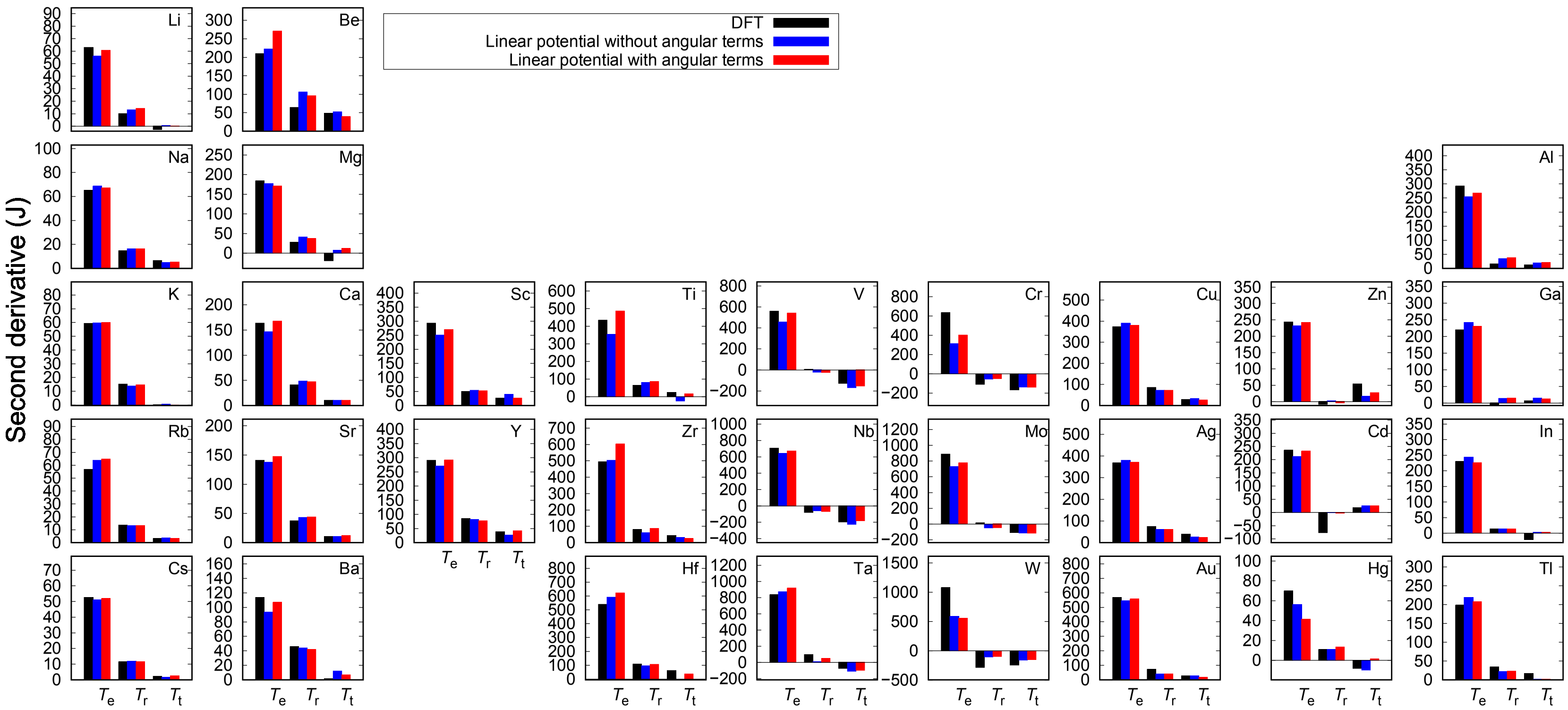


Figure 4.4: Second derivatives of energies respective to matrices expressed as Eqn.(4.18), (4.19), and (4.20) of FCC structures obtained from DFT calculations, linear potentials without angular terms, and linear potentials angular terms.

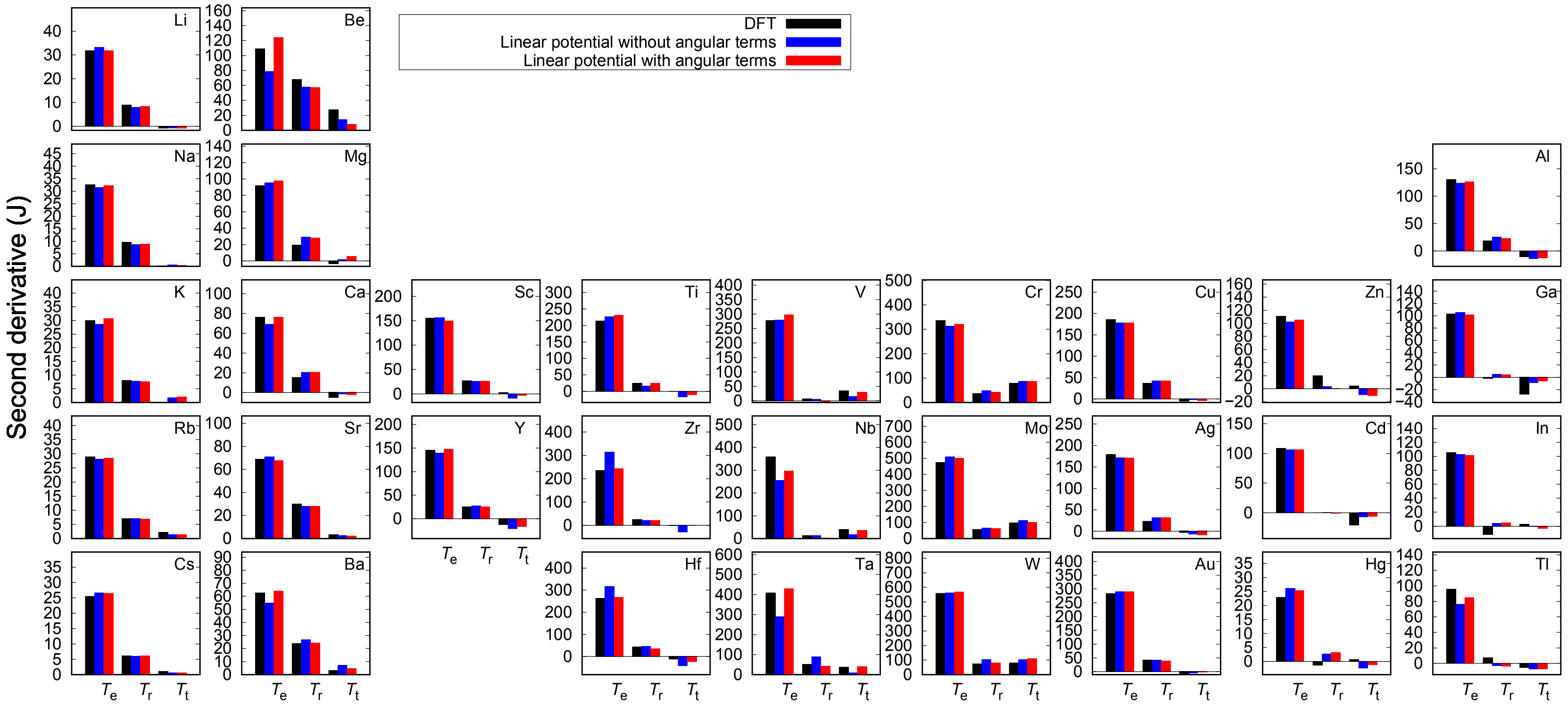


Figure 4.5: Second derivatives of energies respective to matrices expressed as Eqn.(4.18), (4.19), and (4.20) of BCC structures obtained from DFT calculations, linear potentials without angular terms, and linear potentials angular terms.

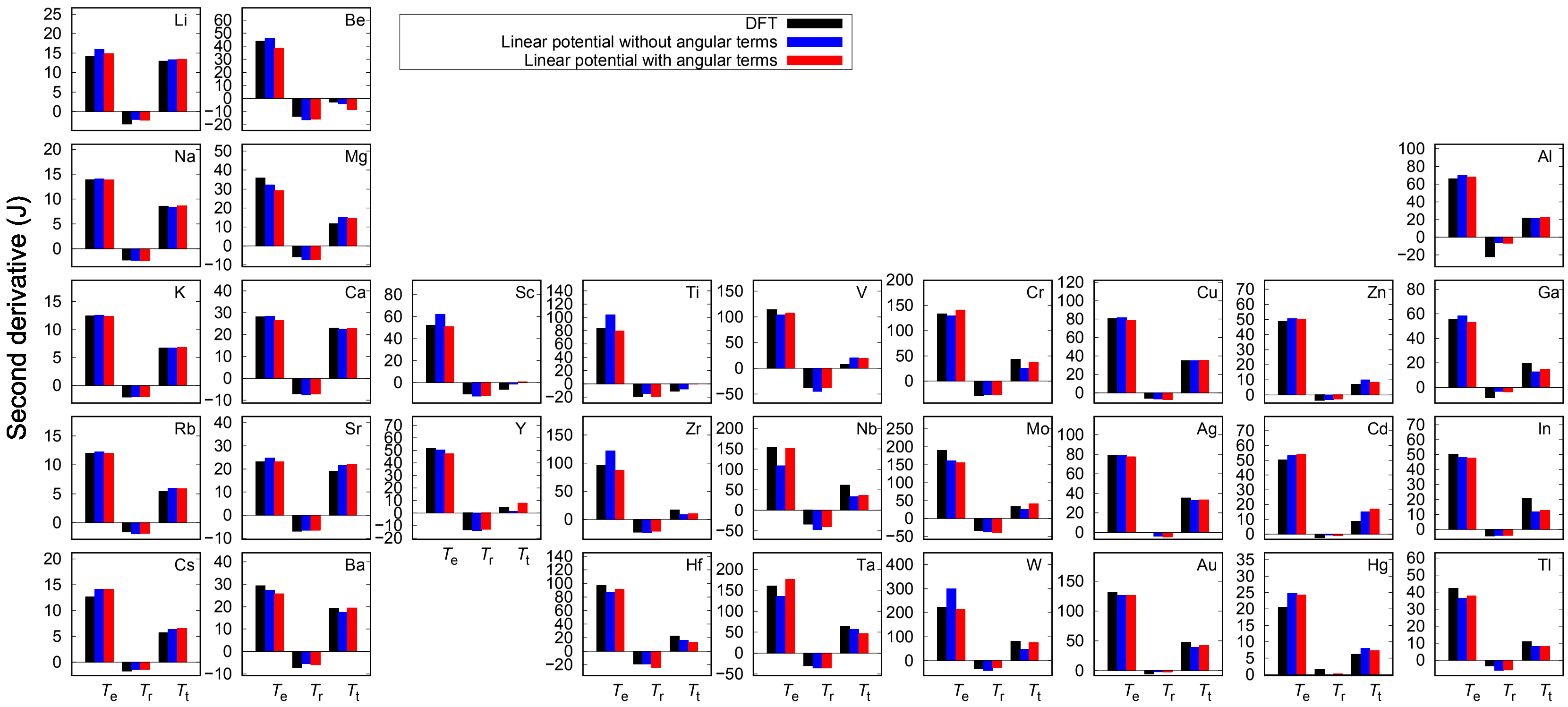


Figure 4.6: Second derivatives of energies respective to matrices expressed as Eqn.(4.18), (4.19), and (4.20) of SC structures obtained from DFT calculations, linear potentials without angular terms, and linear potentials angular terms.

