

---

# A Materials Foundation Model via Hybrid Invariant-Equivariant Architectures

---

Keqiang Yan<sup>1\*</sup>, Montgomery Bohde<sup>1\*</sup>, Andrii Kryvenko<sup>1\*</sup>, Ziyu Xiang<sup>1</sup>, Kaiji Zhao<sup>1</sup>, Siya Zhu<sup>1</sup>, Saagar Kolachina<sup>1</sup>, Doğuhan SarıTÜRK<sup>1</sup>, Jianwen Xie<sup>2</sup>, Raymundo Arróyave<sup>1</sup>, Xiaoning Qian<sup>1,3</sup>, Xiaofeng Qian<sup>1</sup>, Shuiwang Ji<sup>1</sup>

<sup>1</sup>Texas A&M University

<sup>2</sup>Lambda Inc.

<sup>3</sup>Brookhaven National Laboratory  
{keqiangyan, sji}@tamu.edu

## Abstract

Machine learning interatomic potentials (MLIPs) can predict energy, force, and stress of materials and enable a wide range of downstream discovery tasks. A key design choice in MLIPs involves the trade-off between invariant and equivariant architectures. Invariant models offer computational efficiency but may not perform as well, especially when predicting high-order outputs. In contrast, equivariant models can capture high-order symmetries, but are computationally expensive. In this work, we propose HIENet, a hybrid invariant-equivariant materials interatomic potential model that integrates both invariant and equivariant message passing layers, while provably satisfying key physical constraints. HIENet achieves state-of-the-art performance with considerable computational speedups over prior models. Experimental results on both common benchmarks and downstream materials discovery tasks demonstrate the efficiency and effectiveness of HIENet.

## 1 Introduction

The discovery of materials with desired properties underpins a wide range of technological advancements [1, 2, 3, 4, 5, 6]. However, traditional materials discovery relies heavily on time-consuming and costly trial-and-error experimental methods. Computational approaches, particularly those leveraging advanced quantum mechanical methods including density functional theory (DFT), have accelerated this process [7]. Despite their benefits, these methods come with significant computational costs, making simulating systems with a large number of atoms extremely expensive.

Recent progress in machine learning interatomic potentials (MLIPs) offers a promising path forward by enabling the prediction of energies, forces, and stresses of materials while achieving significant speedups compared to traditional DFT methods. However, existing MLIP models still face a fundamental trade-off: invariant models are computationally efficient but struggle with high-order property predictions and incorporating physical constraints, while equivariant models can better capture high-order interactions but are computationally expensive.

An additional design choice is whether to enforce model predictions to adhere to key physical constraints detailed in Sec. 3.1. Recent works have tried to learn these physical constraints, such as EquiformerV2 [8], which enforces global symmetry operations but not the other physical laws, and ORB [9], which doesn't impose any constraints on model predictions. While it is more computation-

---

\*Equal contribution.

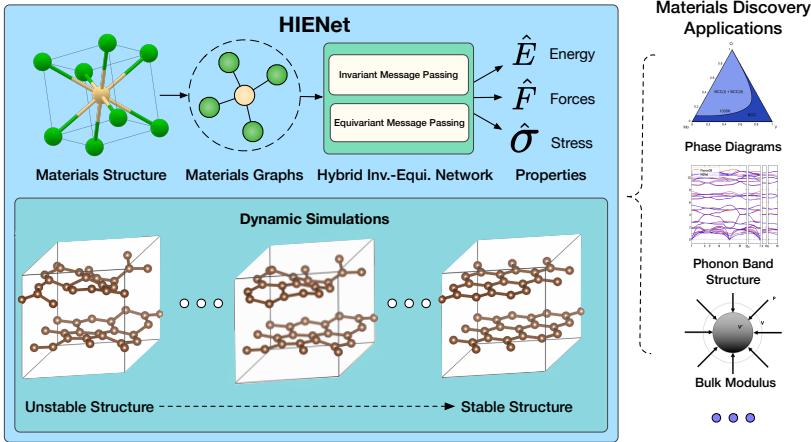


Figure 1: **HIENet overview.** The model converts material structures into graph representations and processes them through a hybrid architecture combining invariant and equivariant message passing networks to predict physical properties. The model supports accurate dynamic simulations (bottom) and enables diverse materials science applications (right).

ally expensive to enforce these physical constraints, it is also necessary for MLIPs to perform well, especially on downstream discovery tasks beyond energy, force, and stress prediction.

In this work, we propose HIENet, a materials MLIP that satisfies key physical constraints for energy, force, and stress predictions while integrating both invariant and equivariant designs to achieve state-of-the-art (SOTA) performance with considerable computational speedups compared to existing models. An overview of HIENet is provided in Figure 1. Unlike prior approaches that rely exclusively on either invariant or equivariant layers, HIENet balances these strategies to leverage the scalability of invariant layers while utilizing equivariant layers to effectively capture high-order interactions. Moreover, in contrast to existing models like EquiformerV2 [8] and Orb [9], HIENet rigorously satisfies physical constraints, including  $O(3)$  equivariance for force and stress, and adheres to physical conservation laws through physics-informed derivative-based methods. Experimental results on common benchmarks including Materials Project Trajectory, Matbench Discovery, and downstream materials discovery tasks including evaluations on phonons, bulk moduli, *ab initio* molecular dynamics, and alloys as detailed in Sec. 5.1, 5.2, 5.3 and Appendix B,C demonstrate the efficiency and effectiveness of HIENet. Additional ablations in Sec. 5.5 further demonstrate that our hybrid invariant-equivariant approach scales well with increasing model capacity and works with different equivariant layer architectures, providing powerful insights into future MLIP designs.

## 2 Preliminaries

**Problem definition.** The core task in developing MLIPs is to learn a mapping from materials atomic structures to quantum mechanical properties. Specifically, given a crystal structure, we aim to predict three quantities; the total energy  $E$ , the forces acting on each atom  $\mathbf{F} = \{\mathbf{F}_i \in \mathbb{R}^3, 1 \leq i \leq n\}$ , where  $n$  denotes number of atoms in a cell, and the stress tensor  $\boldsymbol{\sigma} \in \mathbb{R}^{3 \times 3}$ , which governs cell deformation. While these properties are directly useful for many applications such as structural relaxation and predicting thermodynamic stability, from these we can derive many other important material properties such as phonon band structures and bulk moduli as shown in Sec. 5.3. We also provide preliminaries about molecular dynamics simulation of materials in Appendix A.

**Crystal structures.** Unlike regular molecules, crystals are periodic in nature and are characterized as three-dimensional lattices with infinitely repeating unit cells. Adopting the notation of [10], a crystal structure can be described as a triple  $\mathbf{M} = (\mathbf{Z}, \mathbf{P}, \mathbf{L})$ , which represents both atomic and geometric information. The atomic composition is denoted by  $\mathbf{Z} = [z_1, z_2, \dots, z_n] \in \mathbb{R}^n$ , where  $z_i$  represents the atomic number of  $i$ -th atom in the unit cell. The arrangement of these atoms in the Euclidean space is given by 3D coordinates  $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n] \in \mathbb{R}^{3 \times n}$ . The periodicity of the unit cell is specified by the lattice matrix  $\mathbf{L} = [\ell_1, \ell_2, \ell_3] \in \mathbb{R}^{3 \times 3}$ , whose columns are the three lattice vectors.

### 3 Hybrid Invariant-Equivariant Networks

We propose Hybrid Invariant-Equivariant Network (HIENet), a materials interatomic potential model that integrates both invariant and equivariant message passing layers. HIENet is carefully designed to satisfy important physical constraints detailed in Sec. 3.1, consisting of physics-informed geometric crystal graphs detailed in Sec. 3.2, a hybrid invariant-equivariant network design detailed in Sec. 3.3, and physics-informed property predictions detailed in Sec. 3.4. All together, HIENet achieves state-of-the-art performance on common benchmarks and downstream materials discovery tasks while significantly improving computational efficiency compared to prior models. Additionally, HIENet satisfies all desirable physical constraints with mathematical proofs as shown in Sec. 3.5.

#### 3.1 Physical Constraints for MLIPs

As detailed in Sec. 2, MLIPs are developed to predict energy, forces, and stress of materials atomic systems. While some MLIPs such as EqV2 and ORB can achieve moderate performance on some tasks without satisfying key physical constraints, such models under-perform on many important materials discovery tasks such as phonon frequency calculations, bulk moduli prediction, and molecular dynamics simulations. As such, in order for MLIPs to generalize well and have robust performance across downstream tasks, it is essential that model predictions satisfy key physical constraints.

**Rototranslational Symmetries.** Crystal structures exhibit inherent symmetry under global rotations, translations, and reflections. To respect these symmetries, the predicted energy must be  $E(3)$  invariant, while forces and stress must be  $O(3)$  equivariant. We formalize these requirements as follows:

**Definition 3.1** ( $O(3)$  Equivariance). A MLIP produces  $O(3)$  equivariant predictions if, for a crystal structure  $\mathbf{M} = (\mathbf{Z}, \mathbf{P}, \mathbf{L})$ , its predicted energy  $\hat{E}$ , forces  $\hat{\mathbf{F}} = (\hat{\mathbf{F}}_1, \dots, \hat{\mathbf{F}}_n)$ , and stress tensor  $\hat{\sigma}$  transform under any rotation  $\mathbf{R} \in \mathbb{R}^{3 \times 3}$ ,  $|\mathbf{R}| = \pm 1$  and translation  $\mathbf{b} \in \mathbb{R}^3$  as follows:

$$\begin{aligned}\hat{E}(\mathbf{Z}, \mathbf{P}, \mathbf{L}) &= \hat{E}(\mathbf{Z}, \mathbf{RP} + \mathbf{b}, \mathbf{RL}) \\ \hat{\mathbf{F}}_i(\mathbf{Z}, \mathbf{P}, \mathbf{L}) &= \mathbf{R}^\top \hat{\mathbf{F}}_i(\mathbf{Z}, \mathbf{RP} + \mathbf{b}, \mathbf{RL}), \\ \hat{\sigma}(\mathbf{Z}, \mathbf{P}, \mathbf{L}) &= \mathbf{R}^\top \hat{\sigma}(\mathbf{Z}, \mathbf{RP} + \mathbf{b}, \mathbf{RL}) \mathbf{R}.\end{aligned}$$

Importantly, there is a distinction between invariant/equivariant layers and invariant/equivariant predictions. For example, CHGNet [11] exclusively uses  $E(3)$  invariant message passing layers, yet CHGNet force and stress predictions are  $O(3)$  equivariant because they use gradient-based force and stress calculations. When we refer to a model being  $O(3)$  equivariant, we are referring to the outputs, not the individual layers, unless otherwise specified.

