

A Materials Foundation Model via Hybrid Invariant-Equivariant Architectures

Keqiang Yan^{1*} Montgomery Bohde^{1*} Andrii Kryvenko^{1*} Ziyu Xiang² Kaiji Zhao³ Siya Zhu³
 Saagar Kolachina³ Doğuhan SarıTÜRK³ Jianwen Xie⁴ Raymundo Arróyave^{3,5,6} Xiaoning Qian^{1,2,7}
 Xiaofeng Qian^{3,2,8} Shuiwang Ji¹

Abstract

Materials foundation models can predict energy, force, and stress of materials and enable a wide range of downstream discovery tasks. A key design choice involves the trade-off between invariant and equivariant architectures. Invariant models offer computational efficiency but may not perform well 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 foundation model that integrates both invariant and equivariant message passing layers. HIENet is designed to achieve superior 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 (de Pablo et al., 2019; Stach et al., 2021), including improving semiconductor (Shafian et al., 2025), enabling more efficient renewable energy storage (Lv et al., 2022), developing new sensing technologies (Zheng et al., 2021), and enabling

*Equal contribution ¹Department of Computer Science & Engineering, Texas A&M University, College Station, TX, USA
²Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX, USA ³Department of Materials Science & Engineering, Texas A&M University, College Station, TX, USA ⁴Lambda Inc., USA ⁵J. Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX, USA ⁶Wm Michael Barnes '64 Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, USA ⁷Computing & Data Sciences, Brookhaven National Laboratory, Upton, NY, USA ⁸Department of Physics & Astronomy, Texas A&M University, College Station, TX, USA. Correspondence to: Keqiang Yan <keqiangyan@tamu.edu>, Shuiwang Ji <sji@tamu.edu>.

many other engineering applications (Miracle & Thoma, 2024). However, traditional materials discovery relies heavily on the expertise, intuition and innovation of materials scientists, often employing time-consuming and costly trial-and-error experimental methods. Over the past two decades, computational approaches, particularly those leveraging advanced quantum mechanical methods like first-principles density functional theory (DFT), have accelerated this process (Zhang et al., 2023). Despite their benefits, these methods come with significant computational costs, scaling from $O(n_e^3)$ to $O(n_e^7)$, where n_e is number of electrons in a materials system. This steep scaling renders high-throughput screening across the vast space of materials challenging. Simulating systems with a large number of atoms with quantum mechanical methods becomes extremely expensive with current approaches.

Recent progress in materials foundation models offers a promising path forward by enabling the prediction of energies, forces, and stresses of materials, achieving significant speedups compared to traditional DFT methods. However, these existing foundation 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 capture high-order interactions better but are computationally expensive, making scaling to larger model sizes challenging and expensive. Additionally, model predictions must adhere to key physical constraints: energy should remain invariant under global symmetry operations such as translation, rotation, and reflection, while forces and stresses must be equivariant under the same operations. Moreover, physical laws impose additional requirements, including force equilibrium $\sum \mathbf{F}_i = \mathbf{0}$ when no external influences applied, force conservation $\mathbf{F} = -\nabla E$ (E is total energy), and stress tensor symmetry $\sigma_{ij} = \sigma_{ji}$.

In this work, we propose HIENet, a materials foundation model that satisfies key physical constraints for energy, force, and stress predictions while integrating invariant and equivariant designs to achieve state-of-the-art (SOTA) performance with significant computational speedups compared to existing models. An overview of HIENet is provided in Figure 1. Unlike prior approaches that rely ex-

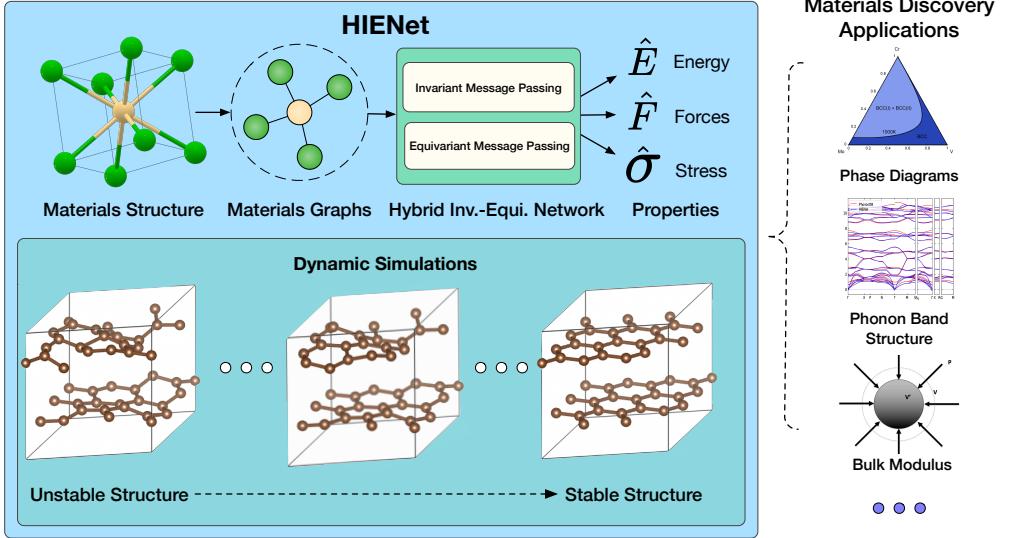


Figure 1. HIENet overview and materials property prediction pipeline. The model converts material structures into graph representations and processes them through a hybrid architecture combining invariant and equivariant message passing networks to predict physical properties. The model supports accurate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

clusively on either invariant or equivariant layers, HIENet balances these strategies to improve both performance and efficiency. Specifically, HIENet leverages the scalability of invariant designs to increase model size and capacity while utilizing equivariant designs to effectively capture high-order interactions and symmetries, which are critical for learning forces accurately, as demonstrated by our experiments. Moreover, in contrast to existing models like EquiformerV2 (Liao et al., 2024) and Orb (Neumann et al., 2024), 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 and Matbench Discovery, and downstream materials discovery tasks including evaluations on phonon band structures, bulk modulus, *ab initio* molecular dynamics, and alloys as detailed in Sec. 4.4, 4.5, and 4.6 demonstrate the efficiency and effectiveness of HIENet.

2. Preliminaries and Related Work

The development of machine learning force fields (MLFFs) requires carefully balancing physical constraints, computational efficiency, and model expressivity. Key challenges arise from the need to preserve symmetries while enabling fast predictions of material properties. In this section, we first formalize the fundamental prediction tasks and outline the unique challenges of crystal structures and their dynamics, followed by an examination of existing MLFF approaches and their limitations.

2.1. Problem Definition

The core task in developing MLFFs is to learn a mapping from atomic structures to quantum mechanical properties while preserving fundamental physical symmetries. Given a crystal structure, we aim to predict three quantities; the total energy E , indicating system stability, atomic forces $\mathbf{F} = \{\mathbf{F}_i \in \mathbb{R}^3, 1 \leq i \leq n\}$, where n denotes number of atoms in a cell, driving structural evolution, and the stress tensor $\boldsymbol{\sigma} \in \mathbb{R}^{3 \times 3}$ governing cell deformation. These predictions must satisfy key physical constraints: energy should be invariant under global symmetry operations such as translation, rotation, and reflection, while forces and stress must be equivariant under these operations. Additionally, physical conservation laws require force equilibrium $\sum \mathbf{F}_i = \mathbf{0}$ when no influences applied, force conservation $\mathbf{F} = -\nabla E$, and stress tensor symmetry $\sigma_{ij} = \sigma_{ji}$.

