Supplementary Information for "Efficient calculations of impurity diffusivity in metals by linearized multi-band embedded atom method potentials"

Sehyeok Park and Takuji Oda*

Department of Nuclear Engineering, Seoul National University

*oda@snu.ac.kr

<Contents>

S1. Additional details of the LMB-EAM construction

(1) Setup for a single-component system

In the LMB-EAM, as described in the main body with Eqs. 1-5, the potential energy of atom-i, E_i, is written as follows:

$$E_i = E_{i,2b} + E_{i,eam} + c_i = \frac{1}{2}\sum_{j\neq i}^{r_{ij}<r_{cut1}} \phi\left(r_{ij}\right) + \sum_{n=1}^{N_{EAM}} \left\{ f_n\left(\rho_{n,i}\right)\right\} + c_i$$
$$= \frac{1}{2}\sum_{j\neq i}^{r_{ij}<r_{cut1}} \sum_{m=0}^{N_{bs1}-1} a_m \cos\left(m\pi \frac{r_{ij}}{r_{cut1}}\right) + \sum_{n=1}^{N_{EAM}} \sum_{m=0}^{N_{bs2}-1} b_{n,m} \cos\left(m\pi\rho_{n,i}\right) + c_i \quad , \quad \text{Eq. S1-1}$$

$$\rho_{n,i} = \sum_{j=1}^{r_{ij}<r_{n,cut2}} g_n\left(r_{ij}\right), \quad \text{where} \quad g_n\left(r_{ij}\right) = \begin{cases} \dfrac{1}{\alpha\rho_{n,MAX}}\left(r_{n,cut2} - r_{ij}\right)^m & \text{for } r_{ij} \le r_{n,cut2} \\ 0 & \text{for } r_{ij} \ge r_{n,cut2} \end{cases} . \quad \text{Eq. S1-2}$$

The definition of each variable is explained in the main body.

For a given configuration, all structural parameters such as r_{ij} and $\rho_{n,i}$ can be calculated. Accordingly, the energy of each atom can be expressed as a linear combination of basis set coefficients

$\{a_m\}$, $\{b_{n,m}\}$ and c. To be specific, the energy of the i-th atom in the j-th configuration can be written down as

$$E_{i,j} = \sum_{m=0}^{N_{bs1}-1} d_{2b,i,j,m} a_m + \sum_{n=1}^{N_{EAM}} \sum_{m=0}^{N_{bs2}-1} d_{eam,i,j,n,m} b_{n,m} + c_i , \quad \text{Eq. S1-3}$$

where $\{d_{2b,i,j,m}\}$ and $\{d_{eam,i,j,n,m}\}$ are values that can be calculated for a given configuration by Eqs. S1-1 and S1-2.

As in Eq. 12 of the main body, the total model error, Y, is expressed as a weighted sum of the energy, force, and stress squared errors from the first-principles (FP) calculation results used as the training data with regularization terms as

$$\begin{aligned} Y\left(\{a_m\},\{b_{n,m}\},c\right) &= Y_{fit} + Y_{reg} \\ &= \sum_{j=1}^{N_{conf}} \left\{ w_e \left(E_{PM}^{j} - E_{FP}^{j}\right)^2 + \frac{w_f}{3N_{atom,j}} \sum_{k=1}^{N_{atom,j}} \sum_{l=1}^{3} \left(F_{PM}^{j,k,l} - F_{FP}^{j,k,l}\right)^2 + \frac{w_s}{6} \sum_{k=1}^{6} \left(S_{PM}^{j,k} - S_{FP}^{j,k}\right)^2 \right\}, \\ &\quad + \sum_{j=0}^{2} \lambda_{2b,j} \Delta_{2b,j} + \sum_{k=1}^{N_{EAM}} \sum_{j=0}^{2} \lambda_{eam,k,j} \Delta_{eam,k,j} \end{aligned} \quad \text{Eq. S1-4}$$

where Y_{fit} is composed of the first 3 terms about the fitting errors and Y_{reg} is composed of the last 2 terms about the regularization. Our goal is to minimize Y by optimizing the model parameters $\{a_m\}$, $\{b_{n,m}\}$, and c, which can be done by solving a system of linear equations that can be established as the partial derivatives of Y with respect to the model parameters, as explained in the main body.

