Supporting Information

Performance and Cost Assessment of Machine Learning Interatomic Potentials

Yunxing Zuo,[†] Chi Chen,[†] Xiangguo Li,[†] Zhi Deng,[†] Yiming Chen,[†] Jörg Behler,[‡] Gábor Csányi,[¶] Alexander V. Shapeev,[§] Aidan P. Thompson,[‖] Mitchell A. Wood,[‖] and Shyue Ping Ong[*,†]

†*Department of NanoEngineering, University of California San Diego, 9500 Gilman Dr, Mail Code 0448, La Jolla, CA 92093-0448, United States*

‡*Universitat Göttingen, Institut für Physikalische Chemie, Theoretische Chemie, Tammannstraße 6, 37077 Göttingen, Germany*

¶*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom*

§*Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Building 3, Moscow, 143026, Russia*

‖*Center for Computing Research, Sandia National Laboratories, Albuquerque, New Mexico 87185, United States*

E-mail: ongsp@eng.ucsd.edu

Detailed formalism of ML-IAPs and Optimized parameters

Behler-Parrinello neural network

The Behler-Parrinello neural network potentials (NNPs)[1–3] express the potential energy surface (PES) in terms of the atom-centered symmetry functions (ACSF)[4] using fully-connected feed-forward neural networks.

The ACSF converts the local atomic environment to numeric vectors that fulfill rotational and translational invariance.[4,5] For atom i in the structure, the radial ACSF $G_i^{\mathrm{atom,rad}}$, which provides information about pair correlations between the atoms, is given as follows:

$$G_i^{\mathrm{atom,rad}} = \sum_{j \neq i}^{N_{\mathrm{atom}}} e^{-\eta(R_{ij}-R_s)^2} \cdot f_c(R_{ij}), \tag{1}$$

where η determines the width of the Gaussian basis, and R_s is the position shift over all neighboring atoms within the cutoff radius R_c. The cutoff function f_c, which ensures a smooth decay in value and slope at cutoff radius R_c, is given as follows:

$$f_c(R_{ij}) = \begin{cases} 0.5 \cdot [\cos(\frac{\pi R_{ij}}{R_c}) + 1], & \text{for } R_{ij} \leq R_c \\ 0.0, & \text{for } R_{ij} > R_c. \end{cases} \tag{2}$$

The angular ACSF takes the following form:

$$G_i^{\mathrm{atom,ang}} = 2^{1-\zeta} \sum_{j,k \neq i}^{N_{\mathrm{atom}}} (1 + \lambda \cos\theta_{ijk})^{\zeta} \cdot e^{-\eta'(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}) \cdot f_c(R_{jk}), \tag{3}$$

where the summation loops over neighbors j and k, θ_{ijk} is the angle, and ζ, η', and λ are three parameters that determine the shape of angular symmetry functions.

Symmetry functions determined by an appropriate choice of hyperparameters (η, R_s, etc.) should reflect the effective relationship between atomic representations and corresponding properties (*e.g.*, energy and force components). A Kalman filter algorithm was used to

optimize the neural network weights, and a grid search was conducted to optimize the hyperparameters so that the fitted model can reproduce not only the basic energy and force values but also the elastic constants. All training structures were equally weighted while energy weight and force weight for each structure were fixed to 1.0 and 10.0, respectively. The optimized hyperparameters for all elemental systems studied in this work are listed in **Table S1** - **Table S6**.

Table S1: Hyperparameters for optimized Ni NNP. The units of R_c, R_s, η and η' are Å, Å, Å^{-2} and Å^{-2}, respectively.

R_c				3.9					
Energy weight				1.0					
Force weight				10.0					
NN layers				(24, 24)					
$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
Ni	Ni	0.0	0.036	Ni	Ni	Ni	1	1	0.036
Ni	Ni	0.0	0.071	Ni	Ni	Ni	1	2	0.036
Ni	Ni	0.0	0.179	Ni	Ni	Ni	1	4	0.036
				Ni	Ni	Ni	1	16	0.036
				Ni	Ni	Ni	-1	1	0.036
				Ni	Ni	Ni	-1	2	0.036
				Ni	Ni	Ni	-1	4	0.036
				Ni	Ni	Ni	-1	16	0.036
				Ni	Ni	Ni	1	1	0.071
				Ni	Ni	Ni	1	2	0.071
				Ni	Ni	Ni	1	4	0.071
				Ni	Ni	Ni	1	16	0.071
				Ni	Ni	Ni	-1	1	0.071
				Ni	Ni	Ni	-1	2	0.071
				Ni	Ni	Ni	-1	4	0.071
				Ni	Ni	Ni	-1	16	0.071
				Ni	Ni	Ni	1	1	0.179
				Ni	Ni	Ni	1	2	0.179
				Ni	Ni	Ni	1	4	0.179
				Ni	Ni	Ni	1	16	0.179
				Ni	Ni	Ni	-1	1	0.179
				Ni	Ni	Ni	-1	2	0.179

Continued on next page

Table S1 – continued from previous page

$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
				Ni	Ni	Ni	-1	4	0.179
				Ni	Ni	Ni	-1	16	0.179

Table S2: Hyperparameters for optimized Cu NNP. The units of R_c, R_s, η and η' are Å, Å, Å^{-2} and Å^{-2}, respectively.

R_c				4.1					
Energy weight				1.0					
Force weight				10.0					
NN layers				(8, 8)					
$G_i^{\mathrm{atom,rad}}$				$G_i^{\mathrm{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
Cu	Cu	3.0	0.036	Cu	Cu	Cu	1	1	0.036
Cu	Cu	3.0	0.071	Cu	Cu	Cu	1	2	0.036
Cu	Cu	3.0	0.179	Cu	Cu	Cu	1	4	0.036
				Cu	Cu	Cu	1	16	0.036
				Cu	Cu	Cu	-1	1	0.036
				Cu	Cu	Cu	-1	2	0.036
				Cu	Cu	Cu	-1	4	0.036
				Cu	Cu	Cu	-1	16	0.036
				Cu	Cu	Cu	1	1	0.071
				Cu	Cu	Cu	1	2	0.071
				Cu	Cu	Cu	1	4	0.071
				Cu	Cu	Cu	1	16	0.071
				Cu	Cu	Cu	-1	1	0.071
				Cu	Cu	Cu	-1	2	0.071
				Cu	Cu	Cu	-1	4	0.071
				Cu	Cu	Cu	-1	16	0.071
				Cu	Cu	Cu	1	1	0.179
				Cu	Cu	Cu	1	2	0.179
				Cu	Cu	Cu	1	4	0.179
				Cu	Cu	Cu	1	16	0.179
				Cu	Cu	Cu	-1	1	0.179
				Cu	Cu	Cu	-1	2	0.179

Continued on next page

Table S2 – continued from previous page

$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
				Cu	Cu	Cu	-1	4	0.179
				Cu	Cu	Cu	-1	16	0.179

Table S3: Hyperparameters for optimized Li NNP. The units of R_c, R_s, η and η' are Å, Å, Å^{-2} and Å^{-2}, respectively.

R_c				5.2					
Energy weight				1.0					
Force weight				10.0					
NN layers				(24, 24)					
$G_i^{\mathrm{atom,rad}}$				$G_i^{\mathrm{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
Li	Li	0.0	0.036	Li	Li	Li	1	1	0.036
Li	Li	1.0	0.036	Li	Li	Li	1	2	0.036
Li	Li	2.0	0.036	Li	Li	Li	1	4	0.036
Li	Li	3.0	0.036	Li	Li	Li	1	16	0.036
Li	Li	4.0	0.036	Li	Li	Li	-1	1	0.036
Li	Li	0.0	0.071	Li	Li	Li	-1	2	0.036
Li	Li	1.0	0.071	Li	Li	Li	-1	4	0.036
Li	Li	2.0	0.071	Li	Li	Li	-1	16	0.036
Li	Li	3.0	0.071	Li	Li	Li	1	1	0.071
Li	Li	4.0	0.071	Li	Li	Li	1	2	0.071
Li	Li	0.0	0.179	Li	Li	Li	1	4	0.071
Li	Li	1.0	0.179	Li	Li	Li	1	16	0.071
Li	Li	2.0	0.179	Li	Li	Li	-1	1	0.071
Li	Li	3.0	0.179	Li	Li	Li	-1	2	0.071
Li	Li	4.0	0.179	Li	Li	Li	-1	4	0.071
Li	Li	0.0	0.357	Li	Li	Li	-1	16	0.071
Li	Li	1.0	0.357	Li	Li	Li	1	1	0.179
Li	Li	2.0	0.357	Li	Li	Li	1	2	0.179
Li	Li	3.0	0.357	Li	Li	Li	1	4	0.179
Li	Li	4.0	0.357	Li	Li	Li	1	16	0.179
Li	Li	0.0	0.714	Li	Li	Li	-1	1	0.179
Li	Li	1.0	0.714	Li	Li	Li	-1	2	0.179

Continued on next page

Table S3 – continued from previous page

$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
Li	Li	2.0	0.714	Li	Li	Li	-1	4	0.179
Li	Li	3.0	0.714	Li	Li	Li	-1	16	0.179
Li	Li	4.0	0.714	Li	Li	Li	1	1	0.357
Li	Li	0.0	1.786	Li	Li	Li	1	2	0.357
Li	Li	1.0	1.786	Li	Li	Li	1	4	0.357
Li	Li	2.0	1.786	Li	Li	Li	1	16	0.357
Li	Li	3.0	1.786	Li	Li	Li	-1	1	0.357
Li	Li	4.0	1.786	Li	Li	Li	-1	2	0.357
Li	Li	0.0	3.571	Li	Li	Li	-1	4	0.357
Li	Li	1.0	3.571	Li	Li	Li	-1	16	0.357
Li	Li	2.0	3.571	Li	Li	Li	1	1	0.714
Li	Li	3.0	3.571	Li	Li	Li	1	2	0.714
Li	Li	4.0	3.571	Li	Li	Li	1	4	0.714
Li	Li	0.0	7.142	Li	Li	Li	1	16	0.714
Li	Li	1.0	7.142	Li	Li	Li	-1	1	0.714
Li	Li	2.0	7.142	Li	Li	Li	-1	2	0.714
Li	Li	3.0	7.142	Li	Li	Li	-1	4	0.714
Li	Li	4.0	7.142	Li	Li	Li	-1	16	0.714
Li	Li	0.0	17.855	Li	Li	Li	1	1	1.786
Li	Li	1.0	17.855	Li	Li	Li	1	2	1.786
Li	Li	2.0	17.855	Li	Li	Li	1	4	1.786
Li	Li	3.0	17.855	Li	Li	Li	1	16	1.786
Li	Li	4.0	17.855	Li	Li	Li	-1	1	1.786
				Li	Li	Li	-1	2	1.786
				Li	Li	Li	-1	4	1.786
				Li	Li	Li	-1	16	1.786
				Li	Li	Li	1	1	3.571
				Li	Li	Li	1	2	3.571

Continued on next page

Table S3 – continued from previous page

$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
				Li	Li	Li	1	4	3.571
				Li	Li	Li	1	16	3.571
				Li	Li	Li	-1	1	3.571
				Li	Li	Li	-1	2	3.571
				Li	Li	Li	-1	4	3.571
				Li	Li	Li	-1	16	3.571
				Li	Li	Li	1	1	7.142
				Li	Li	Li	1	2	7.142
				Li	Li	Li	1	4	7.142
				Li	Li	Li	1	16	7.142
				Li	Li	Li	-1	1	7.142
				Li	Li	Li	-1	2	7.142
				Li	Li	Li	-1	4	7.142
				Li	Li	Li	-1	16	7.142
				Li	Li	Li	1	1	17.855
				Li	Li	Li	1	2	17.855
				Li	Li	Li	1	4	17.855
				Li	Li	Li	1	16	17.855
				Li	Li	Li	-1	1	17.855
				Li	Li	Li	-1	2	17.855
				Li	Li	Li	-1	4	17.855
				Li	Li	Li	-1	16	17.855

Table S4: Hyperparameters for optimized Mo NNP. The units of R_c, R_s, η and η' are Å, Å, Å^{-2} and Å^{-2}, respectively.