3D crystal structures. Unlike regular molecules, crystals are periodic in nature and are characterized as three-dimensional lattices with indefinitely repeating unit cells. Adopting the notation of Yan et al. (2024), a crystal structure can be described as a triple $\mathcal{M} = (\mathbf{Z}, \mathbf{P}, \mathbf{L})$, 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 each 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 $\mathbf{P} = [p_1, p_2, \dots, p_n] \in \mathbb{R}^{3 \times n}$. 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. These vectors can be used to form an infinite crystal structure that

can be expressed formally as a pair of two infinite sets: $\hat{\mathbf{P}} = \{\hat{\mathbf{p}}_i \mid \hat{\mathbf{p}}_i = \mathbf{p}_i + k_1 \boldsymbol{\ell}_1 + k_2 \boldsymbol{\ell}_2 + k_3 \boldsymbol{\ell}_3, k_1, k_2, k_3 \in \mathbb{Z}, i \in \mathbb{Z}, 1 \leq i \leq n\}$, $\hat{\mathbf{Z}} = \{\hat{z}_i \mid \hat{z}_i = z_i, i \in \mathbb{Z}, 1 \leq i \leq n\}$, where $\hat{\mathbf{P}}$ represents atomic positions in the infinite crystal structure, and $\hat{\mathbf{Z}}$ defines their corresponding atomic numbers.

Molecular dynamics simulation and structural optimization for materials. MLFFs can be applied to perform molecular dynamics simulation and optimize the structure of materials. Molecular dynamics simulation is an important method in computational materials science which provides insights about structural, chemical, and thermodynamic properties and allows for in-depth mechanistic understanding and materials discovery. Molecular dynamics simulation 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 σ . MLFFs can also be applied to optimize structures through iterative energy minimization. Details on the molecular dynamics simulation and the structural optimization are provided in Appendix A.1.

2.2. Related Work

In this section, we focus on machine learning-based interatomic potentials and provide related works in conventional computation methods in Appendix A.2.

Recent advances in materials property prediction models (Xie & Grossman, 2018; Choudhary & DeCost, 2021; Yan et al., 2022; Lin et al., 2023; Choudhary et al., 2024; Yan et al., 2024) and the availability of high-quality materials dynamics datasets (Chen & Ong, 2022; Deng et al., 2023; Barroso-Luque et al., 2024) generated using DFT-based algorithms have facilitated the development of materials foundation models capable of predicting energy, force, and stress. Among these models, purely invariant designs, such as M3GNet (Chen & Ong, 2022), CHGNet (Deng et al., 2023), Orb (Neumann et al., 2024), and EScAIP (Qu & Krishnapriyan, 2024), are computationally efficient, but typically do not perform as well, particularly for predicting higher order tensors such as force and stress. While some invariant models such as Orb (Neumann et al., 2024), and EScAIP (Qu & Krishnapriyan, 2024) are able to achieve competitive performance on benchmarks, this does not necessarily translate to downstream tasks, as shown in Sec. 4.6. In contrast, equivariant designs, including MACE (Batatia et al., 2023), SevenNet (Park et al., 2024), and EquiformerV2 (Barroso-Luque et al., 2024), respect physical constraints like equivariance. However, they are computationally expensive due to the extensive use of tensor product operations, which limits their scalability. Additionally, some equivariant models still do not obey other

important physical laws, undermining their performance on realistic material system simulations as seen in Sec. 4.

Different from these prior foundation models, our proposed HIENet detailed in Sec. 3 satisfies all key physical constraints while combining 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 materials foundation model design.

3. Hybrid Invariant-Equivariant Networks

Our HIENet model architecture is based on the ComFormer model (Yan et al., 2024), with several important changes to the model architecture and crystal graph construction. We emphasize that our purpose in this work is to develop a materials foundation model and demonstrate the importance of including both invariant and equivariant message passing layers, not to analyze the specific designs of these layers, as this has been extensively studied by previous works (Gasteiger et al., 2021; Batatia et al., 2022a; Liao et al., 2024). A detailed diagram of the HIENet architecture can be seen in Figure 2. In this section we only provide an overview of the key differences from ComFormer and justifications for these changes. See Appendix B for a more detailed description of our HIENet model architecture.

3.1. Physical Constraints

In order for MLFF models to be practical and robust for downstream tasks, they need to ensure various physical symmetries. Specifically, for force prediction, models need to ensure forces form a conservative vector field and that the forces on each atom sum to zero excluding friction or external forces. For stress prediction, they must ensure the stress tensor σ is symmetric. Finally, models need to ensure that each of these predictions transforms appropriately under physical transformations of the crystal system.

While models such as EquiformerV2 (Barroso-Luque et al., 2024) and Orb (Neumann et al., 2024) are able to achieve good performance on benchmarks, these results do not translate to downstream molecular dynamics stimulation tasks because the model predictions do not obey basic physical laws, as shown in Sec. 4.

Gradient-Based Calculations. 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:

$$\hat{\mathbf{F}}_i = -\nabla_{\mathbf{p}_i} \hat{E} \quad (1)$$

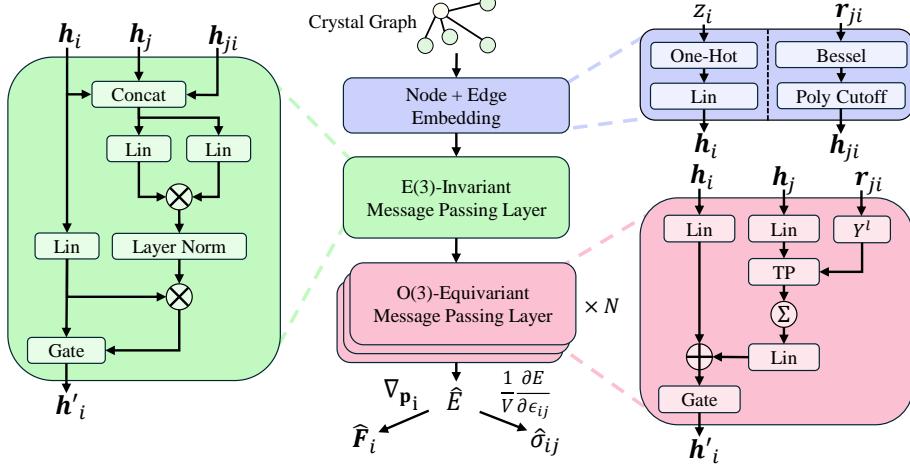


Figure 2. HIENet Model Architecture. We construct an $O(3)$ equivariant crystal graph representation and embed node and edge features. We then apply an invariant message passing layer followed by several equivariant message passing layers before predicting the total energy, \hat{E} and using physical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

where $\nabla_{\mathbf{p}_i}$ represents the gradient with respect to the position vector \mathbf{p}_i .

Proposition 3.1. *HIENet predictions $\hat{\mathbf{F}}_i$ form a conservative vector field.*

Proposition 3.2. *HIENet predictions satisfy force equilibrium $\sum_{i=1}^N \hat{\mathbf{F}}_i = \mathbf{0}$ when no external influences applied.*

Proofs of Props. 3.1, 3.2 are in Appendix C. We compute the stress tensor as:

$$\hat{\sigma}_{ij} = \frac{1}{V} \frac{\partial \hat{E}}{\partial \epsilon_{ij}} \quad (2)$$

where ϵ is the lattice strain tensor and V is the volume of the unit cell. We ensure that $\hat{\boldsymbol{\sigma}}$ will be symmetric by first symmetrizing the strain matrix $\epsilon_{\text{sym}} = \frac{1}{2}(\epsilon + \epsilon^\top)$.

O(3) Equivariance. Beyond physical constraints on forces and stress, we also want force and stress tensors to transform appropriately under rotations and reflections of the input material. HIENet achieves $O(3)$ equivariance by using the gradient-based approach to compute force and stress. Prior works (Chen & Ong, 2022; Deng et al., 2023) have shown that even invariant models can produce equivariant force and stress predictions using this method. By relying on these physics-informed calculations rather than directly predicting forces and stress, our model strictly satisfies all required physical constraints.