In the following, we present the forms of the partial derivatives of Y more specifically and confirm that the partial derivatives can form a system of linear equations.

(2) The energy term

We first focus on the energy fitting error term in Y_{fit}:

$$\begin{aligned} w_e \sum_{j=1}^{N_{conf}} \left(E_j^{PM} - E_j^{FP}\right)^2 &= w_e \sum_{j=1}^{N_{conf}} \left(\left(\sum_{i=1}^{N_{atom,j}} E_{i,j} \right) - E_j^{FM} \right)^2 \\ &= w_e \sum_{j=1}^{N_{conf}} \left(\sum_{i=1}^{N_{atom,j}} \left(\sum_{m=0}^{N_{bs1}-1} d_{2b,i,j,m} a_m + \sum_{n=1}^{N_{EAM}} \sum_{m=0}^{N_{bs2}-1} d_{eam,i,j,n,m} b_{n,m} + c_i \right) - E_j^{FP} \right)^2 \end{aligned} . \quad \text{Eq. S1-5}$$

The partial derivatives of Eq. S1-5 are written as

$$\frac{\partial}{\partial a_m} w_e \sum_{j=1}^{N_{conf}} \left(E_j^{PM} - E_j^{FP}\right)^2$$
$$= 2w_e \sum_{j=1}^{N_{conf}} \left(\sum_{i=1}^{N_{atom,j}} \left(\sum_{m=0}^{N_{bs1}-1} d_{2b,i,j,m} a_m + \sum_{n=1}^{N_{EAM}} \sum_{m=0}^{N_{bs2}-1} d_{eam,i,j,n,m} b_{n,m} + c \right) - E_j^{FP} \right) \times \sum_{j=1}^{N_{conf}} \sum_{i=1}^{N_{atom,j}} \sum_{m=0}^{N_{bs1}-1} d_{2b,i,j,m} = 0$$
$$\frac{\partial}{\partial b_{n,m}} w_e \sum_{j=1}^{N_{conf}} \left(E_j^{PM} - E_j^{FP}\right)^2$$
$$= 2w_e \sum_{j=1}^{N_{conf}} \left(\sum_{i=1}^{N_{atom,j}} \left(\sum_{m=0}^{N_{bs1}-1} d_{2b,i,j,m} a_m + \sum_{n=1}^{N_{EAM}} \sum_{m=0}^{N_{bs2}-1} d_{eam,i,j,n,m} b_{n,m} + c \right) - E_j^{FP} \right) \times \sum_{j=1}^{N_{conf}} \sum_{i=1}^{N_{atom,j}} \sum_{n=1}^{N_{EAM}} \sum_{m=0}^{N_{bs2}-1} d_{eam,i,j,n,m} = 0$$
$$\frac{\partial}{\partial c} w_e \sum_{j=1}^{N_{conf}} \left(E_j^{PM} - E_j^{FP}\right)^2$$
$$= 2w_e \sum_{j=1}^{N_{conf}} \left(\sum_{i=1}^{N_{atom,j}} \left(\sum_{m=0}^{N_{bs1}-1} d_{2b,i,j,m} a_m + \sum_{n=1}^{N_{EAM}} \sum_{m=0}^{N_{bs2}-1} d_{eam,i,j,n,m} b_{n,m} + c \right) - E_j^{FP} \right) \times \sum_{j=1}^{N_{conf}} \sum_{i=1}^{N_{atom,j}} 1 = 0$$

. Eq. S1-6

Each equation in Eq. S1-6 is linear with respect to $\{a_m\}$, $\{b_{n,m}\}$, and c.

The key point here is that, since the atom energy is described as a linear combination of $\{a_m\}$, $\{b_{n,m}\}$, and c, as shown in Eq. S1-3, the system energy is also described as a linear combination. Accordingly, as shown in Eq. S1-6, the energy fitting error term included in the problem to minimize Y also only contain linear terms.