R_c				5.2					
Energy weight				1.0					
Force weight				10.0					
NN layers				(16, 16)					
$G_i^{\mathrm{atom,rad}}$				$G_i^{\mathrm{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
Mo	Mo	0.0	0.036	Mo	Mo	Mo	1	1	0.036
Mo	Mo	1.0	0.036	Mo	Mo	Mo	1	2	0.036
Mo	Mo	2.0	0.036	Mo	Mo	Mo	1	4	0.036
Mo	Mo	3.0	0.036	Mo	Mo	Mo	1	16	0.036
Mo	Mo	4.0	0.036	Mo	Mo	Mo	-1	1	0.036
Mo	Mo	0.0	0.071	Mo	Mo	Mo	-1	2	0.036
Mo	Mo	1.0	0.071	Mo	Mo	Mo	-1	4	0.036
Mo	Mo	2.0	0.071	Mo	Mo	Mo	-1	16	0.036
Mo	Mo	3.0	0.071	Mo	Mo	Mo	1	1	0.071
Mo	Mo	4.0	0.071	Mo	Mo	Mo	1	2	0.071
Mo	Mo	0.0	0.179	Mo	Mo	Mo	1	4	0.071
Mo	Mo	1.0	0.179	Mo	Mo	Mo	1	16	0.071
Mo	Mo	2.0	0.179	Mo	Mo	Mo	-1	1	0.071
Mo	Mo	3.0	0.179	Mo	Mo	Mo	-1	2	0.071
Mo	Mo	4.0	0.179	Mo	Mo	Mo	-1	4	0.071
				Mo	Mo	Mo	-1	16	0.071
				Mo	Mo	Mo	1	1	0.179
				Mo	Mo	Mo	1	2	0.179
				Mo	Mo	Mo	1	4	0.179
				Mo	Mo	Mo	1	16	0.179
				Mo	Mo	Mo	-1	1	0.179
				Mo	Mo	Mo	-1	2	0.179

Continued on next page

Table S4 – continued from previous page

$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
				Mo	Mo	Mo	-1	4	0.179
				Mo	Mo	Mo	-1	16	0.179

Table S5: Hyperparameters for optimized Si NNP. The units of R_c, R_s, η and η' are Å, Å, Å^{-2} and Å^{-2}, respectively.

R_c				5.5					
Energy weight				1.0					
Force weight				10.0					
NN layers				(24, 24)					
$G_i^{\mathrm{atom,rad}}$				$G_i^{\mathrm{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
Si	Si	4.0	0.036	Si	Si	Si	1	1	0.036
Si	Si	4.0	0.071	Si	Si	Si	1	2	0.036
Si	Si	4.0	0.179	Si	Si	Si	1	4	0.036
				Si	Si	Si	1	16	0.036
				Si	Si	Si	-1	1	0.036
				Si	Si	Si	-1	2	0.036
				Si	Si	Si	-1	4	0.036
				Si	Si	Si	-1	16	0.036
				Si	Si	Si	1	1	0.071
				Si	Si	Si	1	2	0.071
				Si	Si	Si	1	4	0.071
				Si	Si	Si	1	16	0.071
				Si	Si	Si	-1	1	0.071
				Si	Si	Si	-1	2	0.071
				Si	Si	Si	-1	4	0.071
				Si	Si	Si	-1	16	0.071
				Si	Si	Si	1	1	0.179
				Si	Si	Si	1	2	0.179
				Si	Si	Si	1	4	0.179
				Si	Si	Si	1	16	0.179
				Si	Si	Si	-1	1	0.179
				Si	Si	Si	-1	2	0.179

Continued on next page

Table S5 – continued from previous page

$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
				Si	Si	Si	-1	4	0.179
				Si	Si	Si	-1	16	0.179

Table S6: Hyperparameters for optimized Ge NNP. The units of R_c, R_s, η and η' are Å, Å, Å^{-2} and Å^{-2}, respectively.

R_c				5.6					
Energy weight				1.0					
Force weight				10.0					
NN layers				(24, 24)					
$G_i^{\mathrm{atom,rad}}$				$G_i^{\mathrm{atom,ang}}$					
Reference	Neighbor	R_s	η	Reference	Neighbor 1	Neighbor 2	λ	ζ	η'
Ge	Ge	4.0	0.036	Ge	Ge	Ge	1	1	0.036
Ge	Ge	4.0	0.071	Ge	Ge	Ge	-1	1	0.036
Ge	Ge	4.0	0.179	Ge	Ge	Ge	1	1	0.071
Ge	Ge	4.0	0.357	Ge	Ge	Ge	-1	1	0.071
Ge	Ge	4.0	0.714	Ge	Ge	Ge	1	1	0.179
Ge	Ge	4.0	3.571	Ge	Ge	Ge	-1	1	0.179
Ge	Ge	4.0	7.142	Ge	Ge	Ge	1	1	0.357
Ge	Ge	4.0	17.855	Ge	Ge	Ge	-1	1	0.357
				Ge	Ge	Ge	1	1	0.714
				Ge	Ge	Ge	-1	1	0.714
				Ge	Ge	Ge	1	1	1.786
				Ge	Ge	Ge	-1	1	1.786
				Ge	Ge	Ge	1	1	3.571
				Ge	Ge	Ge	-1	1	3.571
				Ge	Ge	Ge	1	1	7.142
				Ge	Ge	Ge	-1	1	7.142
				Ge	Ge	Ge	1	1	17.855
				Ge	Ge	Ge	-1	1	17.855

Gaussian approximation potentials

The Gaussian Approximation Potentials (GAPs),[6–8] implemented in the QUIP program package,[9] applies Gaussian process regression (GPR) to interpolate the atomic energy based on the spatial distribution of the neighboring atoms. The atomic energy function is given as follows:

$$\epsilon(\boldsymbol{R}) = \sum_k b_k \boldsymbol{K}(\boldsymbol{R}, \boldsymbol{R}_k), \tag{4}$$

where $\boldsymbol{R}$ represents the geometry of neighboring atoms within cutoff radius R_c, and k indexes a set of reference data points $\boldsymbol{R}_k$ that serve as a basis on which the atomic energy function is expanded. $\boldsymbol{K}$ is a kernel function that captures the variation of the energy function in terms of changing neighboring configurations.

In this work, the kernel function used is the "Smooth Overlap of Atomic Positions" (SOAP),[7,10–12] which represents the rotationally integrated overlap of neighbor densities. The atomic configuration of a central atom i is represented by its neighbor density as a sum of Gaussians centered over all neighboring atoms j within a cutoff distance R_c,

$$\rho_i(\boldsymbol{R}) = \sum_j f_c(R_{ij}) \cdot \exp(-\frac{|\boldsymbol{R} - \boldsymbol{R}_{ij}|^2}{2\sigma_{\text{atom}}^2}), \tag{5}$$

where f_c is a cutoff function ensures smooth decay in value and slope at cutoff radius R_c, and σ_{atom} is a smearing parameter. This neighbor density is then expanded in terms of a basis of spherical harmonics $Y_{lm}(\hat{\boldsymbol{R}})$, and radial functions $g_n(R)$:

$$\rho_i(\boldsymbol{R}) = \sum_{nlm} c_{nlm}\ g_n(R) Y_{lm}(\hat{\boldsymbol{R}}), \tag{6}$$

and the rotationally invariant spherical power spectrum of atom i is then expressed in terms of expansion coefficients c_{nlm} as follows:

$$p_{n_1 n_2 l}(\boldsymbol{R}_i) = \sum_{m=-l}^{l} c^*_{n_1 lm} c_{n_2 lm}, \tag{7}$$

Finally, the SOAP kernel is written as a dot product of power spectrums and raised to a small integer power as follows:

$$\boldsymbol{K}(\boldsymbol{R}, \boldsymbol{R'}) = \sum_{n_1 n_2 l} (p_{n_1 n_2 l}(\boldsymbol{R}) p_{n_1 n_2 l}(\boldsymbol{R'}))^{\zeta}, \tag{8}$$

where ζ controls the raised power of dot products of spherical power spectrum. Normalization is then carried out to ensure that the kernel of each atomic environment with itself is unity.

In the potential fitting process, all training structures were performed evenly regardless of their categories, and the regularization parameters (i.e. smearing parameters of energies and forces) corresponding to the expected errors determined how energies and forces were weighted. Thus the choice of these hyperparameters has a significant influence on the performance of GAP models.[13,14] Here, we have optimized five hyperparameters via grid search: cutoff radius, regularization parameters of energy σ_{energy} and force σ_{force}, and pair of maximum powers for radial components as well as angular components in spherical harmonics expansion $(n_{\text{max}}, l_{\text{max}})$. The positive integer power of SOAP kernel ζ has been fixed to 4, in line with previous works.[12–15] A full set of optimized and fixed hyperparameters for each elemental system are listed in **Table S7**.

Table S7: Hyperparameters of the GAP model across different chemistries.

	Ni	Cu	Li	Mo	Si	Ge
cutoff radius (Å)	3.9	4.1	4.8	5.2	5.1	5.4
σ_{energy} (eV/atom)	0.01	0.01	0.01	0.01	0.01	0.01
σ_{force} (eV/Å)	0.05	0.05	0.05	0.05	0.05	0.05
$(n_{\text{max}}, l_{\text{max}})$	(8, 8)	(8, 8)	(8, 8)	(8, 8)	(8, 8)	(8, 8)
ζ	4	4	4	4	4	4

Spectral neighbor analysis potential

The bispectrum of the neighbor density and spectral neighbor analysis potential (SNAP) formalism has been extensively studied in the previous works.[7,16] The basic concept of the bispectrum formalism is to project the 3D local atomic neighbor density into a set of coefficients that satisfy the invariant properties by expansion on a spherical harmonics basis.

In SNAP, the atomic neighbor density around a central atom i at position $\boldsymbol{R}$ is defined as follows:

$$\rho_i(\mathbf{R}) = \delta(\mathbf{R}) + \sum_{R_{ij}<R_c} f_c(R_{ij}) \cdot \omega_j \cdot \delta(\mathbf{R} - \mathbf{R}_{ij}), \tag{9}$$

where R_{ij} is the position of neighbor atom j relative to central atom i, and ω_j is the dimensionless weight to discriminate atom types. The cutoff function f_c ensures that the neighbor atomic density decreases smoothly to zero at the cutoff radius R_c.

The angular distribution of neighbor density function can be projected onto spherical harmonic functions $Y_m^l(\theta, \phi)$. In the bispectrum approach, the radial distribution is converted into an additional polar angle θ_0 defined by $\theta_0 = \theta_0^{\max} \frac{r}{R_c}$. Thus the density function can be represented in the 3-sphere (θ, ϕ, θ_0) coordinates instead of (θ, ϕ, r). The density function on 3-sphere can then be expanded with 4-dimensional hyperspherical harmonics $U_{m,m'}^j$, as follows:

$$\rho_i(\mathbf{R}) = \sum_{j=0}^{\infty} \sum_{m,m'=-j}^{j} u_{m,m'}^j U_{m,m'}^j, \tag{10}$$

where the coefficients $u_{m,m'}^j$ are obtained as the inner product of the neighbor density function with the basic function given by the following:

$$u_{m,m'}^j = U_{m,m'}^j(0,0,0) + \sum_{R_{ij}<R_c} f_c(R_{ij}) \cdot \omega_j \cdot U_{m,m'}^j(\theta, \phi, \theta_0). \tag{11}$$

The bispectrum components $B_{j_1,j_2,j}$ can then obtained via following:

$$B_{j_1,j_2,j} = \sum_{m_1,m_1'=-j_1}^{j_1} \sum_{m_2,m_2'=-j_2}^{j_2} \sum_{m,m'=-j}^{j} (u^j_{m,m'})^* \times H^{jmm'}_{\substack{j_1 m_1 m_1' \\ j_2 m_2 m_2'}} u^{j_1}_{m_1,m_1'} u^{j_2}_{m_2,m_2'}, \tag{12}$$

where constants $H^{jmm'}_{\substack{j_1 m_1 m_1' \\ j_2 m_2 m_2'}}$ are coupling coefficients satisfying the conditions $\|j_1 - j_2\| \leq j \leq \|j_1 + j_2\|$.

In the SNAP formalism, the energy E_{SNAP}, force $\boldsymbol{F}^j_{\text{SNAP}}$, and stresses $\boldsymbol{\sigma}^j_{\text{SNAP}}$ are expressed as linear functions of the bispectrum components[16,17] as follows:

$$E_{\text{SNAP}} = \beta_0 N + \boldsymbol{\beta} \cdot \sum_{i=1}^{N} \boldsymbol{B}^i, \tag{13}$$

$$\boldsymbol{F}^j_{\text{SNAP}} = -\boldsymbol{\beta} \cdot \sum_{i=1}^{N} \frac{\partial \boldsymbol{B}^i}{\partial \boldsymbol{r}_j}, \tag{14}$$

$$\boldsymbol{\sigma}^j_{\text{SNAP}} = -\boldsymbol{\beta} \cdot \sum_{i=1}^{N} \boldsymbol{r}_j \otimes \sum_{i=1}^{N} \frac{\partial \boldsymbol{B}^i}{\partial \boldsymbol{r}_j}, \tag{15}$$

where β_0 and the vector $\boldsymbol{\beta}$ are the coefficients derived from fitting with the database of quantum electronic structure calculations.