**Physical Plausibility.** Beyond symmetry considerations, MLIPs must satisfy several key physical laws to be reliable for downstream applications. These include force conservation, force equilibrium, and stress tensor symmetry. We define these constraints formally as follows:

**Definition 3.2** (Force Conservation and Equilibrium). Forces must form a conservative vector field derived from the potential energy surface, and in the absence of external influences, the sum of forces on all atoms is zero:

$$\mathbf{F} = -\nabla_{\mathbf{P}} E, \quad \sum_{i=1}^n \mathbf{F}_i = \mathbf{0}, \tag{1}$$

**Definition 3.3** (Stress Tensor Symmetry). The predicted stress tensor must be symmetric:

$$\sigma_{ij} = \sigma_{ji} \quad \forall i, j \in 1, 2, 3 \tag{2}$$

In addition to these physical laws, the potential energy surface must be continuously differentiable to enable accurate downstream property calculations requiring higher-order derivatives.

Our HIENet model rigorously enforces all the outlined symmetry and physical constraints through the carefully designed geometric crystal graphs, model architecture, and gradient-based force and stress computation, with details provided in Sec. 3.2, 3.3, 3.4, and 3.5.

#### 3.2 Geometric Graph Representations Satisfying Physical Constraints

For a crystal structure,  $\mathbf{M} = (\mathbf{Z}, \mathbf{P}, \mathbf{L})$  we construct an  $O(3)$  equivariant crystal graph  $G = (V, E)$  that preserves physical symmetries inherent in crystal structures. Specifically, for a crystal structure

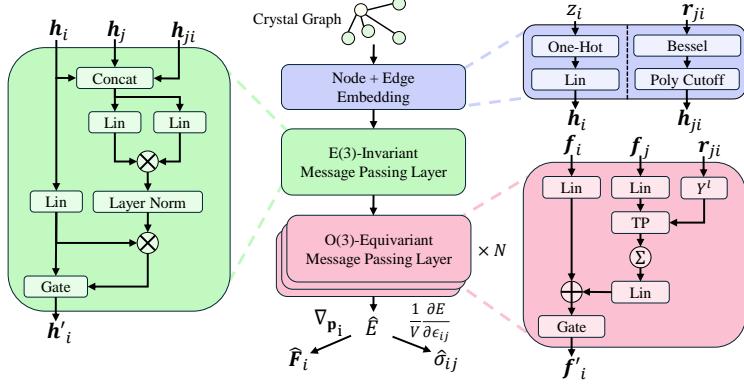


Figure 2: **HIENet Model Architecture.** We construct  $O(3)$  equivariant crystal graphs. We then apply an invariant message passing layer followed by several equivariant message passing layers before predicting the total energy,  $\hat{E}$  and using physical laws to compute  $\hat{\mathbf{F}}$ ,  $\hat{\sigma}$ .

$\mathbf{M} = (\mathbf{Z}, \mathbf{P}, \mathbf{L})$ , each atom  $i$  in the unit cell and all its periodic duplicates are represented by a single node  $i \in V$  with node features  $\mathbf{h}_i = \mathbf{W}_{\text{emb}} \mathbf{z}_i$ , where  $\mathbf{W}_{\text{emb}} \in \mathbb{R}^{d \times n_z}$  is a learnable embedding matrix and  $\mathbf{z}_i$  is the one-hot encoding of atomic number  $z_i$ . An edge will be built from node  $j$  to  $i$  when the Euclidean distance between a periodic duplicate  $j'$  of  $j$  and  $i$  satisfies

$$\|\mathbf{r}_{j'i}\|_2 = \|\mathbf{p}_j + k_1 \ell_1 + k_2 \ell_2 + k_3 \ell_3 - \mathbf{p}_i\|_2 \leq R_{\text{cut}}, \quad k_1, k_2, k_3 \in \mathbb{Z}, \quad (3)$$

where  $R_{\text{cut}}$  is a fixed cutoff radius.

Edge features  $\mathbf{h}_{ji}$  are then embedded using radial Bessel basis functions with a polynomial envelope [12] function  $f_{\text{poly}}$ :

$$\mathbf{h}_{ji} = \frac{2 \sin\left(\frac{n\pi}{R_{\text{cut}}} \|\mathbf{r}_{ji}\|_2\right)}{R_{\text{cut}} \|\mathbf{r}_{ji}\|_2} f_{\text{poly}}(\|\mathbf{r}_{ji}\|_2, R_{\text{cut}}). \quad (4)$$

**Importance of using envelope function.** It is worth noting that the smooth envelope is crucial for energy conservation and computing physically meaningful force predictions. It ensures that the energy and its derivatives smoothly decay to zero at the cutoff boundary.

**Importance of constructing  $O(3)$  equivariant crystal graphs.** Unlike previous work [13, 10], our constructed crystal graphs are  $O(3)$  equivariant due to the use of a fixed cutoff radius and edge vectors  $\mathbf{r}_{ji}$  that are  $O(3)$  equivariant by definition. It is important to note that  $O(3)$  equivariance of the input crystal graphs is a necessary condition for achieving  $O(3)$  equivariant predictions in MLIPs. As shown in our experiments in Appendix E, violating this equivariance in the graph construction leads to a measurable drop in MLIP performance.

### 3.3 Hybrid Invariant-Equivariant Design

Our HIENet model consists of several invariant and equivariant message passing layers which iteratively update node features for each atom.

**E(3) Invariant Layer.** In the invariant message passing layers, we update node features  $\mathbf{h}_i$  using a graph transformer mechanism. Specifically, we compute key  $\mathbf{k}_{ji}$ , query  $\mathbf{q}_{ji}$ , and value  $\mathbf{v}_{ji}$  vectors as:

$$\mathbf{k}_{ji} = \mathbf{W}_K(\mathbf{h}_i \parallel \mathbf{h}_j \parallel \mathbf{h}_{ij}), \quad \mathbf{q}_{ji} = \mathbf{W}_Q(\mathbf{h}_i \parallel \mathbf{h}_j \parallel \mathbf{h}_{ij}), \quad \mathbf{v}_{ji} = \Phi(\mathbf{h}_i \parallel \mathbf{h}_j \parallel \mathbf{h}_{ij}), \quad (5)$$

where  $\Phi$  is an MLP with SiLU nonlinearities [14] and  $\parallel$  denotes vector concatenation. We then compute attention scores and aggregate the values over neighboring nodes to update the node features:

$$\mathbf{h}'_i = \varphi(\mathbf{h}_i) + (1 - \varphi(\mathbf{h}_i)) \sum_{j \in \mathcal{N}_i} \mathbf{v}_{ji} \odot \sigma\left(\frac{\mathbf{q}_{ji} \odot \mathbf{k}_{ji}}{\sqrt{d}}\right), \quad (6)$$

where  $\odot$  represents element-wise multiplication,  $\sigma$  is the sigmoid activation function, and  $\varphi$  is an MLP with a sigmoid activation in the final layer that acts as a learnable gating mechanism.

**O(3) Equivariant Layer.** The HIENet equivariant layers build node features  $\mathbf{f}_{i,\ell}$  for each rotation order  $\ell \leq L_{\max}$ . In practice we use  $L_{\max} = 3$ . In the first equivariant message passing layer, we initialize  $\mathbf{f}_{i,0} = \mathbf{h}_i$ , the output of the previous invariant layer. In each equivariant layer, we embed edge vectors  $\mathbf{r}_{ji}$  using spherical harmonics  $Y^l(\frac{\mathbf{r}_{ji}}{\|\mathbf{r}_{ji}\|})$  and update the equivariant features as:

$$\mathbf{f}'_{i,\ell} = \frac{1}{|\mathcal{N}_i|} \sum_{l=0}^{L_{\max}} \sum_{j \in \mathcal{N}_i} \mathbf{TP}_\ell \left( \mathbf{W} \mathbf{f}_i, Y^l \left( \frac{\mathbf{r}_{ji}}{\|\mathbf{r}_{ji}\|} \right) \right) \quad (7)$$

where  $\mathbf{TP}_\ell$  is the standard tensor product operation yielding outputs with rotation order  $\ell$ . We further add a skip connection and gate mechanism to output updated node features:

$$\mathbf{f}'_i = \psi(\mathbf{W}_{\text{skip}} \mathbf{f}_i + \mathbf{W}_E \mathbf{f}'_i) \quad (8)$$

where  $\psi$  is an equivariant gate activation function defined as:

$$\psi(\mathbf{f}_i) = \phi(\mathbf{f}_{i,0}) \oplus \left( \bigoplus_{0 < l \leq L_{\max}} \phi(\mathbf{f}_{i,0}) \mathbf{f}_{i,l} \right) \quad (9)$$

where  $\phi$  is the SiLU activation function and  $\oplus$  is the direct sum operation.

**Hybrid MLIPs.** Our HIENet model consists of one or more invariant message passing layers followed by several equivariant layers. The final equivariant layer outputs are aggregated across the graph and the energy is predicted as  $E = \sum_i \mathbf{W}_e \mathbf{f}_{i,0}$ . In practice, we find that one invariant message passing layer achieves a good balance between performance and efficiency. We provide additional details and ablations on the model design in Appendix F and E.

### 3.4 Physics Informed Property Predictions

In order to ensure that our force and stress predictions obey the aforementioned physical constraints, we use gradient-based methods to compute force and stress. Specifically, our model directly predicts the total energy,  $\hat{E}$  and we compute the force acting on atom  $i$  as  $\hat{\mathbf{F}}_i = -\nabla_{\mathbf{p}_i} \hat{E}$ , where  $\nabla_{\mathbf{p}_i}$  represents the gradient with respect to the position vector  $\mathbf{p}_i$ . This approach automatically guarantees that:

**Proposition 3.4.** *HIENet predictions  $\hat{\mathbf{F}}_i$  form a conservative vector field.*

**Proposition 3.5.** *HIENet predictions satisfy force equilibrium  $\sum_{i=1}^N \hat{\mathbf{F}}_i = \mathbf{0}$  when no external influences are applied.*

Similarly, we compute the stress tensor through strain derivatives  $\hat{\sigma}_{ij} = \frac{1}{V} \frac{\partial \hat{E}}{\partial \epsilon_{ij}}$ , where  $\epsilon$  is the lattice strain tensor and  $V$  is the volume of the unit cell. We ensure that  $\hat{\sigma}$  will be symmetric by first symmetrizing the strain matrix  $\epsilon_{\text{sym}} = \frac{1}{2}(\epsilon + \epsilon^\top)$ . All together, our approach guarantees that:

**Proposition 3.6.** *HIENet predictions are O(3) equivariant as defined in Sec. 3.1.*

### 3.5 Proofs of Satisfying Physical Constraints

While previous works [15, 16, 17, 11, 18] used gradient-based calculations, none of these works prove that their proposed methods satisfy desired physical laws. In this section, we rigorously prove each of the previously stated propositions and show that HIENet satisfies all desirable physical constraints.