(3) Regularization terms

As described with Eqs. 6-11 of the main body, the regularization terms for the two-body and EAM functions are expressed as

$$\Delta_{2b,k} \equiv \int_0^{r_{cut}} \left(\phi^{(k)}\right)^2 dr = \frac{r_{cut1}}{2} \sum_{m=1}^{N_{bs1}-1} \left(\frac{m\pi}{r_{cut1}}\right)^{2k} {a_m}^2 , \quad \text{Eq. S1-7}$$

$$\Delta_{eam,n,k} \equiv \int_0^{r_{cut}} \left(f_n^{(k)}(\rho)\right)^2 d\rho = \frac{1}{2} \sum_{m=1}^{N_{bs2}-1} (m\pi)^{2k} b_{n,m}^2 . \quad \text{Eq. S1-8}$$

In this study, we consider up to the 2nd order regularization for both two-body and EAM functions. Thus, the regularization terms in Y, denoted as Y_{reg}, are expressed as

$$Y_{reg} = \sum_{j=0}^{2} \lambda_{2b,j} \Delta_{2b,j} + \sum_{k=1}^{N_{EAM}} \sum_{j=0}^{2} \lambda_{eam,k,j} \Delta_{eam,k,j} . \quad \text{Eq. S1-9}$$

Using Eqs. S1-7 and S1-8, the derivatives of Y_{reg} with respect to $\{a_m\}$, $\{b_{n,m}\}$, and c are expressed as

$$\frac{\partial}{\partial a_m} Y_{reg} = r_{cut1} \sum_{m=1}^{N_{bs1}-1} \left(\sum_{k=0}^{2} \left(\frac{m\pi}{r_{cut1}}\right)^{2k} a_m \right)$$
$$\frac{\partial}{\partial b_{n,m}} Y_{reg} = \sum_{m=1}^{N_{bs2}-1} \left(\sum_{k=0}^{2} (m\pi)^{2k} b_{n,m} \right) \quad \text{Eq. S1-10}$$
$$\frac{\partial}{\partial c} Y_{reg} = 0$$

Since each equation of Eq. S1-10 is linear with respect to $\{a_m\}$, $\{b_{n,m}\}$, and c, the inclusion of the regularization terms still preserves the linearity of the problem to minimize Y.

(4) Energy and regularization terms in the multi-component systems

As an example of a multi-component system, we consider a system composed of two elements, A and B here. Assuming that atom i is of element A, the energy of the atom i in the LMB-EAM is written as

$$E_i = E_{2B,i} + E_{EAM,i} = \frac{1}{2}\left\{ \sum_{j^A \neq i^A}^{<r_{cut1-AA}} \phi_{AA}\left(r_{i^A j^A}\right) + \sum_{j^B}^{<r_{cut1-AB}} \phi_{AB}\left(r_{i^A j^B}\right) \right\} + \sum_{n=1}^{N_{EAM}} f_{n,A}\left(\rho_{n,i^A}\right)$$
$$\rho_{n,i^A} = \sum_{j^A \neq i^A}^{r_{ij}<r_{n,cut2-AA}} g_{n,AA}\left(r_{i^A j^A}\right) + \sum_{j^B}^{r_{ij}<r_{n,cut2-AB}} g_{n,AB}\left(r_{i^A j^B}\right) \quad , \quad \text{Eq. S1-11}$$

where the superscript A or B is added to each atom index i and j to clarify the element type of each atom. For the cutoff distances r_{cut1} and r_{cut2}, we can set different values for different combinations as $r_{cut1-AA}$, $r_{cut1-AB}$, $r_{cut2-AA}$, and $r_{cut2-AB}$. As in the single-component case, the ϕ and f functions are expanded by truncated trigonometric series and the g functions are pre-determined. Accordingly, even when we apply the LMB-EAM formalism to a multi-component system, the problem of minimizing Y can still be established as a linear problem.

Meanwhile, as seen in Eqs. S1-7 and S1-8, the regularization term is built for each ϕ and f function. Therefore, for the regularization term, whether the system is single-component or multi-component does not affect the linearity.