In our previous works,[18–20] we have developed a robust alternating two-step fitting process to optimize the hyperparameters (i.e., data weights corresponding to different categories of training structures, cutoff radius, and orders of bispectrum components $J_{\max}$), and to obtain the model coefficients for SNAP as well. In each inner iteration, the fitting of a linear model is performed. In the outer loop, the model fitted in the inner iteration is then used to predict materials properties (e.g., elastic tensors), and the differences between predicted and reference values are then used to optimize hyperparameters via the differential evolution optimization algorithm implemented in Scipy package.[21] The full set of optimized SNAP model hyperparameters and model coefficients for all elemental systems are provided in

Table S8 and **Table S9**, respectively.

Table S8: Hyperparameters of the SNAP model optimized for each elemental system.

	Ni	Cu	Li	Mo	Si	Ge
cutoff radius (Å)	4.1	4.1	5.1	4.6	4.9	5.5
$J_{\max}$	4	4	4	3	4	4
Energy weight of Elastic group	9703.497	9584.654	0.245	4868.530	2958.711	9213.581
Force weight of Elastic group	4.357	90.351	42.605	98.477	62.225	55.845
Energy weight of Surface group	2164.024	1281.721	9919.166	1367.747	7882.800	4131.218
Force weight of Surface group	89.207	88.080	0.443	0.340	70.138	57.763
Energy weight of AIMD group	8870.845	6897.657	674.874	1376.755	2918.824	661.208
Force weight of AIMD group	22.822	4.774	2.241	2.613	21.200	3.645
Energy weight of Vacancy group	780.361	682.488	1542.464	2726.165	9307.700	3266.749
Force weight of Vacancy group	45.451	5.695	20.344	8.045	26.388	1.231

Table S9: Model coefficients of the SNAP model for each elemental system.

				Ni	Cu	Li	Mo	Si	Ge
k	$2j_1$	$2j_2$	$2j$	β_k^{Ni}	β_k^{Cu}	β_k^{Li}	β_k^{Mo}	β_k^{Si}	β_k^{Ge}
0				−13.981	−12.560	−3.443	−17.762	−5.6640	−8.0321
1	0	0	0	0.0145	0.0086	0.0029	0.0218	0.0000	0.0036
2	1	0	1	0.0101	0.0137	0.0041	0.0034	0.0181	−0.0013
3	1	1	2	0.1205	0.0639	0.0184	0.3091	0.0589	−0.0746
4	2	0	2	−0.0069	−0.0094	0.0063	0.0290	−0.0160	−0.0148
5	2	1	3	0.2776	0.1906	0.0562	0.6210	0.2494	−0.0045
6	2	2	2	0.1109	0.0975	0.0065	0.2611	0.0307	0.0249
7	2	2	4	0.1105	0.0810	0.0218	0.1901	0.0790	0.0239
8	3	0	3	0.0103	0.0071	−0.0019	0.0173	0.0378	0.0116
9	3	1	4	0.1234	0.0929	0.0281	0.2779	0.0801	0.0727
10	3	2	3	0.1304	0.0880	0.0177	0.2888	0.1145	0.1424
11	3	2	5	0.1353	0.1192	0.0350	0.1009	0.0877	0.0367
12	3	3	4	0.0614	0.0653	0.0178	0.0940	−0.0145	0.0890
13	3	3	6	0.0248	0.0314	0.0118	0.0116	−0.0168	0.0046
14	4	0	4	0.0051	0.0028	0.0010	0.0181	0.0111	0.0046
15	4	1	5	0.0962	0.0919	0.0181	0.1319	0.0156	−0.0040
16	4	2	4	0.0488	0.0532	0.0036	0.0947	0.0268	0.0720
17	4	2	6	0.0992	0.0988	0.0273	0.0970	0.0487	0.0229
18	4	3	5	0.0818	0.0837	0.0204	0.0680	−0.0649	0.0370
19	4	3	7	0.0200	0.0190	0.0094		0.0029	−0.0024
20	4	4	4	0.0227	0.0284	0.0030	0.0091	−0.0199	−0.0060
21	4	4	6	0.0316	0.0387	0.0037	−0.0173	−0.0179	0.0150
22	4	4	8	0.0015	0.0043	0.0030		0.0031	0.0006
23	5	0	5	−0.0004	0.0018	−0.0004	0.0033	0.0120	0.0063
24	5	1	6	0.0584	0.0542	0.0102	0.0688	0.0037	0.0208
25	5	2	5	0.0419	0.0486	0.0034	0.0628	0.0301	0.0532
26	5	2	7	0.0565	0.0508	0.0098		0.0213	0.0115

Continued on next page

Table S9 – continued from previous page

				Ni	Cu	Li	Mo	Si	Ge
k	$2j_1$	$2j_2$	$2j$	β_k^{Ni}	β_k^{Cu}	β_k^{Li}	β_k^{Mo}	β_k^{Si}	β_k^{Ge}
27	5	3	6	0.0560	0.0667	0.0115	−0.0080	−0.0364	−0.0140
28	5	3	8	0.0004	0.0044	0.0030		−0.0083	0.0056
29	5	4	5	0.0328	0.0447	0.0065	0.0162	−0.0499	−0.0343
30	5	4	7	0.0289	0.0253	0.0043		−0.0011	0.0178
31	5	5	6	0.0137	0.0222	0.0008	−0.0412	−0.0319	−0.0118
32	5	5	8	0.0049	0.0039	−0.0005		−0.0031	0.0092
33	6	0	6	−0.0070	−0.0059	−0.0015	−0.0033	0.0026	0.0028
34	6	1	7	0.0316	0.0337	0.0049		0.0044	0.0151
35	6	2	6	0.0367	0.0392	0.0002	0.0584	0.0325	0.0359
36	6	2	8	0.0231	0.0191	0.0013		0.0058	0.0081
37	6	3	7	0.0317	0.0304	0.0060		−0.0049	−0.0205
38	6	4	6	0.0170	0.0250	0.0024	−0.0017	−0.0164	−0.0007
39	6	4	8	0.0079	0.0044	0.0001		0.0021	0.0083
40	6	5	7	0.0137	0.0175	−0.0009		−0.0272	−0.0241
41	6	6	6	0.0048	0.0086	0.0007	0.0008	−0.0083	−0.0002
42	6	6	8	0.0057	0.0011	0.0000		−0.0062	−0.0035
43	7	0	7	−0.0037	0.0016	−0.0011		0.0006	0.0008
44	7	1	8	0.0072	−0.0095	0.0029		0.0055	0.0024
45	7	2	7	0.0168	0.0077	−0.0015		0.0211	0.0436
46	7	3	8	0.0086	0.0053	0.0029		−0.0073	−0.0048
47	7	4	7	0.0025	0.0038	0.0005		−0.0029	−0.0015
48	7	5	8	0.0007	0.0009	0.0000		−0.0113	−0.0047
49	7	6	7	0.0043	0.0017	0.0006		−0.0088	−0.0075
50	7	7	8	0.0008	−0.0033	0.0000		−0.0069	0.0055
51	8	0	8	−0.0008	0.0015	−0.0004		0.0012	−0.0009
52	8	2	8	0.0025	−0.0081	−0.0005		0.0071	0.0170
53	8	4	8	−0.0008	0.0019	0.0005		0.0010	−0.0124
54	8	6	8	−0.0009	−0.0023	−0.0003		0.0069	0.0038

Continued on next page

Table S9 – continued from previous page

				Ni	Cu	Li	Mo	Si	Ge
k	$2j_1$	$2j_2$	$2j$	β_k^{Ni}	β_k^{Cu}	β_k^{Li}	β_k^{Mo}	β_k^{Si}	β_k^{Ge}
55	8	8	8	0.0008	0.0020	0.0003		−0.0034	−0.0014

Quadratic Spectral neighbor analysis potential

In the initial formulation of SNAP, the linear form of the PES to bispectrum relationship limits the maximum complexity of energy functions to a four-body effect, which may in turn have an impact on its predictive power. Recently Wood and Thompson[22] proposed a quadratic extension of SNAP (qSNAP) approach. The quadratic contributions to the energy can be viewed as a kind of embedding energy in analogy with the embedded atom method (EAM).[23,24] In qSNAP, the SNAP potential is extended via the addition of an embedding energy term as follows:

$$E^i_{\text{SNAP}} = \boldsymbol{\beta} \cdot \boldsymbol{B}^i + F(\rho_i), \tag{16}$$

where $F(\rho_i)$ represents the energy of embedding atom i into the electron density contributed by its neighboring atoms. The "host" electron density of embedding atom i can be expressed as a linear function of the bispectrum components as follows:

$$\rho_i = \boldsymbol{a} \cdot \boldsymbol{B}^i, \tag{17}$$

and the embedding energy can be expressed as a Taylor expansion based on a reference structure with density ρ_0 as follows:

$$F(\rho) = F_0 + (\rho - \rho_0)F' + \frac{1}{2}(\rho - \rho_0)^2 F'' + \cdots . \tag{18}$$

Thus, the modified SNAP energy is then given by the following expression:

$$\begin{aligned} E^i_{\text{SNAP}} &= \boldsymbol{\beta} \cdot \boldsymbol{B}^i + \frac{1}{2}F''(\boldsymbol{a} \cdot \boldsymbol{B}^i)^2 \\ &= \boldsymbol{\beta} \cdot \boldsymbol{B}^i + \frac{1}{2}(\boldsymbol{B}^i)^T \cdot \boldsymbol{\alpha} \cdot \boldsymbol{B}^i, \end{aligned} \tag{19}$$

where $\boldsymbol{\alpha} = F''\boldsymbol{a} \otimes \boldsymbol{a}$ is a symmetric $K \times K$ matrix. Essentially the quadratic extension carries all distinct pairwise products of bispectrum components, and expands the maximum complexity of energy functions to seven-body effects.[22]

In this work, the same two-step fitting approach is used for both the SNAP and qSNAP models. The optimized qSNAP hyperparameters (model coefficients not shown) for all elemental systems are listed in **Table S10**.

Table S10: Hyperparameters of the qSNAP model for each elemental system.

	Ni	Cu	Li	Mo	Si	Ge
cutoff radius (Å)	3.8	3.9	5.1	5.2	4.8	4.9
$J_{\max}$	3	3	3	3	4	3
Energy weight of Elastic group	2314.362	5963.055	0.905	9064.057	9991.648	6815.844
Force weight of Elastic group	66.103	8.517	42.253	57.214	91.932	97.832
Energy weight of Surface group	9733.283	2150.778	6958.904	1548.461	3719.473	2039.495
Force weight of Surface group	76.581	64.385	36.476	9.500	17.942	18.233
Energy weight of AIMD group	1138.766	6781.248	938.597	1501.509	4430.681	722.697
Force weight of AIMD group	5.515	50.632	72.997	7.777	82.268	48.418
Energy weight of Vacancy group	9946.955	9466.920	3303.361	7386.760	8999.320	8087.987
Force weight of Vacancy group	59.829	5.217	33.934	86.292	52.238	65.467

Moment tensor potentials

The Moment Tensor Potential (MTP) model and its formalism have been studied in previous works.[25–27] The fundamental idea of MTP is to construct a contracted rotationally invariant representation of the local atomic environment in a tensorial sense and build a linear correlation between potential energy and atomic representation based on the assumption that the total energy can be partitioned into individual atomic environment contributions.

As is denoted previously,[25,26] the potential energy of the atomic environment of central atom i can be linearly expanded on a set of basis functions $B(\boldsymbol{R})$,

$$V_i(\boldsymbol{R}) = \sum_l \beta_l B(\boldsymbol{R}), \tag{20}$$

The basis functions $B(\boldsymbol{R})$, in turn, depend on a series of moment tensor descriptors over all neighbor atoms j,

$$M_{\mu,\nu}(\boldsymbol{R}) = \sum_j f_\mu(R_{ij}) \underbrace{\boldsymbol{R}_{ij} \otimes \cdots \otimes \boldsymbol{R}_{ij}}_{\nu \text{ times}}, \tag{21}$$

where the functions f_μ are the radial distribution of atomic configuration and the terms $\boldsymbol{R}_{ij} \otimes \cdots \otimes \boldsymbol{R}_{ij}$ are tensors of rank ν entailing angular information about the atomic configuration. These descriptors are rotationally invariant in a tensorial sense (or, to be more precise, rotationally covariant).

