

## Лекция 12

### Оценка производительности и вычислительной стоимости машинного обучения межатомных потенциалов

#### ПОДРОБНЫЙ ФОРМАЛИЗМ ML-IAPS И ОПТИМИЗИРОВАННЫЕ ПАРАМЕТРЫ

##### Нейронная сеть А. Белера-Парринелло

Потенциалы нейронной сети Белера-Парринелло (NNP) [1–3] отображают поверхность потенциальной энергии (PES) в терминах атомно-центрированных функций симметрии (ACSF) [4] с использованием полносвязных нейронных сетей с прямой связью.

ACSF преобразует локальное атомное окружение в числовые векторы, которые удовлетворяют вращательной и трансляционной инвариантности [4, 5]. Для атома  $i$  в структуре радиальный ACSF  $G_i^{\text{atom,rad}}$  который предоставляет информацию о парных корреляциях между атомами, задается следующим образом:

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

где  $\eta$  определяет ширину гауссова базиса, а  $R_s$  - сдвиг положения по всем соседние атомы в пределах радиуса обрезания  $R_c$ . Функция отсечки  $f_c$ , которая обеспечивает плавное уменьшение значения и наклона на радиусе отсечки  $R_c$ , задается следующим образом:

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

Угловая САУ имеет следующий вид:

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

где циклы суммирования по соседям  $j$  и  $k$ ,  $\theta_{ijk}$  - угол, а  $\zeta$ ,  $\eta^J$  и  $\lambda$  - три параметра, которые определяют форму функций угловой симметрии.

Функции симметрии, определяемые соответствующим выбором гиперпараметров ( $\eta$ ,  $R_s$  и т. д.), должны отражать эффективную взаимосвязь между представлениями атомов и соответствующими свойствами (например, компонентами энергии и силы). Алгоритм фильтра Калмана использовался для оптимизации весов нейронной сети, и был проведен поиск по сетке для оптимизации гиперпараметров, чтобы подобранная модель могла воспроизводить не только базовую энергию и значения силы, но и упругие постоянные. Оптимизированные параметры для всех элементных систем, изученных в этой работе, перечислены в Таблицах S1 - Таблицах S6.

ТАБЛИЦА S1: Параметры функции симметрии для оптимизированного Ni NNP. Единицы  $R_c$ ,  $R_s$ ,  $\eta$  и  $\eta^J$

суть  $^\circ A$ ,  $^\circ A$ ,  $^\circ A - 2$  и  $^\circ A - 2$  соответственно

$R_c$				3.9					
$G_i^{atom,rad}$				$G_i^{atom,ang}$					
Reference	Neighbor	$R_s$	$\eta$	Reference	Neighbor 1	Neighbor 2	$\lambda$	$\zeta$	$\eta'$
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
	Ni	Ni	Ni	-1	4	0.179
	Ni	Ni	Ni	-1	16	0.179

ТАБЛИЦА S2: Параметры функции симметрии для оптимизированного Cu NNP. Единицы  $R_c$ ,  $R_s$ ,  $\eta$  и  $\eta^J$ , суть  $^\circ\text{\AA}$ ,  $^\circ\text{\AA}$ ,  $^\circ\text{\AA}^{-2}$  и  $^\circ\text{\AA}^{-2}$  соответственно.

$R_c$				4.1					
$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	$R_s$	$\eta$	Reference	Neighbor 1	Neighbor 2	$\lambda$	$\zeta$	$\eta'$
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
				Cu	Cu	Cu	-1	4	0.179
				Cu	Cu	Cu	-1	16	0.179

ТАБЛИЦА S3: Параметры функции симметрии для оптимизированного Li NNP. Единицы  $R_c$ ,  $R_s$ ,  $\eta$  и  $\eta^J$ , суть  $^\circ\text{\AA}$ ,  $^\circ\text{\AA}$ ,  $^\circ\text{\AA}^{-2}$  и  $^\circ\text{\AA}^{-2}$  соответственно.

TABLE S3 – continued from previous page

$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	$R_s$	$\eta$	Reference	Neighbor 1	Neighbor 2	$\lambda$	$\zeta$	$\eta'$
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
				Li	Li	Li	1	4	3.571
				Li	Li	Li	1	16	3.571
				Li	Li	Li	-1	1	3.571

Continued on next page

TABLE S3 – continued from previous page

$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	$R_s$	$\eta$	Reference	Neighbor 1	Neighbor 2	$\lambda$	$\zeta$	$\eta'$
				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

ТАБЛИЦА S4: Параметры функции симметрии для оптимизированного Mo NNP. Единицы  $R_c$ ,  $R_s$ ,  $\eta$  и  $\eta'$ , суть  $^\circ\text{\AA}$ ,  $^\circ\text{\AA}$ ,  $^\circ\text{\AA}^{-2}$  и  $^\circ\text{\AA}^{-2}$  соответственно.

TABLE S4: Symmetry function parameters for optimized Mo NNP. The units of  $R_c$ ,  $R_s$ ,  $\eta$  and  $\eta'$  are  $\text{\AA}$ ,  $\text{\AA}$ ,  $\text{\AA}^{-2}$  and  $\text{\AA}^{-2}$ , respectively.

$R_c$				5.2					
$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	$R_s$	$\eta$	Reference	Neighbor 1	Neighbor 2	$\lambda$	$\zeta$	$\eta'$
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
				Mo	Mo	Mo	-1	4	0.179
				Mo	Mo	Mo	-1	16	0.179

ТАБЛИЦА S5: Параметры функции симметрии для оптимизированного Si NNP. Единицы  $R_c$ ,  $R_s$ ,  $\eta$  и  $\eta'$ , суть  $^\circ\text{\AA}$ ,  $^\circ\text{\AA}$ ,  $^\circ\text{\AA}^{-2}$  и  $^\circ\text{\AA}^{-2}$  соответственно.

TABLE S5: Symmetry function parameters for optimized Si NNP. The units of  $R_c$ ,  $R_s$ ,  $\eta$  and  $\eta'$  are  $\text{\AA}$ ,  $\text{\AA}$ ,  $\text{\AA}^{-2}$  and  $\text{\AA}^{-2}$ , respectively.

$R_c$				5.2					
$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	$R_s$	$\eta$	Reference	Neighbor 1	Neighbor 2	$\lambda$	$\zeta$	$\eta'$
Si	Si	0.0	0.036	Si	Si	Si	1	1	0.036
Si	Si	0.0	0.071	Si	Si	Si	1	1	0.071
Si	Si	0.0	0.179	Si	Si	Si	1	1	0.179
Si	Si	0.0	0.357	Si	Si	Si	1	1	0.357
Si	Si	0.0	0.714	Si	Si	Si	1	1	0.714
Si	Si	0.0	1.786	Si	Si	Si	1	1	1.786
Si	Si	0.0	3.571	Si	Si	Si	1	1	3.571
Si	Si	0.0	7.142	Si	Si	Si	1	1	7.142
Si	Si	0.0	17.855	Si	Si	Si	1	1	17.855
				Si	Si	Si	-1	1	0.036
				Si	Si	Si	-1	1	0.071
				Si	Si	Si	-1	1	0.179
				Si	Si	Si	-1	1	0.357
				Si	Si	Si	-1	1	0.714
				Si	Si	Si	-1	1	1.786
				Si	Si	Si	-1	1	3.571
				Si	Si	Si	-1	1	7.142
				Si	Si	Si	-1	1	17.855



ТАБЛИЦА S6: Параметры функции симметрии для оптимизированного Ge NNP. Единицы  $R_c$ ,  $R_s$ ,  $\eta$  и  $\eta'$ , суть  $^\circ\text{\AA}$ ,  $^\circ\text{\AA}$ ,  $^\circ\text{\AA}^{-2}$  и  $^\circ\text{\AA}^{-2}$  соответственно.

TABLE S6: Symmetry function parameters for optimized Ge NNP. The units of  $R_c$ ,  $R_s$ ,  $\eta$  and  $\eta'$  are  $\text{\AA}$ ,  $\text{\AA}$ ,  $\text{\AA}^{-2}$  and  $\text{\AA}^{-2}$ , respectively.

$R_c$				5.1					
$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	$R_s$	$\eta$	Reference	Neighbor 1	Neighbor 2	$\lambda$	$\zeta$	$\eta'$
Ge	Ge	0.0	0.036	Ge	Ge	Ge	1	1	0.036
Ge	Ge	1.0	0.036	Ge	Ge	Ge	-1	1	0.036
Ge	Ge	2.0	0.036	Ge	Ge	Ge	1	1	0.071
Ge	Ge	3.0	0.036	Ge	Ge	Ge	-1	1	0.071
Ge	Ge	4.0	0.036	Ge	Ge	Ge	1	1	0.179
Ge	Ge	0.0	0.071	Ge	Ge	Ge	-1	1	0.179
Ge	Ge	1.0	0.071	Ge	Ge	Ge	1	1	0.357
Ge	Ge	2.0	0.071	Ge	Ge	Ge	-1	1	0.357
Ge	Ge	3.0	0.071	Ge	Ge	Ge	1	1	0.714
Ge	Ge	4.0	0.071	Ge	Ge	Ge	-1	1	0.714
Ge	Ge	0.0	0.179	Ge	Ge	Ge	1	1	1.786
Ge	Ge	1.0	0.179	Ge	Ge	Ge	-1	1	1.786
Ge	Ge	2.0	0.179	Ge	Ge	Ge	1	1	3.571
Ge	Ge	3.0	0.179	Ge	Ge	Ge	-1	1	3.571
Ge	Ge	4.0	0.179	Ge	Ge	Ge	1	1	7.142
Ge	Ge	0.0	0.357	Ge	Ge	Ge	-1	1	7.142
Ge	Ge	1.0	0.357	Ge	Ge	Ge	1	1	17.855
Ge	Ge	2.0	0.357	Ge	Ge	Ge	-1	1	17.855
Ge	Ge	3.0	0.357						
Ge	Ge	4.0	0.357						
Ge	Ge	0.0	0.714						
Ge	Ge	1.0	0.714						
Ge	Ge	2.0	0.714						
Ge	Ge	3.0	0.714						
Ge	Ge	4.0	0.714						