Importantly, we use a different graph construction than ComFormer, and exclude additional periodic encodings that cause ComFormer to be $SO(3)$ equivariant. $O(3)$ equivariance is important for model predictions to be physically meaningful as the model predictions should rotate accordingly under reflections of the crystal system. Additionally, the underlying DFT algorithm is $O(3)$ equivariant. We pro-

vide an ablation study on $O(3)$ vs $SO(3)$ equivariance in Appendix D.

3.2. Hybrid Invariant-Equivariant Design

Scalability. The key difference of HIENet compared to previous works is that we apply both $E(3)$ invariant and $O(3)$ equivariant message passing layers. We find that using both invariant and equivariant layers improves performance and enables more efficient scaling of MLFF models. In Sec. 4.7 we provide a thorough analysis and ablations to support this observation.

Computational Efficiency. To further improve the scalability of our model, we remove several operations from ComFormer that we found did not significantly impact performance. Specifically, we found that edge convolution layers did not impact the performance. We also redesigned the eComFormer equivariant message passing layer to use only one tensor product operation instead of two and added skip connections to enable easier optimization of deeper models. Additional ablations of our model compared to ComFormer can be found in Appendix D.

4. Experimental Evaluations

In this section, we evaluate HIENet through a series of comprehensive experiments. We first assess its performance on the Materials Project Trajectory (MPtrj) dataset (Deng et al., 2023) and Matbench Discovery benchmark (Riebesell et al., 2023) in Sec. 4.1 and 4.2. We then analyze its computational efficiency compared to existing approaches in Sec. 4.3. To demonstrate practical utility, we evaluate HIENet on phonon band structure and bulk modulus calculations in Sec. 4.4, *ab initio* molecular dynamics simulations in Sec. 4.5, and alloy

systems in Sec. 4.6. Finally, we provide detailed ablation studies on the benefits of combining invariant and equivariant layers in Sec. 4.7. In these experiments we show that HIENet achieves superior performance while significantly improving computational efficiency compared to competing models. Detailed model settings and training details can be found in Appendix F.

4.1. Materials Project Trajectory Dataset

We train and evaluate our HIENet foundation model on the MPtrj dataset (Deng et al., 2023), which contains 1.58M crystal structures. We split the dataset and use 95% of the data for training and 5% for validation following (Batatia et al., 2023). In order to have a fair comparison, we only compare with models trained on this dataset and without any auxiliary losses. As seen in Table 1, 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 1. Mean absolute errors on train and validation splits for models trained on the MPtrj dataset. Inv. and Eqv. denote whether the model uses invariant or equivariant message passing layers, respectively. Best performing model in **bold** and second best underlined. MACE-MP-0 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-l3i5	✗	✓	<u>8.3</u>	<u>29</u>	<u>2.33</u>
HIENet	✓	✓	5.91	20.76	1.95
Validation					
CHGNet	✓	✗	33	79	3.51
MACE-MP-0	✗	✓	20	45	-
eqV2	✗	✓	<u>12.4</u>	<u>32.22</u>	<u>2.48</u>
HIENet	✓	✓	6.77	24.82	2.31

4.2. Evaluation on Matbench Discovery

We further evaluate our model on the Matbench Discovery benchmark (Riebesell et al., 2023), a comprehensive testbed to benchmark model performance on crystal stability predictions and structure optimizations. 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. To have a fair comparison with other methods, we only compare with models trained on the MPtrj dataset and without auxiliary losses, referred to as ‘compliant models’ on the Matbench Discovery leaderboard. HIENet performance can be seen in Table 2. HIENet performs best or second best on six out of

seven metrics and has the best DAF, Precision, RMSE and R^2 predictions.

Table 2. Model performance on the Unique Prototype split of the Matbench Discovery benchmark. MAE and RMSE are in meV/atom. Best model in **bold** and second best underlined.

Model	HIENet	eqV2	ORB	SevenNet-l3i5	MACE
F1 ↑	0.761	0.77	<u>0.765</u>	0.76	0.669
DAF ↑	4.75	4.64	<u>4.70</u>	4.63	3.78
Precision ↑	0.726	0.709	<u>0.719</u>	0.708	0.577
Accuracy ↑	<u>0.922</u>	0.926	<u>0.922</u>	0.92	0.878
MAE ↓	<u>44</u>	42	45	48	57
RMSE ↓	<u>86</u>	<u>87</u>	91	<u>87</u>	101
R^2 ↑	0.781	<u>0.778</u>	0.756	0.776	0.697

4.3. Efficiency Evaluation

In addition to demonstrating improved performance on MPtrj and Matbench discovery datasets, we show that HIENet is more computationally efficient than competing models. This is highly important for downstream materials discovery applications such as structural relaxation and random structure search, which require thousands of forward passes of the model. As seen in Table 3, HIENet, is 90% faster than SevenNet-l3i5 and over 140% faster than eqV2, while having better performance than both models. Both EquiformerV2 and SevenNet use equivariant message passing layers only, limiting throughput and scalability.

Table 3. Number of parameters and inference throughput of HIENet compared with baseline methods. Throughput evaluated using random samples from the MPtrj dataset on a single Nvidia A100 GPU with batch size 1.

Model	Number of Parameters	Throughput ↑ (Samples / sec.)
eqV2	31,207,434	9.4
SevenNet-l3i5	1,171,327	11.9
HIENet	7,860,155	22.6

4.4. Evaluation on Phonons and Bulk Modulus

To demonstrate a more realistic task, we evaluate the ability of HIENet to calculate phonon frequencies and phonon band structures. 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. Understanding the behavior of phonons is important for condensed matter physics, materials science, mechanical engineering, etc. 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. These atomic forces can be efficiently calculated with foundation models. Here we perform a phonon band structure calculation workflow using Phonopy (Togo et al., 2023; Togo, 2023) on a set of 78 materials, derived from a list of Materials Project (MP) structures from (Riebesell & Naik, 2024). These MP structures have reference phonon band structures documented in both the Materials Project Database (Jain et al., 2013) and the Togo PhononDB Database (Togo et al., 2023; Togo, 2023). The reason for using the PhononDB Database as well as further details regarding the Phonopy workflow is explained in Appendix E.1.

Figure 3 shows the results for four materials systems: Si, CdTe, Cs_2KInF_6 , and GaAgS_2 . It shows that HIENet-predicted phonon band structure of Si exhibits reasonable accuracy, and the phonon band structures for CdTe, Cs_2KInF_6 , and GaAgS_2 are in very good agreement with the DFT results from the PhononDB database across the entire frequency range and high-symmetry k-paths. Furthermore, the phonon band structure of Cs_2KInF_6 contains negative phonon frequencies, indicating the dynamical instability of this crystal structure despite that it is locally stable. Impressively, the result from our HIENet model agrees with the reference PhononDB data extremely well even in this negative frequency regime across all high-symmetry pathways. Table 4 depicts the MAE, MSE, and RMSE of frequency calculation among the models trained on the MPtrj dataset. For all metrics, we see that HIENet exhibits the lowest error. HIENet model can be highly valuable for predicting a material’s thermal conductivity and structure stability.

Model efficacy on zero-shot prediction of material properties was further evaluated through calculations of the fourth-order elastic tensor and the corresponding VRH average bulk modulus K_{VRH} (Hill, 1952), compared against the data in the Materials Project. A sample validation set was generated by first querying the Materials Project Database (Jain et al., 2013) for entries that had elasticity reference with number of atom sites ranging from 1 to 6. This query resulted in a total of 7,601 MP entries, of which the first 2,000 were evaluated as part of this task. Elastic tensors and bulk moduli were computed using the MatCalc’s Elasticity module (Liu et al., 2024) with our HIENet model and other models including MACE-MP-0 (Batatia et al., 2023), SevenNet-0, SevenNet-13i5 (Park et al., 2024), CHGNet (Deng et al., 2023), and EquiformerV2 (Barroso-Luque et al., 2024). Calculation details of K_{VRH} are provided in Appendix E.2.