The choice of μ and ν provides the balance between computational complexity and computational efficiency of MTP.[25] The basis functions $B(\boldsymbol{R})$ are formulated by different ways of contracting the moment tensors $M_{\mu,\nu}$ to a scalar. Each contraction can be encoded by a symmetric $m \times m$ matrix α where diagonal elements α_{ij} demonstrate dimensions of the ith and jth tensors being contracted. In this work we have used the implementation described in previous work[28] which for a single component reduces to learning the optimal radial functions f_μ instead of fixing them to be a universal radial basis.

With grid search method, we have carefully examined the choice of three different hy-

perparameters: cutoff radius, the number of polynomial powers and the number of free parameters by evaluating the performance on reproducing the basic properties, e.g., elastic constants and phonon spectrum, and fixed the hyperparameters energy weight as 1.0 and force weight as 0.01. All training structures were performed evenly regardless of their categories. A full set of optimized and fixed hyperparameters for each elemental system has been presented in **Table S11**. The maximum number of iterations has been set to 500, 1000 and 2000 to ensure the convergence of the fitting process for 16 or lower polynomial powers, 20 polynomial powers and 24 polynomial powers, respectively.

Table S11: Hyperparameters of the MTP model across different chemistries.

	Ni	Cu	Li	Mo	Si	Ge
cutoff radius (Å)	4.0	3.9	5.1	5.2	4.7	5.1
# of polynomial powers	20	20	16	20	24	24
# of free parameters	329	329	125	329	913	913
Energy weight	1.00	1.00	1.00	1.00	1.00	1.00
Force weight	0.01	0.01	0.01	0.01	0.01	0.01

Radial distribution function analysis of AIMD simulations

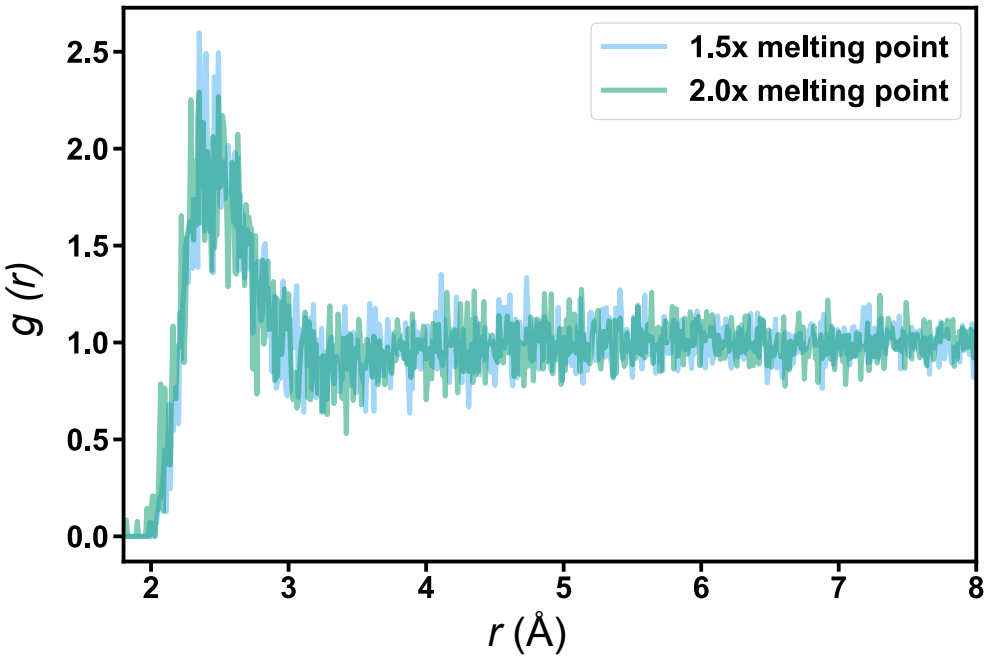


Figure S1: The radial distribution function (RDF) of Si in high temperature (1.5×, 2.0× of the melting point) AIMD simulations.

Normalized prediction accuracy in energies

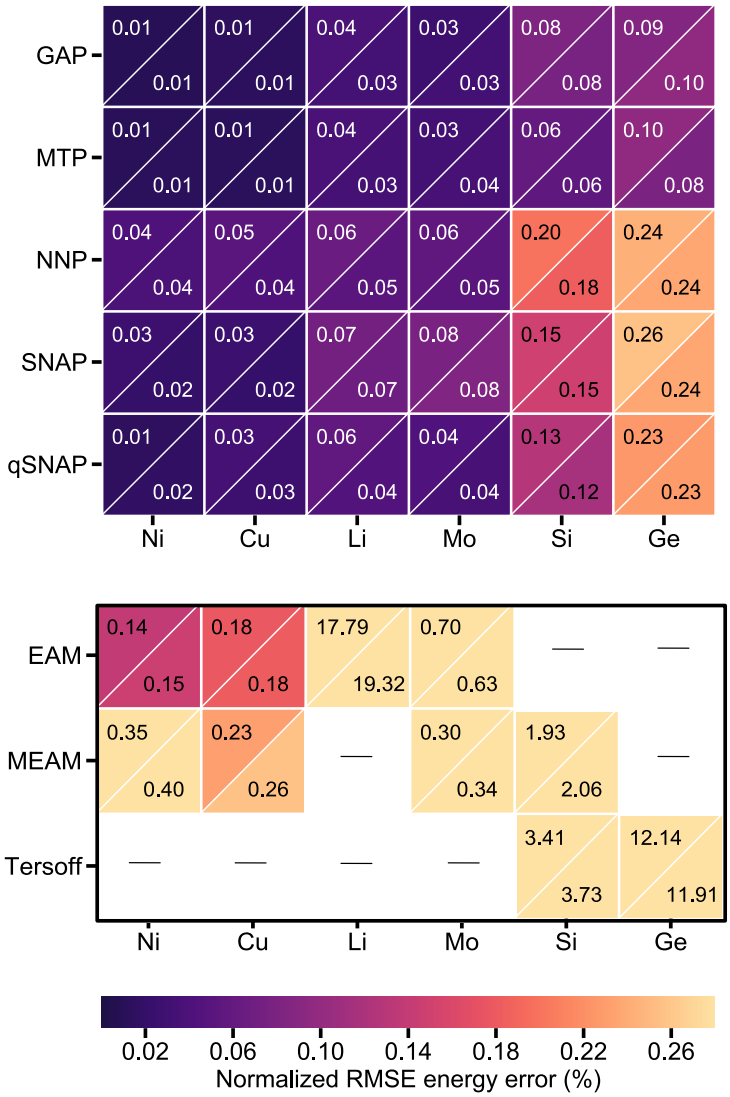


Figure S2: Root-mean-square errors in predicted energies normalized to total energy per atom for all four ML-IAPs as well as traditional IAPs (EAM,[29,30] MEAM,[31–33] Tersoff[34,35]). The upper left and lower right triangles within each cell represent training and test errors, respectively.

Learning curve of different ML-IAPs

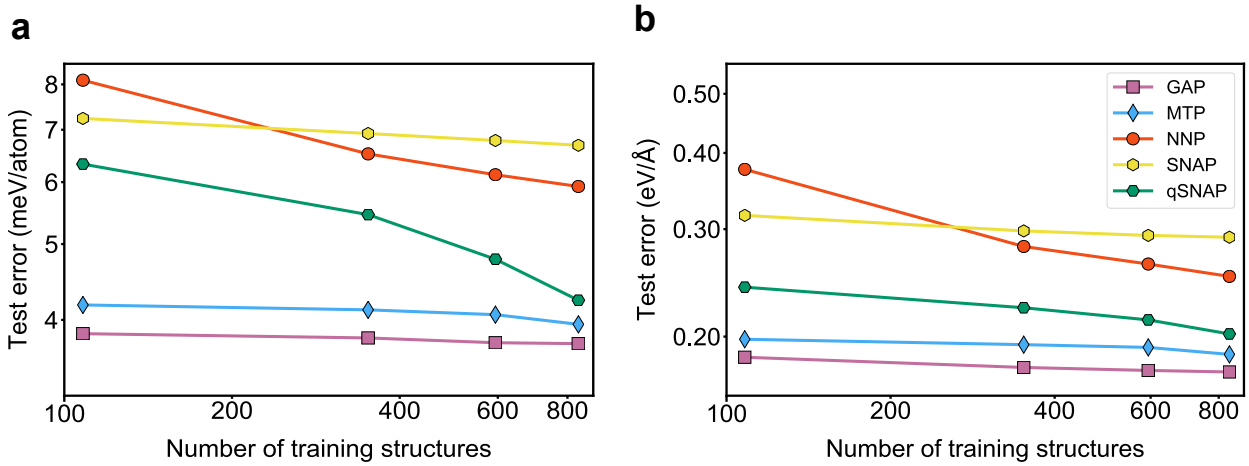


Figure S3: Log-log plot of RMSEs in (a) energies (b) forces versus the the size of the training data for the ML-IAP Mo models. The slope of the line represents the learning rate of different ML-IAPs.

Augmented Si GAP Models

We constructed two additional Si GAP models by augmenting the features of the original GAP model with the following:

- "Close" model: features close to the wurtzite polymorph by euclidean distance less than 0.0844, resulting in 15 additional points;
- "Exact" model: features corresponding to the wurtzite polymorph, resulting in an additional 4 data points.

We performed principal component analysis (PCA) to assess the coverage of active features for the original model (Figure S4a), the "close" model (Figure S4b) and the "exact" model (Figure S4c), and evaluate the energetic differences between the wurtzite and diamond for these three models (Figure S4d). The original model exhibited poor extrapolation because the polymorph phase region was not covered by active features (see Figure S4a). The "close" model exhibits a substantial improvement in accuracy due to the addition of features near the polymorph region (Figure S4b), while the "exact" model exhibits an even larger improvement in accuracy. These results highlight the sensitivity of the Gaussian process model to the reference data.