Continued on next page



TABLE S6 – continued from previous page

$G_i^{\text{atom,rad}}$				$G_i^{\text{atom,ang}}$					
Reference	Neighbor	$R_s$	$\eta$	Reference	Neighbor 1	Neighbor 2	$\lambda$	$\zeta$	$\eta'$
Ge	Ge	0.0	1.786						
Ge	Ge	1.0	1.786						
Ge	Ge	2.0	1.786						
Ge	Ge	3.0	1.786						
Ge	Ge	4.0	1.786						
Ge	Ge	0.0	3.571						
Ge	Ge	1.0	3.571						
Ge	Ge	2.0	3.571						
Ge	Ge	3.0	3.571						
Ge	Ge	4.0	3.571						
Ge	Ge	0.0	7.142						
Ge	Ge	1.0	7.142						
Ge	Ge	2.0	7.142						
Ge	Ge	3.0	7.142						
Ge	Ge	4.0	7.142						
Ge	Ge	0.0	17.855						
Ge	Ge	1.0	17.855						
Ge	Ge	2.0	17.855						
Ge	Ge	3.0	17.855						
Ge	Ge	4.0	17.855						

## Б. Потенциалы гауссовой аппроксимации

Потенциалы приближения Гаусса (GAP) [6–8], реализованные в программном пакете QUIP [9], применяют регрессию гауссовского процесса (GPR) для интерполяции атомной энергии на основе пространственного распределения соседних атомов. Функция атомной энергии задается следующим образом:

$$\epsilon(\mathbf{R}) = \sum_{\mathbf{k}} b_k K(\mathbf{R}, \mathbf{R}_k), \quad (4)$$

где  $R$  представляет геометрию соседних атомов в пределах радиуса отсечки  $R_c$ , а  $k$  индексирует набор контрольных точек  $R_k$ , которые служат в качестве основы для расширения функции атомной энергии.  $K$  - это функция ядра, которая фиксирует изменение функции энергии с точки зрения изменения соседних конфигураций.

В этой работе используется функция ядра - «Smooth Overlap of Atomic Positions» (SOAP) [7, 10–12], которая представляет интегрированное по вращению перекрытие плотностей соседей. Атомная конфигурация центрального атома  $i$  представлена его плотностью соседей в виде суммы

гауссианов с центрами над всеми соседними атомами  $j$  в пределах расстояния отсечки  $R_c$ ,

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

где  $f_c$  - функция отсечки, обеспечивающая плавный спад значения и наклона на радиусе отсечки  $R_c$ , а  $\sigma_{\text{atom}}$  - параметр размытия. Эта плотность соседей затем разлагается на основе сферических гармоник  $Y_{lm}(\hat{\mathbf{R}})$  и радиальных функций  $g_n(R)$ :

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

и вращательно-инвариантный сферический спектр мощности атома  $i$  затем выражается через коэффициенты разложения  $c_{nlm}$  следующим образом:

$$p_{n_1 n_2 l}(\mathbf{R}_i) = \sum_{m=-l}^l c_{n_1 l m}^* c_{n_2 l m}, \quad (7)$$

Наконец, ядро SOAP записывается как скалярное произведение спектров мощности и возводится в малую целую степень следующим образом:

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

	Ni	Cu	Li	Mo	Si	Ge
cutoff radius (Å)	3.9	4.1	4.8	5.2	5.1	5.4
$\sigma_{\text{energy}}$ (eV/atom)	0.01	0.01	0.01	0.01	0.01	0.01
$\sigma_{\text{force}}$ (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)

ТАБЛИЦА S7. Ключевые параметры модели GAP оптимизированы для каждой элементарной системы

где  $\zeta = 4$  в настоящей работе. Затем выполняется нормализация, чтобы гарантировать, что ядро каждой атомарной среды с самим собой является единым целым.

Выбор гиперпараметров, таких как радиус обрезания и параметры регуляризации в линейной системе, оказывает значительное влияние на производительность моделей GAP [13, 14]. Оптимизированные параметры для каждой элементарной системы, определенные с помощью поиска по сетке, перечислены в Таблице S7.

### С. Потенциал спектрального анализа соседей

Биспектр формализма плотности соседей и потенциала спектрального анализа соседей (SNAP) подробно изучался в предыдущих работах [7, 15]. Основная концепция формализма биспектров состоит в проецировании трехмерной локальной плотности атомных соседей в набор коэффициентов, которые удовлетворяют инвариантным свойствам, путем разложения на основе сферических гармоник.

В SNAP плотность соседей вокруг центрального атома  $i$  в позиции  $\mathbf{R}$  определяется

следующим образом:

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

где  $\mathbf{R}_{ij}$  - положение соседнего атома  $j$  относительно центрального атома  $i$ , а  $\omega_j$  - безразмерный вес для различения типов атомов. Функция отсечки  $f_c$

обеспечивает плавное уменьшение плотности соседних атомов до нуля на радиусе отсечки  $R_c$ .

Угловое распределение функции плотности соседей можно спроецировать на сферическую гармоническую функцию  $Y_m^l(\theta, \phi)$ . В биспектральном подходе радиальное распределение преобразуется в дополнительный полярный угол  $\theta_0$ , определяемый как  $\theta_0 = \theta_0^{\max} \frac{r}{R_c}$ . Таким образом, функция плотности может быть представлена в координатах 3-сферы  $(\theta, \phi, \theta_0)$  вместо  $(\theta, \phi, r)$ . Функция плотности на 3-сфере может быть расширена с помощью 4-мерных гиперсферических гармоник  $U_{m,m'}^j$ , в качестве следует:

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

где коэффициенты  $u_{m,m'}^j$  получаются как скалярное произведение функции плотности соседей с базовой функцией, заданной следующим образом:

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

Компоненты биспектра  $B_{j_1,j_2,j}$  могут быть получены следующим образом:

$$B_{j_1,j_2,j} = \sum_{m_1,m'_1=-j_1}^{j_1} \sum_{m_2,m'_2=-j_2}^{j_2} \sum_{m,m'=-j}^j (u_{m,m'}^j)^* \times H_{j_1 m_1 m'_1}^{j m m'} u_{m_1,m'_1}^{j_1} u_{m_2,m'_2}^{j_2}, \quad (12)$$

где константы  $H_{j_1 m_1 m'_1}^{j m m'}$  коэффициенты связи, удовлетворяющие условиям:

$$\|j_1 - j_2\| \leq j \leq \|j_1 + j_2\|.$$

В формализме SNAP энергия  $E_{\text{SNAP}}$ , сила  $\mathbf{F}_{\text{SNAP}}^j$ , и stresses  $\sigma_{\text{SNAP}}^j$

выражены как линейные функции компонент биспектра [15, 16] следующим образом:

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

$$F_{\text{SNAP}}^j = -\beta \cdot \sum_{i=1}^N \frac{\partial B^i}{\partial r_j}, \quad (14)$$

$$\sigma_{\text{SNAP}}^j = -\beta \cdot \sum_{i=1}^N r_j \otimes \sum_{i=1}^N \frac{\partial B^i}{\partial r_j}, \quad (15)$$

где  $\beta_0$  и вектор  $\beta$  - коэффициенты, полученные при согласовании с базой данных расчетов квантовой электронной структуры.

В наших предыдущих работах [17–19] мы разработали надежный чередующийся двухэтапный процесс подбора для получения оптимальных гиперпараметров для SNAP. На каждой внутренней итерации выполняется подгонка линейной модели. Во внешнем цикле модель, подобранная во внутренней итерации, затем используется для прогнозирования свойств материалов (например, тензоров упругости), а различия между прогнозируемыми и эталонными значениями затем используются для оптимизации гиперпараметров с помощью дифференциала.

алгоритм эволюционной оптимизации реализован в пакете Scipy [20].

Оптимизированные гиперпараметры и коэффициенты модели SNAP для всех элементарных систем представлены в таблице S8.