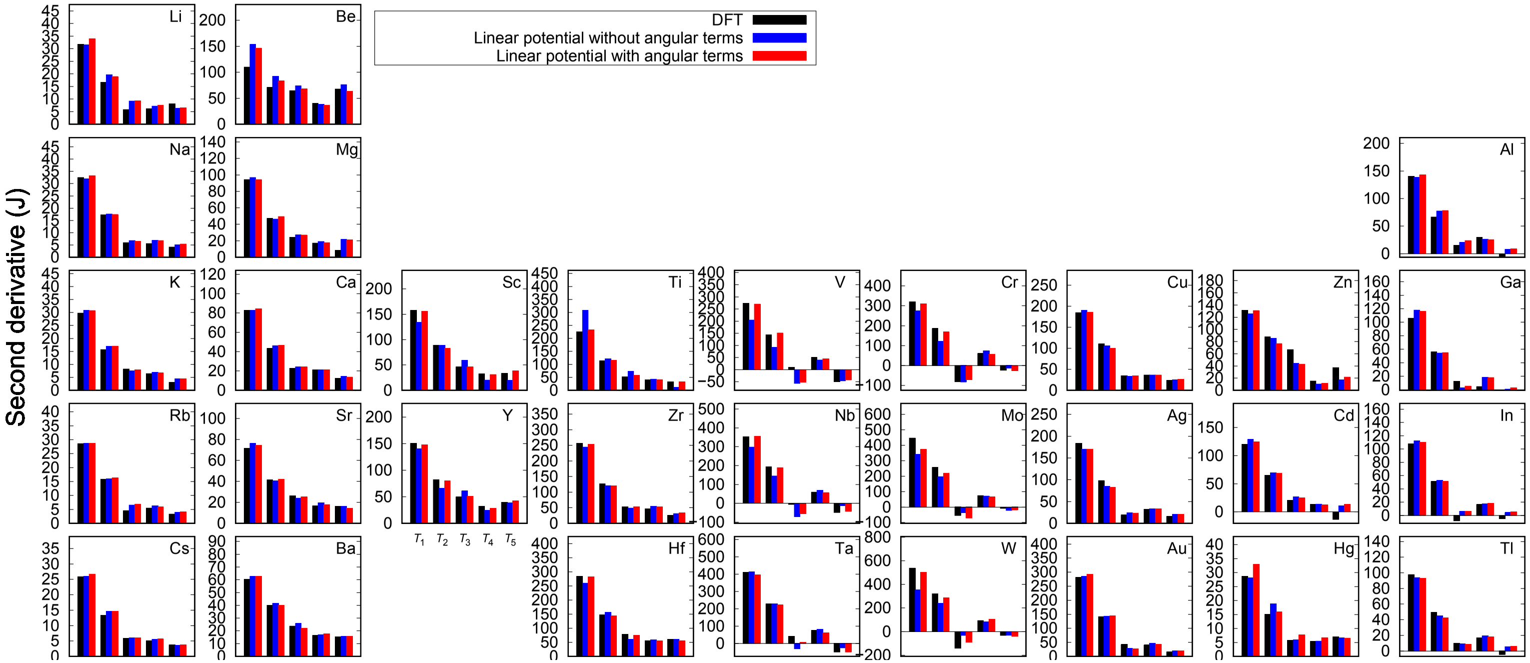


Figure 4.7: Second derivatives of energies respective to matrices expressed as Eqn.(4.24), (4.25), (4.26), (4.27), and (4.28) of HCP structures obtained from DFT calculations and linear potentials.

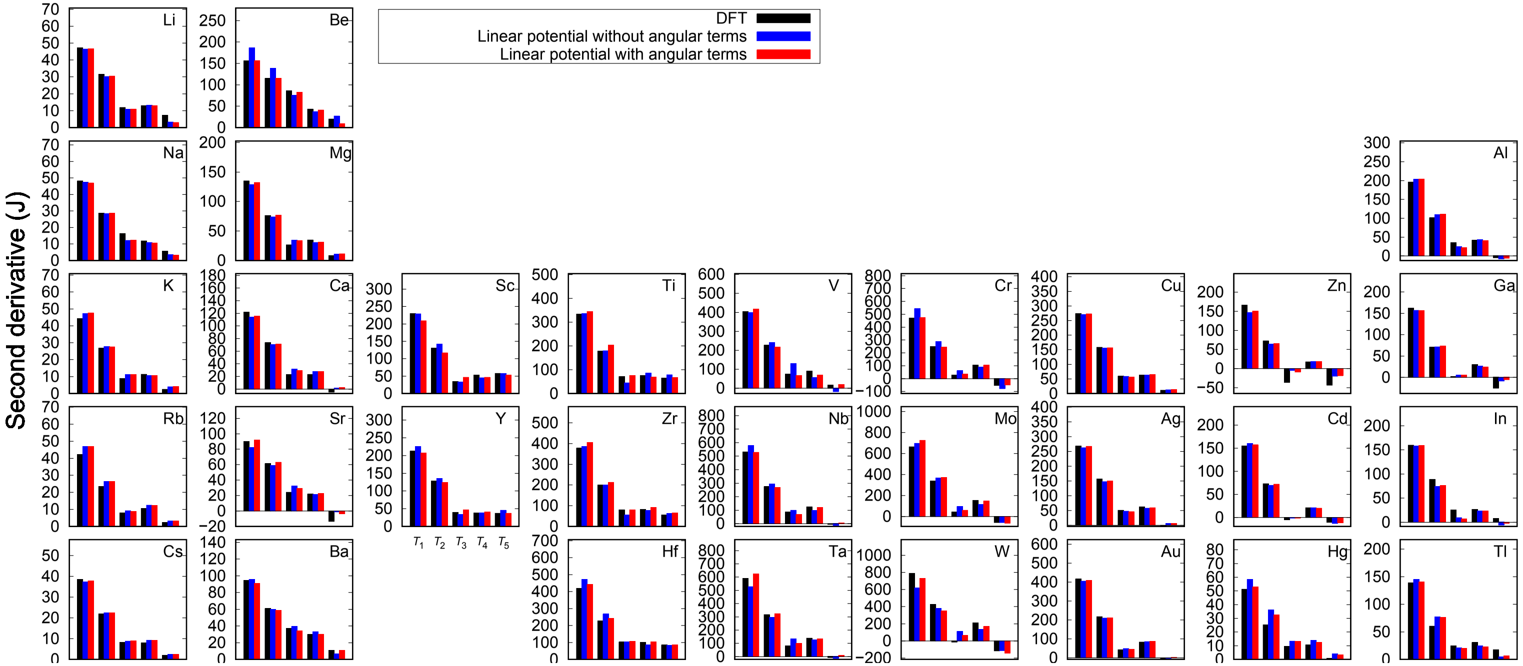


Figure 4.8: Second derivatives of energies respective to matrices expressed as Eqn.(4.24), (4.25), (4.26), (4.27), and (4.28) of ω structures obtained from DFT calculations and linear potentials.

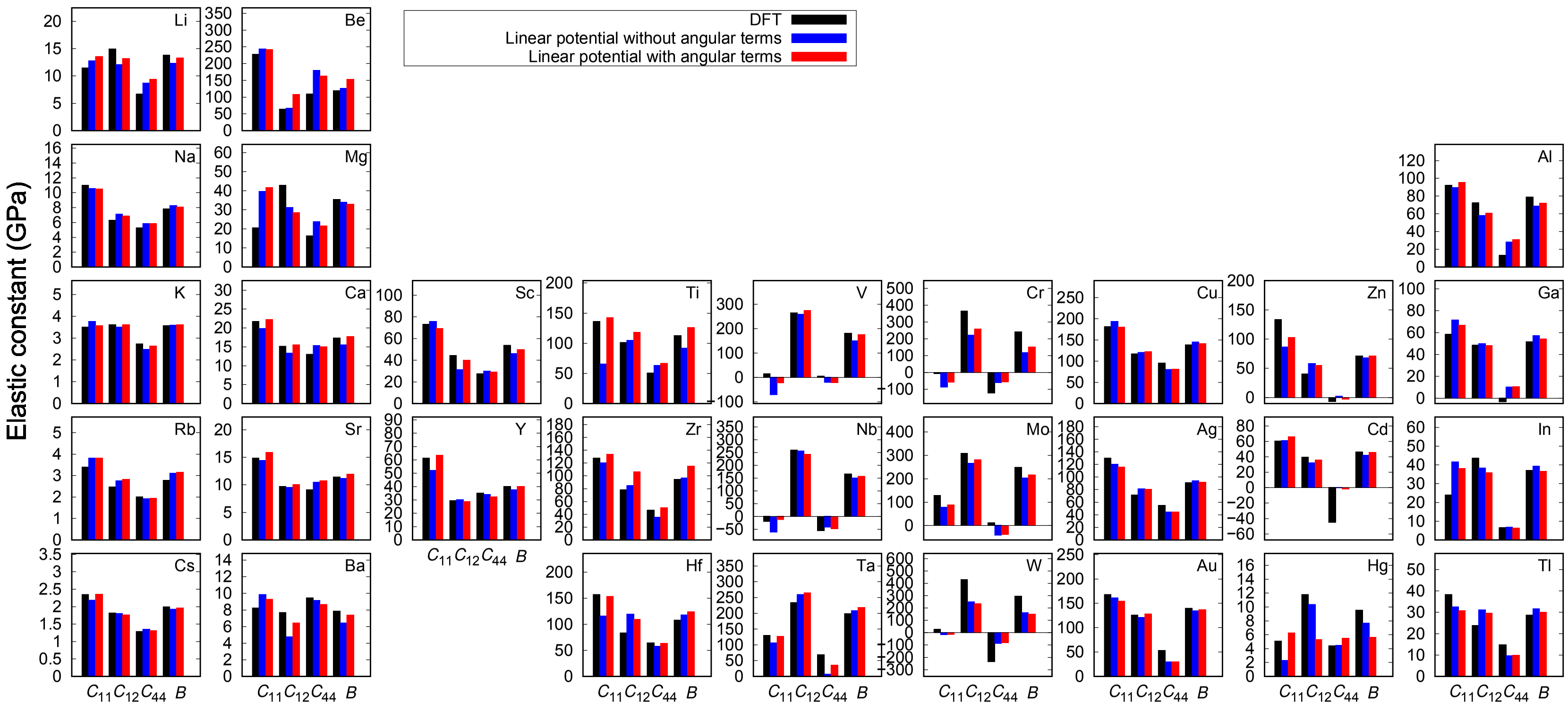


Figure 4.9: Elastic constants of FCC structures obtained from DFT calculations, linear potentials without angular terms, and linear potentials with angular terms.