Parity plots for each model was shown in Figure 4 using Pymatviz (Riebesell et al., 2022). We find that HIENet exhibits generally stronger capability, with the lowest overall MAE and highest overall R^2 compared to other models trained on the MPtrj dataset.

Additionally, in both the phonon band and the bulk modulus

Table 4. Error in phonon frequency prediction of various models from target values in the PhononDB Database. MAE and MSE are computed against each q-point, and RMSE is taken as the root of the MSE over all q-points and bands. Reported values are averaged across all 78 materials. Best performing model in **bold** and second best underlined.

Model	HIENet	MACE	SevenNet-13i5	CHGNet	eqV2
MAE (THz)	0.316	0.529	<u>0.325</u>	1.359	1.359
MSE (THz ²)	0.340	0.837	<u>0.358</u>	4.21	4.65
RMSE (THz)	0.447	0.699	<u>0.455</u>	1.604	1.657

calculations, we observe that EquiformerV2 underperforms other models despite having good performance on the MPtrj and Matbench Discovery benchmarks. This aligns with our intuitions as EquiformerV2 does not enforce important physical constraints.

4.5. Evaluation on *Ab Initio* Molecular Dynamics

To further examine the ability of HIENet to predict energy, force, and stress, we evaluated on *ab initio* molecular dynamics (AIMD) simulations and compared the results with those of SevenNet-13i5, MACE-MP-0, CHGNet, and EquiformerV2 models. Here we evaluate HIENet and other foundation models by comparing the energy, force, and stress predicted by foundation models with the reference data from AIMD simulations. In this section, we generate the dataset for diamond cubic silicon (Si). Details on AIMD dataset generation are provided in Appendix A.3.

For each configuration, we collect energy, forces, and stress. We then use our model and other models to predict these properties and compare with reference data. It is worth noting that CHGNet uses corrected energies as its training target (Jain et al., 2011; Wang et al., 2021). Therefore, when comparing CHGNet’s predictions with DFT results, it is necessary to apply Materials Project energy corrections to the DFT results to ensure consistency.

Figure 5 presents the evaluation results for the Si system comprising 4,000 configurations. For energy prediction, HIENet significantly outperforms other models, which exhibit noticeable shifts in predictions for either high-energy configurations or low-energy configurations. In contrast, HIENet achieves the highest accuracy across the entire energy range. For force prediction, HIENet performs slightly worse than EquiformerV2, but outperforms all other models. While MACE-MP-0 and CHGNet introduce errors in large-force scenarios, SevenNet-13i5 shows slightly lower overall accuracy compared to HIENet. For stress prediction, HIENet exhibits a more pronounced advantage over all other models. Furthermore, unlike the other models, HIENet maintains high accuracy even for high-stress configurations. In summary, HIENet has robust and powerful performance

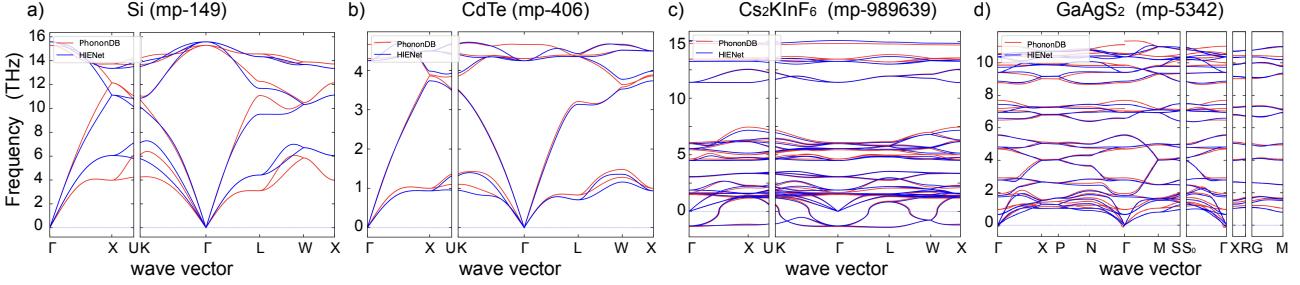


Figure 3. Phonon band structures for a) Si, b) CdTe, c) Cs_2KInF_6 , and d) GaAgS_2 calculated using HIENet compared with the reference data in the PhononDB database.

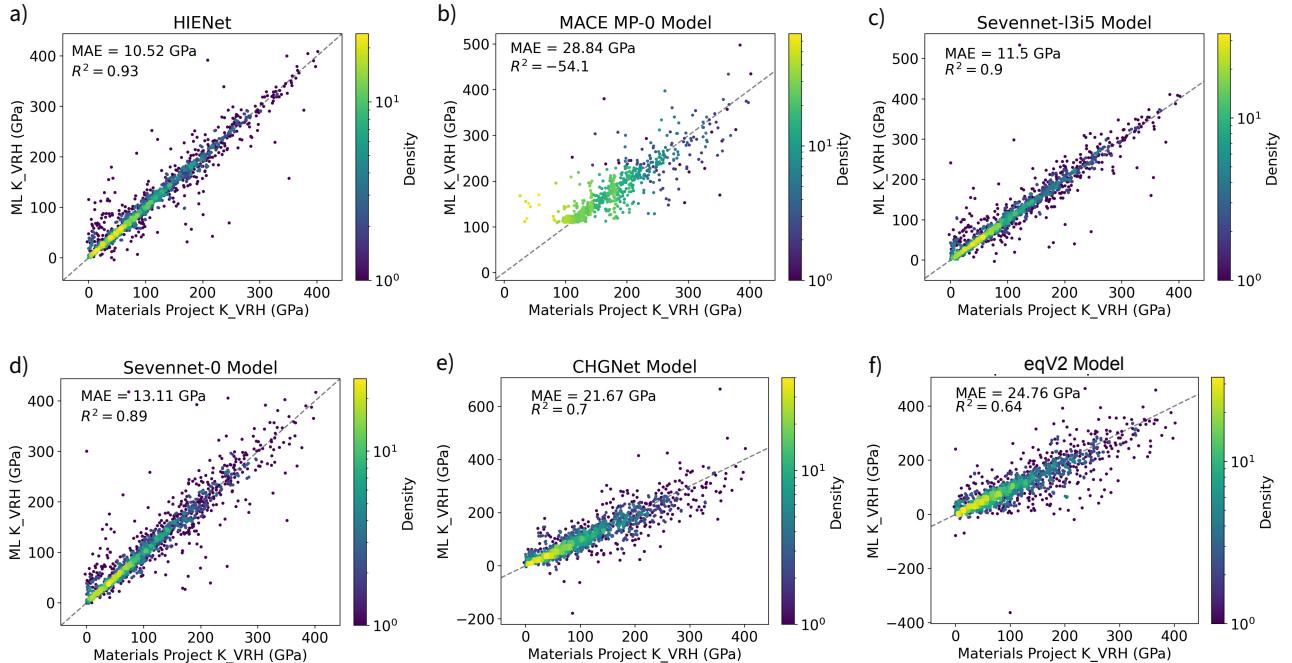


Figure 4. Comparison of bulk modulus K_{VRH} calculated by a) HIENet, b) MACE-MP-0, c) SevenNet-I3i5, d) SevenNet-0, e) CHGNet, and f) EquiformerV2 with the reference data in the Materials Project database.

on *ab initio* molecular dynamics (AIMD) simulations and outperforms existing foundation models.

4.6. Evaluations on Alloys

We also evaluate our model by deploying it within the Alloy Theoretic Automated Toolkit (ATAT) (Van De Walle et al., 2002) framework to calculate phase diagrams, following the approach outlined in (Zhu et al., 2025). 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 (Arróyave, 2022).