*Proof of Proposition 3.4.* By definition, a vector field  $\mathbf{v} : \mathbb{R} \rightarrow \mathbb{R}^n$  is conservative if there exists a continuously differentiable scalar field  $\varphi$  such that  $\mathbf{v} = \nabla \varphi$ .

We compute forces as  $\hat{\mathbf{F}} = -\nabla_{\mathbf{P}} \hat{E}$ , so HIENet predictions form a conservative force-field as long as the energy  $E$  is continuously differentiable with respect to the atom positions,  $\mathbf{p}_i$ . The composition of continuously differentiable functions is continuously differentiable, so it suffices for each operation in HIENet to be continuously differentiable.

Clearly linear operations are continuously differentiable, and we exclusively use sigmoid and SiLU [14] activations throughout the model. The spherical harmonics  $Y^l(\frac{\mathbf{r}_{ji}}{\|\mathbf{r}_{ji}\|})$  are continuously

differentiable within the domain of possible interatomic distances, i.e. for  $\|r_{ji}\| \neq 0$ . Finally, the edge embedding function that produces  $\mathbf{h}_{ji}$  is continuously differentiable for  $\|r_{ji}\| \neq 0$  because  $f_{\text{poly}}$  is specifically chosen so that it is continuously differentiable and decays to 0 at  $R_{\text{cut}}$  [12].

Because each operation in HIENet is continuously differentiable, the force predictions therefore form a conservative vector field, as desired.  $\square$

*Proof of Proposition 3.5.* We define edge force as  $\hat{\mathbf{F}}_{ji} = -\frac{\partial \hat{E}}{\partial \mathbf{r}_{ji}}$ .

Forces acting on each atom can be decomposed as:

$$\hat{\mathbf{F}}_i = -\frac{\partial \hat{E}}{\partial \mathbf{p}_i} = -\sum_{j \in \mathcal{N}_i} \left( \frac{\partial \hat{E}}{\partial \mathbf{r}_{ji}} \frac{\partial \mathbf{r}_{ji}}{\partial \mathbf{p}_i} + \frac{\partial \hat{E}}{\partial \mathbf{r}_{ij}} \frac{\partial \mathbf{r}_{ij}}{\partial \mathbf{p}_i} \right) \quad (10)$$

$$= -\sum_{j \in \mathcal{N}_i} \left( \frac{\partial \hat{E}}{\partial \mathbf{r}_{ji}} - \frac{\partial \hat{E}}{\partial \mathbf{r}_{ij}} \right) = \sum_{j \in \mathcal{N}_i} (\hat{\mathbf{F}}_{ji} - \hat{\mathbf{F}}_{ij}) \quad (11)$$

Summing over all atoms:

$$\sum_{i=1}^n \hat{\mathbf{F}}_i = \sum_{i=1}^n \sum_{j \in \mathcal{N}_i} (\hat{\mathbf{F}}_{ji} - \hat{\mathbf{F}}_{ij}) = \sum_{(i,j) \in \mathcal{E}} (\hat{\mathbf{F}}_{ji} - \hat{\mathbf{F}}_{ij}) = 0 \quad (12)$$

Therefore the forces acting on each atom sum to 0 as desired.  $\square$

*Proof of Proposition 3.6.* First, the radius-based graph construction  $\vartheta_{\text{graph}}$  described in Sec. 3.2 is O(3) equivariant. This is because it only depends on the relative positions between atoms and the resulting displacement vectors  $\mathbf{r}_{ji}$  will rotate accordingly.

Second, the proposed HIENet message passing layers  $\vartheta_{\text{HIENet}}$  are E(3) invariant for the final energy prediction. The invariant message passing layers are E(3) invariant by construction because they only operate the magnitude  $\|\mathbf{r}_{ji}\|$  and for the equivariant message passing layers, we only extract the final  $l = 0$  features, which are invariant. Because of this:

$$\hat{E}(\mathbf{Z}, \mathbf{RP} + \mathbf{b}, \mathbf{RL}) = \vartheta_{\text{HIENet}}(\vartheta_{\text{graph}}(\mathbf{Z}, \mathbf{RP} + \mathbf{b}, \mathbf{RL})) = \vartheta_{\text{HIENet}}(\mathbf{R}\vartheta_{\text{graph}}(\mathbf{Z}, \mathbf{P}, \mathbf{L})) \quad (13)$$

$$= \vartheta_{\text{HIENet}}(\vartheta_{\text{graph}}(\mathbf{Z}, \mathbf{P}, \mathbf{L})) = \hat{E}(\mathbf{Z}, \mathbf{P}, \mathbf{L}), \quad (14)$$

Based on the physics informed property predictions described in Sec. 3.4, we then have:

$$\hat{\mathbf{F}}_i(\mathbf{Z}, \mathbf{RP} + \mathbf{b}, \mathbf{RL}) = -\nabla_{\mathbf{R}\mathbf{p}_i} \hat{E}(\mathbf{Z}, \mathbf{P}, \mathbf{L}) = -\mathbf{R}\nabla_{\mathbf{p}_i} \hat{E}(\mathbf{Z}, \mathbf{P}, \mathbf{L}) = \mathbf{R}\hat{\mathbf{F}}_i(\mathbf{Z}, \mathbf{P}, \mathbf{L}), \quad (15)$$

$$\hat{\boldsymbol{\sigma}}(\mathbf{Z}, \mathbf{RP} + \mathbf{b}, \mathbf{RL}) = \frac{1}{V} \nabla_{\mathbf{R}\epsilon_{ij}} \mathbf{R}^T \hat{E} = \frac{1}{V} \mathbf{R} \nabla_{\epsilon_{ij}} \hat{E} \mathbf{R}^T = \mathbf{R}\hat{\boldsymbol{\sigma}}(\mathbf{Z}, \mathbf{P}, \mathbf{L})\mathbf{R}^T, \quad (16)$$

therefore energy, force, and stress each transform appropriately under rototranslation and HIENet predictions are O(3) equivariant.  $\square$

## 4 Related Work

In this section, we focus on materials MLIPs and provide related works on conventional computation methods in Appendix A.2. Recent advances in MLIPs [15, 19, 20, 21, 22, 23] and the availability of high-quality materials dynamics datasets [17, 24, 25] generated using DFT-based algorithms have facilitated the development of powerful materials MLIPs [26, 27, 13, 28, 29, 10]. Among these MLIPs, models with only invariant layers, such as M3GNet [17], CHGNet [24], Orb [9], and EScAIP [30], are computationally efficient but struggle to produce physically meaningful and robust predictions, especially on downstream tasks. In contrast, models with purely equivariant layers, including MACE [31], SevenNet [32], and EquiformerV2 [25], are more powerful, but also more computationally expensive. Their extensive use of tensor product operations limits their scalability. Moreover, even some equivariant models, such as EquiformerV2, violate force conservation, undermining its utility in realistic materials tasks as seen in Sec. 5.

Different from these existing MLIPs, our proposed HIENet satisfies all key physical constraints and combines the scalability and efficiency of invariant designs with the robustness and symmetry-capturing capabilities of equivariant designs. This novel integration offers a promising direction for the next generation of MLIPs design.

Table 1: Model performance on the Unique Prototype split of the Matbench Discovery benchmark. DAF is the Discovery Acceleration Factor from [33] which measures model performance to classify thermodynamic stability. MAE and RMSE are in meV/atom. RMSD is the root mean squared displacement between predicted and reference structures after relaxation. Missing results from corresponding model marked by -.

Model	HIENet	EquiformerV2	ORB	SevenNet-l3i5	MACE	CHGNet
DAF $\uparrow$	<b>4.879</b>	4.64	<u>4.70</u>	4.63	3.78	3.361
MAE $\downarrow$	<b>40</b>	<u>42</u>	45	48	57	63
RMSE $\downarrow$	<b>73</b>	87	91	<u>87</u>	101	103
$R^2 \uparrow$	<b>0.84</b>	<u>0.778</u>	0.756	0.776	0.697	0.689
F1 $\uparrow$	0.761	<b>0.77</b>	<u>0.765</u>	0.76	0.669	0.613
Accuracy $\uparrow$	<b>0.93</b>	<b>0.93</b>	0.92	0.92	0.88	0.851
Precision $\uparrow$	<b>0.749</b>	0.709	<u>0.719</u>	0.708	0.577	0.514
RMSD $\downarrow$	<b>0.084</b>	-	0.101	<u>0.085</u>	0.091	0.095

## 5 Experimental Evaluations

In this section, we evaluate HIENet’s overall modeling capacity as a MLIP. We assess its performance on the widely used Matbench Discovery benchmark [33] and Materials Project Trajectory (MPtrj) dataset [24] in Sec. 5.1 and 5.2. We then provide evaluations on important downstream materials discovery tasks in Sec. 5.3, where we find that models which do not satisfy physical constraints perform poorly. In Sec. 5.4 we compare the computational efficiency of HIENet to top performing equivariant models and demonstrate that HIENet is able to achieve SOTA performance across all tasks while still providing considerable computational speedups. Finally, in Sec. 5.5 we provide ablations studies to demonstrate the robustness and generality of our hybrid invariant-equivariant network design. We provide additional downstream materials discovery evaluations on *ab initio* molecule dynamics simulations and phase diagram prediction for alloy design in Appendix B and Appendix C.

**Experimental setup.** We train our HIENet on the MPtrj dataset [11], which contains 1.58M crystal structures. We split the dataset and use 95% for training and 5% for validation following [31]. For fair comparison, we compare with models trained on this dataset and without any auxiliary data or training objectives. Specifically, we compare against state-of-the-art methods including EquiformerV2 [8], ORB [9], SevenNet [18], MACE [31], and CHGNet [11]. Of these baselines, all but ORB have equivariant force and stress predictions, and all but EquiformerV2 and ORB satisfy the physical constraints listed in Sec. 3.1. More detailed model settings and training details can be found in Appendix F. In all tables, we mark best performing model in **bold** and second best in underlined.

### 5.1 Evaluation on Matbench Discovery

Matbench Discovery benchmark [33] is a comprehensive testbed for crystalline materials structure optimizations and stability predictions. Notably, the Matbench Discovery benchmark structures come from a different distribution from the MPtrj training dataset, thus posing an out-of-distribution (OOD) generalization problem. As shown in Table 1, HIENet performs best on **six out of seven** metrics and has a significant performance gain on the Discovery Acceleration Factor (DAF). Additionally, we observe that while ORB performs well on all of the energy-related metrics, it has the worst performance of all models on RMSD, a metric that measures the models ability to accurately relax structures to stability. This aligns with out intuitions that downstream tasks such as structural optimization require models to obey physical symmetries.