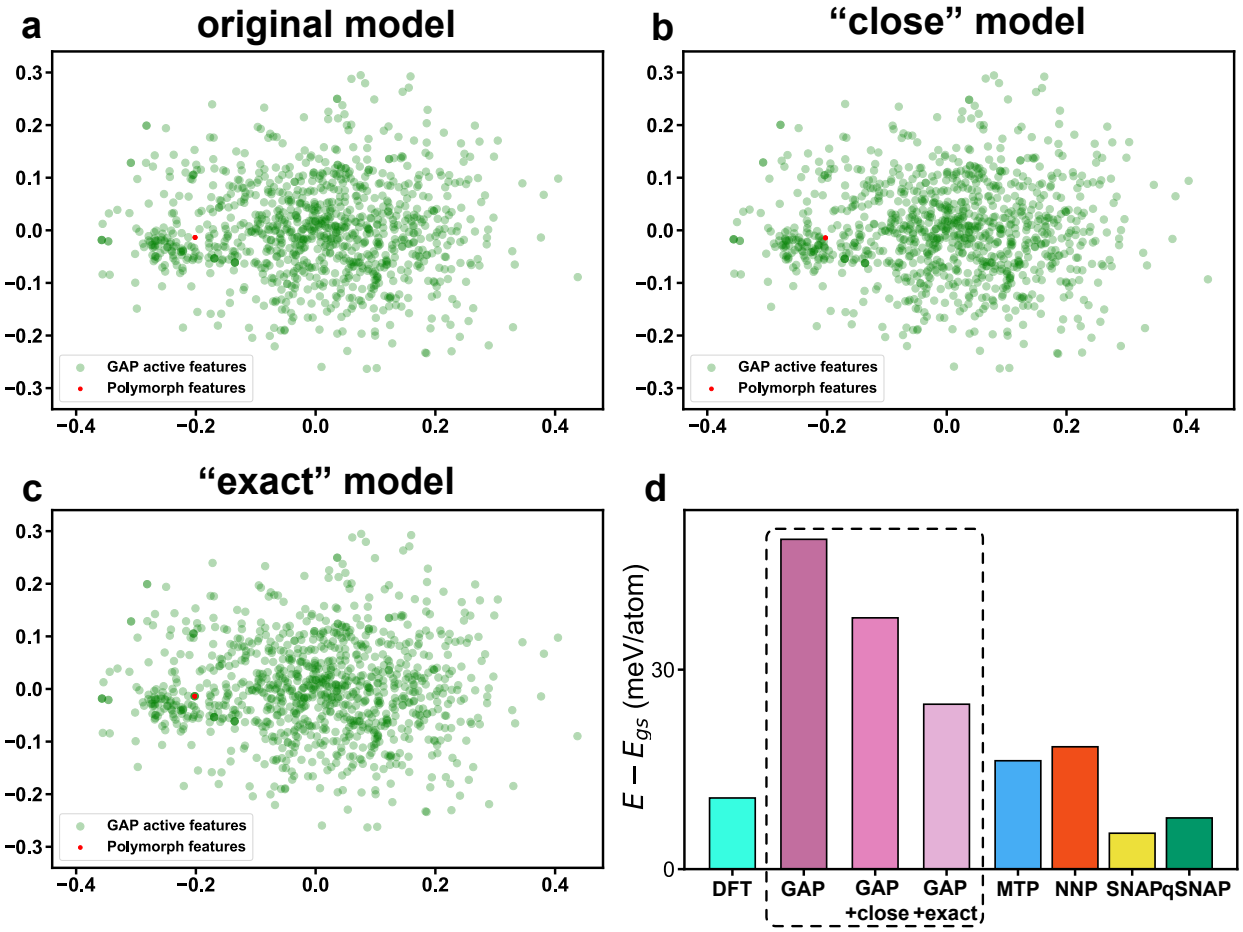
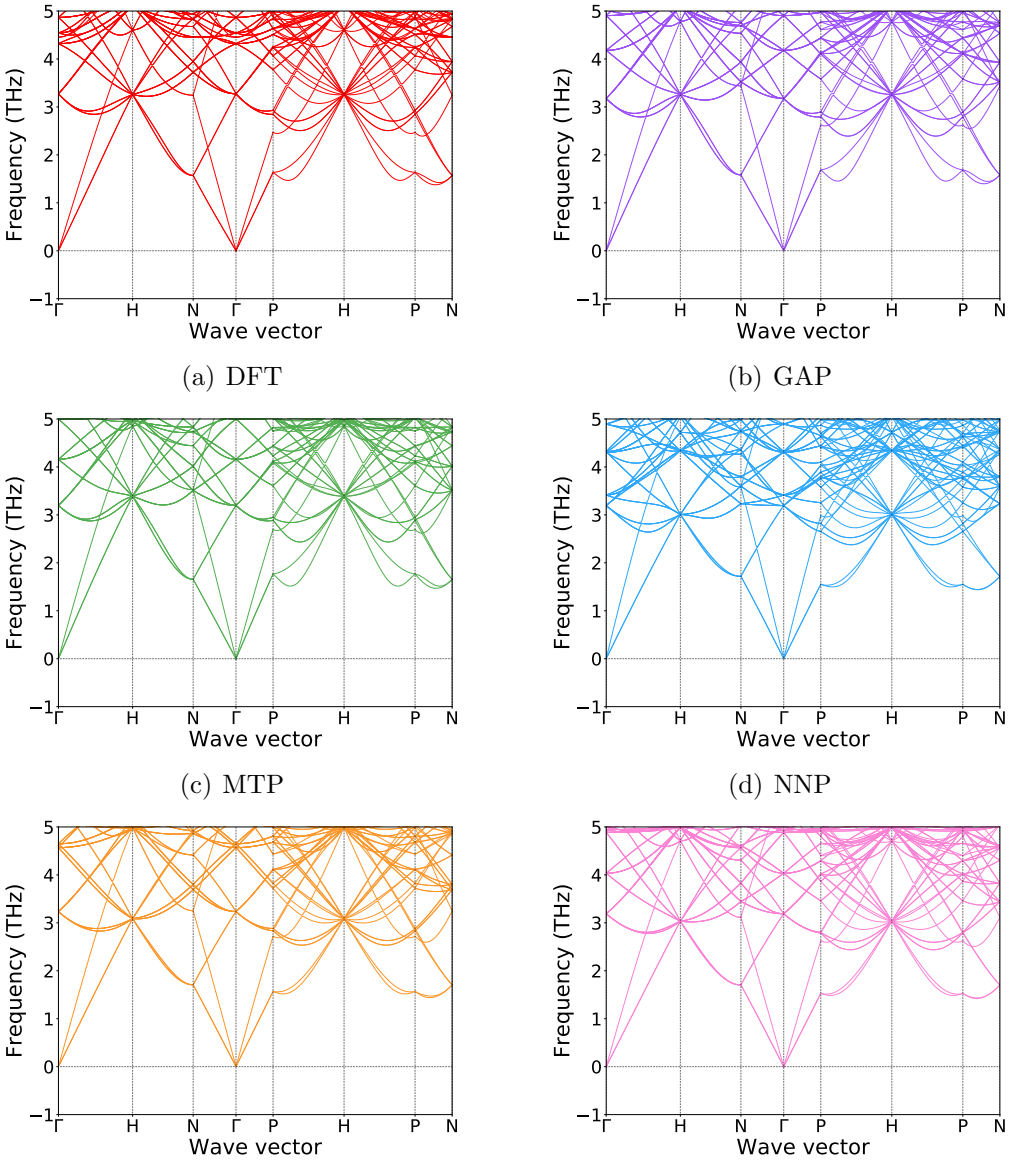


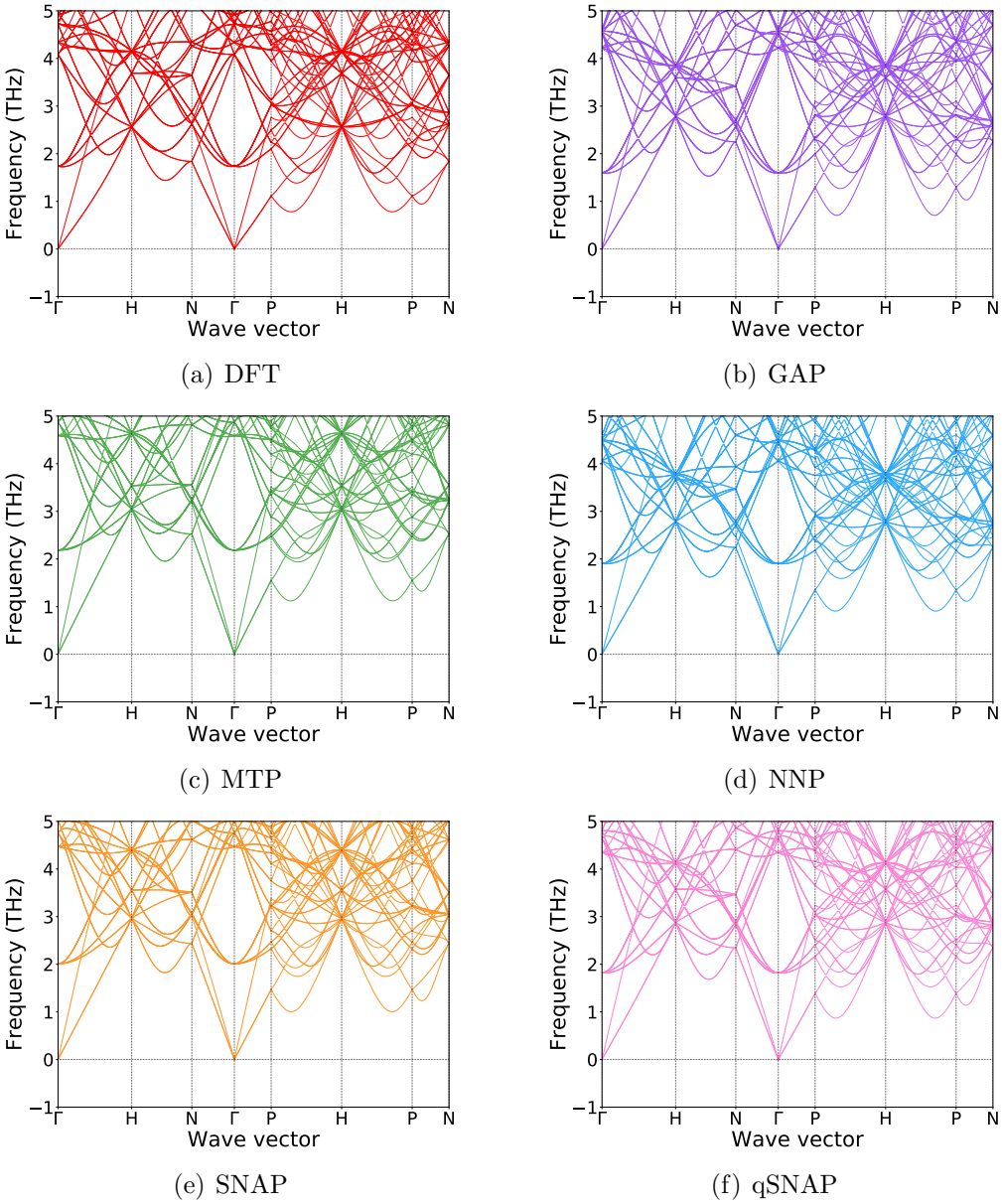
Figure S4: Two-dimensional projection of the principal components of (a) the original model, (b) the "close" model and (c) the "exact" model. (d) Energetic differences between the wurtzite and diamond polymorph for Si of the original GAP model and augmented GAP models compared to DFT and other ML-IAPs.

Phonon dispersion curves



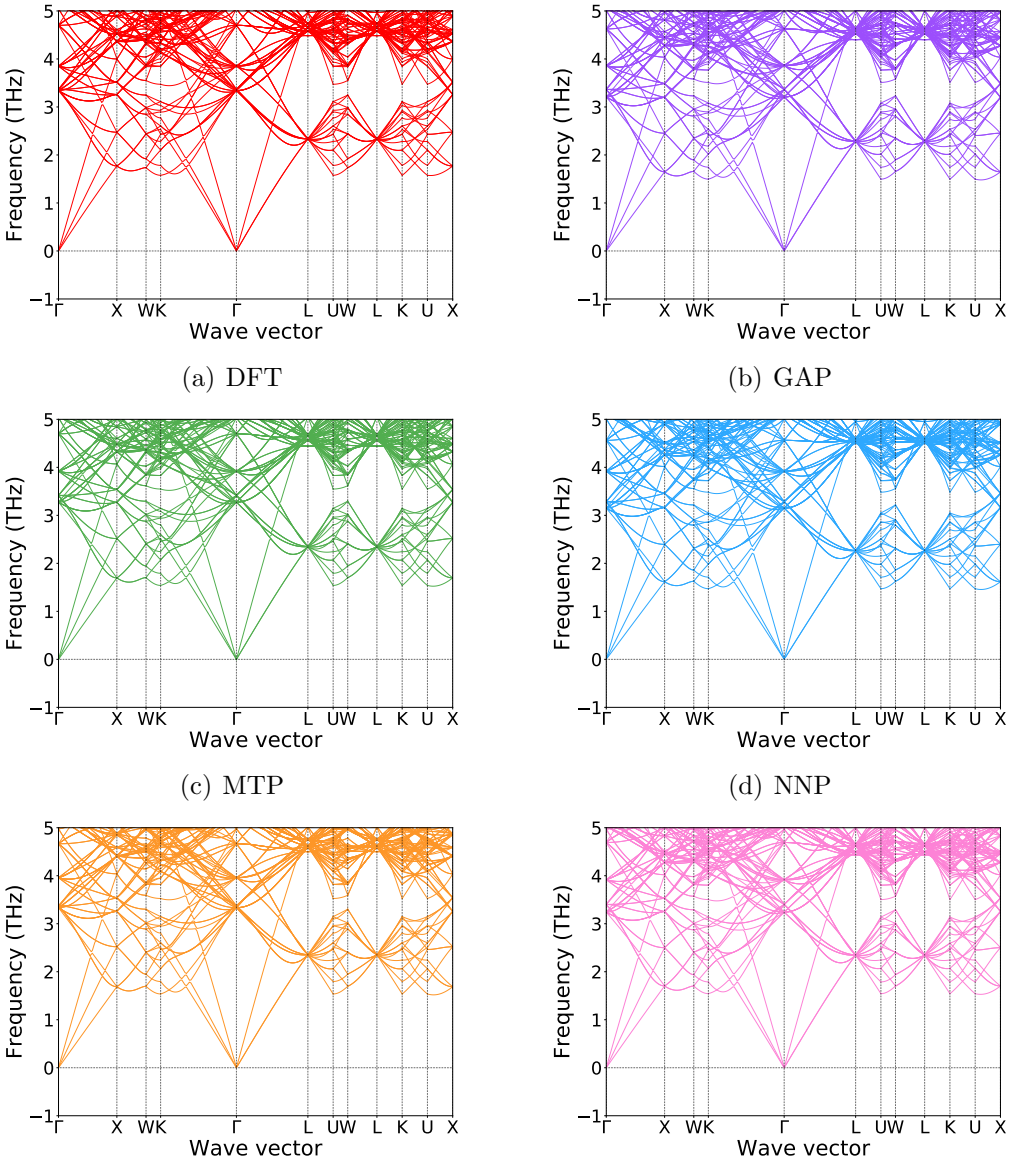
(a) DFT (b) GAP (c) MTP (d) NNP (e) SNAP (f) qSNAP

Figure S5: Phonon dispersion curves for 54-atom bcc Mo supercell via (a) DFT (b) GAP (c) MTP (d) NNP (e) SNAP (f) qSNAP.