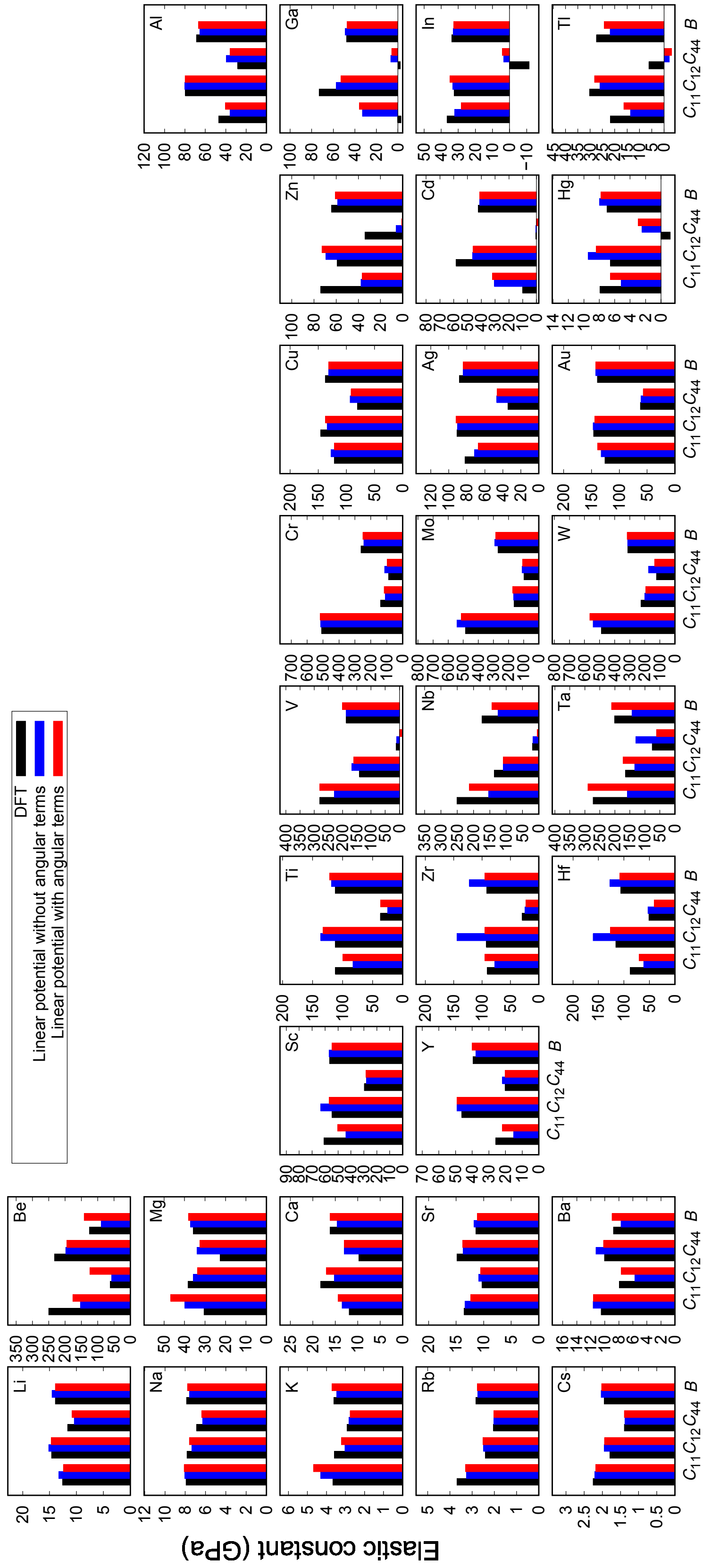


Figure 4.10: Elastic constants of BCC structures obtained from DFT calculations, linear potentials without angular terms, and linear potentials with angular terms.

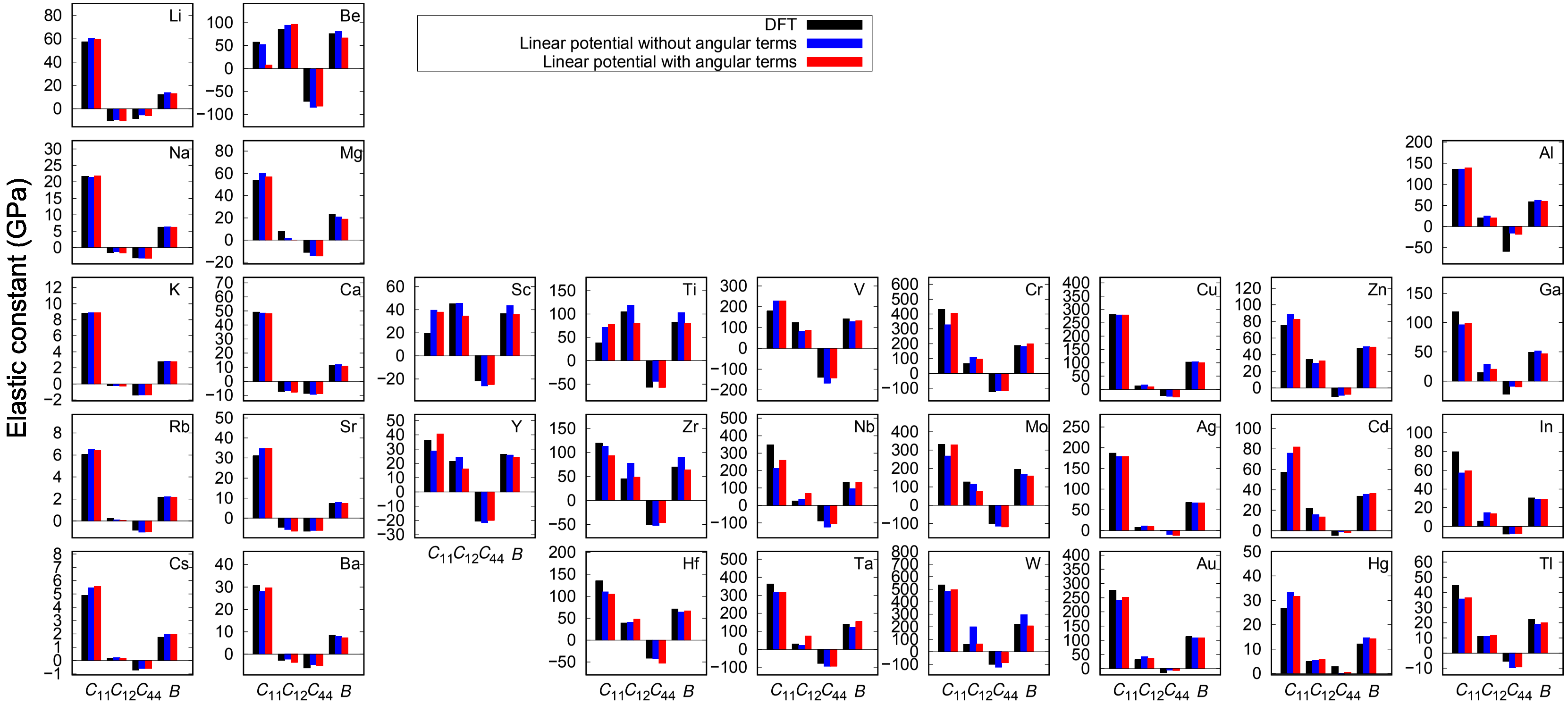


Figure 4.11: Elastic constants of SC structures obtained from DFT calculations, linear potentials without angular terms, and linear potentials with angular terms.

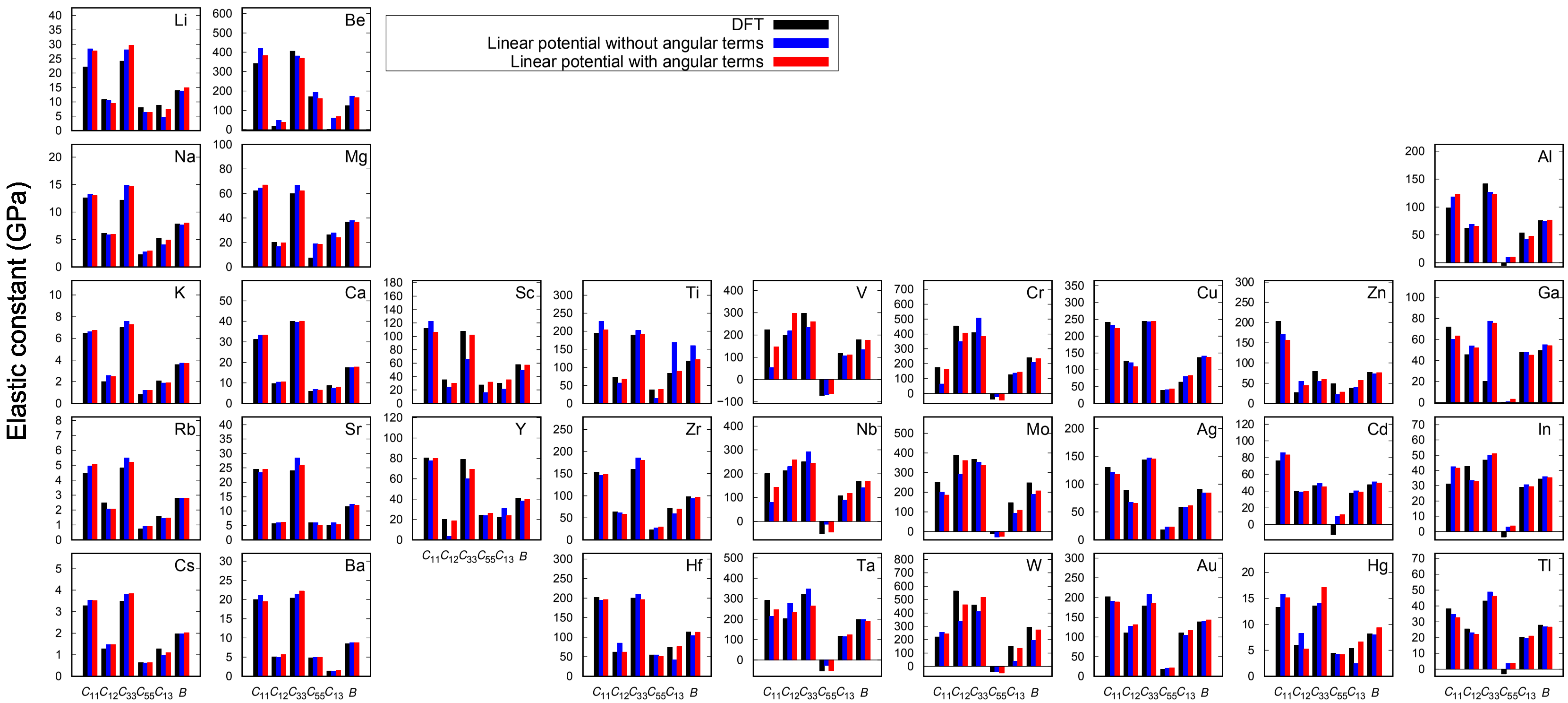


Figure 4.12: Elastic constants of HCP structures obtained from DFT calculations, linear potentials without angular terms, and linear potentials with angular terms.