### 5.2 Materials Project Trajectory Dataset

We then evaluate HIENet’s ability to accurately predict energy, force, and stress on a held-out MPtrj validation set following [24, 31]. SevenNet does not hold-out any validation split and trains their model on the entire 1.58M structures. To compare with SevenNet, we also report HIENet performance on the training split. As seen in Table 2, HIENet achieves state-of-the-art performance across train and validation splits. Notably, HIENet reduces the energy mean absolute error (MAE) by nearly 50% and the force MAE by 23% compared to the previous state-of-the-art EquiformerV2.

Table 3: Error in phonon frequency prediction on structures from [34]. MAE and MSE computed against each q-point, and RMSE taken as the root of MSE over all q-points and bands. Band structures shown in Appendix.

Model	HIENet	EquiformerV2	ORB	SevenNet-13i5	MACE	CHGNet
MAE (THz)	<b>0.316</b>	1.359	1.601	0.325	0.529	1.359
MSE (THz <sup>2</sup> )	<b>0.332</b>	4.65	5.441	0.358	0.837	4.21
RMSE (THz)	<b>0.446</b>	1.657	1.973	0.455	0.699	1.604

Table 4: Error in bulk modulus  $K_{VRH}$  prediction across 1,763 crystal structures sampled from Material Project.

Model	HIENet	EquiformerV2	ORB	SevenNet-13i5	MACE	CHGNet
MAE (GPa) ↓	<b>10.52</b>	24.76	34.7	<u>11.5</u>	28.84	21.67
$R^2 \uparrow$	<b>0.93</b>	0.64	-21.9	<u>0.9</u>	-54.1	0.7

### 5.3 Evaluations on Phonons and Bulk Modulus Prediction

#### Phonon frequency evaluation.

Phonons are collective excitations of atomic vibrations in crystal structures with translational symmetry, playing a crucial role in determining the dynamical stability and thermal conductivity of materials. The calculation of phonon band structure relies on the atomic forces upon displacement of atoms in different phonon modes along high-symmetry paths in the first Brillouin zone. Because of this, it is critical for model predictions to obey physical symmetries and for the forces to be conservative. Here we perform a phonon band structure

calculation workflow using Phonopy [35, 36] on the set of structures from [34]. We provide additional details on the dataset, workflow, and additional results on the phonon band structures in Appendix D.1.

Table 2: MAE on train and validation splits. Inv. and Eqv. denote whether the model uses invariant or equivariant message passing layers. ORB [9] does not report results on MPtrj and MACE does not report stress performance.

Model	Inv.	Eqv.	Energy ↓ (meV/atom)	Forces ↓ (meV/Å)	Stress ↓ (kBar)
Train					
SevenNet-0	✗	✓	11.5	41	2.78
SevenNet-13i5	✗	✓	<u>8.3</u>	<u>29</u>	<u>2.33</u>
HIENet	✓	✓	<b>5.91</b>	<b>20.76</b>	<b>1.95</b>
Validation					
CHGNet	✓	✗	33	79	3.51
MACE	✗	✓	20	45	-
EquiformerV2	✗	✓	<u>12.4</u>	<u>32.22</u>	<u>2.48</u>
HIENet	✓	✓	<b>6.77</b>	<b>24.82</b>	<b>2.31</b>

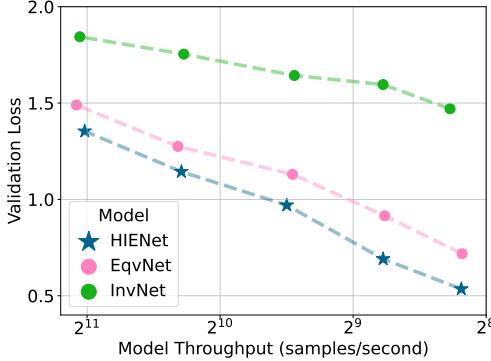
As shown in Table 3, HIENet outperforms all baseline models across all metrics for phonon frequency calculations. Additionally, while EquiformerV2 and ORB have good performance on MPtrj and Matbench Discovery, they have the worst performance among all models on this task. As previously mentioned, this may be because this task requires models predictions to obey physical constraints and be physically meaningful in order for the phonon calculations to be accurate.

**Evaluation on bulk modulus.** Model efficacy on zero-shot prediction of material properties was further evaluated on calculations of the fourth-order elastic tensor and the corresponding VRH average bulk modulus  $K_{VRH}$  [37]. A test set was generated by querying the Materials Project Database [38] for entries with between 1 and 6 sites that also had reported elasticity values. Following [31], we remove entries with highly unphysical bulk modulus reference values less than -50 GPa or greater than 600 GPa as well as those resulting in a calculated singular matrix, resulting in a final evaluation set of 1,763 crystal systems. Elastic tensors and bulk moduli were computed using the MatCalc’s Elasticity module [39]. Additional details on our workflow and dataset are provided in Appendix D.2.

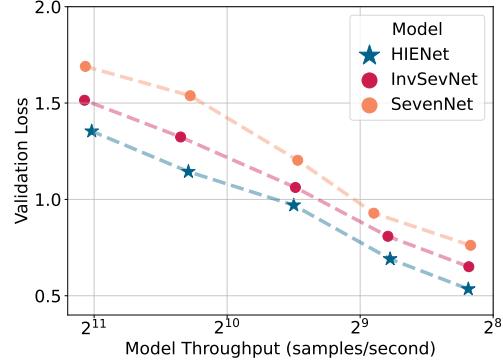
As shown in Table 4, HIENet outperforms all models on MAE and  $R^2$ . In fact, HIENet and SevenNet are the only models capable of achieving reasonable accuracy, demonstrating both the difficulty of this task and the robustness of our model. We provide parity plots for all models in Appendix.

### 5.4 HIENet efficiency

In addition to demonstrating improved performance across all benchmarks and downstream tasks, we show that HIENet is more computationally efficient than competing equivariant models. This is highly important for downstream materials discovery applications such as structural relaxation and



**Figure 3: Hybrid Architecture Ablation.** Model throughput measured on an Nvidia H100 GPU with a batch size of 128. EqvNet uses only equivariant message passing layers and InvNet uses only invariant layers. Validation loss measured on the MPtrj validation set. Hybrid invariant-equivariant models (HIENet) consistently outperform equivariant-only and invariant-only models across all model sizes.



**Figure 4: Hybrid Architecture Generality.** Model throughput measured on an Nvidia H100 GPU with a batch size of 128. InvSevNet represents the SevenNet model [18] with an additional invariant message passing layer. InvSevNet outperforms SevenNet across all tested model sizes, but still performs worse than HIENet.

random structure search, which require thousands of forward passes of the model. As seen in Table 5, HIENet is 90% faster than SevenNet-I3i5 and over 140% faster than EquiformerV2, all while having better performance than both models. Both EquiformerV2 and SevenNet exclusively use equivariant message passing layers, which limits the throughput and scalability of these models. At the same time, models without  $O(3)$  equivariant force and stress predictions, such as ORB, may be faster, but will perform poorly on realistic materials discovery tasks, as shown in Sec. 5.1 and 5.3.

### 5.5 Generality and Robustness of Hybrid Network Design

**Hybrid architecture ablation.** We provide an ablation study to demonstrate that our hybrid invariant-equivariant architecture outperforms invariant-only and equivariant-only models. In Figure 3 we see that HIENet outperforms EqvNet (only equivariant layers) and InvNet (only invariant layers) across a wide range of model sizes. Additionally, we see that InvNet consistently performs poorly, which aligns with our intuitions that equivariant message passing layers are important to capture high-order atomic interactions and accurately predict force and stress.

**Hybrid architecture generality.** To show that our idea of combining invariant and equivariant layers works well across different model designs, we add our invariant layer to SevenNet to form a hybrid model based on their equivariant designs. As shown in Figure 4, this InvSevNet consistently outperforms the base SevenNet even when controlling for model speed. Additionally, we observe that HIENet still outperforms the InvSevNet model across all tested model sizes. Because our hybrid invariant-equivariant works well with different models, we believe this approach provides a general new direction to design powerful and efficient MLIP models.

## 6 Conclusion, Limitations, and Societal Impacts

We propose HIENet, a machine learning interatomic potential for materials, and demonstrate the importance of (1) integrating both invariant and equivariant message-passing layers and (2) satisfying physical constraints to construct powerful and efficient MLIPs. HIENet achieves state-of-the-art performance across a range of benchmarks and downstream applications, while being significantly faster than existing equivariant models. We provide ablation studies to further demonstrate the generality and robustness of our hybrid network designs. Current limitations include (1) we currently focus on

materials discovery tasks, while extensions to molecular and protein science are underexplored and (2) due to limited computing resources, we face challenges in training HIENet on hundred-million-scale datasets, where we believe its full potential could be realized. We plan to explore these directions in the future work. The societal impacts of novel materials discovery may apply to this work.