TABLE S8: Key parameters of the SNAP model optimized for each elemental system

				Ni	Cu	Li	Mo	Si	Ge
cutoff radius (Å)				3.9	3.8	5.1	5.2	4.9	5.0
$k$	$2j_1$	$2j_2$	$2j$	$\beta_k^{\text{Ni}}$	$\beta_k^{\text{Cu}}$	$\beta_k^{\text{Li}}$	$\beta_k^{\text{Mo}}$	$\beta_k^{\text{Si}}$	$\beta_k^{\text{Ge}}$
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

TABLE S8 – continued from previous page

				Ni	Cu	Li	Mo	Si	Ge
cutoff radius (Å)				3.9	4.1	5.1	4.6	4.9	5.5
$k$	$2j_1$	$2j_2$	$2j$	$\beta_k^{\text{Ni}}$	$\beta_k^{\text{Cu}}$	$\beta_k^{\text{Li}}$	$\beta_k^{\text{Mo}}$	$\beta_k^{\text{Si}}$	$\beta_k^{\text{Ge}}$
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
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

---



TABLE S8 – continued from previous page

				Ni	Cu	Li	Mo	Si	Ge
cutoff radius ( $\text{\AA}$ )				3.9	4.1	5.1	4.6	4.9	5.5
$k$	$2j_1$	$2j_2$	$2j$	$\beta_k^{\text{Ni}}$	$\beta_k^{\text{Cu}}$	$\beta_k^{\text{Li}}$	$\beta_k^{\text{Mo}}$	$\beta_k^{\text{Si}}$	$\beta_k^{\text{Ge}}$
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
55	8	8	8	0.0008	0.0020	0.0003		−0.0034	−0.0014

## D. Квадратичный спектральный потенциал анализа соседей

В первоначальной формулировке SNAP линейная форма отношения PES к биспектру ограничивает максимальную сложность энергетических функций до четырехчастичного эффекта, который, в свою очередь, может влиять на

его предсказательную способность. Недавно Вуд и Томпсон [21] предложили квадратичное расширение подхода SNAP (qSNAP). Квадратичные вклады в энергию можно рассматривать как своего рода энергию погружения по аналогии с методом погруженного атома (EAM) [22, 23]. В qSNAP потенциал SNAP расширяется за счет добавления члена энергии внедрения следующим образом:

$$E_{\text{SNAP}}^i = \beta \cdot B^i + F(\rho_i), \quad (16)$$

где  $F(\rho_i)$  представляет собой энергию погружения атома  $i$  в электронную плотность, вносимую его соседними атомами. «Основная» электронная плотность погружающегося атома  $i$  может быть выражена как линейная функция компонентов биспектра следующим образом:

$$\rho_i = a \cdot B^i, \quad (17)$$

а энергия погружения может быть выражена как разложение Тейлора на основе эталонной структуры с плотностью  $\rho_0$  следующим образом:

$$F(\rho) = F_0 + (\rho - \rho_0)F' + \frac{1}{2}(\rho - \rho_0)^2 F'' + \dots \quad (18)$$

Таким образом, модифицированная энергия SNAP определяется следующим выражением:

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

где  $\alpha = F''a \otimes a$  - симметричная матрица размера  $K \times K$ . По сути, квадратичное расширение переносит все различные попарные произведения компонентов биспектра и расширяет максимальную сложность энергетических функций до эффектов семи тел [21].

В этой работе для моделей SNAP и qSNAP используется одинаковый двухэтапный подход. Оптимизированные гиперпараметры qSNAP для всех элементарных систем перечислены в таблице S9.

	Ni	Cu	Li	Mo	Si	Ge
cutoff radius (Å)	3.8	3.9	5.1	5.2	4.8	4.9
$J_{\max}$	6	6	6	6	8	6

ТАБЛИЦА S9. Ключевые параметры модели qSNAP, оптимизированные для каждой элементарной системы

#### Д. Тензорные потенциалы моментов

Модель тензорного потенциала момента (МТП) и ее формализм изучались в предыдущих работах [24–26]. Фундаментальная идея МТП состоит в том, чтобы построить сжатое вращательно-инвариантное представление локальной атомной среды в тензорном смысле и построить линейную корреляцию между потенциальной энергией и атомным представлением на основе предположения, что полная энергия может быть разделена на отдельные атомные среды. взносы.

Как было указано ранее [24, 25], потенциальная энергия атомного окружения центрального атома  $i$  может быть линейно разложена на набор базисных функций  $B(R)$ ,

$$V_i(R) = \sum_l \beta_l B_l(R), \quad (20)$$

Базисные функции  $B(R)$ , в свою очередь, зависят от ряда дескрипторов тензора моментов по всем соседним атомам  $j$ :

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

где функции  $f_{\mu}$  представляют собой радиальное распределение атомной конфигурации, а члены  $R_{ij} \otimes$

$\cdots \otimes R_{ij}$  - тензоры ранга  $\nu$ , несущие угловую информацию об атомной конфигурации. Эти дескрипторы инвариантны относительно вращения в тензорном смысле (или, точнее, ковариантны относительно вращения).

Выбор  $\mu$  и  $\nu$  обеспечивает баланс между вычислительной сложностью и вычислительной эффективностью МТР [24]. Базисные функции  $V(R)$  формулируются разными способами

сжатия тензоров момента  $M_{\mu, \nu}$  до скаляра. Каждое сжатие может быть закодировано симметричной матрицей  $\alpha$   $m \times m$ , где диагональные элементы  $\alpha_{ij}$  демонстрируют размерности  $i$ -го и  $j$ -го тензоров, которые сжимаются. В этой работе мы использовали реализацию, описанную в предыдущей работе [27], которая для одного компонента сводится к изучению оптимального радиального

функции  $f_{\mu}$  вместо того, чтобы фиксировать их как универсальный радиальный базис.

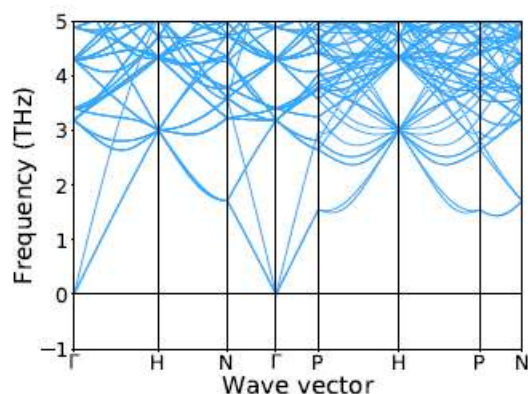
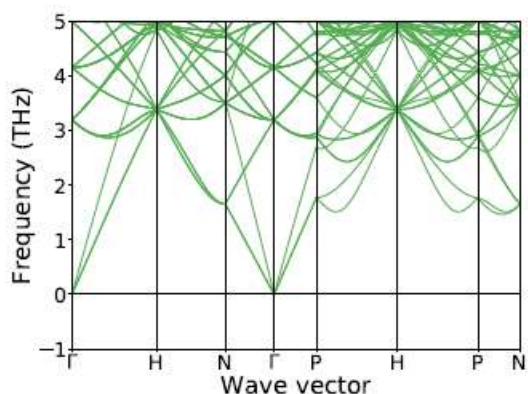
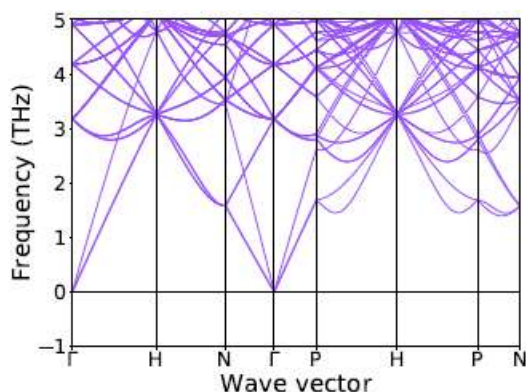
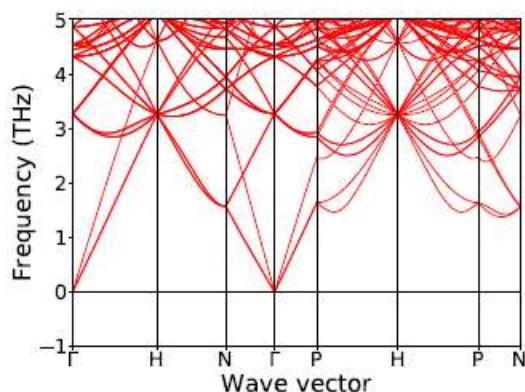
	Ni	Cu	Li	Mo	Si	Ge
cutoff radius (Å) радиус отсечения	4.0	3.9	5.1	5.2	4.7	5.1
# of polynomial powers	20	20	16	20	24	24
К-во степеней полинома						
# of free parameters	329	329	125	329	913	913
К-во бесплатных параметров						
Energy weight	1.00	1.00	1.00	1.00	1.00	1.00
Энергетический вес						
Force weight	0.01	0.01	0.01	0.01	0.01	0.01
Силовая нагрузка						

ТАБЛИЦА S10. Ключевые параметры модели МТР оптимизированы для каждой элементарной системы

Мы внимательно изучили выбор гиперпараметров, включая радиус отсечения, количество полиномиальных степеней, а также количество свободных параметров путем оценки производительности при воспроизведении основных свойств, например, упругих констант и фононного спектра, в то время как взвешенные параметры менее важны для прогнозируемых свойств. Полный набор оптимизированных

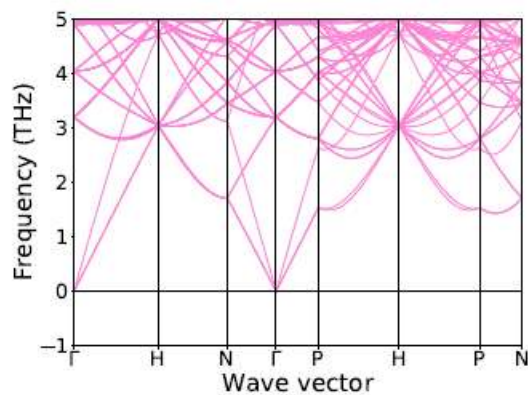
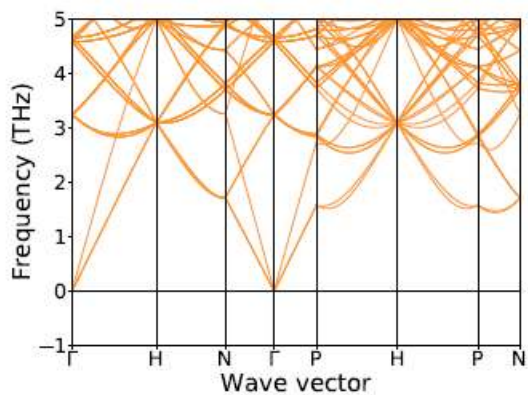
гиперпараметров для каждой элементарной системы представлен в таблице S10. Максимальное количество итераций было установлено равным 500, 1000 и 2000, чтобы гарантировать сходимость процесса аппроксимации для 16 или более низких степеней полинома, 20 полиномиальных степеней и 24 полиномиальных степеней соответственно.