Starting with the simple Au-Pt binary systems, we first generate Special Quasirandom Structures (SQS) (Zunger et al., 1990) 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 various foundation models. Specifically, we used the following foundation models: CHGNet (Deng et al., 2023), MACE (Batzatia et al., 2022b;a), GRACE (Bochkarev et al., 2024), ORB (Neumann et al., 2024), and SevenNet (SevenNet-I3i5) (Batzner et al., 2022; Park et al., 2024). All of these foundation models were trained exclusively on the MPtraj dataset to ensure consistency in training data across models, allowing for a fair comparison of their predictive performance. For all *ab initio* calculations, VASP (Kresse & Hafner, 1993; 1994; Kresse & Furthmüller, 1996a;b) is used with the PBE exchange-correlation functional and PAW pseudopotentials at the level of GGA (Blöchl, 1994; Perdew et al., 1996). 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 all the foundation models. Un-

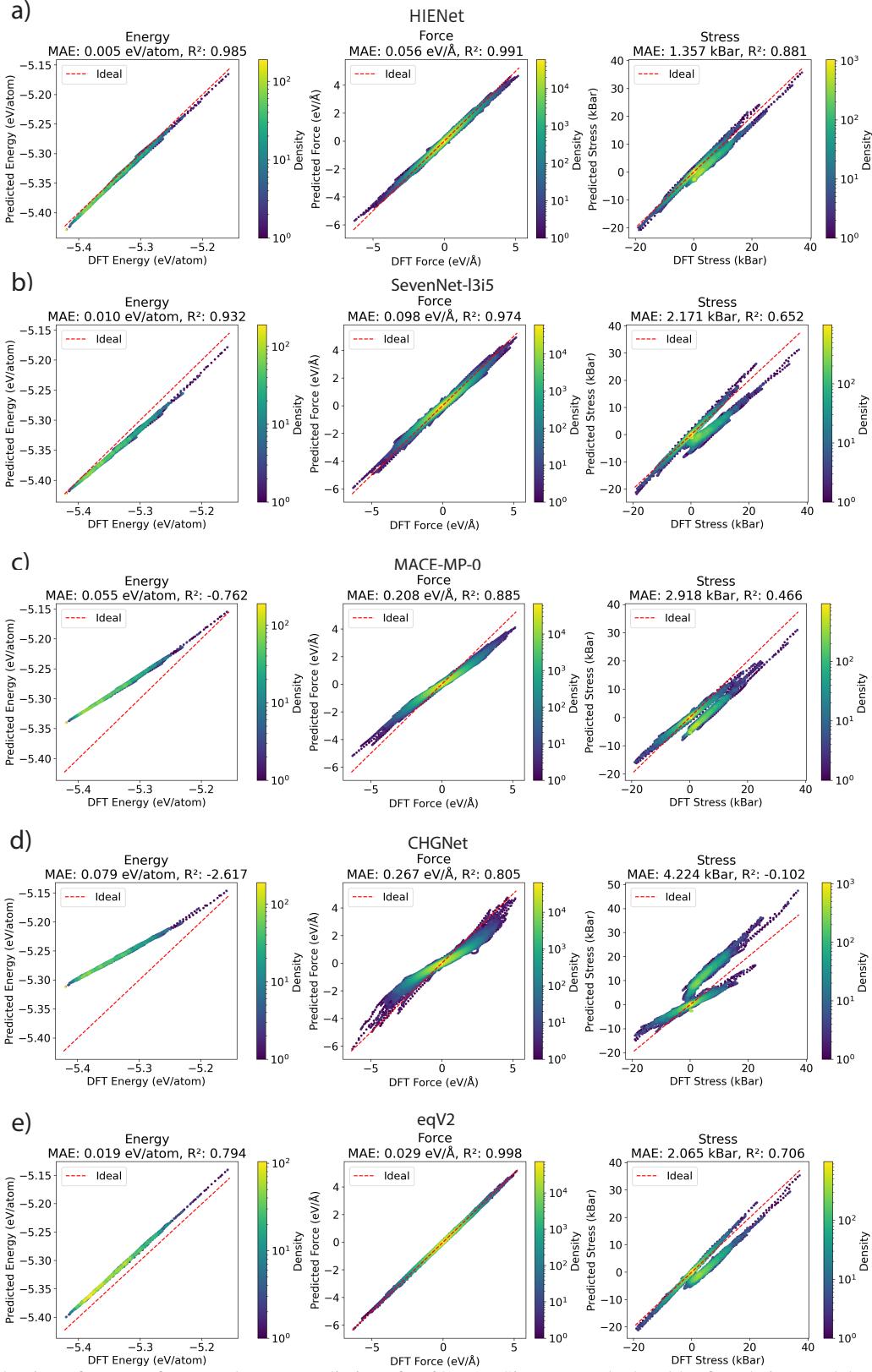


Figure 5. Evaluation of energy, force, and stress predictions for 64-atom Si system calculated by foundation models: a) HIENet, b) SevenNet-I3i5, c) MACE-MP-0, d) CHGNet, and e) eqV2 with respect to the DFT results.

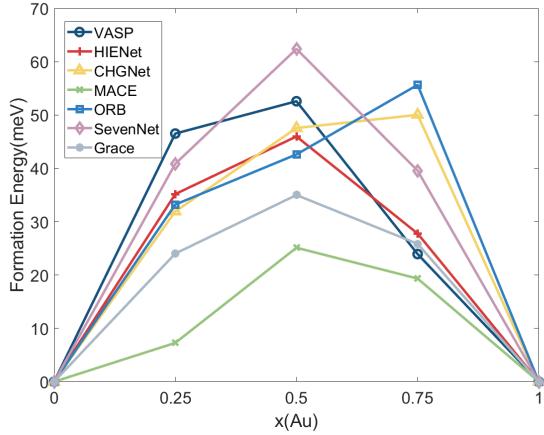


Figure 6. The formation energies per atom of the Au-Pt binary FCC system calculated using various foundation models trained on the MPtrj dataset.

Table 5. Ordering of Au-Pt formation energies and Error (%) calculated with different foundation models (1 indicates lowest formation energy and 3 indicates the highest). Best performing model—relative to ordering and (% error) given by VASP—in **bold**.

Model	$\Delta G(x_{\text{Au}} = 0.25)$	$\Delta G(x_{\text{Au}} = 0.5)$	$\Delta G(x_{\text{Au}} = 0.75)$
CHGNet	1 (-31.4%)	2 (-9.6%)	3 (109.1%)
MACE	1 (-84.4%)	3 (-52.2%)	2 (-19.2%)
SevenNet	2 (-12.2%)	3 (18.6%)	1 (65.3%)
ORB	1 (-28.6%)	2 (-19.0%)	3 (132.5%)
GRACE	1 (-48.4%)	3 (-33.4%)	2 (8.0%)
HIENet	2 (-24.4%)	3 (-12.6%)	1 (16.1%)
VASP	2	3	1

like the results previously reported, we find that models like ORB exhibit much larger errors in formation energy when relying solely on the MPtrj dataset. In contrast, our model shows strong agreement with the first-principles results, even trained with only the MPtrj dataset. In addition, although all the models successfully give a positive formation energy for the SQS's, when predicting the miscibility gap in the phase diagram, most of the models including CHGNet, MACE, ORB, and GRACE 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 5. Such ordering of formation energies is highly important in thermodynamics and materials science, as it governs the stability of the phases, as a necessary (albeit not sufficient) condition for a topologically correct phase diagram is for a model to produce the correct ordering in the energetics of the structures competing for equilibrium (Ober & Van der Ven, 2024).