## References

- [1] Juan J de Pablo, Nicholas E Jackson, Michael A Webb, Long-Qing Chen, Joel E Moore, Dane Morgan, Ryan Jacobs, Tresa Pollock, Darrell G Schlom, Eric S Toberer, et al. New frontiers for the materials genome initiative. *npj Computational Materials*, 5(1):41, 2019.
- [2] Eric Stach, Brian DeCost, A Gilad Kusne, Jason Hattrick-Simpers, Keith A Brown, Kristofer G Reyes, Joshua Schrier, Simon Billinge, Tonio Buonassisi, Ian Foster, et al. Autonomous experimentation systems for materials development: A community perspective. *Matter*, 4(9):2702–2726, 2021.
- [3] Shafidah Shafian, Fitri Norizatie Mohd Salehin, Sojeong Lee, Azlan Ismail, Shuhaida Mohamed Shuhidan, Lin Xie, and Kyungkon Kim. Development of Organic Semiconductor Materials for Organic Solar Cells via the Integration of Computational Quantum Chemistry and AI-Powered Machine Learning. *ACS Applied Energy Materials*, 2025.
- [4] Chade Lv, Xin Zhou, Lixiang Zhong, Chunshuang Yan, Madhavi Srinivasan, Zhi Wei Seh, Chuntai Liu, Hongge Pan, Shuzhou Li, Yonggang Wen, et al. Machine learning: an advanced platform for materials development and state prediction in lithium-ion batteries. *Advanced Materials*, 34(25):2101474, 2022.
- [5] Youbin Zheng, Ning Tang, Rawan Omar, Zhipeng Hu, Tuan Duong, Jing Wang, Weiwei Wu, and Hossam Haick. Smart materials enabled with artificial intelligence for healthcare wearables. *Advanced Functional Materials*, 31(51):2105482, 2021.
- [6] DB Miracle and DJ Thoma. Autonomous research and development of structural materials—an introduction and vision. *Current Opinion in Solid State and Materials Science*, 33:101188, 2024.
- [7] Xuan Zhang, Limei Wang, Jacob Helwig, Youzhi Luo, Cong Fu, Yaochen Xie, Meng Liu, Yuchao Lin, Zhao Xu, Keqiang Yan, Keir Adams, Maurice Weiler, Xiner Li, Tianfan Fu, Yucheng Wang, Haiyang Yu, YuQing Xie, Xiang Fu, Alex Strasser, Shenglong Xu, Yi Liu, Yuanqi Du, Alexandra Saxton, Hongyi Ling, Hannah Lawrence, Hannes Stärk, Shurui Gui, Carl Edwards, Nicholas Gao, Adriana Ladera, Tailin Wu, Elyssa F. Hofgard, Aria Mansouri Tehrani, Rui Wang, Ameya Daigavane, Montgomery Bohde, Jerry Kurtin, Qian Huang, Tuong Phung, Minkai Xu, Chaitanya K. Joshi, Simon V. Mathis, Kamyar Azizzadenesheli, Ada Fang, Alán Aspuru-Guzik, Erik Bekkers, Michael Bronstein, Marinka Zitnik, Anima Anandkumar, Stefano Ermon, Pietro Liò, Rose Yu, Stephan Günnemann, Jure Leskovec, Heng Ji, Jimeng Sun, Regina Barzilay, Tommi Jaakkola, Connor W. Coley, Xiaoning Qian, Xiaofeng Qian, Tess Smidt, and Shuiwang Ji. Artificial intelligence for science in quantum, atomistic, and continuum systems. *arXiv preprint arXiv:2307.08423*, 2023.
- [8] Yi-Lun Liao, Brandon M Wood, Abhishek Das, and Tess Smidt. EquiformerV2: Improved Equivariant Transformer for Scaling to Higher-Degree Representations. In *The Twelfth International Conference on Learning Representations*, 2024.
- [9] Mark Neumann, James Gin, Benjamin Rhodes, Steven Bennett, Zhiyi Li, Hitarth Choubisa, Arthur Hussey, and Jonathan Godwin. Orb: A Fast, Scalable Neural Network Potential, 2024.
- [10] Keqiang Yan, Cong Fu, Xiaofeng Qian, Xiaoning Qian, and Shuiwang Ji. Complete and efficient graph transformers for crystal material property prediction. In *International Conference on Learning Representations*, 2024.
- [11] Bowen Deng, Peichen Zhong, KyuJung Jun, Janosh Riebesell, Kevin Han, Christopher J. Bartel, and Gerbrand Ceder. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, page 1–11, 2023.

- [12] Johannes Gasteiger, Florian Becker, and Stephan Günnemann. GemNet: Universal Directional Graph Neural Networks for Molecules. In *Advances in Neural Information Processing Systems*, 2021.
- [13] Keqiang Yan, Yi Liu, Yuchao Lin, and Shuiwang Ji. Periodic graph transformers for crystal material property prediction. In *The 36th Annual Conference on Neural Information Processing Systems*, pages 15066–15080, 2022.
- [14] Stefan Elfwing, Eiji Uchibe, and Kenji Doya. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural networks*, 107:3–11, 2018.
- [15] Kristof Schütt, Pieter-Jan Kindermans, Huziel Enoc Sauceda Felix, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in neural information processing systems*, 30, 2017.
- [16] Kristof Schütt, Oliver Unke, and Michael Gastegger. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International Conference on Machine Learning*, pages 9377–9388. PMLR, 2021.
- [17] Chi Chen and Shyue Ping Ong. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, 2022.
- [18] Yutack Park, Jaesun Kim, Seungwoo Hwang, and Seungwu Han. Scalable parallel algorithm for graph neural network interatomic potentials in molecular dynamics simulations. *J. Chem. Theory Comput.*, 20(11):4857–4868, 2024.
- [19] Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation- and translation-equivariant neural networks for 3d point clouds, 2018.
- [20] Yi Liu, Limei Wang, Meng Liu, Yuchao Lin, Xuan Zhang, Bora Oztekin, and Shuiwang Ji. Spherical message passing for 3D molecular graphs. In *International Conference on Learning Representations*, 2022.
- [21] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13(1):2453, 2022.
- [22] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in Neural Information Processing Systems*, 35:11423–11436, 2022.
- [23] Zhao Xu, Haiyang Yu, Montgomery Bohde, and Shuiwang Ji. Equivariant graph network approximations of high-degree polynomials for force field prediction. *Transactions on Machine Learning Research*, 2024.
- [24] Bowen Deng, Peichen Zhong, KyuJung Jun, Janosh Riebesell, Kevin Han, Christopher J Bartel, and Gerbrand Ceder. CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 5(9):1031–1041, 2023.
- [25] Luis Barroso-Luque, Muhammed Shuaibi, Xiang Fu, Brandon M Wood, Misko Dzamba, Meng Gao, Ammar Rizvi, C Lawrence Zitnick, and Zachary W Ulissi. Open Materials 2024 (OMat24) Inorganic Materials Dataset and Models. *arXiv preprint arXiv:2410.12771*, 2024.
- [26] Tian Xie and Jeffrey C Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical review letters*, 120(14):145301, 2018.
- [27] Kamal Choudhary and Brian DeCost. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7(1):185, 2021.

- [28] Yuchao Lin, Keqiang Yan, Youzhi Luo, Yi Liu, Xiaoning Qian, and Shuiwang Ji. Efficient approximations of complete interatomic potentials for crystal property prediction. In *Proceedings of the 40th International Conference on Machine Learning*, pages 21260–21287, 2023.
- [29] Kamal Choudhary, Daniel Wines, Kangming Li, Kevin F Garrity, Vishu Gupta, Aldo H Romero, Jaron T Krogel, Kayahan Saritas, Addis Fuhr, Panchapakesan Ganesh, et al. JARVIS-Leaderboard: a large scale benchmark of materials design methods. *npj Computational Materials*, 10(1):93, 2024.
- [30] Eric Qu and Aditi S Krishnapriyan. The Importance of Being Scalable: Improving the Speed and Accuracy of Neural Network Interatomic Potentials Across Chemical Domains. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.
- [31] Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M Elena, Dávid P Kovács, Janosh Riebesell, Xavier R Advincula, Mark Asta, Matthew Avaylon, William J Baldwin, et al. A foundation model for atomistic materials chemistry. *arXiv preprint arXiv:2401.00096*, 2023.
- [32] Yutack Park, Jaesun Kim, Seungwoo Hwang, and Seungwu Han. Scalable Parallel Algorithm for Graph Neural Network Interatomic Potentials in Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation*, 20:4857–4868, 2024.
- [33] Janosh Riebesell, Rhys EA Goodall, Anubhav Jain, Philipp Benner, Kristin A Persson, and Alpha A Lee. Matbench Discovery—An evaluation framework for machine learning crystal stability prediction. *arXiv preprint arXiv:2308.14920*, 2023.
- [34] Janosh Riebesell and Aakaash A Naik. ffonons, 2024.
- [35] Atsushi Togo, Laurent Chaput, Terumasa Tadano, and Isao Tanaka. Implementation strategies in phonopy and phono3py. *J. Phys. Condens. Matter*, 35(35):353001, 2023.
- [36] Atsushi Togo. First-principles phonon calculations with phonopy and phono3py. *J. Phys. Soc. Jpn.*, 92(1):012001, 2023.
- [37] Richard Hill. The elastic behaviour of a crystalline aggregate. *Proceedings of the Physical Society. Section A*, 65(5):349, 1952.
- [38] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1), 2013.
- [39] Runze Liu, Elliott Liu, Janosh Riebesell, Ji Qi, Shyue Ping Ong, and Tsz Wai Ko. MatCalc, January 2024.
- [40] B. J. Alder and T. E. Wainwright. Studies in molecular dynamics. i. general method. *The Journal of Chemical Physics*, 31(2):459–466, 08 1959.
- [41] Magnus Rudolph Hestenes, Eduard Stiefel, et al. *Methods of conjugate gradients for solving linear systems*, volume 49. NBS Washington, DC, 1952.
- [42] Roger Fletcher. *Practical methods of optimization*. John Wiley & Sons, 2000.
- [43] P Hohenberg and WJPR Kohn. Density functional theory (dft). *Phys. Rev*, 136(1964):B864, 1964.
- [44] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.
- [45] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
- [46] Hendrik J. Monkhorst and James D. Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13:5188–5192, Jun 1976.

- [47] Axel Van De Walle, M Asta, and G Ceder. The alloy theoretic automated toolkit: A user guide. *Calphad*, 26(4):539–553, 2002.
- [48] Siya Zhu, Doğuhan SarıTÜRK, and Raymundo Arróyave. Accelerating calphad-based phase diagram predictions in complex alloys using universal machine learning potentials: Opportunities and challenges. *Acta Materialia*, page 120747, 2025.
- [49] Raymundo Arróyave. Phase stability through machine learning. *Journal of Phase Equilibria and Diffusion*, 43(6):606–628, 2022.
- [50] Alex Zunger, S-H Wei, LG Ferreira, and James E Bernard. Special quasirandom structures. *Physical Review Letters*, 65(3):353, 1990.
- [51] Georg Kresse and Jürgen Hafner. Ab initio molecular dynamics for liquid metals. *Physical Review B*, 47(1):558, 1993.
- [52] Georg Kresse and Jürgen Hafner. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous–semiconductor transition in germanium. *Physical Review B*, 49(20):14251, 1994.
- [53] Georg Kresse and Jürgen Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996.
- [54] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169, 1996.
- [55] Peter E Blöchl. Projector augmented-wave method. *Physical Review B*, 50(24):17953, 1994.
- [56] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865, 1996.
- [57] S-L Chen, S Daniel, FCYA Zhang, YA Chang, X-Y Yan, F-Y Xie, R Schmid-Fetzer, and WA Oates. The pandat software package and its applications. *Calphad*, 26(2):175–188, 2002.
- [58] Yoyo Hinuma, Giovanni Pizzi, Yu Kumagai, Fumiyasu Oba, and Isao Tanaka. Band structure diagram paths based on crystallography. *Computational Materials Science*, 128:140–184, 2017.
- [59] Kohei Shinohara Atsushi Togo and Isao Tanaka. Spglib: a software library for crystal symmetry search. *Science and Technology of Advanced Materials: Methods*, 4(1):2384822, 2024.
- [60] David C. Langreth and M. J. Mehl. Beyond the local-density approximation in calculations of ground-state electronic properties. *Phys. Rev. B*, 28:1809–1834, Aug 1983.
- [61] Maarten De Jong, Wei Chen, Thomas Angsten, Anubhav Jain, Randy Notestine, Anthony Gamst, Marcel Sluiter, Chaitanya Krishna Ande, Sybrand Van Der Zwaag, Jose J Plata, et al. Charting the complete elastic properties of inorganic crystalline compounds. *Scientific data*, 2(1):1–13, 2015.
- [62] Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional Message Passing for Molecular Graphs. In *International Conference on Learning Representations*, 2020.
- [63] Ilya Loshchilov and Frank Hutter. Decoupled Weight Decay Regularization. In *International Conference on Learning Representations*, 2019.
- [64] Ilya Loshchilov and Frank Hutter. SGDR: Stochastic Gradient Descent with Warm Restarts. In *International Conference on Learning Representations*, 2022.
- [65] Anton Bochkarev, Yury Lysogorskiy, and Ralf Drautz. Graph atomic cluster expansion for semilocal interactions beyond equivariant message passing. *Physical Review X*, 14(2):021036, 2024.