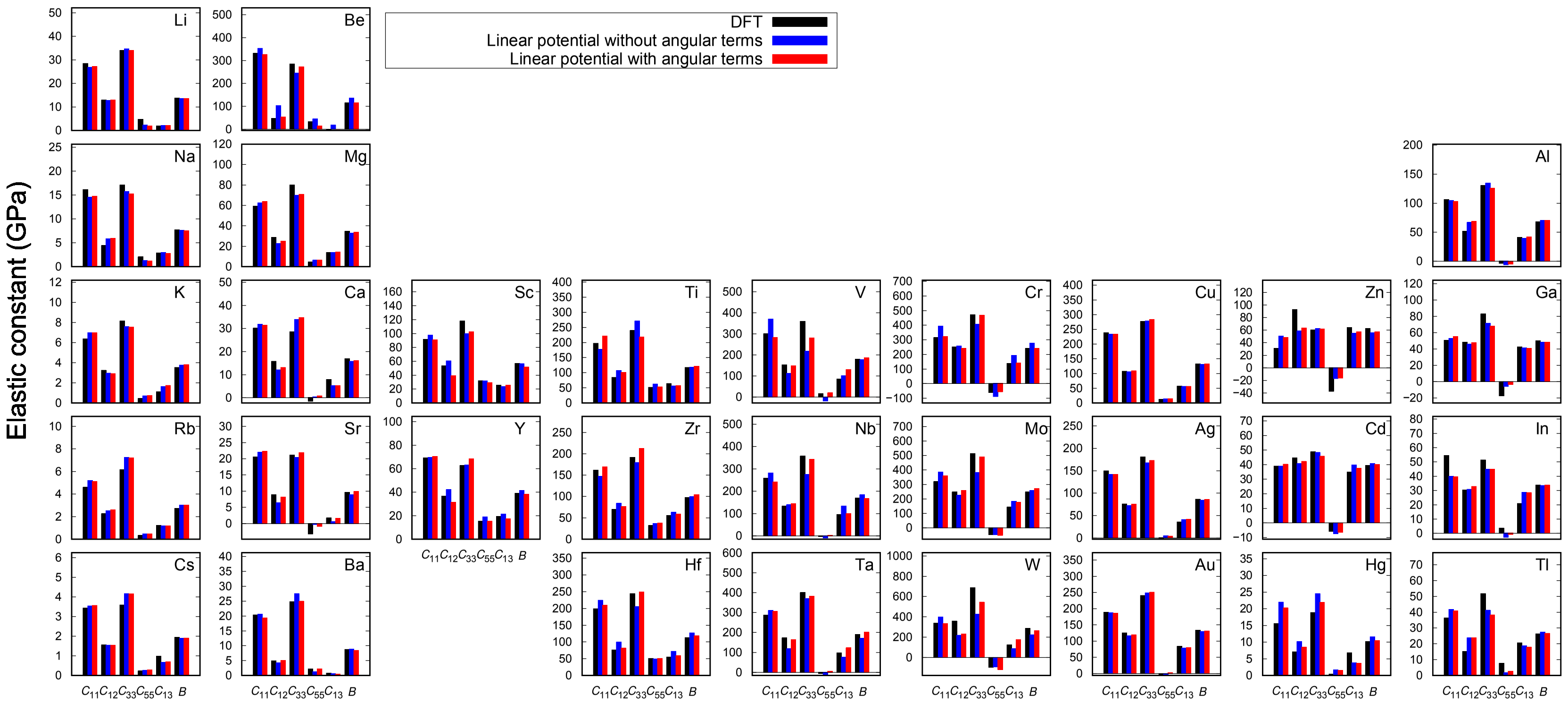


Figure 4.13: Elastic constants of ω structures obtained from DFT calculations, linear potentials without angular terms, and linear potentials with angular terms.

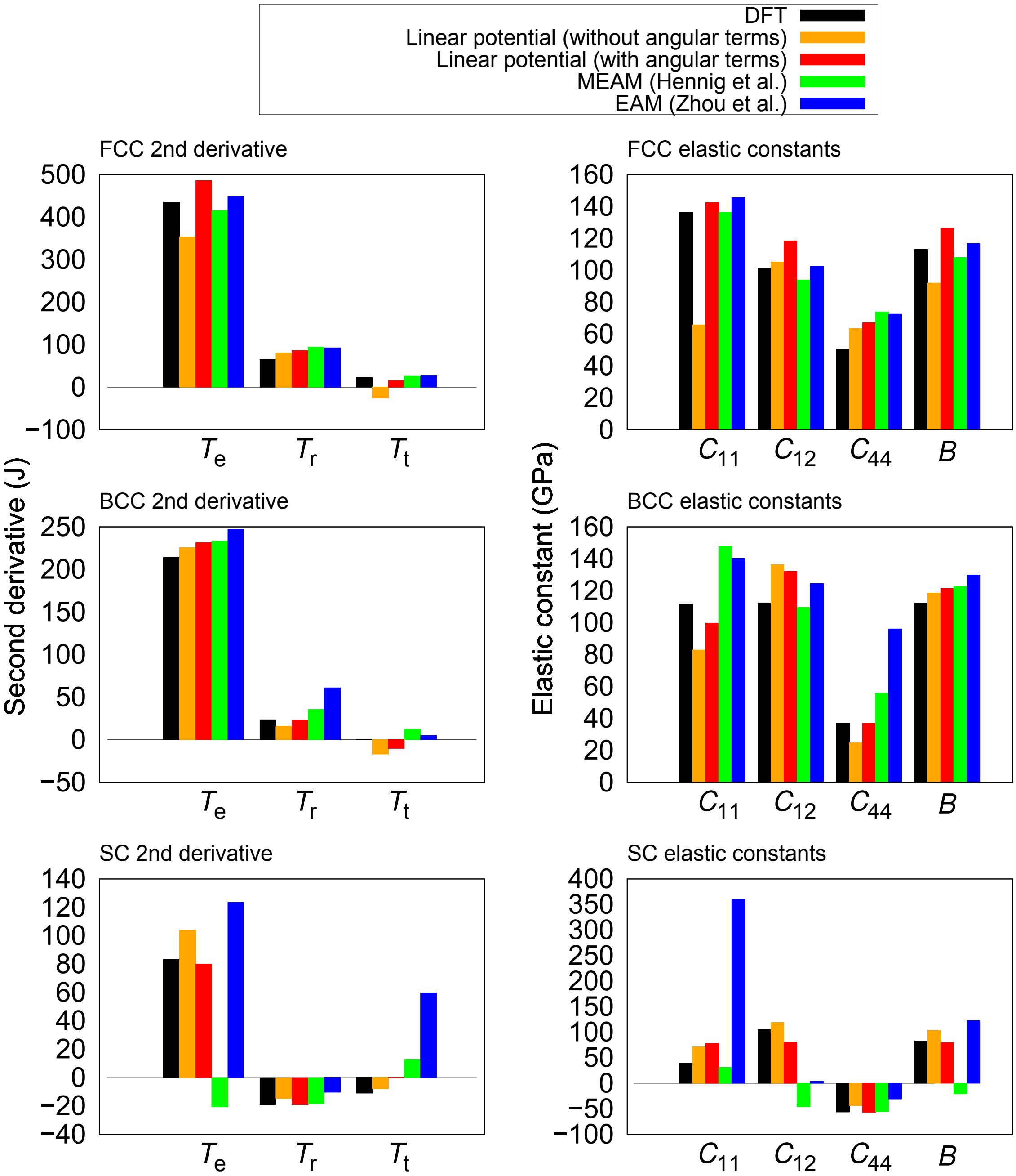


Figure 4.14: Second derivatives of energies respective to matrices expressed as Eqn.(4.18), (4.19), and (4.20) and elastic constants for Ti BCC, FCC, and SC crystals obtained from DFT calculations, linear potentials without angular terms, linear potentials with angular terms, Zhou EAM [55], and Hennig modified EAM [60].