## II. КРИВЫЕ ДИСПЕРСИИ ФОНОНОВ, ВЫЧИСЛЕННЫЕ DFT, GAP, MTP, NNP, SNAP И QSNAP



(c)MTP

(d)NNP

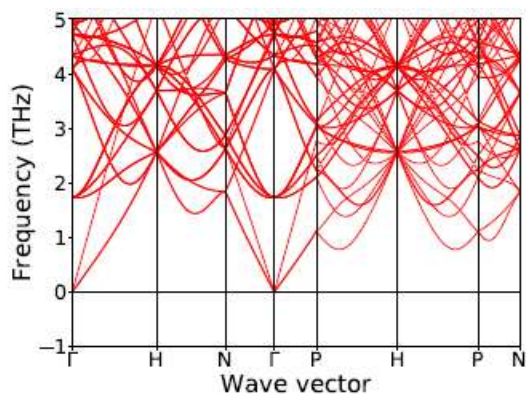


(e)SNAP

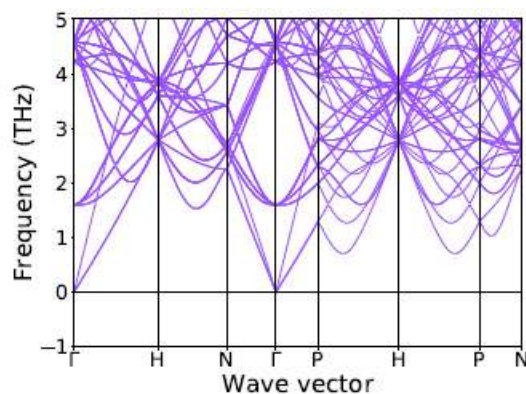
(f)qSNAP



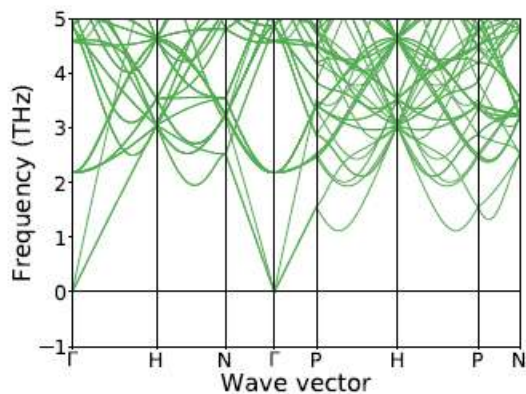
РИС. S1. Кривые дисперсии фононов для 54-атомной сверхъячейки ОЦК Мо с помощью (а) DFT (б) GAP (в) MTP (д) NNP (е) SNAP (ф) qSNAP.



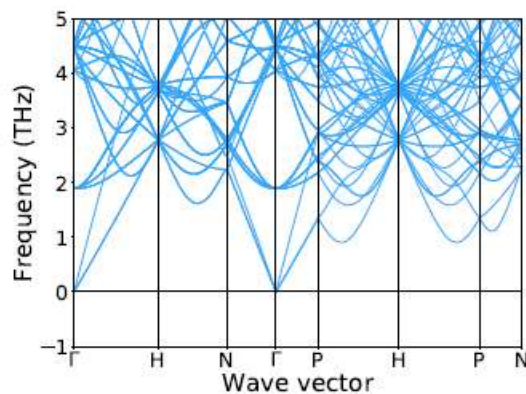
(a)DFT



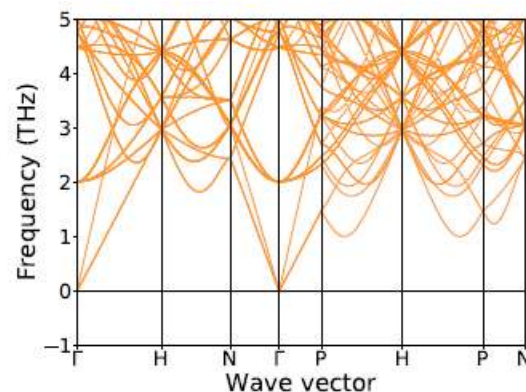
(b)GAP



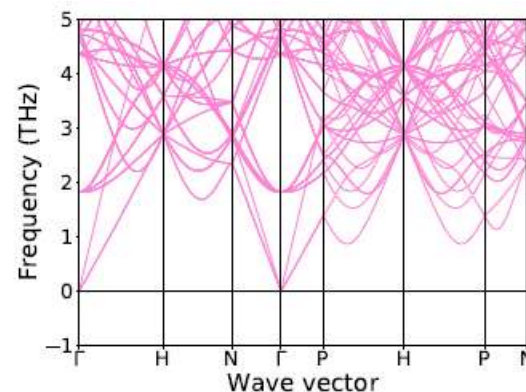
(c)MTP



(d)NNP

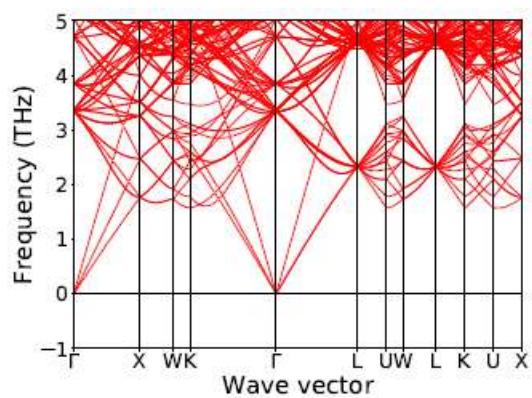


(e)SNAP

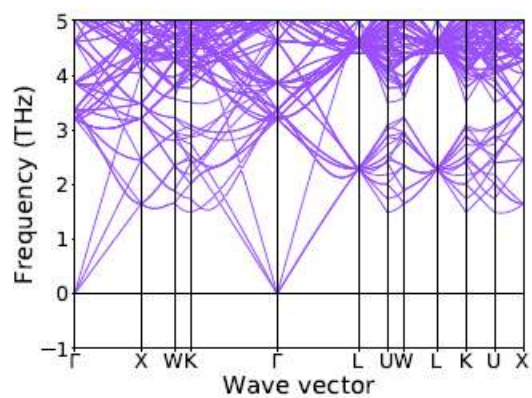


(f)qSNAP

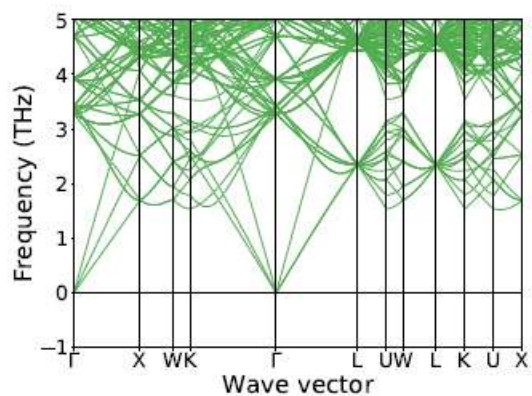
РИС. S2. Кривые дисперсии фононов для 54-атомной ОЦК-сверхъячейки Li с помощью (а) DFT (б) GAP (в) MTP (д) NNP (е) SNAP (ф) qSNAP



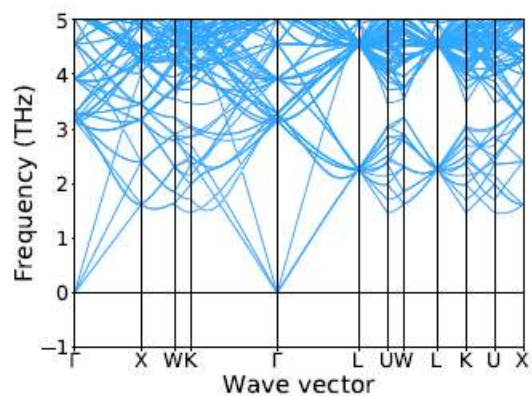
(a)DFT



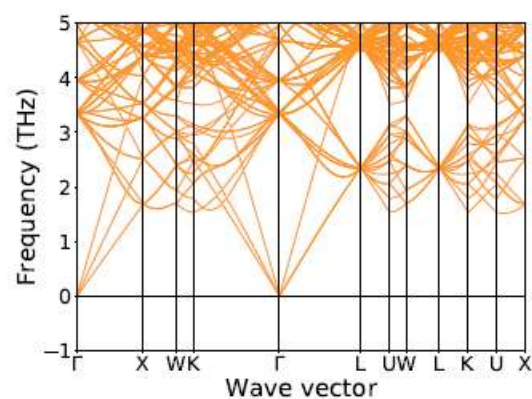
(b)GAP



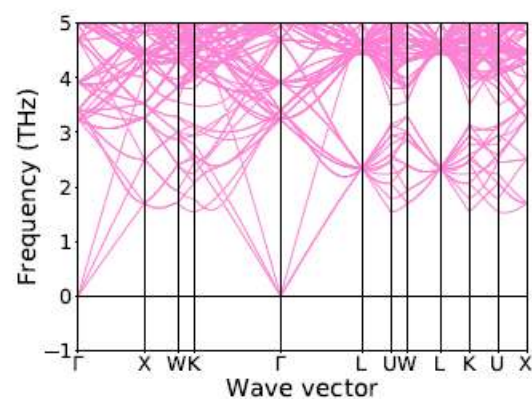
(c)MTP



(d)NNP



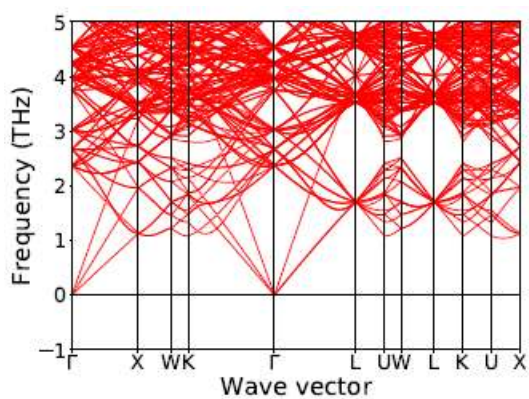
(e)SNAP



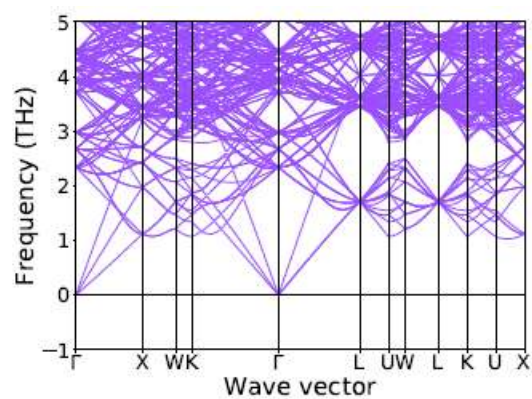
(f)qSNAP

РИС. S3. Кривые дисперсии фононов для 108-атомной сверхъядчейки с ГЦК-Ni (a) DFT (b) GAP (c) MTP (d) NNP (e) SNAP (f) qSNAP