(a) DFT (b) GAP (c) MTP (d) NNP (e) SNAP (f) qSNAP

Figure S6: Phonon dispersion curves for 54-atom bcc Li supercell via (a) DFT (b) GAP (c) MTP (d) NNP (e) SNAP (f) qSNAP.



(a) DFT

(b) GAP

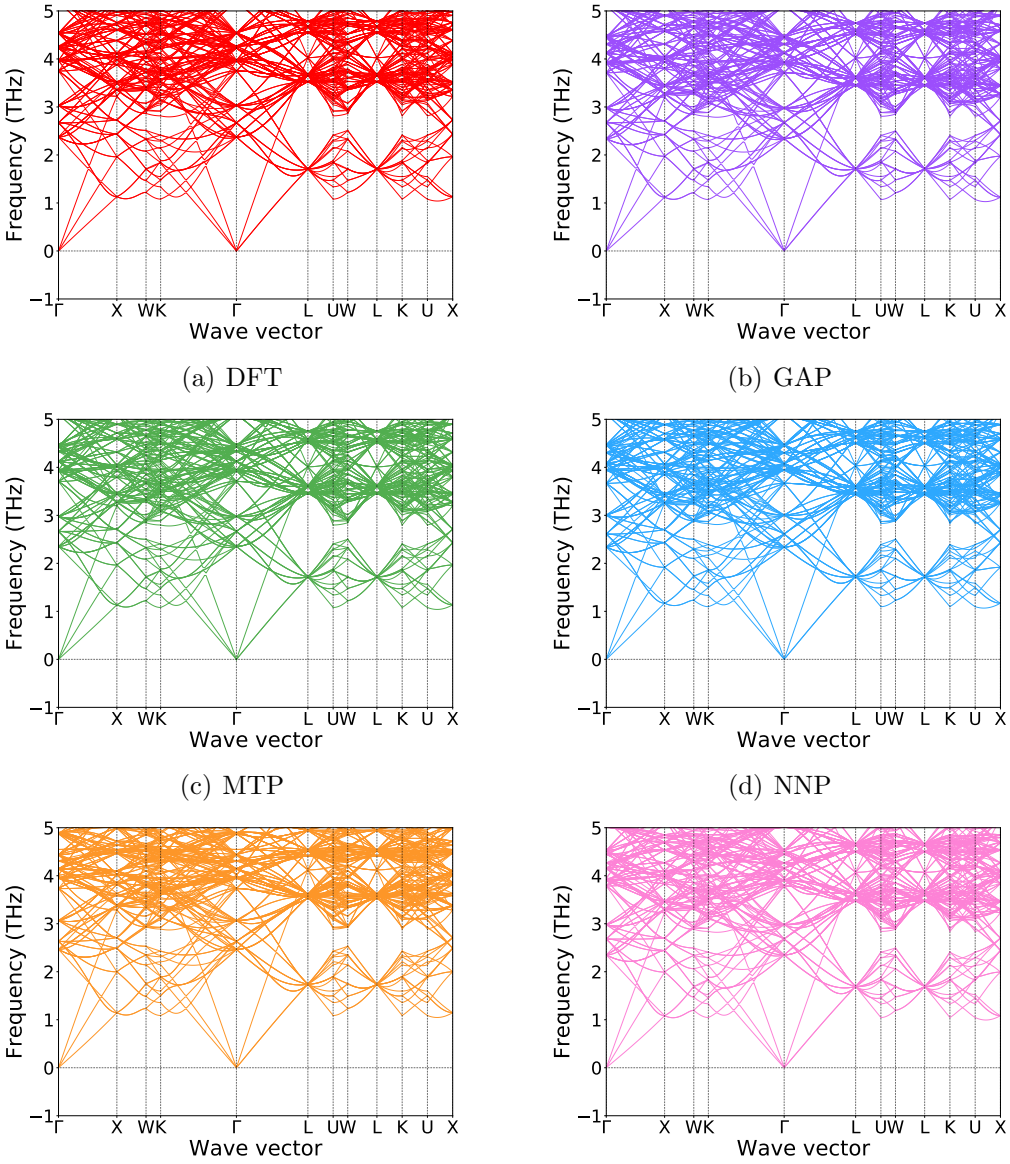
(c) MTP

(d) NNP

(e) SNAP

(f) qSNAP

Figure S7: Phonon dispersion curves for 108-atom fcc Ni supercell via (a) DFT (b) GAP (c) MTP (d) NNP (e) SNAP (f) qSNAP.



(a) DFT

(b) GAP

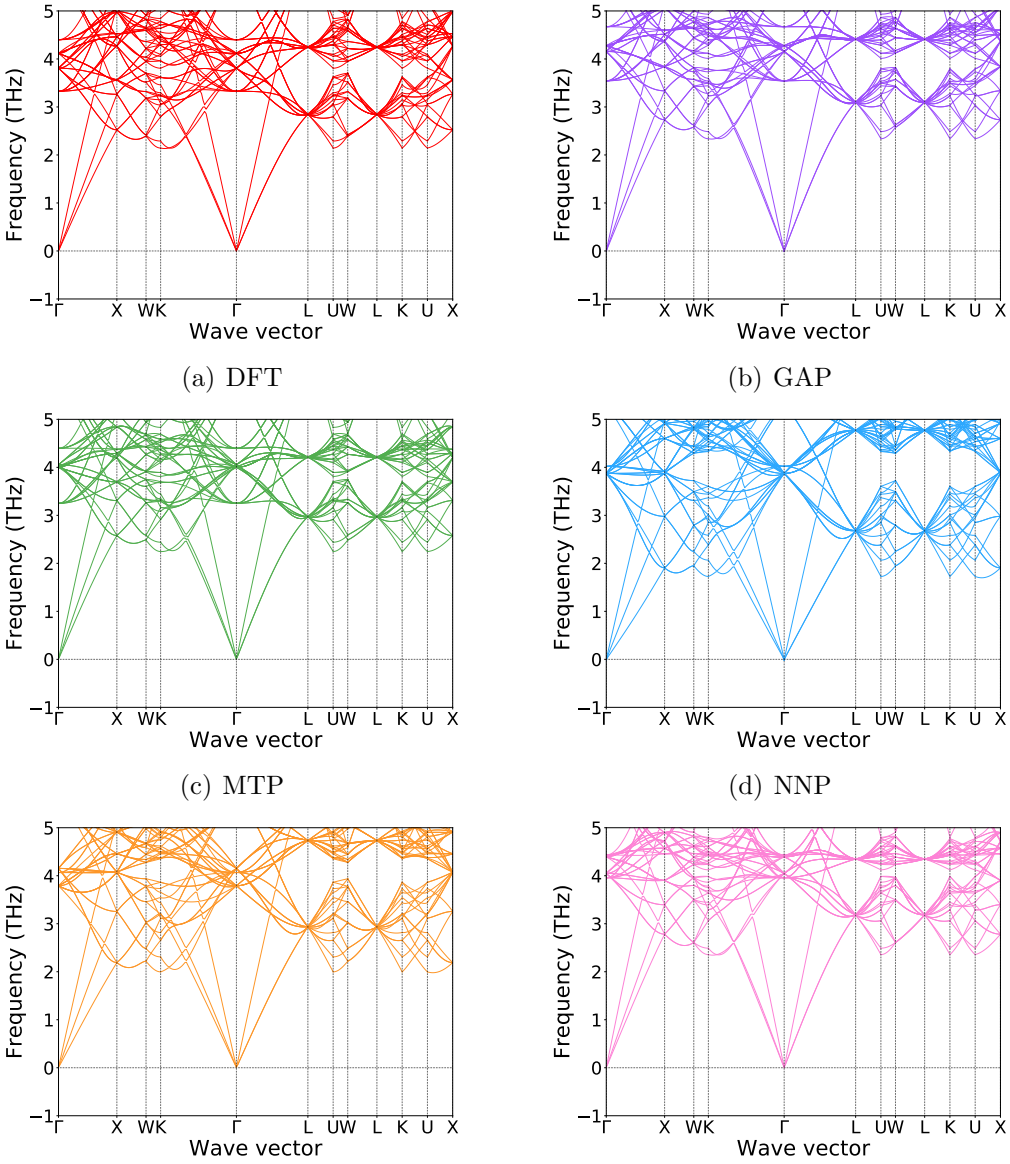
(c) MTP

(d) NNP

(e) SNAP

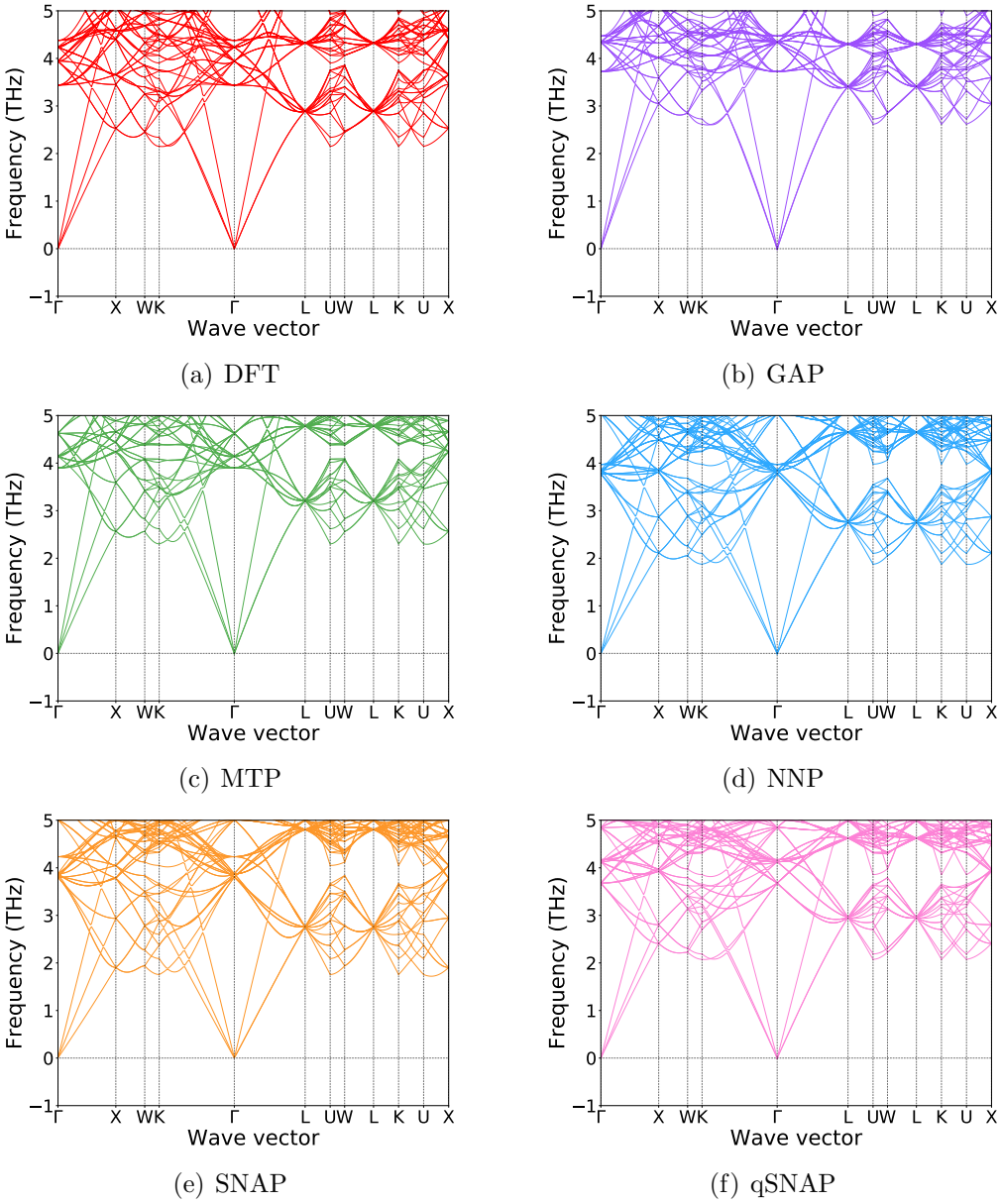
(f) qSNAP

Figure S8: Phonon dispersion curves for 108-atom fcc Cu supercell via (a) DFT (b) GAP (c) MTP (d) NNP (e) SNAP (f) qSNAP.