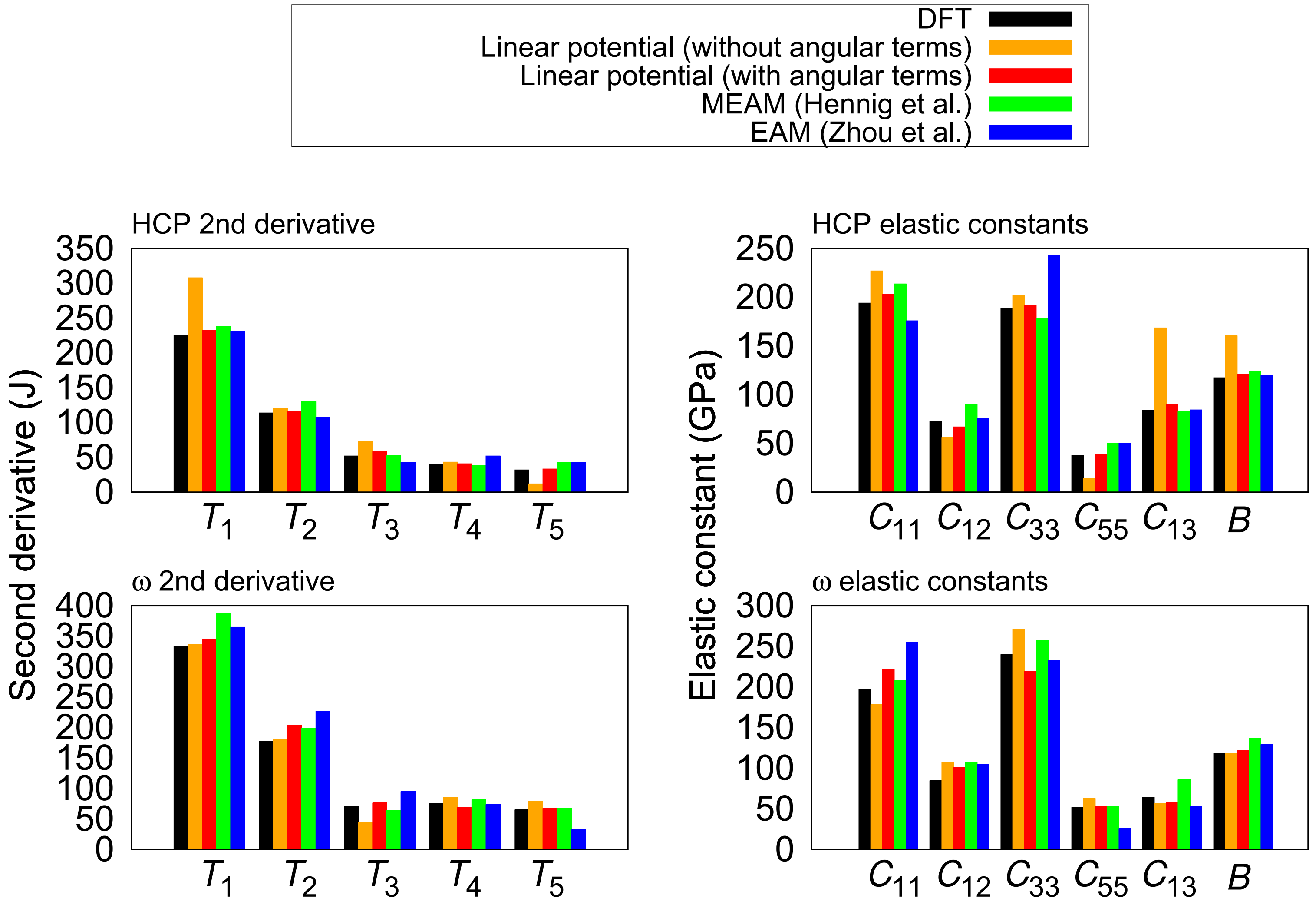


Figure 4.15: Second derivatives of energies respective to matrices expressed as Eqn.(4.24), (4.25), (4.26), (4.27), and (4.28) and elastic constants for Ti HCP and ω crystals obtained from DFT calculations, linear potentials without angular terms, linear potentials with angular terms, Zhou EAM [55], and Hennig modified EAM [60].

4.5.2 Phonon dispersion curves

For MD simulation, prediction of acting force on atoms and estimation of stability of structures play very important roles. Then, I examined the applicability of the interatomic potentials by comparing phonon dispersion curves computed with DFT calculation and interatomic potentials. The phonon dispersion curves were calculated with the supercell approach [57] for FCC, HCP, BCC, simple cubic, ω, and β-Sn structures with the equilibrium lattice constant. For all methods to calculate energies, lattice constants when $\gamma = 0$ are same as those obtained from DFT structure optimization. To evaluate the dynamical matrix, each symmetrically independent atomic position was displaced by 0.01 Å. The forces acting on atoms by the interatomic potential can then be analytically computed. Supercells were made by $4 \times 4 \times 4$ expansion of the conventional unit cells for all 6 structures. The phonon calculations were performed using the PHONOPY code [58].

Figure 4.16, 4.17, 4.18, 4.19, 4.20, and 4.21 show the dispersion curves with FCC, HCP, BCC, simple cubic, ω, and β-tin structures. For the most of typical metals, the both linear potentials with angular terms and linear potentials without angular terms reproduce phonon dispersion curves of DFT well. For some typical metals such as Be and transition metals, the prediction by linear potentials without angular terms is relatively poor and angular terms improve the prediction of phonon dispersion curves. There are some discrepancies between the phonon dispersion curves obtained by linear interatomic potentials even with angular terms and those obtained by DFT calculations. As discussed in Chapter 2, independent $2l+1$ descriptors are needed for each l to describe any interatomic potentials and the discrepancies can be ascribed to the absent of those descriptors.

Figure 4.22 shows the phonon dispersion curves obtained from DFT, linear potentials, and conventional EAM and modified EAM for Ti. While discrepancies between DFT and even modified EAM are found, linear potential with angular terms predicts phonon dispersion curve relatively well.

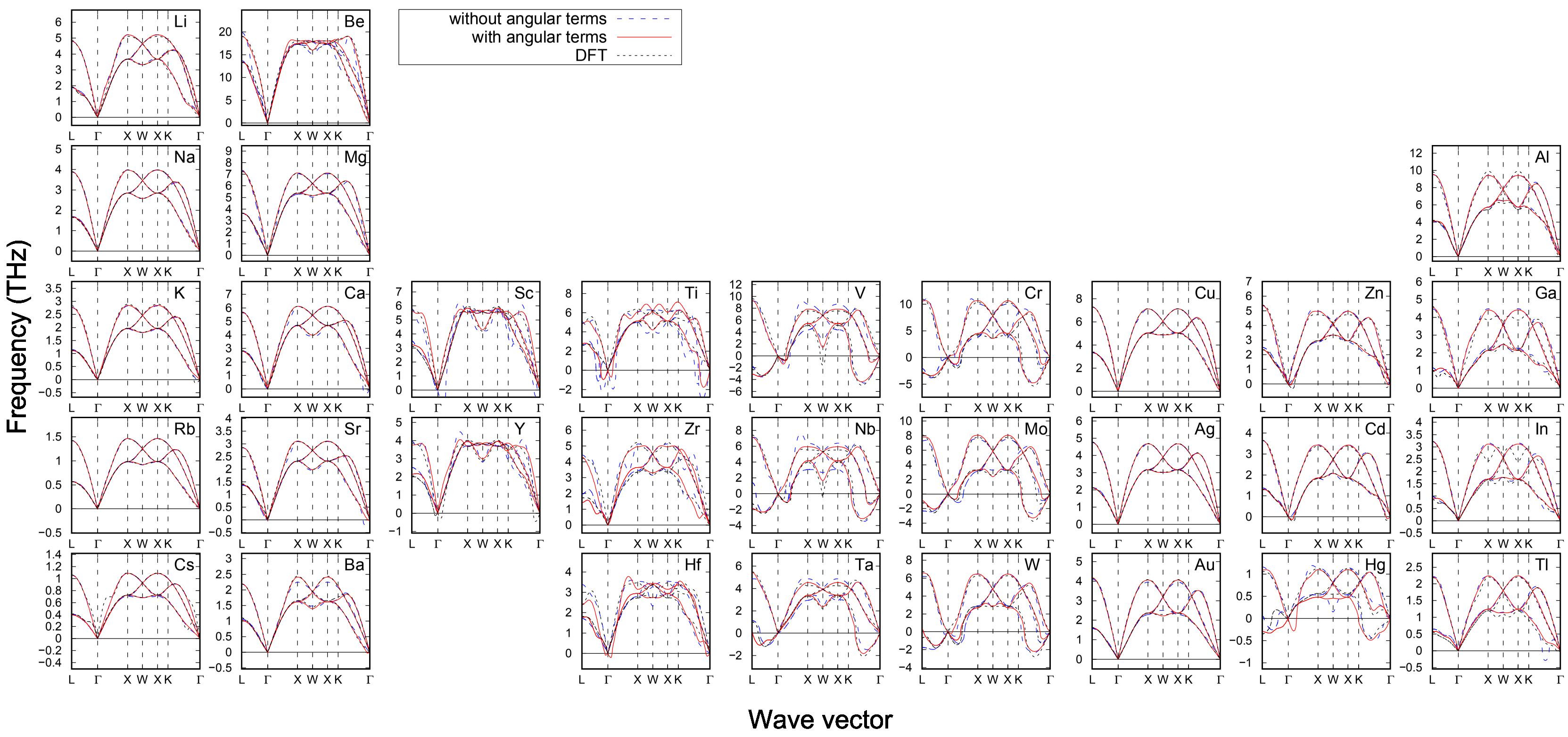


Figure 4.16: Phonon dispersion curves for 31 elemental metals with FCC structures obtained by linear potentials with angular terms, linear potentials with angular terms, and DFT calculations. Negative values indicate imaginary modes.