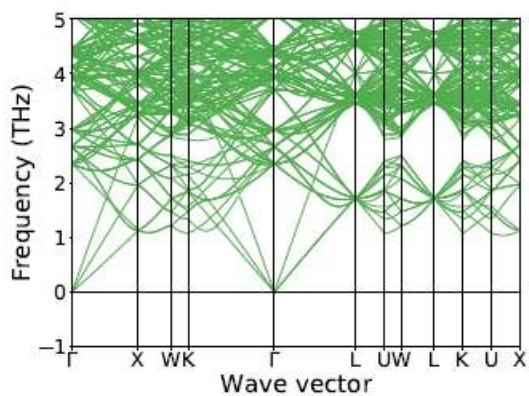




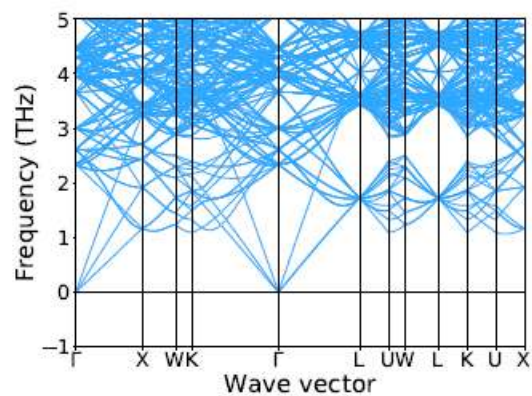
(a) DFT



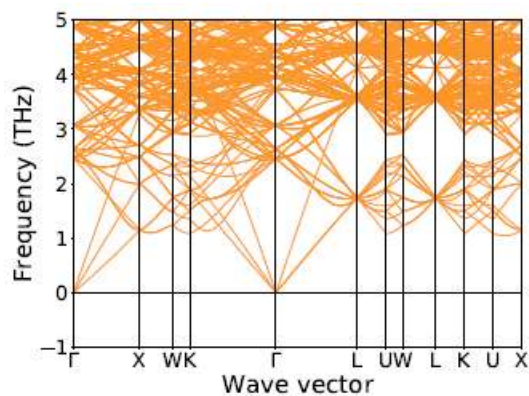
(b) GAP



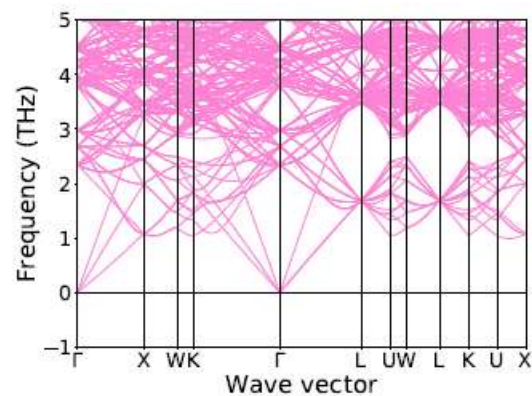
(c) MTP



(d) NNP

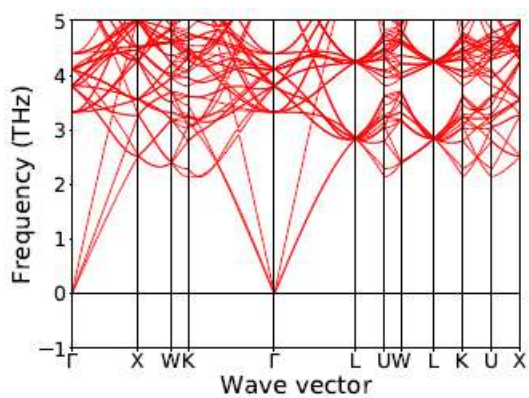


(e) SNAP

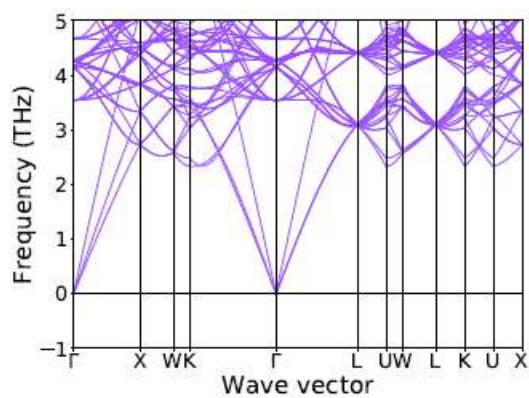


(f) qSNAP

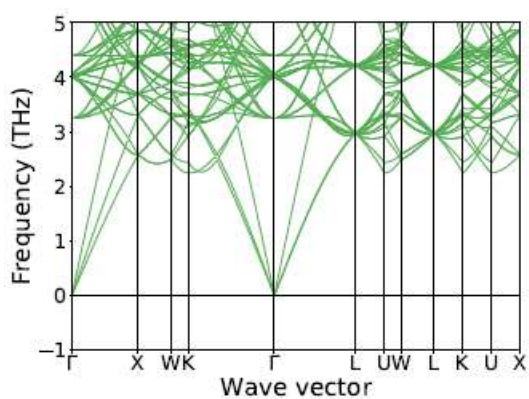
РИС. S4. Кривые дисперсии фононов для 108-атомной сверхъячейки Cu с ГЦК-ячейкой по (a) DFT (б) GAP (в) MTP (d) NNP (e) SNAP (f) qSNAP



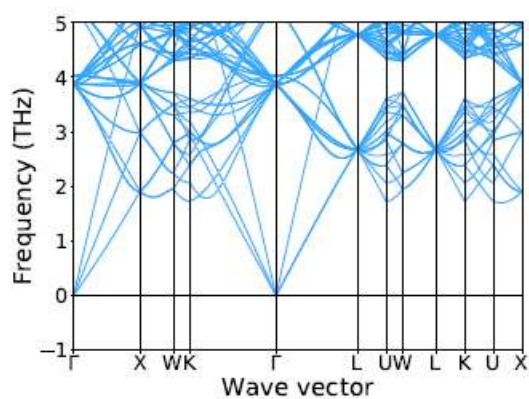
(a)DFT



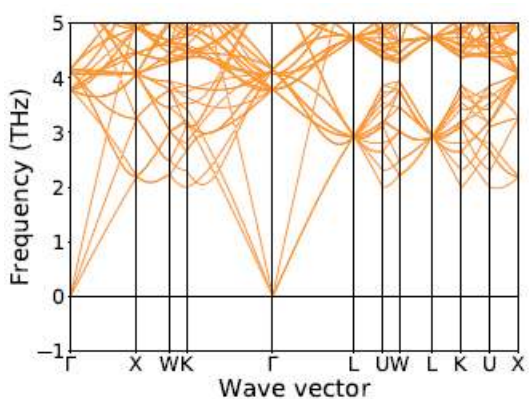
(b)GAP



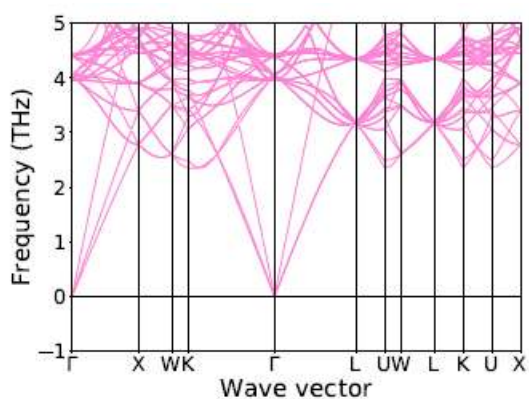
(c)MTP



(d)NNP



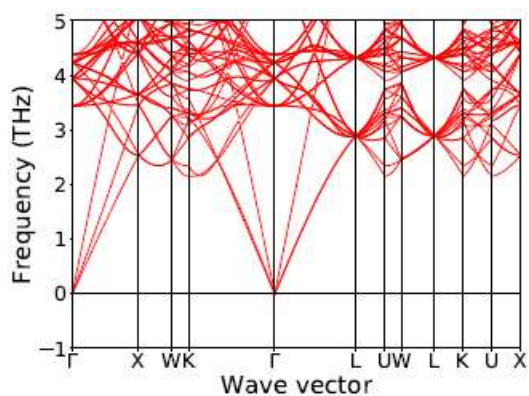
(e)SNAP



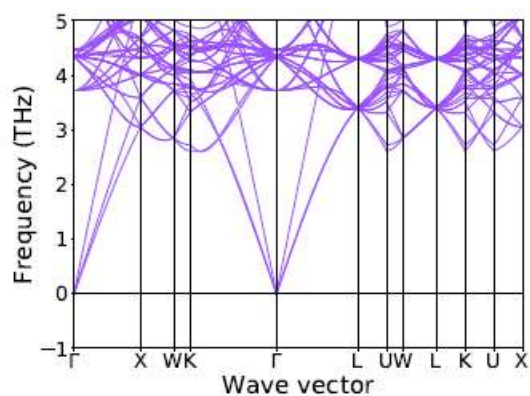
(f)qSNAP

РИС. S5. Кривые дисперсии фононов для сверхъячейки Si с 64 атомами алмаза с помощью (а) DFT (б) GAP (в) MTP (д) NNP (е) SNAP (ф) qSNAP.