To better evaluate the difference of formation energies, we

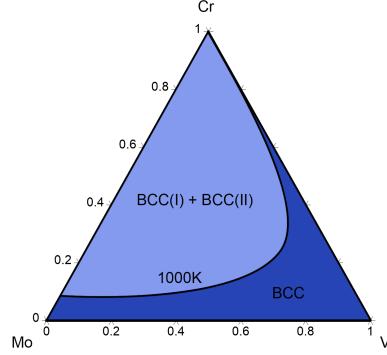


Figure 7. Cr-Mo-V ternary phase diagram at 1,000 K calculated with ATAT and HIENet. Only the BCC phase is included in the calculation. The phase diagram is plotted with the Pandat (Chen et al., 2002) software package.

define the error as:

$$\text{Error} = \frac{\sum_x |\Delta G_{\text{MLIP}}(x) - \Delta G_{\text{FP}}(x)|}{\sum_x |\Delta G_{\text{FP}}(x)|}, \quad (3)$$

where $\Delta G_{\text{MLIP}}(x)$ and $\Delta G_{\text{FP}}(x)$ represent the formation free energy at composition x at 0 K, calculated using foundation models and DFT, respectively. In Table 6, we show the errors of formation energies of Au-Pt, Ag-Pt, Cr-Mo, and Nb-V binary systems and observe that HIENet achieves the best overall performance.

Additionally, we can also use our model 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.

4.7. Combining Invariant and Equivariant Layers

To demonstrate the performance and speed advantage of using both invariant and equivariant layers, we train 3 models, each with 4 layers. HIENet uses one invariant layer and 3 equivariant layers, I-Net uses 4 invariant layers, and E-Net uses 4 equivariant layers. The results in Table 7 align

Table 6. Comparison of errors in formation energies for different binary systems (Au-Pt, Ag-Pt, Cr-Mo, and Nb-V) using various foundation models trained on the MPtrj dataset. The error is computed as a percentage deviation from DFT-calculated formation energy. Best performing model(s) in **bold**.

Model	Au-Pt (FCC)	Ag-Pt (FCC)	Cr-Mo (BCC)	Nb-V (BCC)
CHGNet	37.169%	98.039%	107.154%	52.308%
MACE	57.946%	59.359%	106.197%	30.973%
SevenNet	25.281%	73.840%	83.470%	53.351%
ORB	44.695%	40.664%	142.348%	33.071%
GRACE	34.134%	68.213%	26.548%	29.756%
HIENet	17.717%	60.434%	15.419%	11.146%

Table 7. Mean absolute errors on MPtrj validation set for HIENet compared to invariant-only and equivariant-only baseline models. Models trained for 25 epochs on the MPtrj dataset. Best performing model in **bold** and second best model underlined.

Model	Energy ↓ (meV/atom)	Force ↓ (meV/Å)	Stress ↓ (kBar)	Throughput ↑ (Samples / sec.)
I-Net	28.00	86.05	4.74	98.0
E-Net	<u>14.22</u>	<u>44.12</u>	<u>3.16</u>	24.3
HIENet	12.59	42.12	3.09	30.2

with our intuitions about combining both types of layers. Specifically, I-Net is fast, but has poor prediction accuracy, E-Net has better accuracy than I-Net, but is much slower, and HIENet outperforms both models while still being faster than E-Net.

Additionally, we find that the order in which message passing layers are applied has a significant impact on performance. In Table 8 we investigate different orders of layers and find that the only combination that performs well is applying invariant layer(s) followed by equivariant layers. Equivariant message passing layers are limited in terms of the non-linearities they can apply, so we hypothesize that applying invariant layers before equivariant layers builds more informative node representations that enable the equivariant layers to be more powerful than in equivariant-only models.

Table 8. Mean absolute errors on MPtrj validation set for HIENet compared to models trained with different orders of message passing layers. 'Inv. first' represents the baseline HIENet architecture of applying one invariant layer followed by several equivariant layers, 'Equiv. First' applies several equivariant layers followed by one invariant layer, and 'Mixed' applies alternating invariant and equivariant layers. Models trained for 20 epochs on the MPtrj dataset. Best performing model in **bold**.

MP Layer Ordering	Energy ↓ (meV/atom)	Force ↓ (meV/Å)	Stress ↓ (kBar)
Mixed	50.35	94.62	6.41
Equiv. First	32.26	77.22	4.94
Inv. First	16.26	49.29	3.48

5. Conclusion

We propose HIENet, a foundation model for materials simulations and discovery, and demonstrate the importance of incorporating both invariant and equivariant message passing layers to build powerful and efficient MLFF models. We show that HIENet achieves state-of-the-art performance on a variety of benchmarks and downstream applications, while being significantly faster than competing models. Finally, we provide several ablation studies to further demonstrate

the significance of our contributions and provide insights on designing powerful hybrid invariant-equivariant networks.

Acknowledgments

K.Y., M.B., A.K. and S.J. acknowledge partial support from National Science Foundation (NSF) under grant IIS-2243850, ARPA-H under grant 1AY1AX000053, and National Institutes of Health under grant U01AG070112. K.Z., S.K. and X.F.Q. acknowledge partial support from NSF under awards CMMI-2226908 and DMR-2103842 and the donors of ACS Petroleum Research Fund under Grant 65502-ND10. R.A. and D.S. acknowledge support from ARO through Grant No. W911NF-22-2-0117. R.A., X.N.Q. and S.Z. acknowledge NSF through Grant No. 2119103 (DMREF). Z. X. and X.N.Q. acknowledge support from NSF through grants SHF-2215573 and IIS-2212419. Some computations were carried out at the Texas A&M High-Performance Research Computing (HPRC) facility. We gratefully acknowledge the support of Lambda Inc. for providing computing resources.

References

- Alder, B. J. and Wainwright, T. E. Studies in molecular dynamics. i. general method. *The Journal of Chemical Physics*, 31(2):459–466, 08 1959. ISSN 0021-9606. doi: 10.1063/1.1730376.
- Arróyave, R. Phase stability through machine learning. *Journal of Phase Equilibria and Diffusion*, 43(6):606–628, 2022.
- Atsushi Togo, K. S. and Tanaka, I. Spglib: a software library for crystal symmetry search. *Science and Technology of Advanced Materials: Methods*, 4(1):2384822, 2024. doi: 10.1080/27660400.2024.2384822. URL <https://doi.org/10.1080/27660400.2024.2384822>.
- Barroso-Luque, L., Shuaibi, M., Fu, X., Wood, B. M., Dzamba, M., Gao, M., Rizvi, A., Zitnick, C. L., and Ulissi, Z. W. Open Materials 2024 (OMat24) Inorganic Materials Dataset and Models. *arXiv preprint arXiv:2410.12771*, 2024.
- Batatia, I., Batzner, S., Kovács, D. P., Musaelian, A., Simm, G. N. C., Drautz, R., Ortner, C., Kozinsky, B., and Csányi, G. The Design Space of E(3)-Equivariant Atom-Centered Interatomic Potentials, 2022a.
- Batatia, I., Kovacs, D. P., Simm, G., Ortner, C., and Csányi, G. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in Neural Information Processing Systems*, 35: 11423–11436, 2022b.