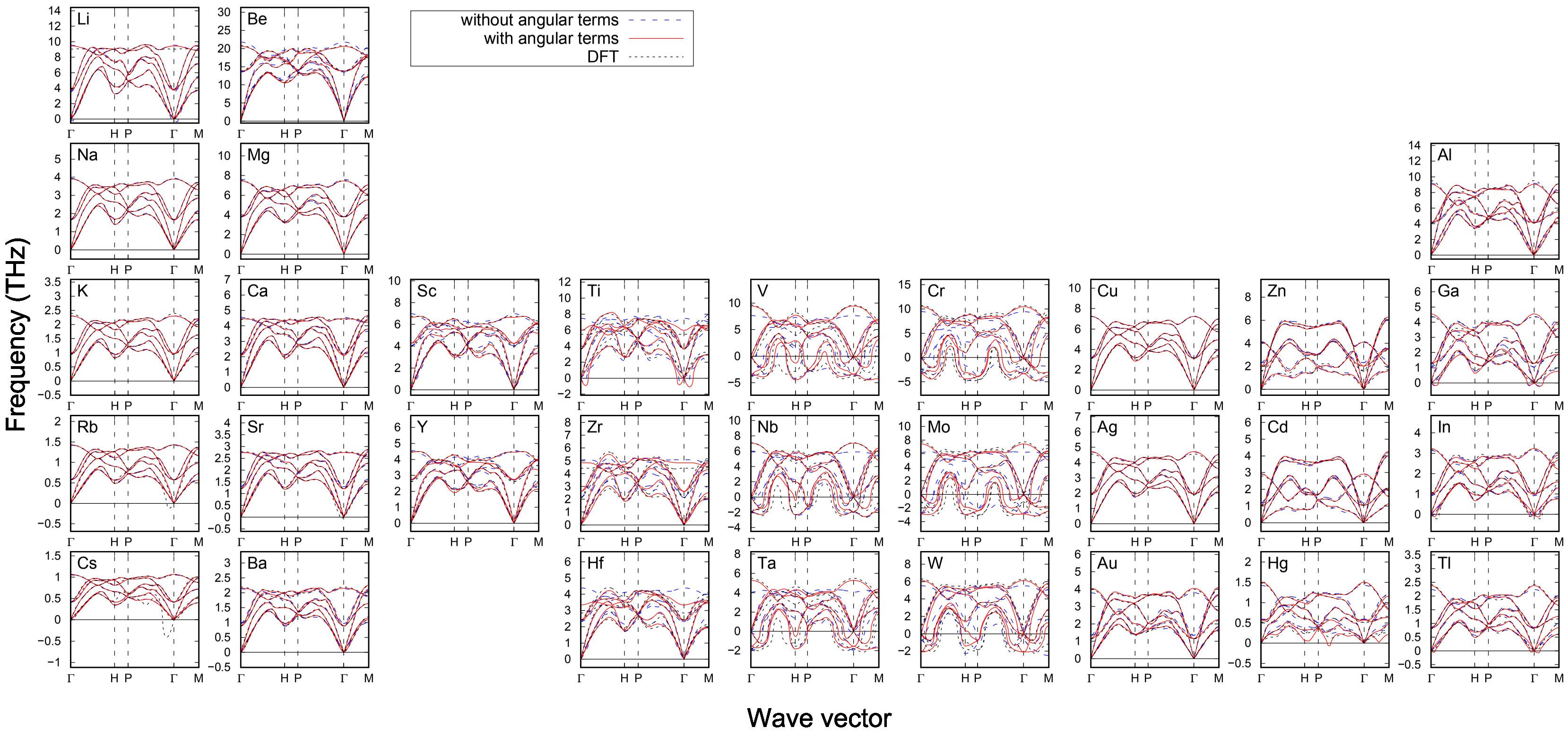


Figure 4.17: Phonon dispersion curves for 31 elemental metals with HCP structures obtained by linear potentials with angular terms, linear potentials with angular terms, and DFT calculations. Negative values indicate imaginary modes.

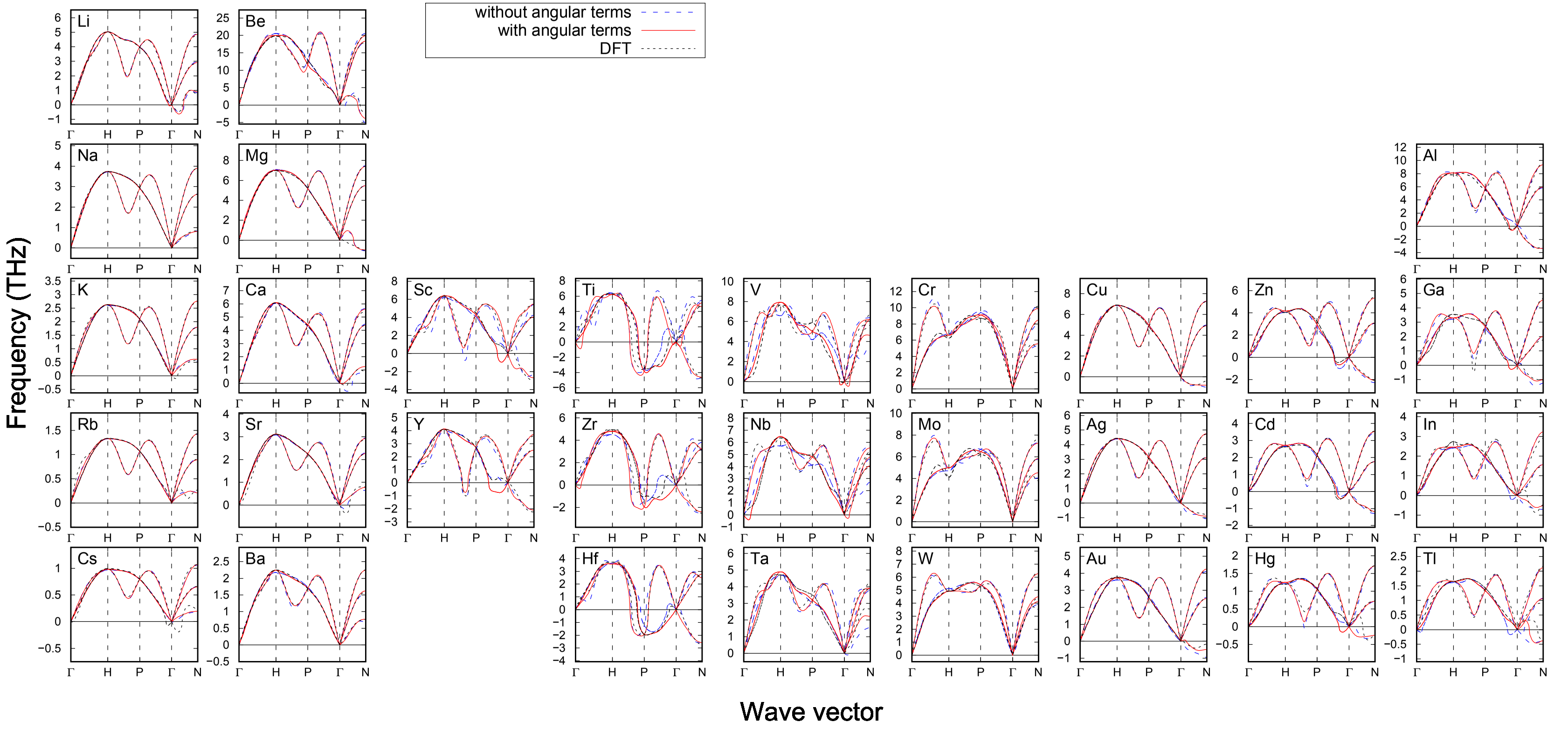


Figure 4.18: Phonon dispersion curves for 31 elemental metals with BCC structures obtained by linear potentials with angular terms, linear potentials with angular terms, and DFT calculations. Negative values indicate imaginary modes.

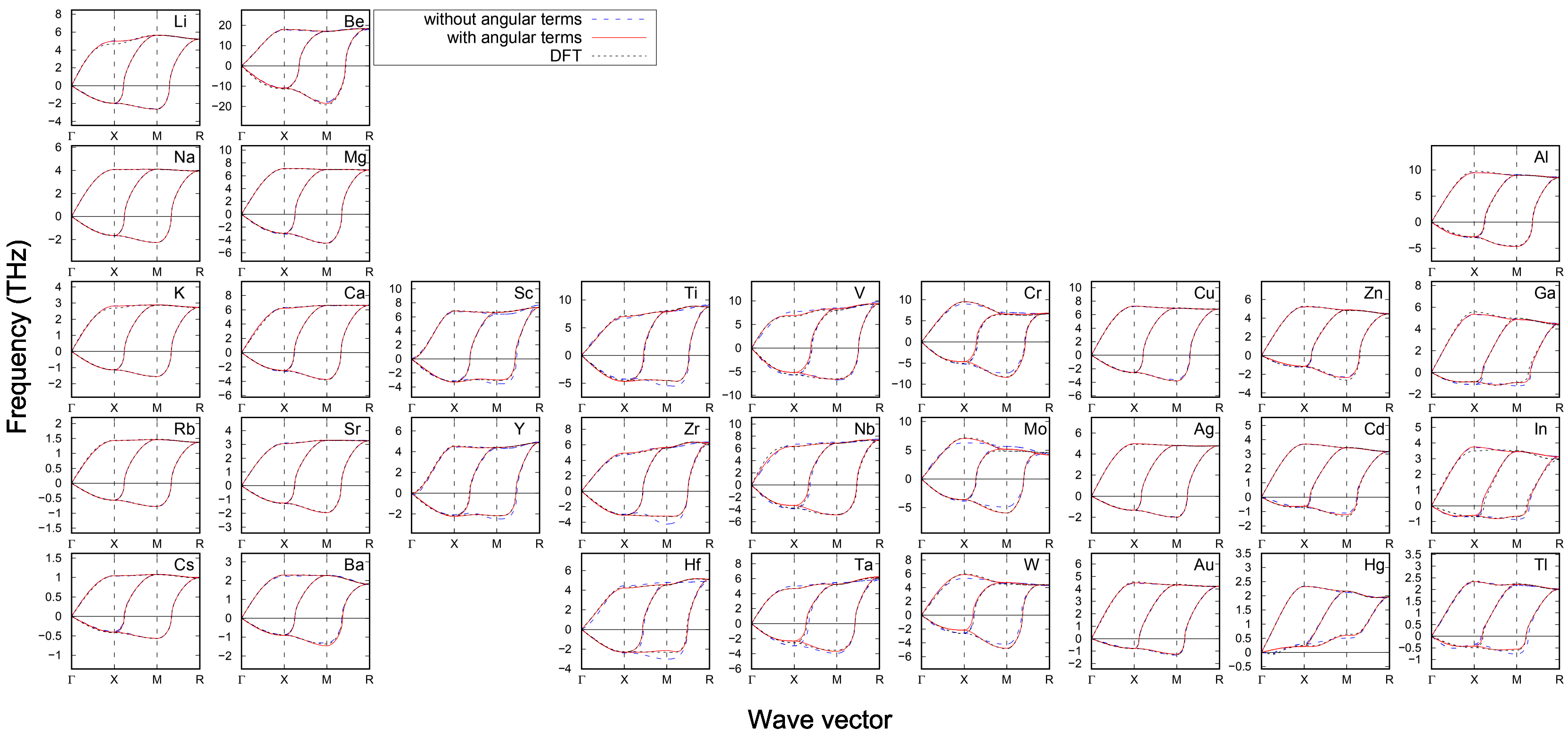


Figure 4.19: Phonon dispersion curves for 31 elemental metals with simple cubic structures obtained by linear potentials with angular terms, linear potentials with angular terms, and DFT calculations. Negative values indicate imaginary modes.