## A Molecular Dynamics Simulation

### A.1 Molecular Dynamics Simulations and Structural Optimization of Materials

**Molecular dynamics simulation.** Molecular dynamics (MD) simulation [40] is an important computational method to compute structural, chemical, and thermodynamic properties, which allows for in-depth mechanistic understanding and materials discovery. MD simulation essentially solves Newton’s equations of motion for both atomic positions and cell parameters of a material system under a specific thermodynamic ensemble. Specifically, the simulation workflow relies on iterative computation of the total system energy  $E$ , atomic forces  $\mathbf{F}_i$ , and stress tensor  $\sigma$ . For a given starting structure configuration,  $E$ ,  $\mathbf{F}_i$ , and  $\sigma$  can be calculated using classical methods or machine learning interatomic potentials. The acceleration, velocity, and position of atoms can be subsequently determined over a time step through numerical integration methods such as the Velocity-Verlet algorithm under a thermodynamic ensemble. The atomic forces of the new structure will then be updated for the next time step. By iterative numerical integration, the system will evolve under the thermodynamic ensemble and interatomic interactions determined by the force field. Stress also plays a crucial role in MD simulations when controlling pressure, such as in an NPT ensemble (i.e. under the constant number of particle, constant pressure, and constant temperature condition). In order to obtain statistically averaged physical quantities, such calculation needs to be performed iteratively for many time steps, hence computational efficiency becomes critical.

**Structural optimization.** Different from molecular dynamics, structural optimization usually aims to relax the structure and/or cell parameters to their ground state or metastable state. It also involves the calculation of energy, force and stress, which are subsequently used by optimization algorithms or optimizers to update the structure, such as Conjugate Gradient algorithm (CG) [41] and Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) [42]. This process is repeated until the final convergence criteria is reached,

### A.2 Conventional Computation Methods

Several kinds of simulation techniques are widely used in computational materials science at various scales, such as Density Functional Theory (DFT) [43, 44], MD simulations [40], and Monte Carlo (MC) simulations [45]. DFT is a quantum mechanical method that can be used to simulate material systems at the electronic level. Its key principle is that the ground-state energy of a system can be expressed as a functional of electron density, which reduces  $3N_e$ -dimensional interacting many-body system down to a fictitious 3-dimensional non-interacting system. However, DFT is computationally expensive and is limited to small systems. MD simulation method has already been elaborated in Appendix A.1 where a force field is required for calculating energy, force, and stress. There are two types of MD simulations depending on the underlying force field: *ab initio* MD (AIMD) simulations where atomic forces are calculated by quantum mechanical method such as DFT, and classical MD simulations where empirical force fields are used to calculate atomic forces. AIMDs are relatively more accurate but computationally expensive, limiting its application to small systems. Classical MD simulations are computationally efficient and can handle large systems, but very often they either lack the accuracy required for highly precise simulations, or cannot be transferred to different simulation conditions. MC simulations are based on statistical mechanics which rely on iterative energy calculations and configuration sampling and updates. Another key challenge is that both classical MD and MC simulations depend on the availability of empirical force fields for the system of interest. Therefore, it is highly desirable to develop machine learning interatomic potentials that can provide accurate and efficient calculations of energy, force, and stress of arbitrary materials system, which will significantly advance materials science, physics and chemistry and allow for studying fundamental mechanism and discovering new materials.

## B Evaluations on *Ab Initio* Molecular Dynamics

As mentioned in Appendix A.1, *Ab initio* molecular dynamics (AIMD) simulation is an incredibly important application of machine learning interatomic potentials (MLIPs). Here we evaluate HIENet and baseline MLIPs on AIMD simulations.

To evaluate MLIPs performance, we generate a testing dataset consisting of silicon (Si) systems containing 64 atoms in a  $2 \times 2 \times 2$  supercell. A  $\Gamma$ -centered Monkhorst–Pack k-point sampling

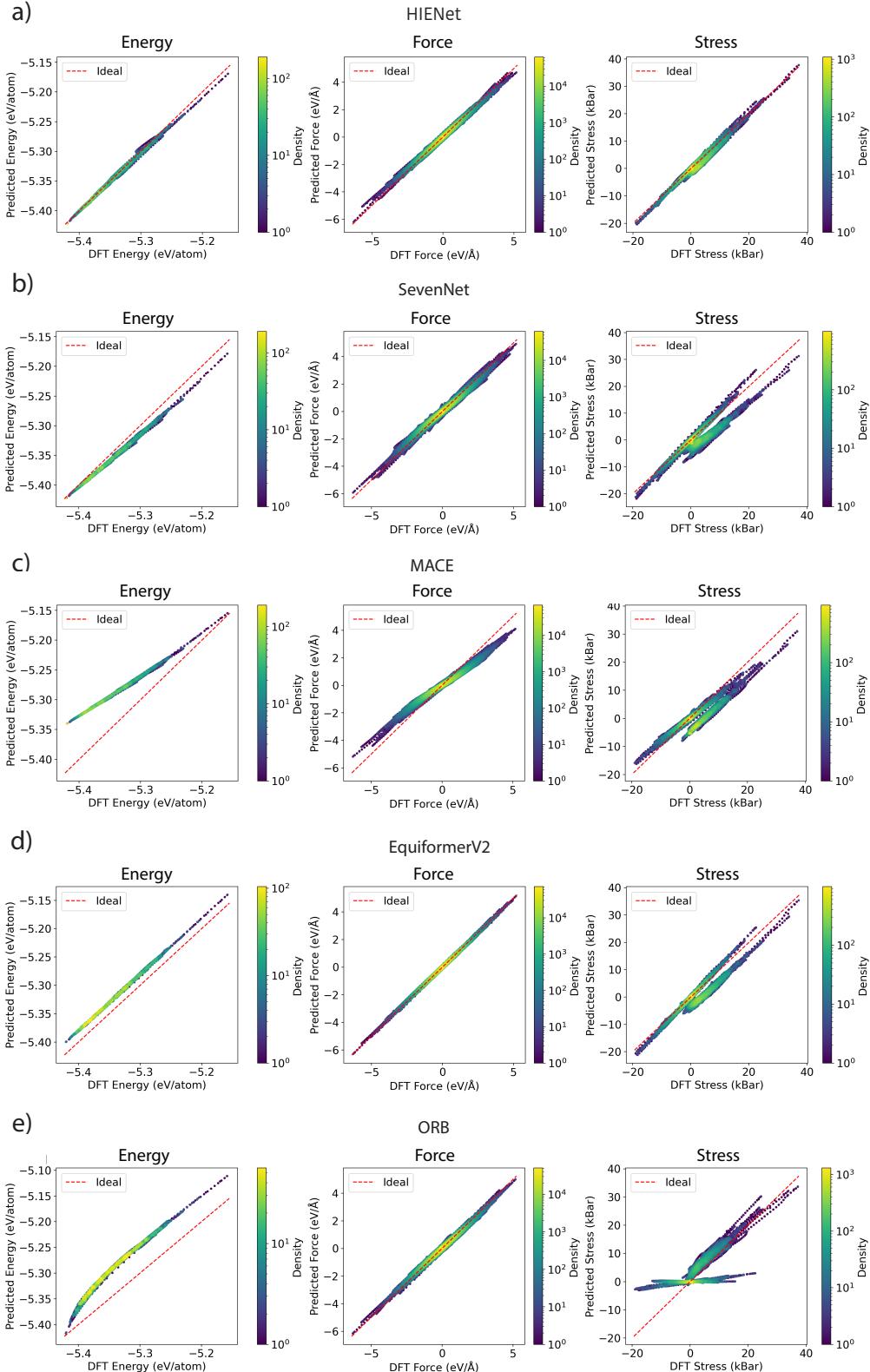


Figure 5: Evaluation of energy, force, and stress predictions for 64-atom Si system calculated by foundation models: a) HIE Net, b) SevenNet-l3i5, c) MACE-MP-0, d) CHGNet, and e) eqV2\_31M\_mp with respect to the DFT results.

Table 6: MLIP prediction accuracy across 4,000 configurations of Si systems. Energy MAE is in meV/atom, Force MAE is in meV/ $\text{\AA}$ , and Stress MAE is in kBar. Best performing model for each metric in **bold** and second best underlined.

		HIENet	EquiformerV2	ORB	SevenNet-l3i5	MACE	CHGNet
Energy	MAE	<b>2</b>	19	48	<u>10</u>	55	79
	$R^2$	<b>0.995</b>	0.794	-0.344	<u>0.932</u>	-0.762	-2.617
Force	MAE	<u>62</u>	<b>29</b>	65	98	208	267
	$R^2$	<u>0.989</u>	<b>0.998</b>	0.988	0.974	0.885	0.805
Stress	MAE	<b>0.995</b>	<u>2.065</u>	2.693	2.171	2.918	4.224
	$R^2$	<b>0.938</b>	<u>0.706</u>	0.507	0.652	0.466	-0.102

grid of  $2 \times 2 \times 2$  [46] was used. AIMD simulations were performed in the NVT ensemble with a Nosé-Hoover thermostat at four temperatures of 300, 500, 700, and 900 K with time step of 1 fs for 1,000 steps at each temperature. In total, 4,000 configurations were generated for model evaluation. Since configurations are sampled from a variety of temperatures, this task represents an out-of-distribution generalization problem compared to the MPtrj training dataset. We select Si systems because it is a representative material of great interest and importance to the semiconductor industry.

AIMD simulations were conducted using DFT as implemented in VASP with the PBE exchange-correlation energy functional. A plane-wave basis set with a cutoff energy of 520 eV was used to ensure numerical accuracy in the simulations. To ensure consistency between training and evaluation, all input settings were generated using the MPRelaxSet class.

For each system configuration, we compute MLIP energy, forces, and stress and compare with DFT reference data. As shown in Table 6, HIENet achieves vastly better accuracy on energy and stress performance compared to baseline models, though EquiformerV2 has better accuracy on force predictions. Parity plots for each model are shown in Fig. 5, where we observe that HIENet consistently performs well across all system configurations.