(a) DFT (b) GAP (c) MTP (d) NNP (e) SNAP (f) qSNAP

Figure S9: Phonon dispersion curves for 64-atom diamond Si supercell via (a) DFT (b) GAP (c) MTP (d) NNP (e) SNAP (f) qSNAP.



(a) DFT

(b) GAP

(c) MTP

(d) NNP

(e) SNAP

(f) qSNAP

Figure S10: Phonon dispersion curves for 64-atom diamond Ge supercell via (a) DFT (b) GAP (c) MTP (d) NNP (e) SNAP (f) qSNAP.

Trade-off between prediction accuracy and computational cost

All timings were performed using LAMMPS calculations on a single CPU core of Intel i7-6850k 3.6 GHz.

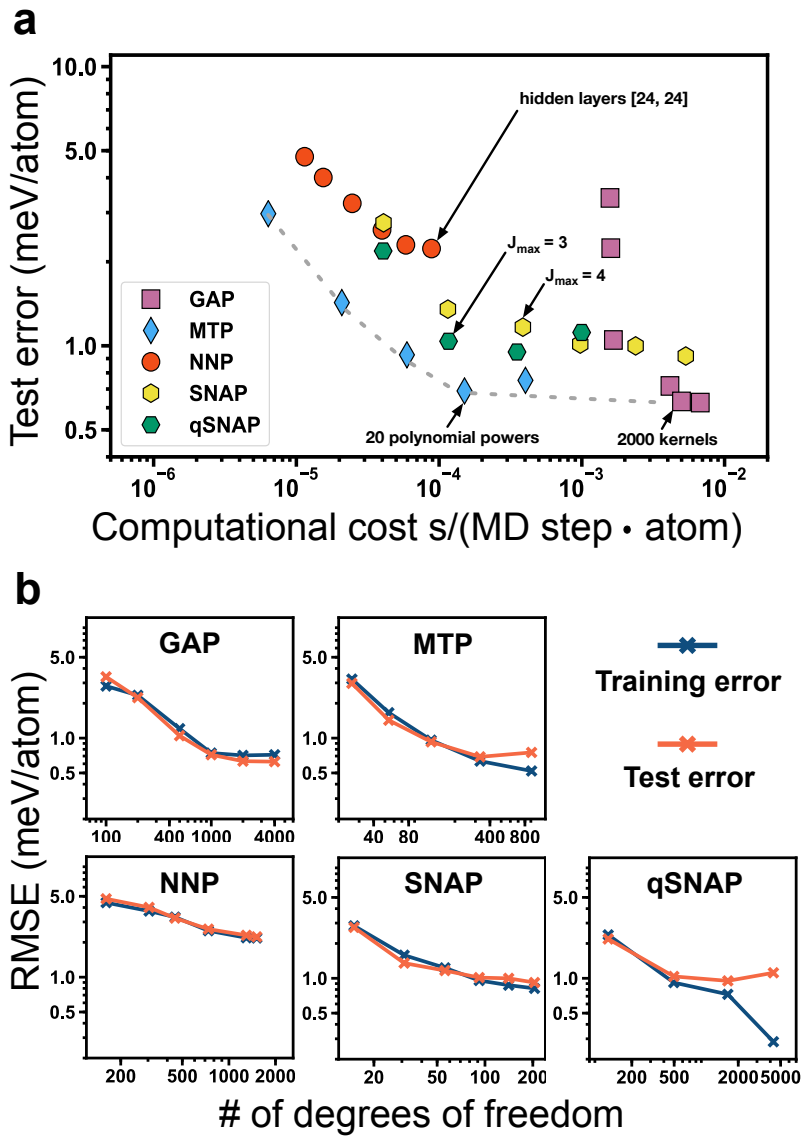


Figure S11: **a** Test error versus computational cost for fcc Ni system. An dashed Pareto frontier represents boundary enclosing points of different ML-IAPs. **b** Comparisons between training error and test error versus the number of degrees of freedom for each ML-IAP.

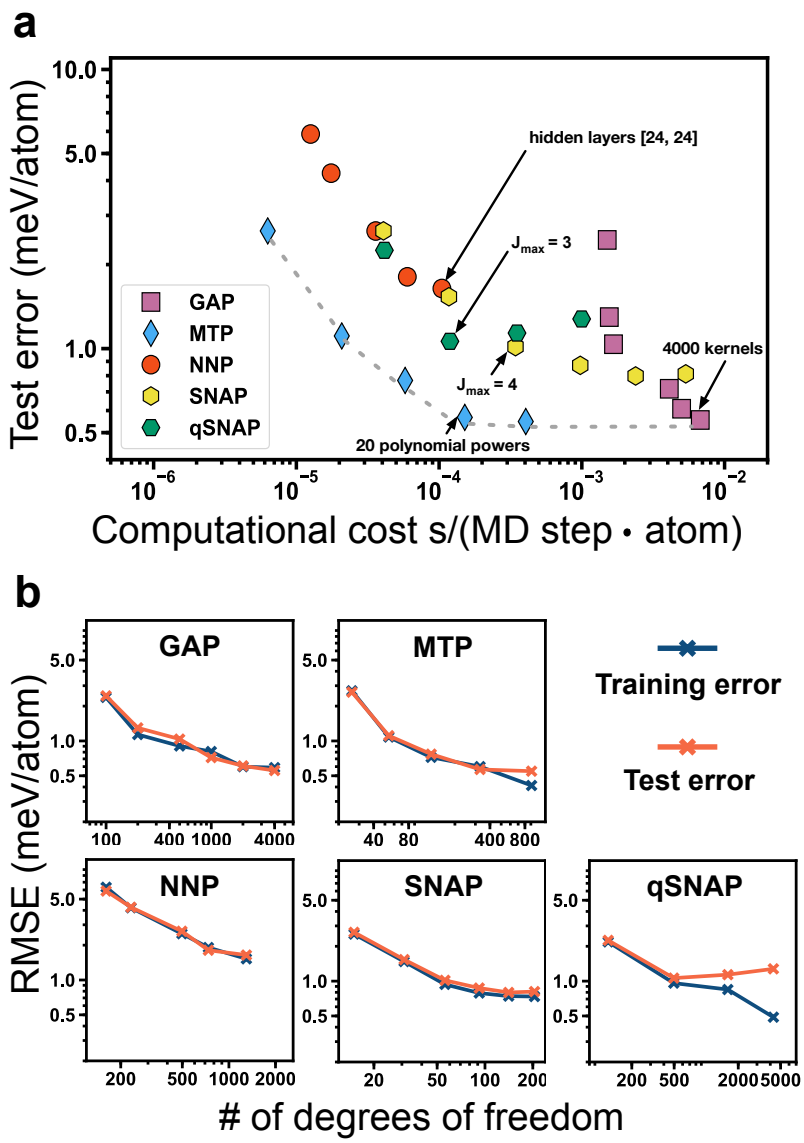


Figure S12: **a** Test error versus computational cost for fcc Cu system. An dashed Pareto frontier represents boundary enclosing points of different ML-IAPs. **b** Comparisons between training error and test error versus the number of degrees of freedom for each ML-IAP.

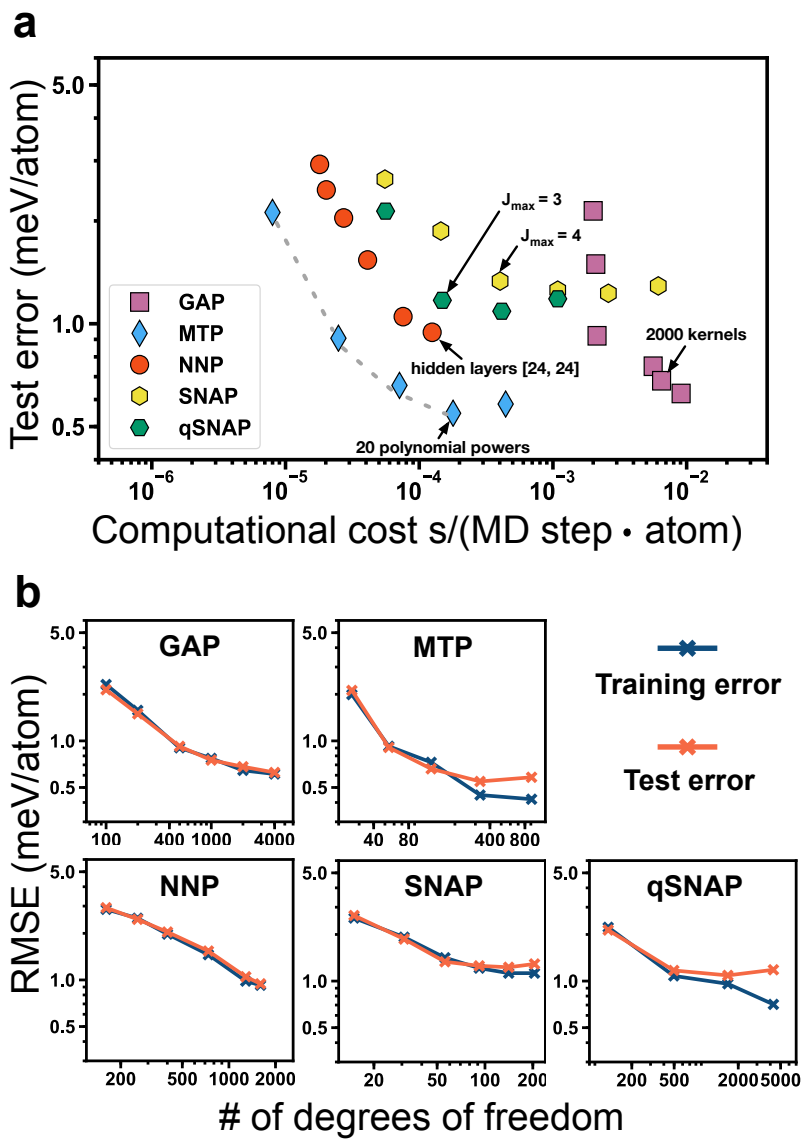


Figure S13: **a** Test error versus computational cost for bcc Li system. An dashed Pareto frontier represents boundary enclosing points of different ML-IAPs. **b** Comparisons between training error and test error versus the number of degrees of freedom for each ML-IAP.

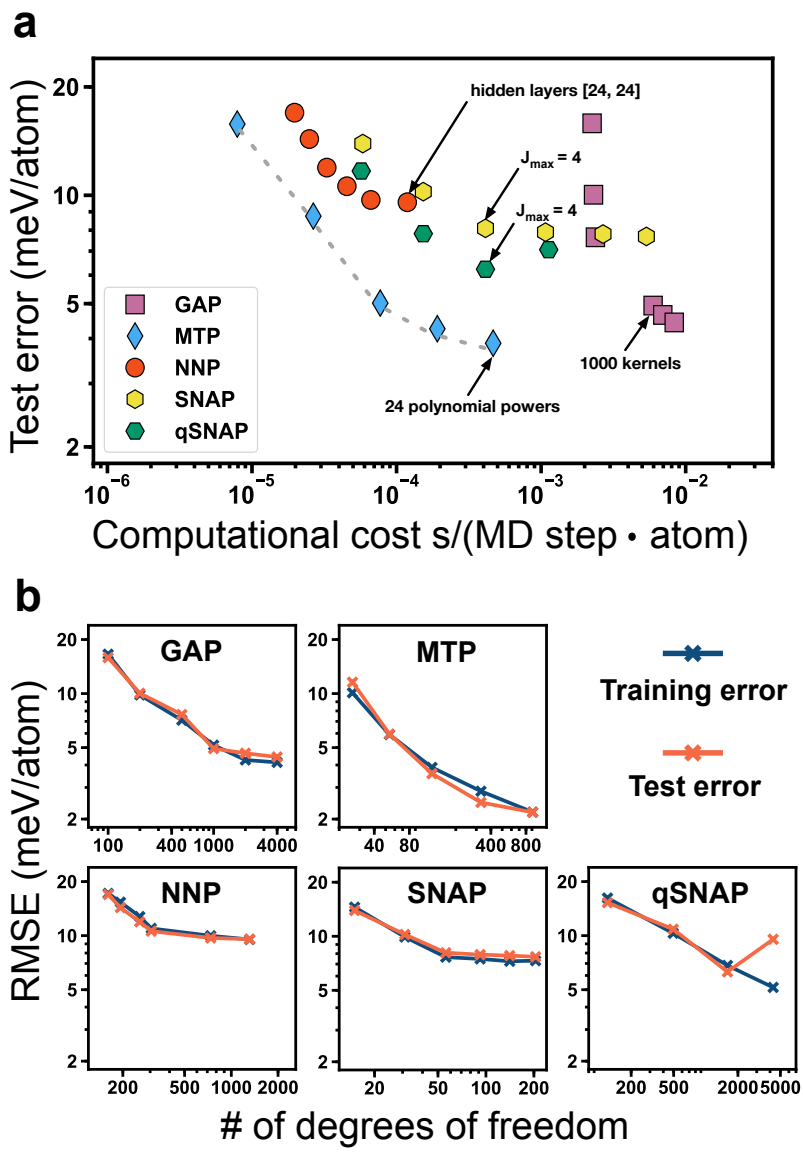


Figure S14: **a** Test error versus computational cost for diamond Si system. An dashed Pareto frontier represents boundary enclosing points of different ML-IAPs. **b** Comparisons between training error and test error versus the number of degrees of freedom for each ML-IAP.