(5) The Force term

For the force term, as a test case, we use a system composed of two elements (A and B). For a given configuration, the force term acting on atom k, which is of element A, in the x direction due to the two-body and EAM functions is expressed as

$$-\frac{\partial E}{\partial x_{k^A}} = -\frac{\partial E_{2b}}{\partial x_{k^A}} - \frac{\partial E_{eam}}{\partial x_{k^A}}$$
$$-\frac{\partial E_{2b}}{\partial x_{k^A}} = \sum_{j^A \neq k^A}^{<r_{cut1-AA}} \phi'_{AA}\left(r_{k^A j^A}\right)\frac{\Delta x_{k^A j^A}}{r_{k^A j^A}} + \frac{1}{2}\sum_{j^B}^{<r_{cut1-AB}} \left[\phi'_{AB}\left(r_{k^A j^B}\right) + \phi'_{BA}\left(r_{k^A j^B}\right)\right]\frac{\Delta x_{k^A j^B}}{r_{k^A j^B}}$$
$$-\frac{\partial E_{eam}}{\partial x_{k^A}} = \sum_{j^A \neq k^A}\left(\left[f'_A\left(\rho_{k^A}\right) + f'_A\left(\rho_{j^A}\right)\right] g'_{AA}\left(r_{k^A j^A}\right)\frac{\Delta x_{k^A j^A}}{r_{k^A j^A}} \right) \quad , \quad \text{Eq. S1-12}$$
$$+\sum_{j^B}\left(f'_A\left(\rho_{k^A}\right) g'_{AB}\left(r_{k^A j^B}\right)\frac{\Delta x_{k^A j^B}}{r_{k^A j^B}} \right) + \sum_{j^B}\left\{ f'_B\left(\rho_{j^B}\right) g'_{BA}\left(r_{k^A j^B}\right)\frac{\Delta x_{k^A j^B}}{r_{k^A j^B}} \right\}$$

where $\Delta x_{k^A j^A} = x_{j^A} - x_{k^A}$, etc. The derivatives of ϕ and f functions are expressed as

$$\phi'\left(r_{ij}\right)=\sum_{m=0}^{N_{bs1}-1}\left(-\frac{m\pi a_m}{r_{cut1}}\right)\sin\left(m\pi\frac{r_{ij}}{r_{cut1}}\right),\qquad \text{Eq. S1-13}$$

$$f_n'\left(\rho_{n,i}\right)=\sum_{m=0}^{N_{bs2}-1}\left(-b_{n,m}m\pi\right)\sin\left(m\pi\rho_{n,i}\right).\qquad \text{Eq. S1-14}$$

Since these functions are also linear with respect to the model parameters, the inclusion of the force terms still preserves the linearity of the problem to minimize Y.

(6) The stress term

The force expression can be simplified as

$$-\frac{\partial E_{2b}}{\partial x_i}=\frac{1}{2}\sum_{j\neq i}^{<r_c}\left[f_{ij}'\left(r_{ij}\right)+f_{ji}'\left(r_{ji}\right)\right]\frac{x_j-x_i}{r_{ij}}=\sum_{j\neq i}^{<r_c}f_{ij}'\left(r_{ij}\right)\frac{x_j-x_i}{r_{ij}},\qquad \text{Eq. S1-15}$$

and

$$-\frac{\partial E_{eam}}{\partial x_i}=\sum_{j\neq i}^{<r_c}\left(\left[f_i'\left(\rho_i\right)g_{ij}'\left(r_{ij}\right)+f_j'\left(\rho_j\right)g_{ji}'\right]\frac{x_j-x_i}{r_{ij}}\right),\qquad \text{Eq. S1-16}$$

where g_{ji} is the effective density from atom j at atom i position. Using these simplified force expressions, we can express the atomic stress tensor of atom i due to the two-body function ($s_{2b,i}^{ab}$) and the EAM function ($s_{eam,i}^{ab}$) for the ab component, where a and b are x, y or z, as follows:

$$s_{2b,i}^{ab}=\sum_{j\neq i}^{<r_c}\left\{\frac{1}{2}\left(r_j^a-r_i^a\right)\left(\left[f_{ij}'\left(r_{ij}\right)\right]\frac{r_j^b-r_i^b}{r_{ij}}\right)\right\},\qquad \text{Eq. S1-17}$$

$$s_{eam,i}^{ab}=\sum_{j\neq i}^{<r_c}\left\{\frac{1}{2}\left(r_j^a-r_i^a\right)\left(\left[f_i'\left(\rho_i\right)g_{ij}'\left(r_{ij}\right)+f_j'\left(\rho_j\right)g_{ji}'\right]\frac{r_j^b-r_i^b}{r_{ij}}\right)\right\}.\qquad \text{Eq. S1-18}$$

As confirmed in (5), the derivatives of the ϕ and f functions are linear with respect to the model parameters. The system-wise stress tensor is obtained as the sum of the atomic stress tensors, and thus the system stress tensor can also be expressed as linear functions. Therefore, adding the stress terms into the fitting target still maintains the problem to be linear with respect to $\{a_m\}$, $\{b_{n,m}\}$, and c.