## C Evaluations on Alloys

We also evaluate MLIP performance on phase diagram calculations using the Alloy Theoretic Automated Toolkit (ATAT) [47] framework following the approach outlined in [48]. Phase diagrams are graphical representations of the state of materials under arbitrary conditions and accurately predicting them is a necessary condition for the further development of complex materials [49].

Starting with the simple Au-Pt binary systems, we first generate Special Quasirandom Structures (SQS) [50] of FCC Au-Pt with different compositions using ATAT, with 32 atoms in a  $2 \times 2 \times 2$  supercell—the SQS structures are designed to mimic disordered alloys within a certain precision. Then, the relaxation and free energy calculations are carried out using ab initio calculations and MLIPs. For all *ab initio* calculations, VASP [51, 52, 53, 54] is used with the PBE exchange-correlation functional and PAW pseudopotentials at the level of GGA [55, 56]. The k-point density is set to 8,000 k-points per reciprocal atom for all calculations.

In Figure 6, we plot the formation energies of the Au-Pt FCC binary systems calculated by HIENet and baseline MLIPs. We see that HIENet shows strong agreement with first-principles DFT results as our model predictions closely match the true formation energy across all Au concentrations.

In addition, although all the models successfully give a positive formation energy for the SQS’s, predicting the miscibility gap in the phase diagram, most of the models including CHGNet, MACE, ORB, GRACE and EquiformerV2 fail to reproduce the correct ordering of the formation energies:  $\Delta G(x_{\text{Au}} = 0.5) > \Delta G(x_{\text{Au}} = 0.25) > \Delta G(x_{\text{Au}} = 0.75)$ , as shown in Table 7. Such ordering of formation energies is highly important in thermodynamics and materials science, as it governs the stability of the phases.

Finally, we demonstrate how HIENet can be used for multi-element systems. In Figure 7, we present a ternary phase diagram for the Cr-Mo-V system at 1,000 K calculated with ATAT and HIENet. The

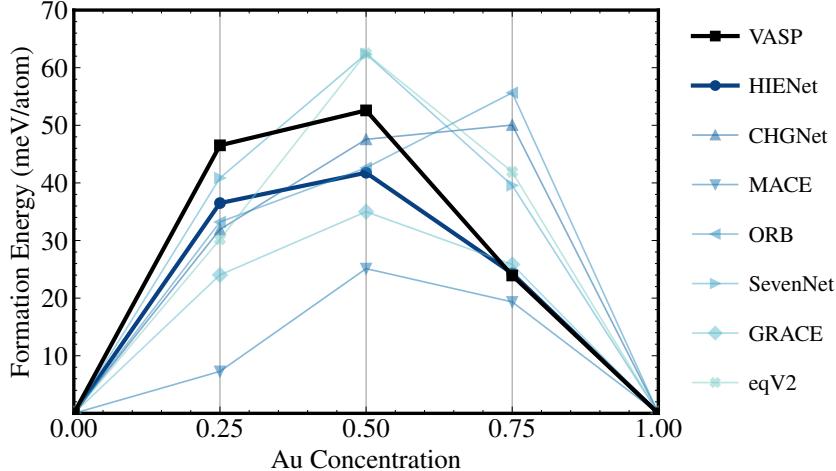


Figure 6: Formation energies per atom of the Au-Pt binary FCC system calculated with models trained on the MPtrj dataset. eqV2 refers to EquiformerV2 and SevenNet is the SevenNet-l3i5 model.

Table 7: Ordering of Au-Pt formation energies calculated with different potentials (1 for the lowest formation energy and 3 for the highest). Ideally, MLIP predictions should match the VASP ordering of formation energies.

Model	Ordering of Formation Energies		
	$\Delta G(x_{\text{Au}} = 0.25)$	$\Delta G(x_{\text{Au}} = 0.5)$	$\Delta G(x_{\text{Au}} = 0.75)$
CHGNet	1	2	3
MACE	1	3	2
ORB	1	2	3
SevenNet-l3i5	2	3	1
GRACE	1	3	2
EquiformerV2	1	3	2
HIENet	2	3	1
VASP	2	3	1

ternary phase diagram calculation correctly identifies the BCC phase miscibility gap in the Cr-Mo region.

## D Evaluations on Phonon and Bulk Modulus

### D.1 Phonon Frequency Evaluation

As the calculations of the Material Project phonon dataset were performed using the PBEsol exchange-correlation energy functional, it would be inconsistent to compare them with the models trained on the data using the Perdew-Burke-Ernzerhof (PBE) [56] exchange-correlation energy functional. PhononDB, a database of phonon calculations including band structure, DOS, and thermal properties for over 10,000 materials evaluated using the PBE functional, provides a more effective reference for comparison, hence was used as the reference for the evaluation as detailed below. Phonon frequencies and corresponding band structures were computed using the Phonopy package via the finite displacement method [35, 36] where MLIPs were employed to compute the dynamical matrices and corresponding phonon band structures of each crystal structure. To ensure direct comparison between PhononDB and calculated data, the Phonopy objects were initialized with the same unit cell and supercell matrices as used in PhononDB calculations. Additionally, the primitive cell matrix was included if defined. Displaced supercells were generated using a default displacement of 0.01 Å and the corresponding forces were evaluated with our model. High-symmetry k-path in the Brillouin zone was computed using SeeK-Path [58, 59]. Using this workflow, the high-symmetry k-paths and the

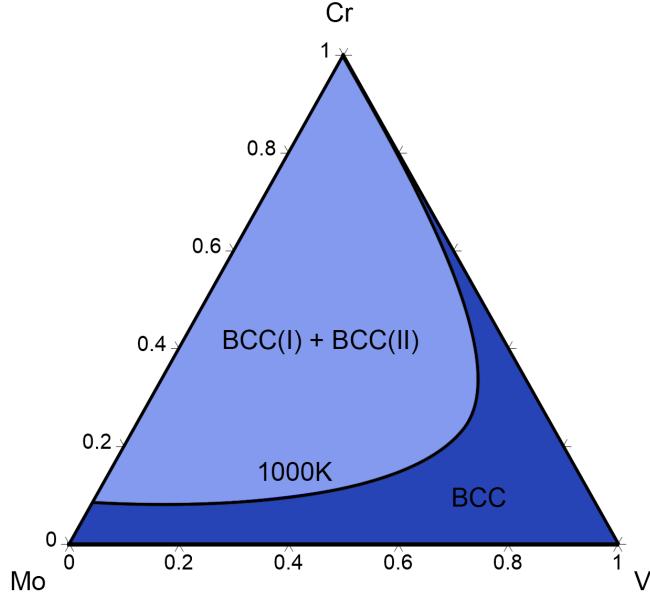


Figure 7: Cr-Mo-V ternary phase diagram at 1,000 K calculated with ATAT and HIENet. Only the BCC phase is included in the calculation. The phase diagram is plotted with the Pandat [57] software package.

sampling grids were identical between the reference phonon band structure from PhononDB and the predicted band structure from our model.

In addition to the frequency evaluation in Table 3, we provide several phonon band structure diagrams calculated using HIENet in Figure 8 for Si, CdTe,  $\text{Cs}_2\text{KInF}_6$ , and  $\text{GaAgS}_2$  systems. We observe that the HIENet-predicted phonon band structures of Si exhibits reasonable accuracy, and the phonon band structures for CdTe,  $\text{Cs}_2\text{KInF}_6$ , and  $\text{GaAgS}_2$  are in very good agreement with the PhononDB DFT results across the entire frequency range and high-symmetry k-paths. Furthermore, the phonon band structure of  $\text{Cs}_2\text{KInF}_6$  contains negative phonon frequencies, indicating the dynamical instability of the crystal structure despite its local stability. Impressively, the HIENet predictions agree with the DFT data extremely well even in this negative frequency regime across all high-symmetry pathways. As such, HIENet can be a powerful MLIP for predicting a materials thermal conductivity and structural stability.

## D.2 Bulk Modulus Evaluation

To compute bulk modulus, we need to calculate the elastic tensor for each crystal. The latter is calculated by first relaxing the input structure to the default force tolerance of 0.1 eV/Å using each MLIP. The relaxed structure is then deformed with strains of  $(\pm 0.005, \pm 0.01)$  applied to normal modes and strains of  $(\pm 0.06, \pm 0.03)$  applied to shear modes for a total of 4 strain magnitudes for each of the 6 strain modes. The resulting stress-strain values are fit linearly to compute the elastic tensor. The reference elastic constants in the Materials Project were calculated using DFT with the PBE functional in the generalized gradient approximation (GGA) [60] as implemented the Vienna Ab-initio Simulation Package (VASP) [54]. For metallic entries, a plane wave cutoff energy of 700 eV with k-point density of 7,000 per reciprocal atom was used. For non-metallic entries such as insulators or semiconductors, a plane wave cutoff energy of 700 eV was once again used with a k-point density of 10,000 per reciprocal atom [61]. In addition to the main results reported in Table 4, we provide parity plots for each model in Fig. 9.

## E Additional Ablation Results

To empirically justify why we use  $O(3)$  equivariant crystal graph representations instead of the geometrically complete but  $SO(3)$  equivariant crystal graphs from [10], we provide an additional

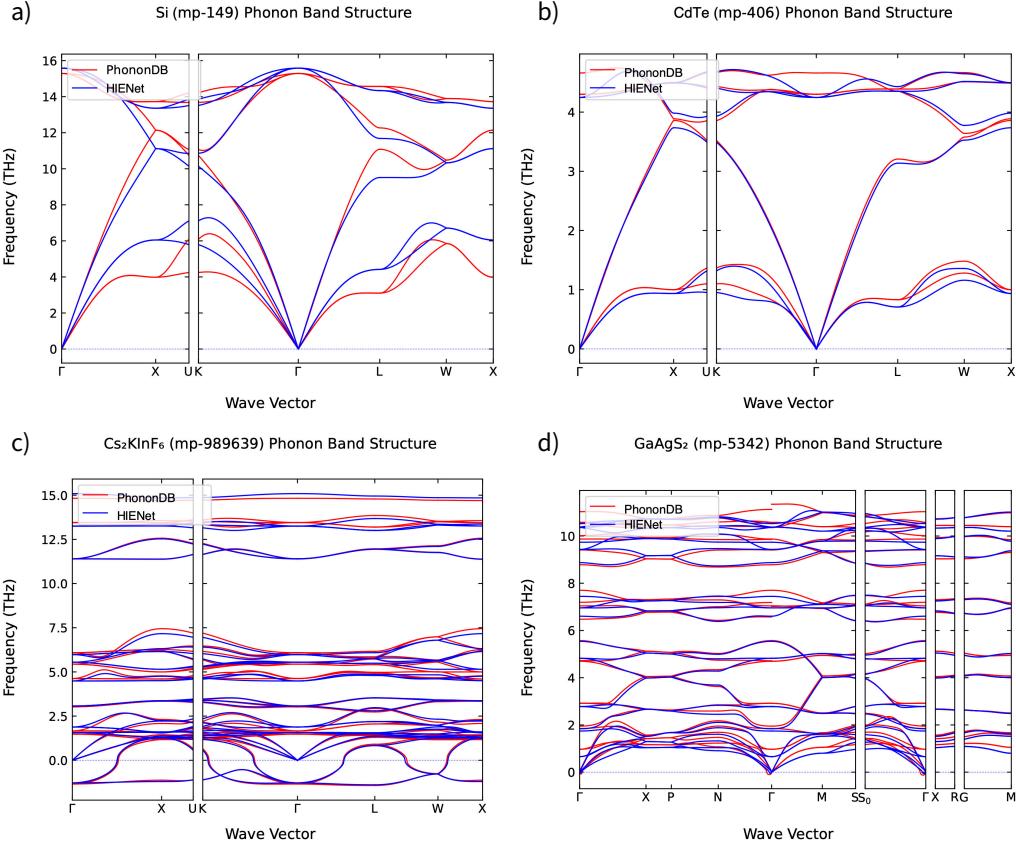


Figure 8: Phonon band structures for a) Si, b) CdTe, c)  $\text{Cs}_2\text{KInF}_6$ , and d)  $\text{GaAgS}_2$  calculated using HIENet compared with reference data in the PhononDB database.