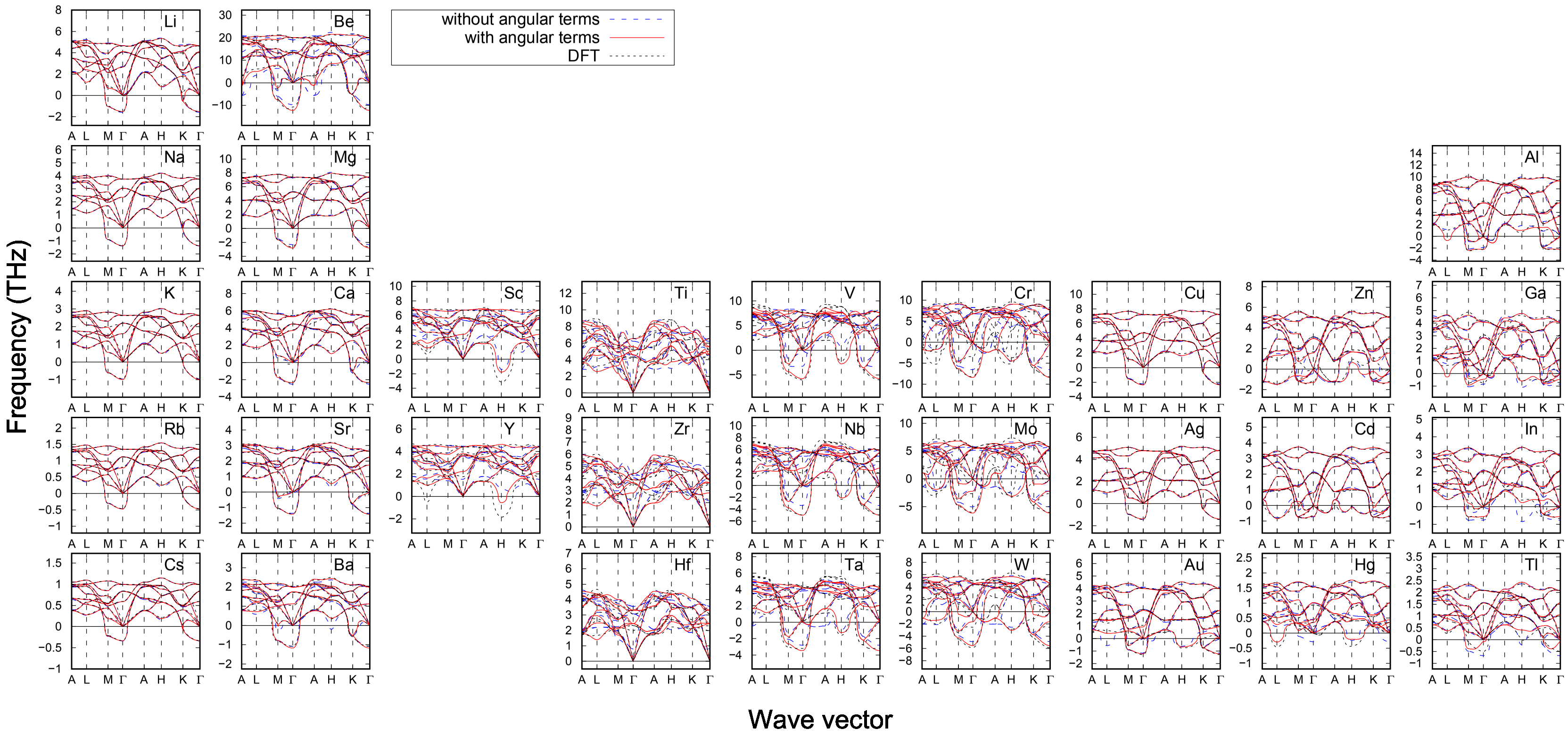


Figure 4.20: Phonon dispersion curves for 31 elemental metals with ω structures obtained by linear potentials with angular terms, linear potentials with angular terms, and DFT calculations. Negative values indicate imaginary modes.

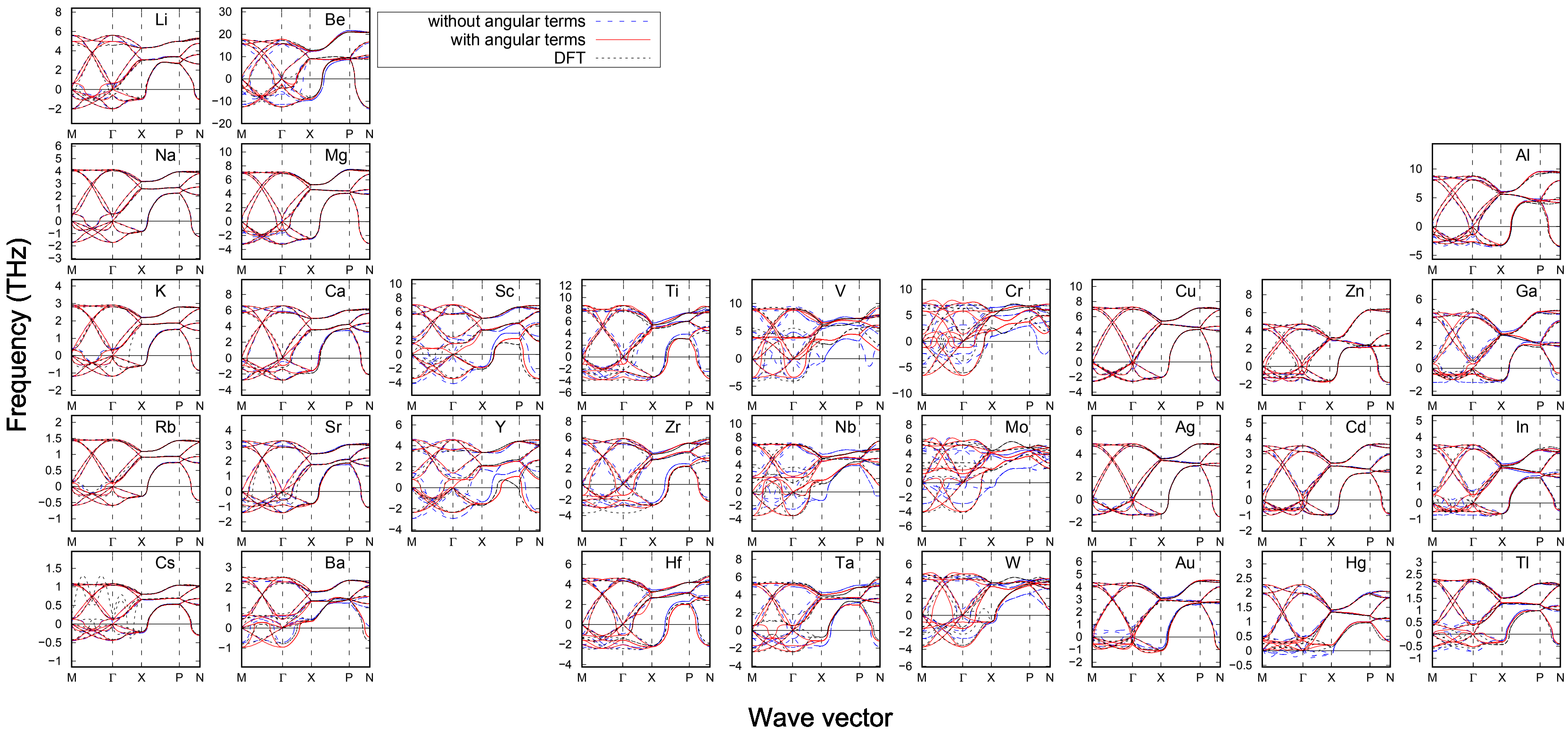


Figure 4.21: Phonon dispersion curves for 31 elemental metals with β-Sn structures obtained by linear potentials with angular terms, linear potentials with angular terms, and DFT calculations. Negative values indicate imaginary modes.

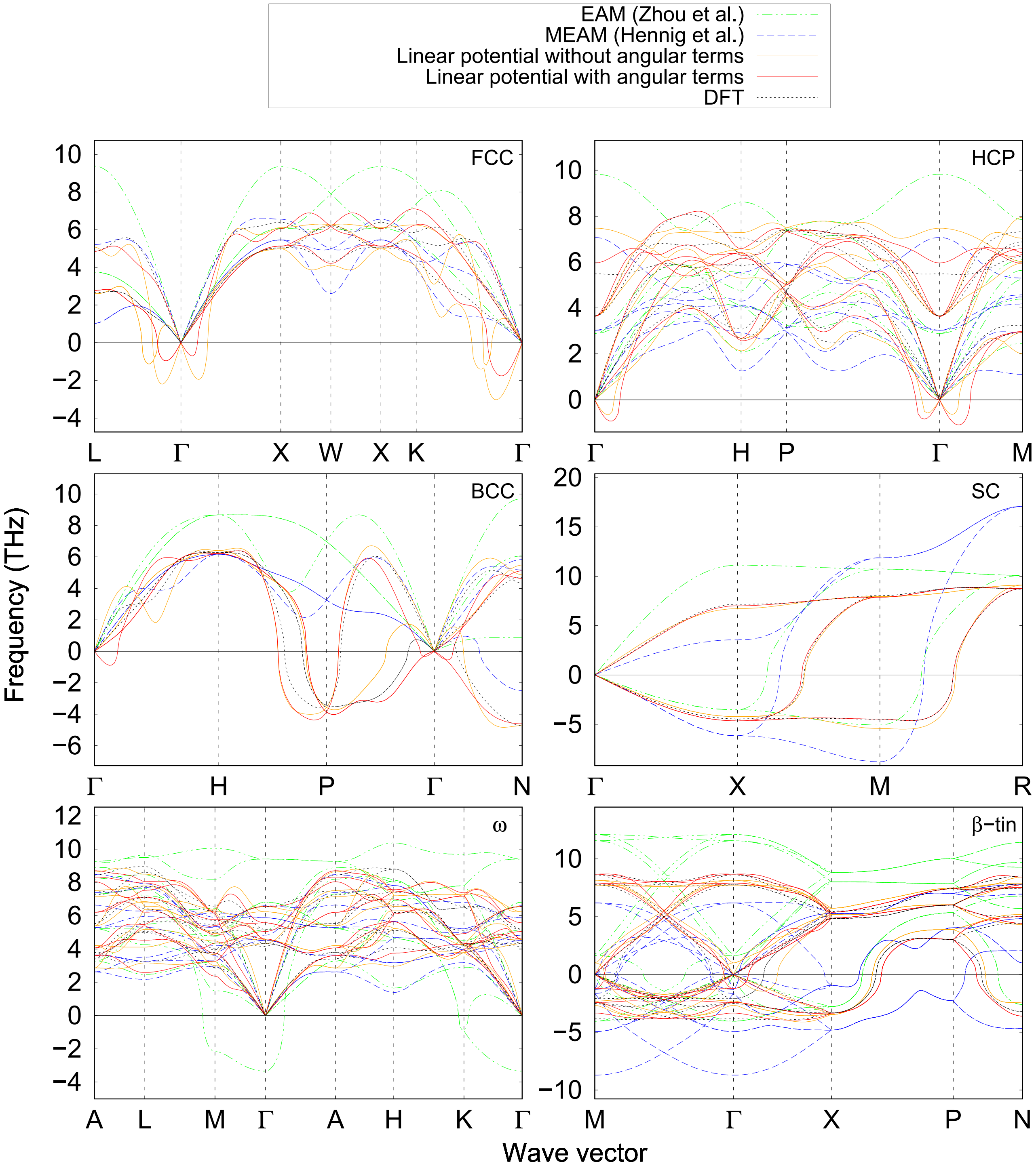


Figure 4.22: Phonon dispersion curves for Ti metals of DFT, linear potential without angular terms, linear potential with angular terms, Zhou EAM [55], and Hennig modified EAM [60].