S2. Details of the training sets used for the LMB-EAMs

The training sets used to construct the LMB-EAMs for H diffusion in bcc-W and O diffusion in liquid Na in this study were taken from previous studies[1,2], where MTPs were generated. Although the details of the training set can be found in these previous studies, for the convenience of the readers, the contents of

the training sets are briefly summarized in Table S1. In total, the training sets contain 936 configurations for the W-H case and 360 configurations for the Na-O case.

Table S1. The contents of the training sets used to construct the LMB-EAMs. QHA stands for the quasi-harmonic approximation. For more details, please refer to Ref. 1 for the W-H case and Ref. 2 for the Na-O case.

Type	The number of atoms in a unit cell	The number of configurations
(a) For the W-H potential		
Lattice deformation by Mote Carlo methods	2 W atoms	122
Geometry optimization calculations for hydrogen at T-site/O-site/Tri-site.	54 W atoms; 1 H atom	180
NEB calculations for H migrations	54 W atoms; 1 H atom	144
QHA with H	54 W atoms; 1 H atom	45
QHA without H	54 H atoms	24
Active learning MD	54 W atoms; 1 H atom	421
(b) For the Na-O potential		
Liquid Na from FPMD	102 Na atoms	100
Liquid Na with an O atom from FPMD	101 Na atoms; 1 O atom	100
Iterative learning MD	101 Na atoms; 1 O atom	160

S3. Details of the convergence test conditions

For the convergence behavior of the fitting error, there are five hyperparameters that can affect the fitting error: (1) N_{bs1}, the number of cosine basis functions for each two-body potential function; (2) N_{bs2}, the number of cosine basis functions for each embedding energy function; (3) N_{EAM}, the number of the embedding energy functions; (4) λ, the regularization weight; (5) the shapes of each g function (cubic or quartic; cutoff distance). In the convergence tests described in the main body, the default value of each parameter was set as follows, and one parameter was changed for each convergence test: $N_{EAM} = 3$ with $r_{1,cut2} = 3.5$ Å , $r_{2,cut2} = 4.75$ Å and $r_{3,cut2} = 6.0$ Å ; $N_{bs1} = 80$; $N_{bs2} = 50$; $\lambda = 0.001$ for all regularization weights; cubic forms for g functions.

- ✓ N_{bs1}, N_{bs2} test conditions: 9 samples with N_{bs} = 10 ~ 100.
- ✓ N_{EAM} test conditions: N_{EAM} = 1 ~ 9 with cubic or quartic shape g function. The test was conducted by combining functions with various $r_{n,cut2}$ values depending on the number of N_{EAM}. $r_{n,cut2}$ values were 3.5, 3.8125, 4.125, 4.4375, 4.75, 5.0625, 5.375, 5.6875, and 6.0 Å. Table S2 shows the number of combinations of embedding energy function.
- ✓ λ test conditions: 11 samples with $\lambda = 1\times10^{-13}$ ~ 100.

The convergence test results are shown in Figure 1 of the main body.

Table S2. The number of combinations of embedding energy function.

N_{EAM}	1	2	3	4	5	6	7	8	9
# of combination	9	36	84	126	126	84	36	9	1

S4. Effects of the regularization parameters for the Na-O case

The root mean square error (RMSE) from the DFT reference data increases as the weight of the regularization term increases, as shown in Fig. 1(d) of the main body. After performing test calculations with various λ values, we decided to use 0.001 for all λ values because a smooth potential shape was realized with a reasonably small increase in the RMSE compared to the case without regularization. Figure S1 shows the effect of the regularization on the shape of the two-body potential function and the embedding energy functions for both the W-H and Na-O cases, complementing Figure 2 in the main text, which shows only the W-H case. As can be seen in these figures, training without regularization does not guarantee the smoothness of the potential shape at the data points that rarely appear in the fitting reference data, demonstrating that regularization significantly improves the smoothness of the potential curve.