Table 8: Mean absolute errors on MPtrj validation set for HIENet with O(3) and SO(3) equivariant crystal graphs. Models trained for 20 epochs on the MPtrj dataset. Best performing model in **bold**.

Equivariance	Energy ↓	Force ↓	Stress ↓
	(meV/atom)	(meV/Å)	(kBar)
SO(3)	19.13	56.12	3.98
O(3)	<b>16.26</b>	<b>49.29</b>	<b>3.48</b>

ablation study in Table 8 where we include the additional periodic encoding from [10]. We observe that SO(3) equivariant HIENet performs slightly worse, which aligns with our intuitions as the underlying DFT algorithm is O(3) equivariant.

## F Model Settings and Experimental Details

### F.1 HIENet Settings

HIENet consists of 1 invariant and 3 equivariant message passing layers. For the invariant message passing layers, we use a hidden dimension of 512 for node features and a single attention head. The equivariant layers use a representation that consists of 512 scalar channels with  $l = 0$ , 128 vector channels with  $l = 1$ , 64 higher-order tensor channels with  $l = 2$ , and 32 higher-order tensor channels with  $l = 3$ . We use 8 radial Bessel basis functions for distance encoding and a polynomial envelope [62] with  $p = 6$ . We use SiLU and sigmoid activation functions [14] throughout the network

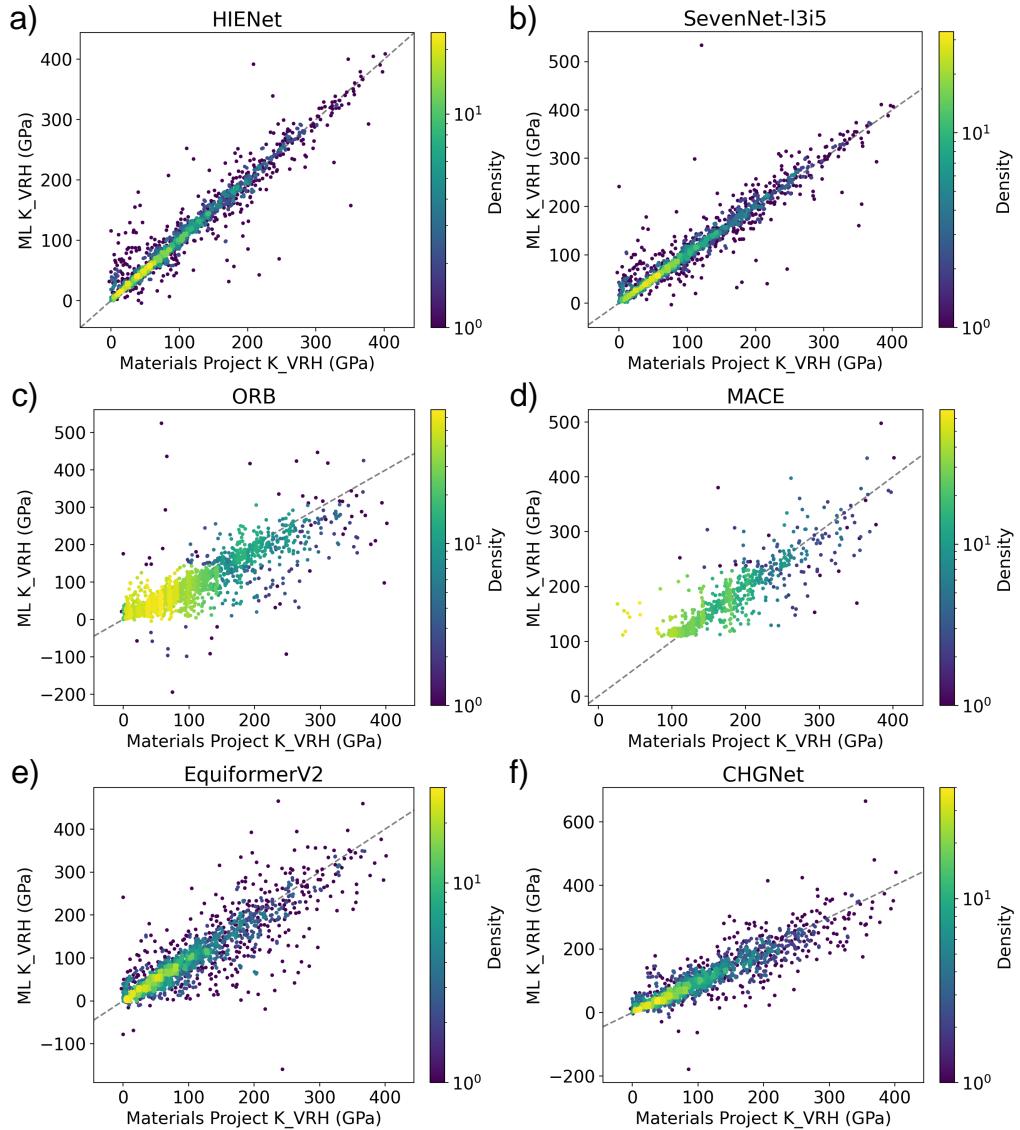


Figure 9: Comparison of bulk modulus  $K_{VRH}$  calculated by a) HIENet, b) SevenNet-I3i5, c) ORB, d) MACE, e) EquiformerV2, and f) CHGNet with the reference data in the Materials Project database.

to ensure smooth and continuously differentiable gradients. To prevent overfitting, we regularize the model by applying dropout to the MLPs that operate on scalar features in both invariant and equivariant message passing layers. Specifically, we employ a dropout rate of  $p_{\text{attn}} = 0.1$  for MLPs involved in attention calculations, while using a lower rate of  $p = 0.06$  for all other MLPs in the network. Additionally, we scale the input energies by the root mean square (RMS) of forces from the training dataset and shift by element-wise reference energies from the same dataset.

Following [31], we split the Materials Project Trajectory (MPtrj) Dataset [24] into training (95%) and validation (5%) sets. We train the model for 250 epochs on a platform with 2 AMD EPYC 7J13 64-Core Processors (240 cores total), 1.7 TiB DDR4 memory, and 8 NVIDIA A100-SXM4-80GB GPU accelerators. We use a total batch size of 384 (48 per GPU), which results in the model taking 118 minutes per training epoch and 6 minutes per validation epoch.

## F.2 Optimization

We optimize HIENet using the AdamW optimizer [63] with weight decay of 0.001. The learning rate follows a cosine annealing schedule [64] with an initial warm-up phase to stabilize early training.

The loss function combines energy, force, and stress predictions with different weighting factors as:

$$\mathcal{L} = \lambda_E \mathcal{L}_E + \lambda_F \mathcal{L}_F + \lambda_\sigma \mathcal{L}_\sigma \quad (17)$$

where  $\mathcal{L}_E$ ,  $\mathcal{L}_F$ , and  $\mathcal{L}_\sigma$  represent the Huber losses for energy, force, and stress predictions, respectively, with  $\delta = 0.01$ . We set the weighting coefficients  $\lambda_E = 1.0$ ,  $\lambda_F = 1.0$ , and  $\lambda_\sigma = 0.01$ .

To improve model generalization and training stability, we additionally maintain an exponential moving average (EMA) of model parameters with a decay rate of 0.999.

The hyperparameters for both the model architecture and optimization are summarized in Table 9.

Table 9: Hyperparameters used for model training.

Hyperparameter	Value
Optimizer	AdamW
Learning rate scheduler	Cosine Annealing
Maximum learning rate	0.01
Minimum learning rate	0.000005
Warmup epochs	0.1
Warmup factor	0.2
Number of epochs	250
Batch size	48
Weight decay	0.001
Dropout rate, $p$	0.06
Attention dropout rate, $p_{\text{attn}}$	0.1
Energy loss weight, $\lambda_E$	1.0
Force loss weight, $\lambda_F$	1.0
Stress loss weight, $\lambda_\sigma$	0.01
Model EMA Decay	0.999

## F.3 Envelope Function

As mentioned in Sec. 3.2, we use the polynomial envelope function [12]:

$$f_{\text{poly}}(r) = 1 - \frac{(p+1)(p+2)}{2} d^p + p(p+2)d^{p+1} - \frac{p(p+1)}{2} d^{p+2} \quad (18)$$

where  $p \in \mathbb{Z}, 0 < p$ . In practice, we select  $p = 6$ . It is critical to have such an envelope function in order to ensure that the MLIP energy predictions are continuously differentiable with respect to atom positions. The polynomial envelope was selected because the first and second derivatives of  $h_{ji}$  will then go to 0 at the cutoff radius  $R_{\max}$ .

## G Licenses for Existing Assets

We have used datasets including the Materials Project Trajectory (MPtrj) dataset [24] with MIT License and Materials Project Database [38] with the Creative Commons Attribution 4.0 International License. For evaluations, we have used the Matbench Discovery benchmark [33] with MIT License, Phonopy and Togo PhononDB Database [35, 36] with the BSD 3-Clause License, Alloy Theoretic Automated Toolkit (ATAT) [47] with the Creative Commons Attribution-NoDerivatives 4.0 International License, and MatCalc’s Elasticity module [39] with the BSD 3-Clause License. For model comparisons, we included EquiformerV2 [8] with the MIT License, ORB [9] with the Apache License Version 2.0, SevenNet [18] with the GNU General Public License Version 3.0, GRACE [65] with the Academic Software Licence, MACE [31] with the MIT License, and CHGNet [11] with the BSD 3-Clause License.