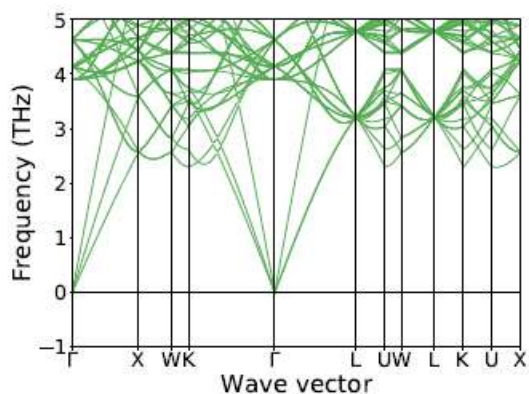




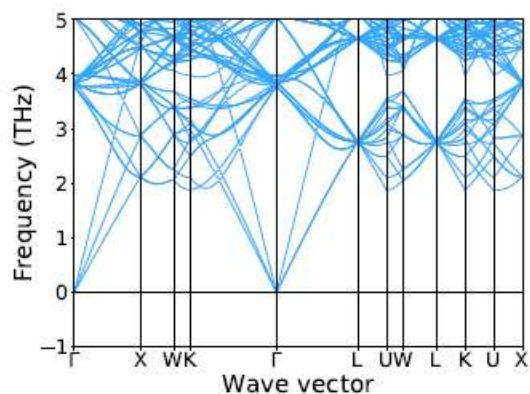
(a)DFT



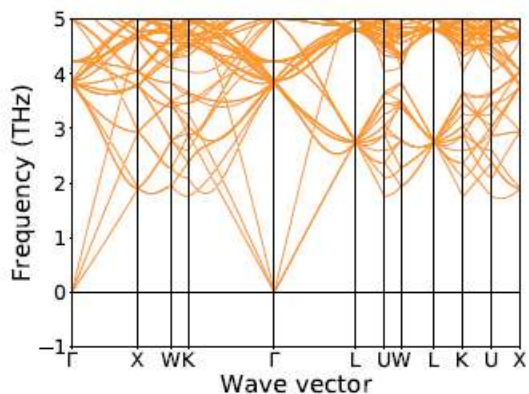
(b)GAP



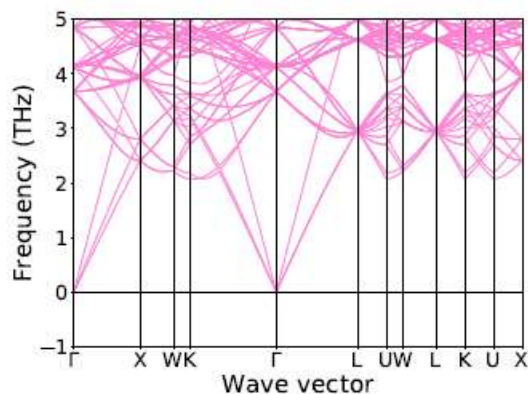
(c)MTP



(d)NNP



(e)SNAP



(f)qSNAP

РИС. S6. Кривые дисперсии фононов для сверхъячейки Ge из 64 атомов алмаза с помощью (а) DFT (б) GAP (в) MTP (д) NNP (е) SNAP (ф) qSNAP.

### III. ТОРГОВЛЯ МЕЖДУ ТОЧНОСТЬЮ ПРОГНОЗИРОВАНИЯ И РАСЧЕТНЫМИ ЗАТРАТАМИ

Все тайминги были выполнены с использованием вычислений LAMMPS на одном ядре процессора Intel i7-6850k 3,6 ГГц.

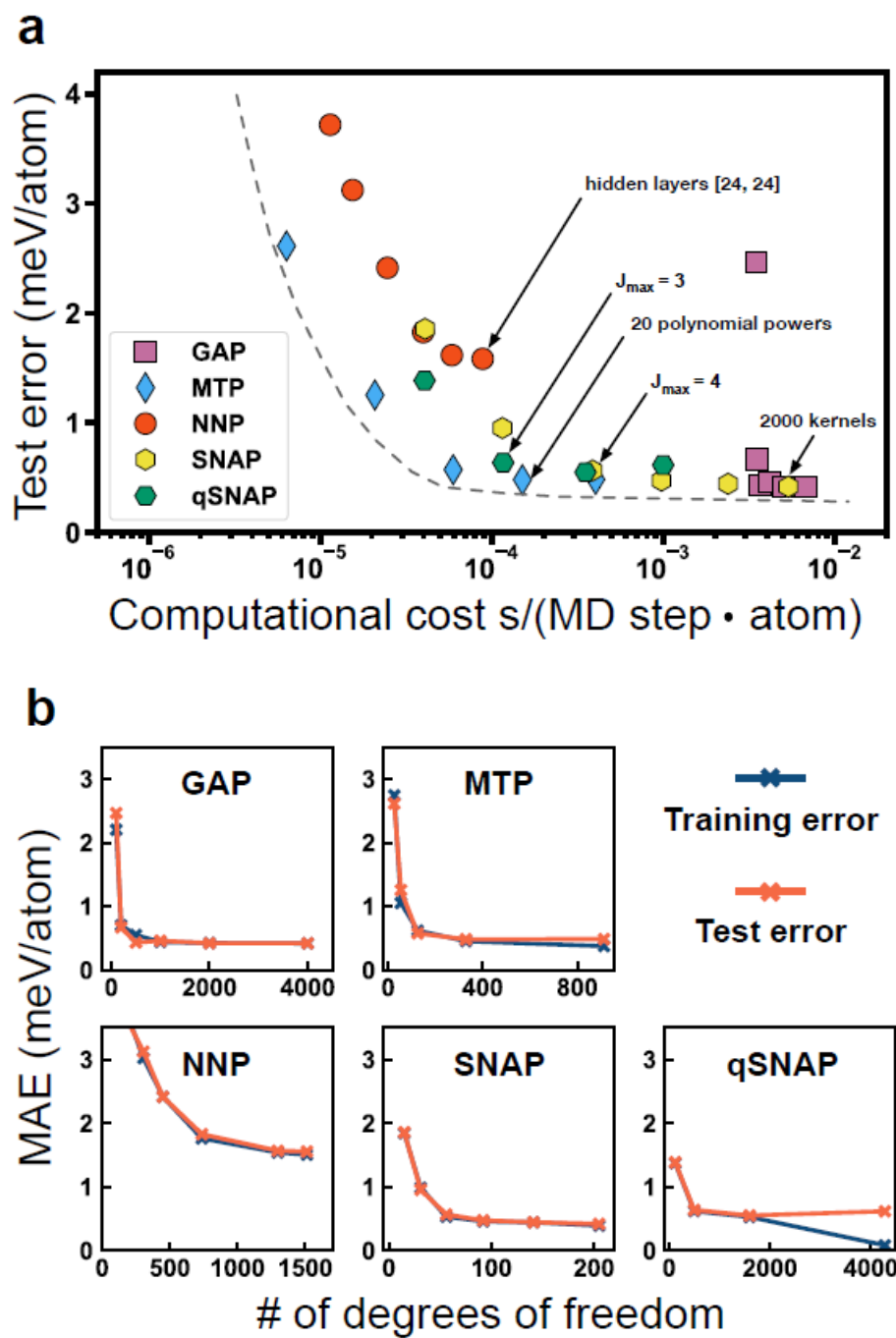


РИС. S7. Ошибка теста в сравнении с вычислительными затратами для системы с ГЦК-никелем. Граница Парето представляет пограничные ограждающие точки разных ML-IAP. b Сравнение ошибки обучения и теста ошибка в зависимости от количества степеней свободы для каждого ML-IAP.