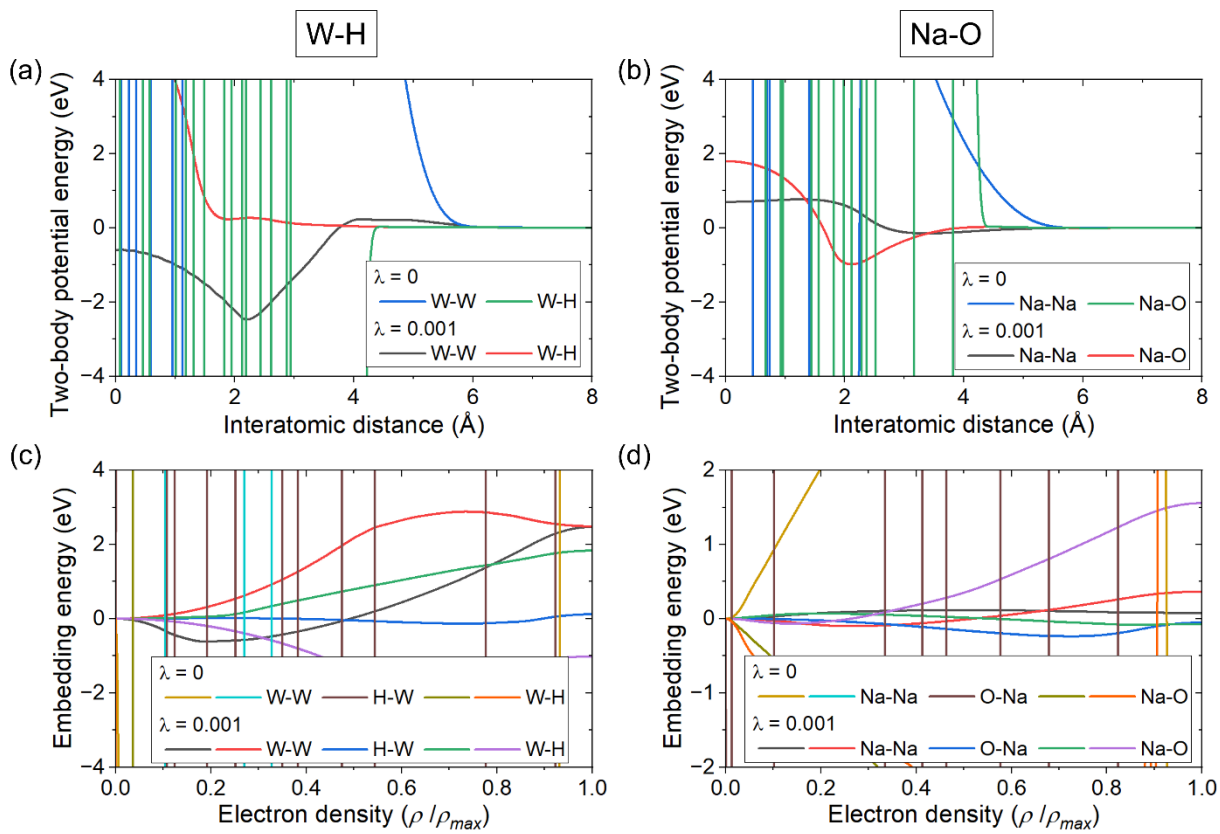

Figure S1. Effects of regularization on the potential energy curves.

S5. Details of the two-body correction (2BC) function for the short-range interaction

The 2BC function, f_{2BC}, is a linear combination of truncated power functions as

$$f_{2BC}(r) = \sum_{i=1}^{N} a_i g_i(r), \quad \text{Eq. S5-1}$$

$$g_i(r) = \begin{cases} \left(r_{cut,i} - r\right)^{n_i} & \text{for } r \le r_{cut,i} \\ 0 & \text{for } r > r_{cut,i} \end{cases}. \quad \text{Eq. S5-2}$$

In the fitting to the target function generated based on the ZBL potential and DFT calculation results, i.e., Eq. 18 of the main body, we adjusted N and optimized $\{a_i, r_{cut,i}\}$ by random variation using a Monte Carlo method. The model parameters obtained for the W-W pair in this study are listed in Table S3.