4.6 Conclusion

In this chapter, I constructed linear potentials for 31 metals using AFS as angular descriptors and calculated prediction errors, elastic constants, and phonon dispersion curves.

Firstly, I constructed linear potentials with angular descriptors for 31 elements and compared accuracy of linear potentials with angular descriptors and that without angular descriptors. For all 31 elements, RMS errors of energies and forces decreased.

Next, I evaluated elastic constants of 31 elements. For the most of transition metals, prediction of elastic constants improved by adopting angular terms. On the contrary, for some elements in upper right of periodical table, there were still large discrepancy between elastic constants obtained by DFT and those obtained by linear potentials with angular terms. For those metals, more training data and/or other descriptors may be needed to predict elastic constants well.

At last, phonon dispersion curves are calculated by linear potentials. For the most of elements, angular descriptors also improved prediction of phonon dispersion curves.

For Ti, these properties are also compared to conventional EAM and modified EAM. In comparison to conventional potentials, prediction of elastic constants was good especially for unstable structures like SC. Phonon dispersion curves were also reproduced by linear potentials better than conventional potentials.

4.7 Appendix: fast implementation of angular terms

As discussed above, AFS can be calculated by both spherical harmonics and bonding angle. However, those two methods has practical problem to implement. If AFS is calculated by spherical harmonics, it is difficult to calculate force acting on atoms because spherical harmonics function is calculated by spherical coordination and the differential of that have singular point on z-axis. Calculation by bonding angle like Eqn.(4.6) has no such singular point, however, that takes much computational costs due to double notation of summation. Therefore, it is desirable to represent Eqn.(4.6) by Cartesian coordination and summation of single index. This can be achieved by formula deformation such as

$$\begin{aligned} \mathrm{AFS}^{(i)}_{l,n,m} = \sum_{j,k} f_n(r_{ij}) f_m(r_{ik}) (\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{ik})^l = \sum_{j,k} f_n(r_{ij}) f_m(r_{jk}) \left(\frac{x_{ij}x_{ik} + y_{ij}y_{ik} + z_{ij}z_{ik}}{r_{ij}r_{ik}} \right)^l = \\ \sum_{a+b+c=l} \frac{l!}{a!b!c!} \sum_{j,k} f_n(r_{ij}) f_m(r_{ik}) \frac{x^a_{ij}y^b_{ij}z^c_{ij}x^a_{ik}y^b_{ik}z^c_{ik}}{r^l_{ij}r^l_{ik}} = \\ \sum_{a+b+c=l} \frac{l!}{a!b!c!} \left(\sum_{j} f_n(r_{ij}) \frac{x^a_{ij}y^b_{ij}z^c_{ij}}{r^l_{ij}} \right) \left(\sum_{j} f_m(r_{ij}) \frac{x^a_{ij}y^b_{ij}z^c_{ij}}{r^l_{ij}} \right) \\ . \end{aligned} \tag{4.34}$$

Chapter 5

General conclusion

In this study, firstly I defined what interatomic potential is and derived the general form of interatomic potentials. Interatomic potentials of linear combination of descriptors by polynomial approximation. I constructed interatomic potential for 31 elements by such models and estimated accuracy of interatomic potentials.

In Chapter 2, firstly I defined interatomic potential as mapping from atomic distribution function to energy and derived the generalized formula of interatomic potential. I also showed the relationship between the generalized formula, conventional interatomic potentials and machine learning interatomic potential (MLIP). It was also revealed that the model of interatomic potential expressed by linear combination of descriptors [24, 31] can be generalized by polynomial approximation of the generalized formula.

In Chapter 3, I constructed by both previous and generalized linear interatomic potential for 31 elemental metals using only radial descriptor. For all elements, compared with the previous model, generalized model constructed more accurate interatomic potentials with less basis functions. Next, I predicted several properties of elements, namely, energies of structures which are not used for construction of interatomic potentials, elastic constants of cubic crystals, and phonon dispersion curves. For typical elements, interatomic potentials constructed in this study reproduced energies, bulk modulus, and phonon dispersion curves obtained from density functional theory (DFT) calculations accurately. The predictions of elastic constants excepting bulk modulus were relatively poor. It may be ascribed to training data set. For transition metals, the predictions of elastic constants and phonon dispersion curves were relatively worse than that for typical elements. For Al and Cu, these properties are also compared to conventional EAMs. The predictions of this work were more accurate than conventional EAMs, especially for unstable structures.

In Chapter 4, linear potentials are constructed including angular Fourier series (AFS) for 31 elements. For all elements, AFS decreased root mean square (RMS) errors of both energies and forces acting on atoms. For also transition metals, interatomic potentials with AFS reproduced bulk modulus, and phonon dispersion curves obtained from DFT calculations accurately. For Ti, these properties are also compared to conventional an EAM and a modified EAM. The predictions of linear were more accurate than conventional potentials.

The interatomic potentials constructed in this study reproduced properties of 31 elements obtained by DFT better than conventional potentials and previous linear interatomic potentials. This fact implies that those potentials are useful to perform reliable atomic scale simulations like molecular dynamics calculation.

Bibliography

[1] V. Yamakov, D. Wolf, S. R. Phillpot, A. K. Mukherjee, and H. Gleiter. Nat. Mater., **1**, 45 (2002).

[2] V. Yamakov, D. Wolf, S. Phillpot, and H. Gleiter. Acta Mater., **50**, 5005 (2002).

[3] M. Muzyk, Z. Pakiela, and K. Kurzydlowski. Scr. Mater., **66**, 219 (2012).

[4] Z. Pei, L.-F. Zhu, M. Frik, S. Sandlbes, J. v. Pezold, H. W. Sheng, C. P. Race, S. Zaefferer, B. Svendsen, D. Raabe, and J. Neugebauer. New. J. Phys., **15**, 043020 (2013).

[5] S. Nosé. J. Chem. Phys., **81**, 511 (1984).

[6] N. Shuichi. Progress of Theoretical Physics Supplement, **103**, 1 (1991).

[7] J. A. McCammon, B. R. Gelin, and M. Karplus. Nature, **267**, 585 (1997).

[8] P. Hohenberg and W. Kohn. Phys. Rev., **136**, B864 (1964).

[9] W. Kohn and L. J. Sham. Phys. Rev., **140**, A1133 (1965).

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

[11] M. Finnis. *Interatomic Forces in Condensed Matter.* Oxford University Press, New York (2003).

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

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

[14] M. S. Daw, S. M. Foiles, and M. I. Baskes. Materials Science Reports, **9**, 251 (1993).

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

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

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

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

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

[20] S. Lorenz, A. Groß, and M. Scheffler. Chem. Phys. Lett., **395**, 210 (2004).

[21] T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren. J. Chem. Phys., **103**, 4129 (1995).

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

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

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

[25] H. Eshet, R. Z. Khaliullin, T. D. Kühne, J. Behler, and M. Parrinello. Phys. Rev. B, **81**, 184107 (2010).

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

[27] N. Artrith and J. Behler. Phys. Rev. B, **85**, 045439 (2012).

[28] G. C. Sosso, G. Miceli, S. Caravati, J. Behler, and M. Bernasconi. Phys. Rev. B, **85**, 174103 (2012).

[29] J. Behler. Phys. Chem. Chem. Phys., **13**, 17930 (2011).

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

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

[32] A. E. Carlsson. Beyond pair potentials in elemental transition metals and semiconductors. In *Solid State Physics*, volume 43, pages 1. Academic Press, Boston (1990).

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

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

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

[36] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. v. Lilienfeld. Phys. Rev. Lett., **108**, 058301 (2012).

[37] K. T. Schütt, H. Glawe, F. Brockherde, A. Sanna, K. R. Müller, and E. K. U. Gross. Phys. Rev. B, **89**, 205118 (2014).

[38] O. A. v. Lilienfeld, R. Ramakrishnan, M. Rupp, and A. Knoll. Int. J. Quantum Chem. (2015).

[39] F. Faber, A. Lindmaa, O. A. v. Lilienfeld, and R. Armiento. Int. J. Quantum Chem. (2015).

[40] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti. Phys. Rev. B, **28**, 784 (1983).

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

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

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

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

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

[46] C. A. Becker, F. Tavazza, Z. T. Trautt, and R. A. B. d. Macedo. Current Opinion in Solid State and Materials Science, **17**, 277 (2013). Frontiers in Methods for Materials Simulations.

[47] `http://www.ctcms.nist.gov/potentials`.

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

[49] `http://lammps.sandia.gov`.

[50] M. Mendelev, M. Kramer, C. Becker, and M. Asta. Philosophical Magazine, **88**, 1723 (2008).

[51] J. M. Winey, A. Kubota, and Y. M. Gupta. Modelling and Simulation in Materials Science and Engineering, **17**, 055004 (2009).

[52] J. M. Winey, A. Kubota, and Y. M. Gupta. Modelling and Simulation in Materials Science and Engineering, **18**, 029801 (2010).

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

[54] Y. Mishin, M. J. Mehl, D. A. Papaconstantopoulos, A. F. Voter, and J. D. Kress. Phys. Rev. B, **63**, 224106 (2001).

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

[56] L. Fast, J. M. Wills, B. Johansson, and O. Eriksson. Phys. Rev. B, **51**, 17431 (1995).

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

[58] A. Togo, F. Oba, and I. Tanaka. Phys. Rev. B, **78**, 134106 (2008).

[59] A. P. Thompson, S. J. Plimpton, and W. Mattson. J. Chem. Phys., **131**, (2009).

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

Acknowledgement

The author would like to express the deepest gratitude to Professor Isao Tanaka at Kyoto University for continuous support and constant encouragement throughout this work. The author wishes to express his gratitude to Associate Professor Atsuto Seko for his continuous helpful discussions and advice. The author would also like to express his gratitude to Assistant Professor Hiroyuki Hayashi for his productive advices. The author would like to express his deepest gratitude to Professor Akira Sakai and Hiroyuki Nakamura for their critical reading of this thesis.

The author sincerely thanks to Dr. Atsushi Togo for generous discussions and advice for first principle calculation. The author acknowledges his great appreciation to Dr. Yoyo Hinuma for helpful advice of supercomputer and calculation. The author is deeply grateful to Dr. Yuji Ikeda for helpful advice of generous discussions and calculation. The author would like to thank to Assistant Professor Yu Kumagai at Tokyo Institute of Technology for advice on the first principle calculations.

The author would like to show his appreciation to Mr. Akifumi Sumitani for his helpful discussions and encouragement in the laboratory life. The author wishes to make grateful acknowledgement to the member of Professor Tanaka's group for helpful discussions throughout this work.

Finally, the author would like to sincerely thank his family for their enduring support and encouragement.