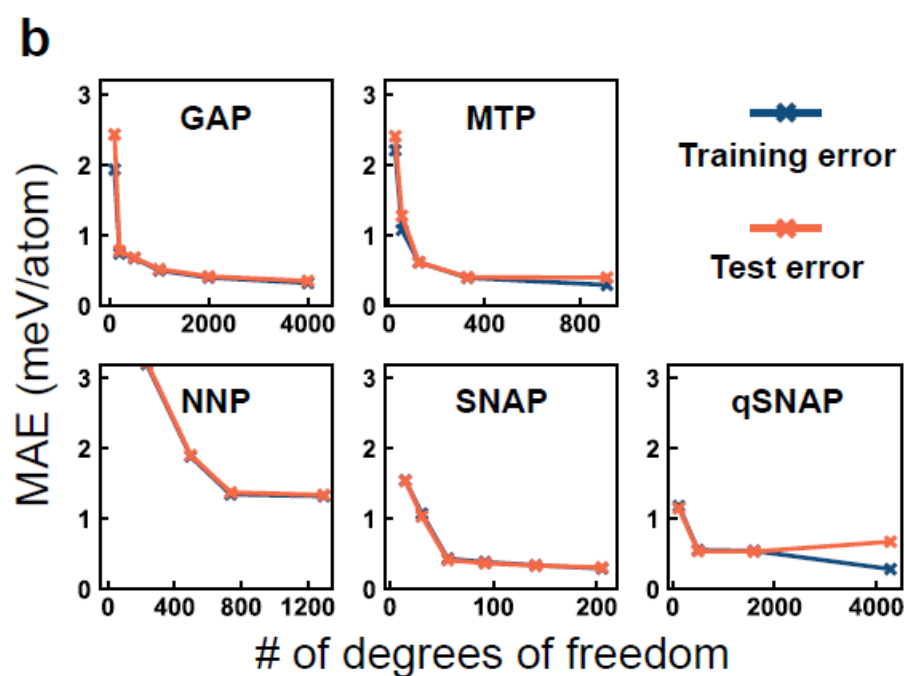
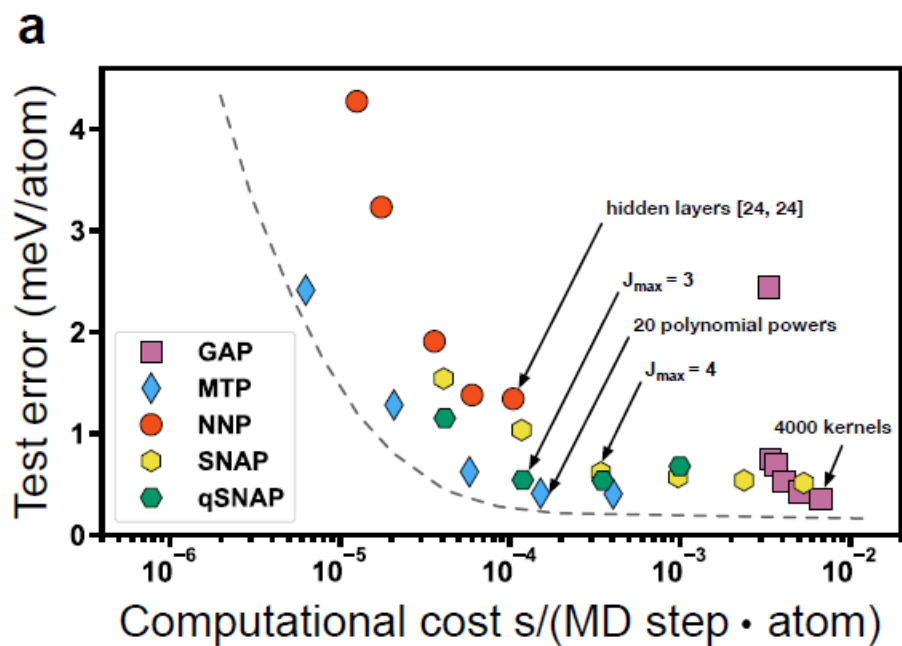


РИС. S8. Ошибка теста в сравнении с вычислительными затратами для системы с ГЦК Cu. Граница Парето представляет собой границу ограждающие точки разных ML-IAP. b Сравнение ошибки обучения и ошибки теста с количество степеней свободы для каждого ML-IAP.

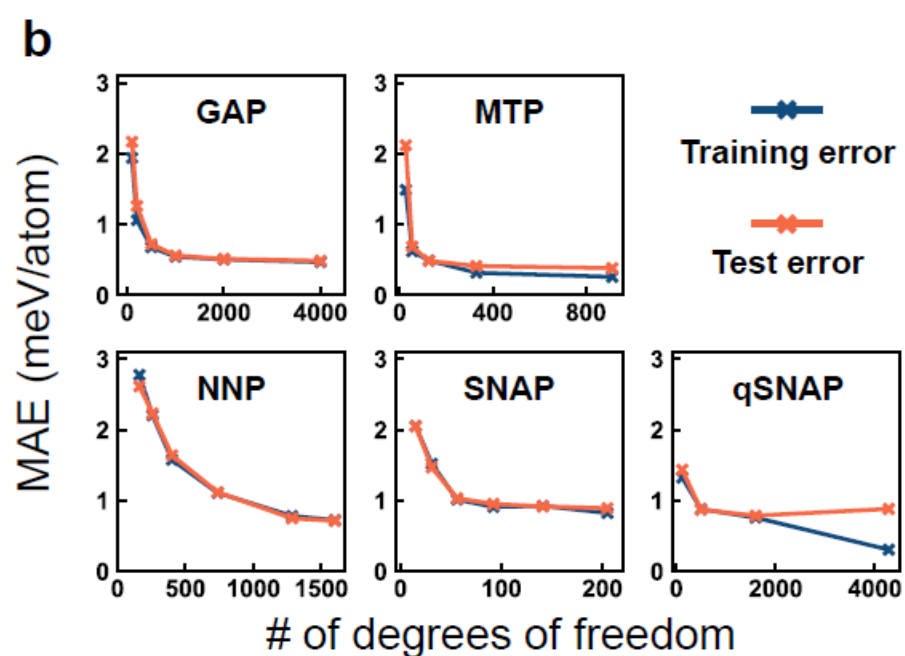
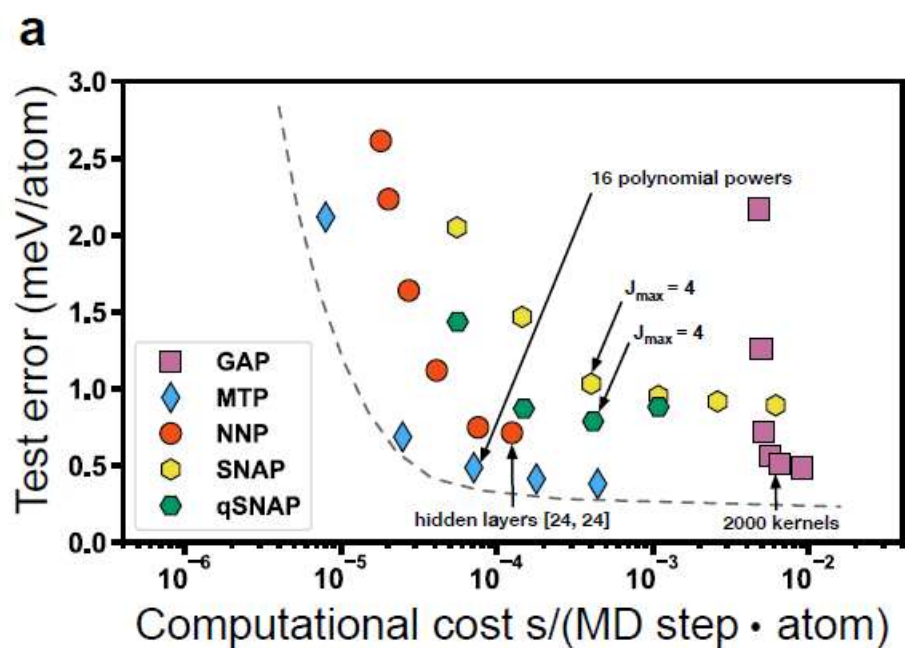


РИС. S9. Ошибка теста в сравнении с вычислительными затратами для системы ВСС Li. Граница Парето представляет собой границу ограждающие точки разных ML-IAP. b Сравнение ошибки обучения и ошибки теста количество степеней свободы для каждого ML-IAP.

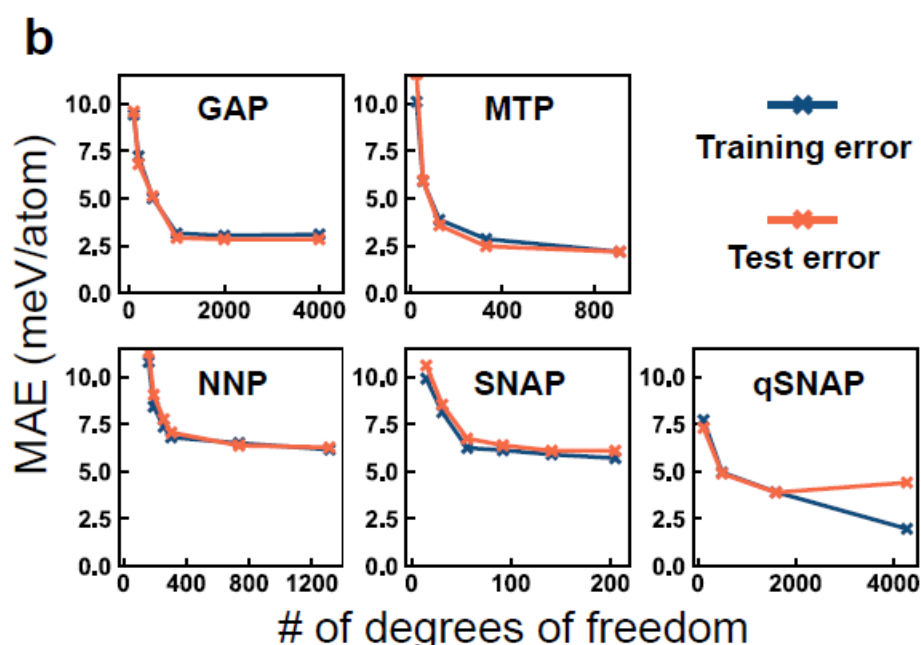
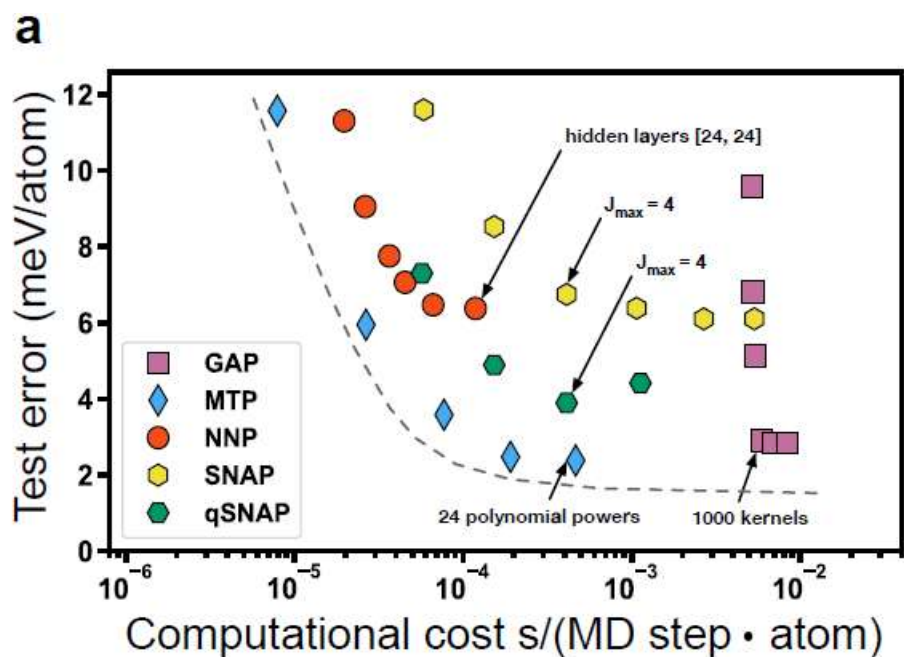