- Batatia, I., Benner, P., Chiang, Y., Elena, A. M., Kovács, D. P., Riebesell, J., Advincula, X. R., Asta, M., Avaylon, M., Baldwin, W. J., et al. A foundation model for atomistic materials chemistry. *arXiv preprint arXiv:2401.00096*, 2023.
- Batzner, S., Musaelian, A., Sun, L., Geiger, M., Mailoa, J. P., Kornbluth, M., Molinari, N., Smidt, T. E., and Kozinsky, B. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13(1):2453, 2022.
- Blöchl, P. E. Projector augmented-wave method. *Physical Review B*, 50(24):17953, 1994.
- Bochkarev, A., Lysogorskiy, Y., and Drautz, R. Graph atomic cluster expansion for semilocal interactions beyond equivariant message passing. *Physical Review X*, 14:021036, Jun 2024.
- Chen, C. and Ong, S. P. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, 2022.
- Chen, S.-L., Daniel, S., Zhang, F., Chang, Y., Yan, X.-Y., Xie, F.-Y., Schmid-Fetzer, R., and Oates, W. The pandat software package and its applications. *Calphad*, 26(2):175–188, 2002.
- Choudhary, K. and DeCost, B. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7(1):185, 2021.
- Choudhary, K., Wines, D., Li, K., Garrity, K. F., Gupta, V., Romero, A. H., Krogel, J. T., Saritas, K., Fuhr, A., Ganesh, P., et al. JARVIS-Leaderboard: a large scale benchmark of materials design methods. *npj Computational Materials*, 10(1):93, 2024.
- De Jong, M., Chen, W., Angsten, T., Jain, A., Notestine, R., Gamst, A., Sluiter, M., Krishna Ande, C., Van Der Zwaag, S., Plata, J. J., et al. Charting the complete elastic properties of inorganic crystalline compounds. *Scientific data*, 2(1):1–13, 2015.
- de Pablo, J. J., Jackson, N. E., Webb, M. A., Chen, L.-Q., Moore, J. E., Morgan, D., Jacobs, R., Pollock, T., Schlom, D. G., Toberer, E. S., et al. New frontiers for the materials genome initiative. *npj Computational Materials*, 5(1):41, 2019.
- Deng, B., Zhong, P., Jun, K., Riebesell, J., Han, K., Bartel, C. J., and Ceder, G. CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 5(9):1031–1041, 2023.
- Elfwing, S., Uchibe, E., and Doya, K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural networks*, 107:3–11, 2018.
- Fletcher, R. *Practical methods of optimization*. John Wiley & Sons, 2000.
- Gasteiger, J., Groß, J., and Günnemann, S. Directional Message Passing for Molecular Graphs. In *International Conference on Learning Representations*, 2020.
- Gasteiger, J., Becker, F., and Günnemann, S. GemNet: Universal Directional Graph Neural Networks for Molecules. In *Advances in Neural Information Processing Systems*, 2021.
- Hestenes, M. R., Stiefel, E., et al. *Methods of conjugate gradients for solving linear systems*, volume 49. NBS Washington, DC, 1952.
- Hill, R. The elastic behaviour of a crystalline aggregate. *Proceedings of the Physical Society. Section A*, 65(5):349, 1952.
- Hinuma, Y., Pizzi, G., Kumagai, Y., Oba, F., and Tanaka, I. Band structure diagram paths based on crystallography. *Computational Materials Science*, 128:140–184, 2017. ISSN 0927-0256. doi: <https://doi.org/10.1016/j.commatsci.2016.10.015>. URL <https://www.sciencedirect.com/science/article/pii/S0927025616305110>.
- Hohenberg, P. and Kohn, W. Density functional theory (dft). *Phys. Rev.*, 136(1964):B864, 1964.
- Jain, A., Hautier, G., Ong, S. P., Moore, C. J., Fischer, C. C., Persson, K. A., and Ceder, G. Formation enthalpies by mixing gga and gga+ u calculations. *Physical Review B—Condensed Matter and Materials Physics*, 84(4):045115, 2011.
- Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1), 2013.
- Kohn, W. and Sham, L. J. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.
- Kresse, G. and Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996a.

- Kresse, G. and Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169, 1996b.
- Kresse, G. and Hafner, J. Ab initio molecular dynamics for liquid metals. *Physical Review B*, 47(1):558, 1993.
- Kresse, G. and Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Physical Review B*, 49(20):14251, 1994.
- Langreth, D. C. and Mehl, M. J. Beyond the local-density approximation in calculations of ground-state electronic properties. *Phys. Rev. B*, 28:1809–1834, Aug 1983. doi: 10.1103/PhysRevB.28.1809. URL <https://link.aps.org/doi/10.1103/PhysRevB.28.1809>.
- Liao, Y.-L., Wood, B. M., Das, A., and Smidt, T. EquiformerV2: Improved Equivariant Transformer for Scaling to Higher-Degree Representations. In *The Twelfth International Conference on Learning Representations*, 2024.
- Lin, Y., Yan, K., Luo, Y., Liu, Y., Qian, X., and Ji, S. Efficient approximations of complete interatomic potentials for crystal property prediction. In *Proceedings of the 40th International Conference on Machine Learning*, pp. 21260–21287, 2023.
- Liu, R., Liu, E., Riebesell, J., Qi, J., Ong, S. P., and Ko, T. W. MatCalc, January 2024. URL <https://github.com/materialsvirtuallab/matcalc>.
- Loshchilov, I. and Hutter, F. Decoupled Weight Decay Regularization. In *International Conference on Learning Representations*, 2019.
- Loshchilov, I. and Hutter, F. SGDR: Stochastic Gradient Descent with Warm Restarts. In *International Conference on Learning Representations*, 2022.
- Lv, C., Zhou, X., Zhong, L., Yan, C., Srinivasan, M., Seh, Z. W., Liu, C., Pan, H., Li, S., Wen, Y., et al. Machine learning: an advanced platform for materials development and state prediction in lithium-ion batteries. *Advanced Materials*, 34(25):2101474, 2022.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
- Miracle, D. and Thoma, D. Autonomous research and development of structural materials—an introduction and vision. *Current Opinion in Solid State and Materials Science*, 33:101188, 2024.
- Monkhorst, H. J. and Pack, J. D. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13: 5188–5192, Jun 1976. doi: 10.1103/PhysRevB.13.5188. URL <https://link.aps.org/doi/10.1103/PhysRevB.13.5188>.
- Neumann, M., Gin, J., Rhodes, B., Bennett, S., Li, Z., Choubisa, H., Hussey, A., and Godwin, J. Orb: A Fast, Scalable Neural Network Potential, 2024.
- Ober, D. E. and Van der Ven, A. Thermodynamically informed priors for uncertainty propagation in first-principles statistical mechanics. *Physical Review Materials*, 8:103803, 2024.
- Park, Y., Kim, J., Hwang, S., and Han, S. Scalable Parallel Algorithm for Graph Neural Network Interatomic Potentials in Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation*, 20:4857–4868, 2024.
- Perdew, J. P., Burke, K., and Ernzerhof, M. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865, 1996.
- Qu, E. and Krishnapriyan, A. S. 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.
- Riebesell, J. and Naik, A. A. ffonons, 2024. URL <https://github.com/janosh/ffonons>.
- Riebesell, J., Yang, H. D., Goodall, R., and Baird, S. G. Py-matviz, October 2022. URL <https://github.com/janosh/pymatviz>.
- Riebesell, J., Goodall, R. E., Jain, A., Benner, P., Persson, K. A., and Lee, A. A. Matbench Discovery—An evaluation framework for machine learning crystal stability prediction. *arXiv preprint arXiv:2308.14920*, 2023.
- Shafian, S., Mohd Salehin, F. N., Lee, S., Ismail, A., Mohamed Shuhidan, S., Xie, L., and Kim, K. 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.
- Stach, E., DeCost, B., Kusne, A. G., Hattrick-Simpers, J., Brown, K. A., Reyes, K. G., Schrier, J., Billinge, S., Buonassisi, T., Foster, I., et al. Autonomous experimentation systems for materials development: A community perspective. *Matter*, 4(9):2702–2726, 2021.
- Togo, A. First-principles phonon calculations with phonopy and phono3py. *J. Phys. Soc. Jpn.*, 92(1):012001, 2023. doi: 10.7566/JPSJ.92.012001.