Table S3. The 2BC model parameters obtained for the W-W pair.

i	n_i	$r_{cut,i}$ (Å)	a_i (eV/Å^n)
1	3	0.60045667414109	-7.6327673159507E+05
2	4	0.89540888871784	2.4886782455988E+03
3	3	1.10009310026657	9.1684635791850E+02
4	4	1.49847881118696	5.6552687174325E+02
5	3	1.70196644957268	2.9172182782360E+02
6	4	2.24736917826690	3.3107817596202E+01
7	3	2.39102821773431	1.3057807293697E+01

S6. How to use the LMB-EAM potentials constructed in the present study

(1) LMB-EAMs

All LMB-EAMs generated in this study are available in the LAMMPS format in the form of an archive file, potential-files.zip, which is also provided as Supplementary Information. They can be used by LAMMPS as follows:

<For the W-H potential>

```
pair_style hybrid/overlay  eam/fs eam/fs eam/fs


pair_coeff * * eam/fs 1 LMB-EAM_W-H_01-03.eam.fs    W H
pair_coeff * * eam/fs 2 LMB-EAM_W-H_02-03.eam.fs    W H
```

```
pair_coeff * * eam/fs 3 LMB-EAM_W-H_03-03.eam.fs    W H
```

<For the Na-O potential>

```
pair_style hybrid/overlay  eam/fs eam/fs eam/fs


pair_coeff * * eam/fs 1 LMB-EAM_Na-O_01-03.eam.fs    Na O
pair_coeff * * eam/fs 2 LMB-EAM_Na-O_02-03.eam.fs    Na O
pair_coeff * * eam/fs 3 LMB-EAM_Na-O_03-03.eam.fs    Na O
```

It should be noted that the H-H interaction and the O-O interaction are not included in these W-H and Na-O potential models, respectively. Thus, for impurity simulations, these potential models can only be used for simulations of a single impurity atom.

The LMB-EAMs for H, N, or I impurity in liquid Na can also be used with the following file.

- ✓ Na-H
  - LMB-EAM_Na-H_01-03.eam.fs
  - LMB-EAM_Na-H_02-03.eam.fs
  - LMB-EAM_Na-H_03-03.eam.fs
- ✓ Na-N
  - LMB-EAM_Na-N_01-03.eam.fs
  - LMB-EAM_Na-N_02-03.eam.fs
  - LMB-EAM_Na-N_03-03.eam.fs
- ✓ Na-I
  - LMB-EAM_Na-I_01-03.eam.fs
  - LMB-EAM_Na-I_02-03.eam.fs
  - LMB-EAM_Na-I_03-03.eam.fs

As the same with the Na-O potential model, impurity-impurity interactions are not included.

(2) LMB-EAM + 2BC for the W-H case

For the W-H potential model, we have additionally prepared a short-range correction potential for the W-W pair, i.e., 2BC. This can be used together with the LMB-EAM using the hybrid/overlay command of LAMMPS as follows.

```
pair_style hybrid/overlay   eam/fs eam/fs eam/fs table linear 5000

pair_coeff * * eam/fs 1 LMB-EAM_W-H_01-03.eam.fs    W H
pair_coeff * * eam/fs 2 LMB-EAM_W-H_02-03.eam.fs    W H
pair_coeff * * eam/fs 3 LMB-EAM_W-H_03-03.eam.fs    W H
pair_coeff 1 1  table   2BC_W-W.dat   COR2B_LMBEAM 2.4
```

References

1 H. Kwon, M. Shiga, H. Kimizuka and T. Oda, Accurate description of hydrogen diffusivity in bcc metals using machine-learning moment tensor potentials and path-integral methods, *Acta Mater*, 2023, **247**, 118739.

2 J. Gil and T. Oda, Accurate and Efficient Calculation of the Solution Enthalpy and Diffusivity of Solutes in Liquid Metals Using Machine Learning Potential, *J Chem Theory Comput*, 2022, **18**, 5568–5576.