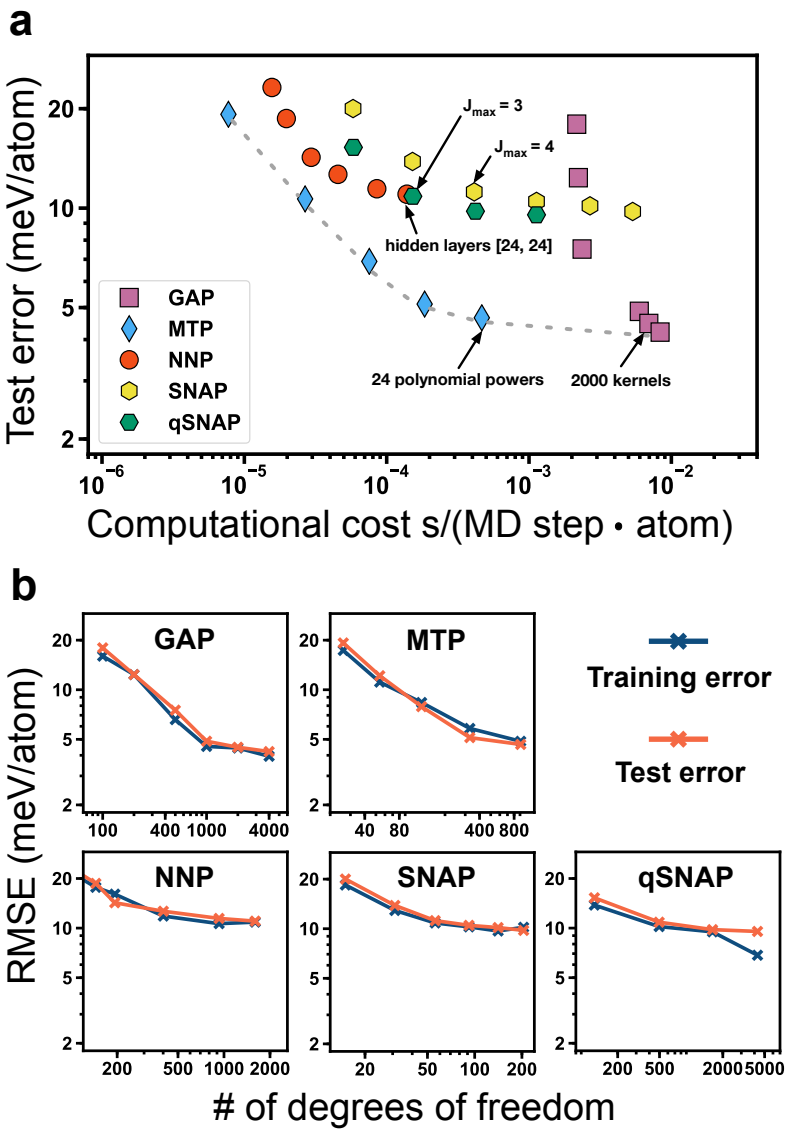


Figure S15: **a** Test error versus computational cost for diamond Ge system. An dashed Pareto frontier represents boundary enclosing points of different ML-IAPs. **b** Comparisons between training error and test error versus the number of degrees of freedom for each ML-IAP.

Linear scaling relationship of ML-IAPs

To investigate the scaling effect on the computational cost of different ML-IAPs, we perform npt MD simulations at a constant temperature of 300 K on structures with different size including $6 \times 6 \times 6$, $10 \times 10 \times 10$, $14 \times 14 \times 14$, and $18 \times 18 \times 18$ supercells, corresponding to 432, 2000, 5488, and 11664 atoms of bulk Mo, respectively. We also attempted to compare the relative computational cost between the ML-IAPs and classical IAPs. As is presented in Table S12, all ML-IAPs show greatly linear scaling relationship with respect to the number of atoms. Even though the MTP, NNP, SNAP, and qSNAP ML-IAPs demonstrate ~ 50 times higher computational cost than that of EAM, they have sufficiently low computational cost in terms of their accuracy compared to the classical IAPs.

Table S12: The scaling effect on computational cost of different ML-IAPs and classical IAPs. The unit of computational cost is seconds per MD step. Timings were performed by LAMMPS calculations on a single CPU core of Intel i7-6850k 3.6 GHz.

No. of atoms	432	2000	5488	11664
GAP	1.455	6.711	18.519	38.461
MTP	0.069	0.308	0.843	1.799
NNP	0.052	0.236	0.674	1.439
SNAP	0.063	0.294	0.809	1.701
qSNAP	0.063	0.299	0.805	1.730
EAM[29]	0.001	0.004	0.009	0.018
MEAM[32]	0.033	0.147	0.400	0.844

References

(1) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(2) Behler, J. Neural Network Potential-Energy Surfaces in Chemistry: A Tool for Large-Scale Simulations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17930.

(3) Behler, J. Constructing High-Dimensional Neural Network Potentials: A Tutorial Review. *Int. J. Quantum Chem.* **2015**, *115*, 1032–1050.

(4) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134*, 074106.

(5) Behler, J. Representing Potential Energy Surfaces by High-Dimensional Neural Network Potentials. *J. Phys.: Condens. Matter* **2014**, *26*, 183001.

(6) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.

(7) Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B* **2013**, *87*, 184115.

(8) Bartók, A. P.; Csányi, G. Gaussian Approximation Potentials: A Brief Tutorial Introduction. *Int. J. Quantum Chem.* **2015**, *115*, 1051–1057.

(9) Bartók, A. P.; Csányi, G. http://www.libatoms.org/ (accessed 2017).

(10) De, S.; Bartók, A. P.; Csányi, G.; Ceriotti, M. Comparing Molecules and Solids across Structural and Alchemical Space. *Phys. Chem. Chem. Phys.* **2016**, *18*, 13754–13769.

(11) Deringer, V. L.; Pickard, C. J.; Csányi, G. Data-Driven Learning of Total and Local Energies in Elemental Boron. *Phys. Rev. Lett.* **2018**, *120*, 156001.

(12) Bartók, A. P.; Kermode, J.; Bernstein, N.; Csányi, G. Machine Learning a General-Purpose Interatomic Potential for Silicon. *Phys. Rev. X* **2018**, *8*, 041048.

(13) Szlachta, W. J.; Bartók, A. P.; Csányi, G. Accuracy and Transferability of Gaussian Approximation Potential Models for Tungsten. *Phys. Rev. B* **2014**, *90*, 104108.

(14) Deringer, V. L.; Csányi, G. Machine Learning Based Interatomic Potential for Amorphous Carbon. *Phys. Rev. B* **2017**, *95*, 094203.

(15) Dragoni, D.; Daff, T. D.; Csányi, G.; Marzari, N. Achieving DFT Accuracy with a Machine-Learning Interatomic Potential: Thermomechanics and Defects in Bcc Ferromagnetic Iron. *Phys. Rev. Mater.* **2018**, *2*, 013808.

(16) Thompson, A.; Swiler, L.; Trott, C.; Foiles, S.; Tucker, G. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *J. Comput. Phys.* **2015**, 15.

(17) Wood, M. A.; Thompson, A. P. Quantum-Accurate Molecular Dynamics Potential for Tungsten. *arXiv:1702.07042* **2017**,

(18) Chen, C.; Deng, Z.; Tran, R.; Tang, H.; Chu, I.-H.; Ong, S. P. Accurate Force Field for Molybdenum by Machine Learning Large Materials Data. *Phys. Rev. Mater.* **2017**, *1*, 043603.

(19) Li, X.-G.; Hu, C.; Chen, C.; Deng, Z.; Luo, J.; Ong, S. P. Quantum-Accurate Spectral Neighbor Analysis Potential Models for Ni-Mo Binary Alloys and Fcc Metals. *Phys. Rev. B* **2018**, *98*, 094104.

(20) Deng, Z.; Chen, C.; Li, X.-G.; Ong, S. P. An Electrostatic Spectral Neighbor Analysis Potential for Lithium Nitride. *npj Comput. Mater.* **2019**, *5*, 75.

(21) Jones, E.; Oliphant, T.; Peterson, P. SciPy: Open source scientific tools for Python http://www.scipy.org/ (accessed 2001).

(22) Wood, M. A.; Thompson, A. P. Extending the Accuracy of the SNAP Interatomic Potential Form. *J. Chem. Phys.* **2018**, *148*, 241721.

(23) Foiles, S. M.; Baskes, M. I.; Daw, M. S. Embedded-Atom-Method Functions for the Fcc Metals Cu, Ag, Au, Ni, Pd, Pt, and Their Alloys. *Phys. Rev. B* **1986**, *33*, 7983–7991.

(24) Daw, M. S.; Baskes, M. I. Semiempirical, Quantum Mechanical Calculation of Hydrogen Embrittlement in Metals. *Phys. Rev. Lett.* **1983**, *50*, 1285–1288.

(25) Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Model. Simul.* **2016**, *14*, 1153–1173.

(26) Podryabinkin, E. V.; Shapeev, A. V. Active Learning of Linearly Parametrized Interatomic Potentials. *Comput. Mater. Sci.* **2017**, *140*, 171–180.

(27) Novoselov, I.; Yanilkin, A.; Shapeev, A.; Podryabinkin, E. Moment Tensor Potentials as a Promising Tool to Study Diffusion Processes. *Comput. Mater. Sci.* **2019**, *164*, 46–56.

(28) Gubaev, K.; Podryabinkin, E. V.; Hart, G. L. W.; Shapeev, A. V. Accelerating high-throughput searches for new alloys with active learning of interatomic potentials. *Comput. Mater. Sci.* **2019**, *156*, 148–156.

(29) Zhou, X. W.; Johnson, R. A.; Wadley, H. N. G. Misfit-Energy-Increasing Dislocations in Vapor-Deposited CoFe/NiFe Multilayers. *Phys. Rev. B* **2004**, *69*, 144113.

(30) Nichol, A.; Ackland, G. J. Property trends in simple metals: An empirical potential approach. *Phys. Rev. B* **2016**, *93*, 184101.

(31) Asadi, E.; Asle Zaeem, M.; Nouranian, S.; Baskes, M. I. Two-Phase Solid–Liquid Coexistence of Ni, Cu, and Al by Molecular Dynamics Simulations Using the Modified Embedded-Atom Method. *Acta Mater.* **2015**, *86*, 169–181.

(32) Park, H.; Fellinger, M. R.; Lenosky, T. J.; Tipton, W. W.; Trinkle, D. R.; Rudin, S. P.; Woodward, C.; Wilkins, J. W.; Hennig, R. G. *Ab Initio* Based Empirical Potential Used to Study the Mechanical Properties of Molybdenum. *Phys. Rev. B* **2012**, *85*, 214121.

(33) Lenosky, T. J.; Sadigh, B.; Alonso, E.; Bulatov, V. V.; de la Rubia, T. D.; Kim, J.; Voter, A. F.; Kress, J. D. Highly optimized empirical potential model of silicon. *Modell. Simul. Mater. Sci. Eng.* **2000**, *8*, 825–841.

(34) Kumagai, T.; Izumi, S.; Hara, S.; Sakai, S. Development of bond-order potentials that can reproduce the elastic constants and melting point of silicon for classical molecular dynamics simulation. *Comput. Mater. Sci.* **2007**, *39*, 457 – 464.

(35) Mahdizadeh, S. J.; Akhlamadi, G. Optimized Tersoff empirical potential for germanene. *J. Mol. Graphics Modell.* **2017**, *72*, 1 – 5.