РИС. S10. а Ошибка теста в сравнении с вычислительными затратами для системы алмаз-Si. Граница Парето представляет пограничные ограждающие точки разных ML-IAP. б Сравнение ошибки обучения и ошибки теста в зависимости от количества степеней свободы для каждого ML-IAP.



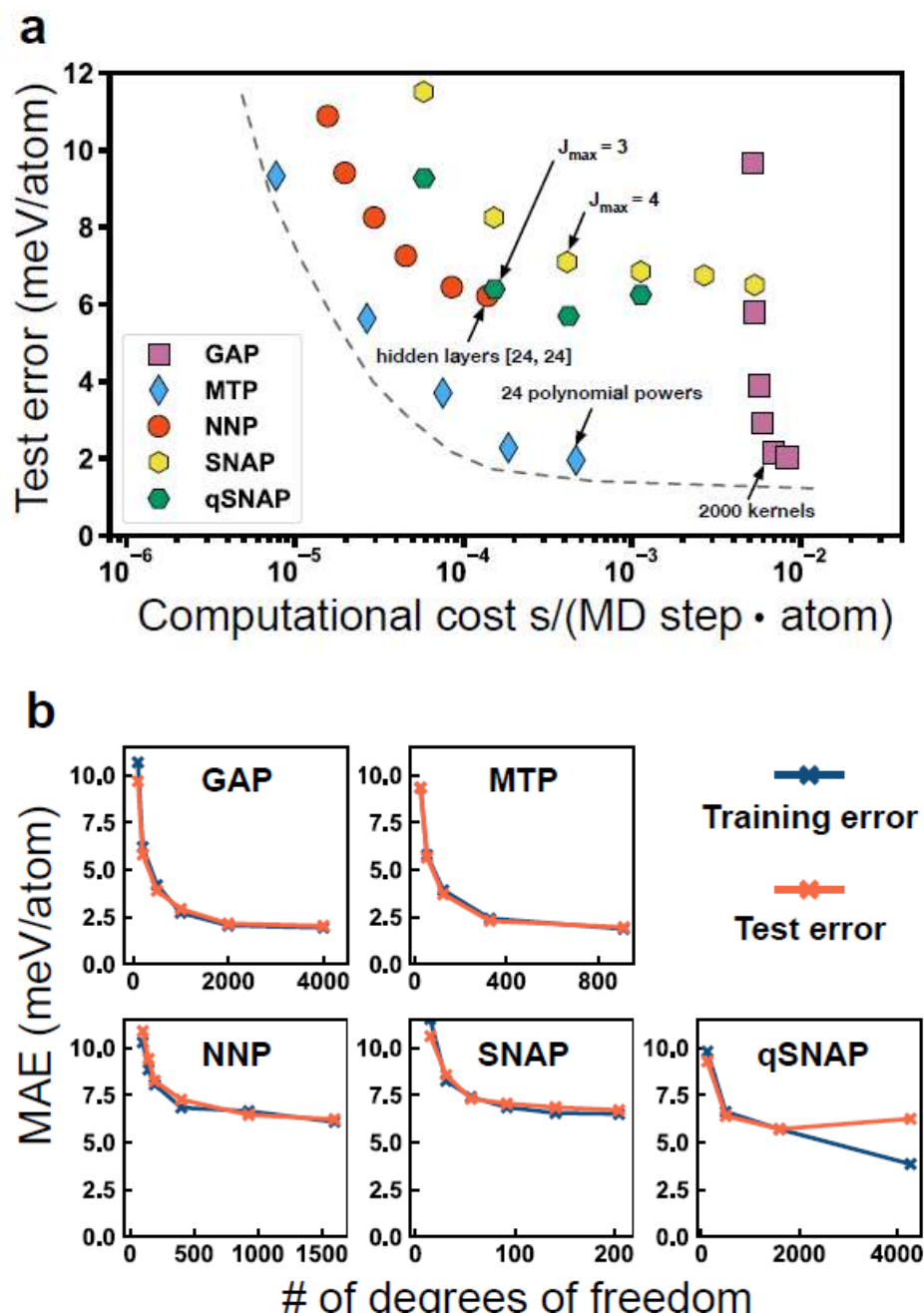


РИС. S11. Ошибка теста в сравнении с вычислительными затратами для алмазной системы Ge. Граница Парето представляет пограничные граничные точки разных ML-IAP. b Сравнение ошибки обучения и ошибки теста в зависимости от количества степеней свободы для каждого ML-IAP.

#### IV. ЛИНЕЙНОЕ МАСШТАБИРУЮЩЕЕ СВЯЗЬ ML-IAPS

Чтобы исследовать влияние масштабирования на вычислительные затраты различных ML-IAP, мы выполняем моделирование npt MD при постоянной температуре 300 К на структурах разного размера, включая  $6 \times 6 \times 6$ ,  $10 \times 10 \times 10$ ,  $14 \times 14 \times 14$ , и  $18 \times 18 \times 18$  сверхъядер, соответствующих 432, 2000, 5488 и 11664 атомам объемного Mo соответственно. Вычислительные

затраты представлены в таблице S11. Все ML-IAP демонстрируют сильно линейное масштабное соотношение по отношению к количеству атомов.

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

Таблица S11. Влияние масштабирования на вычислительные затраты на различных ML IAPS. Единица вычислительных затрат - секунды на шаг MD. Тайминги были выполнены расчетами LAMMPS на одном ядре процессора Intel i7-6850k 3.6 GHz.

- [1] J. Behler and M. Parrinello, Physical Review Letters **98**, 146401 (2007).
- [2] J. Behler, Physical Chemistry Chemical Physics **13**, 17930 (2011).
- [3] J. Behler, International Journal of Quantum Chemistry **115**, 1032 (2015).
- [4] J. Behler, The Journal of Chemical Physics **134**, 074106 (2011).
- [5] J. Behler, Journal of Physics: Condensed Matter **26**, 183001 (2014).
- [6] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Physical Review Letters **104**, 136403 (2010).
- [7] A. P. Bartók, R. Kondor, and G. Csányi, Physical Review B **87**, 184115 (2013).
- [8] A. P. Bartók and G. Csányi, International Journal of Quantum Chemistry **115**, 1051 (2015).
- [9] A. P. Bartók and G. Csányi, .
- [10] S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, Physical Chemistry Chemical Physics **18**, 13754 (2016).
- [11] V. L. Deringer, C. J. Pickard, and G. Csányi, Physical Review Letters **120**, 156001 (2018).
- [12] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, Physical Review X **8**, 041048 (2018).
- [13] W. J. Szlachta, A. P. Bartók, and G. Csányi, Physical Review B **90**, 104108 (2014).
- [14] V. L. Deringer and G. Csányi, Physical Review B **95**, 094203 (2017).
- [15] A. Thompson, L. Swiler, C. Trott, S. Foiles, and G. Tucker, Journal of Computational Physics , 15 (2015).

- [16] M. A. Wood and A. P. Thompson, arXiv:1702.07042 [cond-mat, physics:physics] (2017), arXiv:1702.07042 [cond-mat, physics:physics].
- [17] C. Chen, Z. Deng, R. Tran, H. Tang, I.-H. Chu, and S. P. Ong, Physical Review Materials **1**, 043603 (2017).
- [18] X.-G. Li, C. Hu, C. Chen, Z. Deng, J. Luo, and S. P. Ong, Physical Review B **98**, 094104 (2018).
- [19] Z. Deng, C. Chen, X.-G. Li, and S. P. Ong, npj Computational Materials **5**, 75 (2019).
- [20] E. Jones, T. Oliphant, and P. Peterson, “SciPy: Open source scientific tools for Python,” (2001–), <http://www.scipy.org/>.
- [21] M. A. Wood and A. P. Thompson, The Journal of Chemical Physics **148**, 241721 (2018).
- [22] S. M. Foiles, M. I. Baskes, and M. S. Daw, Physical Review B **33**, 7983 (1986).
- [23] M. S. Daw and M. I. Baskes, Physical Review Letters **50**, 1285 (1983).
- [24] A. V. Shapeev, Multiscale Modeling & Simulation **14**, 1153 (2016).
- [25] E. V. Podryabinkin and A. V. Shapeev, Computational Materials Science **140**, 171 (2017).
- [26] I. Novoselov, A. Yanilkin, A. Shapeev, and E. Podryabinkin, Computational Materials Science **164**, 46 (2019).
- [27] K. Gubaev, E. V. Podryabinkin, G. L. W. Hart, and A. V. Shapeev, Computational Materials Science **156**, 148 (2019), 1806.10567.