- Togo, A., Chaput, L., Tadano, T., and Tanaka, I. Implementation strategies in phonopy and phono3py. *J. Phys. Condens. Matter*, 35(35):353001, 2023. doi: 10.1088/1361-648X/acd831.
- Van De Walle, A., Asta, M., and Ceder, G. The alloy theoretic automated toolkit: A user guide. *Calphad*, 26(4):539–553, 2002.
- Wang, A., Kingsbury, R., McDermott, M., Horton, M., Jain, A., Ong, S. P., Dwaraknath, S., and Persson, K. A. A framework for quantifying uncertainty in dft energy corrections. *Scientific reports*, 11(1):15496, 2021.
- Xie, T. and Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical review letters*, 120(14):145301, 2018.
- Yan, K., Liu, Y., Lin, Y., and Ji, S. Periodic graph transformers for crystal material property prediction. In *The 36th Annual Conference on Neural Information Processing Systems*, pp. 15066–15080, 2022.
- Yan, K., Fu, C., Qian, X., Qian, X., and Ji, S. Complete and efficient graph transformers for crystal material property prediction. In *International Conference on Learning Representations*, 2024.
- Zhang, X., Wang, L., Helwig, J., Luo, Y., Fu, C., Xie, Y., Liu, M., Lin, Y., Xu, Z., Yan, K., Adams, K., Weiler, M., Li, X., Fu, T., Wang, Y., Yu, H., Xie, Y., Fu, X., Strasser, A., Xu, S., Liu, Y., Du, Y., Saxton, A., Ling, H., Lawrence, H., Stärk, H., Gui, S., Edwards, C., Gao, N., Ladera, A., Wu, T., Hofgard, E. F., Tehrani, A. M., Wang, R., Daigavane, A., Bohde, M., Kurtin, J., Huang, Q., Phung, T., Xu, M., Joshi, C. K., Mathis, S. V., Azizzadenesheli, K., Fang, A., Aspuru-Guzik, A., Bekkers, E., Bronstein, M., Zitnik, M., Anandkumar, A., Ermon, S., Liò, P., Yu, R., Günnemann, S., Leskovec, J., Ji, H., Sun, J., Barzilay, R., Jaakkola, T., Coley, C. W., Qian, X., Qian, X., Smidt, T., and Ji, S. Artificial intelligence for science in quantum, atomistic, and continuum systems. *arXiv preprint arXiv:2307.08423*, 2023.
- Zheng, Y., Tang, N., Omar, R., Hu, Z., Duong, T., Wang, J., Wu, W., and Haick, H. Smart materials enabled with artificial intelligence for healthcare wearables. *Advanced Functional Materials*, 31(51):2105482, 2021.
- Zhu, S., Saritürk, D., and Arróyave, R. Accelerating calphad-based phase diagram predictions in complex alloys using universal machine learning potentials: Opportunities and challenges. *Acta Materialia*, pp. 120747, 2025.
- Zunger, A., Wei, S.-H., Ferreira, L., and Bernard, J. E. Special quasirandom structures. *Physical Review Letters*, 65(3):353, 1990.

A. Molecular Dynamics Simulation

A.1. Molecular dynamics simulation and structural optimization of materials

Molecular dynamics simulation. Molecular dynamics (MD) simulation (Alder & Wainwright, 1959) 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 σ . For a given starting structure configuration, E , \mathbf{F}_i , and σ can be calculated by the force field such as classical force field or foundation models. 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 the 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) (Hestenes et al., 1952) and Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) (Fletcher, 2000). 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) (Hohenberg & Kohn, 1964; Kohn & Sham, 1965), MD simulations (Alder & Wainwright, 1959), and Monte Carlo (MC) simulations (Metropolis et al., 1953). 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 foundation models 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.

A.3. *Ab initio* molecular dynamics simulation

Ab initio molecular dynamics (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, with additional AIMD-related settings as detailed below.

The dataset was generated for silicon (Si) system containing 64 atoms in a $2 \times 2 \times 2$ supercell. A Γ -centered Monkhorst–Pack k-point sampling grid of $2 \times 2 \times 2$ (Monkhorst & Pack, 1976) 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.

B. HIENet Architecture

B.1. Crystal Graph Construction

We construct a crystal graph that is O(3) equivariant and invariant to all unit cell transformations. Specifically, each node \mathbf{z}_i in the constructed graph represents atom i and all of its infinite duplicates in repeating unit cells. We then build edges using a radius-based graph construction such that there exists an edge \mathbf{r}_{ji} between nodes i and j if there is a periodic duplicate j' such that:

$$\|\mathbf{p}_j + k'_1 \mathbf{l}_1 + k'_2 \mathbf{l}_2 + k'_3 \mathbf{l}_3 - \mathbf{p}_i\|_2 \leq R_{\text{cut}} \quad (4)$$

where R_{cut} is a fixed cutoff radius. For our model we set $R_{\text{cut}} = 5\text{\AA}$.

B.2. Embedding Layer

We create initial node embeddings \mathbf{h}_i with a linear projection of atomic number one-hot encodings: $\mathbf{h}_i = \mathbf{W}_{\text{emb}} \mathbf{z}_i$. We create edge embeddings \mathbf{h}_{ij} for edge e_{ij} using 8 radial Bessel basis functions:

$$\mathbf{h}_{ij} = \frac{2 \sin\left(\frac{n\pi}{R_{\text{cut}}} \|\mathbf{r}_{ij}\|_2\right)}{R_{\text{cut}} \|\mathbf{r}_{ij}\|_2} f_{\text{poly}}(\|\mathbf{r}_{ij}\|_2, R_{\text{cut}}) \quad (5)$$

where f_{poly} is the polynomial envelope of Gasteiger et al. (2020).

B.3. E(3) Invariant Message Passing

For our invariant message passing, we use a graph transformer layer to update node features \mathbf{h}_i . Specifically, we form key \mathbf{k}_{ji} , and query \mathbf{q}_{ji} vectors as

$$\mathbf{k}_{ji} = \mathbf{W}_K(\mathbf{h}_i || \mathbf{h}_j || \mathbf{h}_{ij}), \quad \mathbf{q}_{ji} = \mathbf{W}_Q(\mathbf{h}_i || \mathbf{h}_j || \mathbf{h}_{ij}) \quad (6)$$

where $||$ denotes vector concatenation. We then build value vectors \mathbf{v}_{ji} and compute attention weights α_{ji} :

$$\mathbf{v}_{ji} = \Phi(\mathbf{h}_i || \mathbf{h}_j || \mathbf{h}_{ij}), \quad \alpha_{ji} = \sigma\left(\frac{\mathbf{q}_{ji} \odot \mathbf{k}_{ji}}{\sqrt{d}}\right) \quad (7)$$

where Φ is an MLP and \odot represents the Hadamard (elementwise) product. Finally, we compute updated node features:

$$\mathbf{h}'_i = \varphi(\mathbf{h}_i) + (1 - \varphi(\mathbf{h}_i)) \sum_{j \in \mathcal{N}_i} \alpha_{ji} \odot \mathbf{v}_{ji} \quad (8)$$

Here $\varphi(\mathbf{h}_i)$ is an MLP with a sigmoid activation that acts as a learnable gating mechanism.

B.4. O(3) Equivariant Message Passing

We first embed edge vectors \mathbf{r}_{ji} using spherical harmonics $Y^l(\frac{\mathbf{r}_{ji}}{\|\mathbf{r}_{ji}\|})$ for each rotation order l up to L_{max} . We then build equivariant features as:

$$\mathbf{f}_i = \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} \mathbf{TP}_0 \left(\mathbf{W} \mathbf{h}_j, Y^0 \left(\frac{\mathbf{r}_{ji}}{\|\mathbf{r}_{ji}\|} \right) \right) + \sum_{l=1}^{L_{\text{max}}} \sum_{j \in \mathcal{N}_i} \mathbf{TP}_0 \left(\mathbf{W} \mathbf{h}_j, Y^l \left(\frac{\mathbf{r}_{ji}}{\|\mathbf{r}_{ji}\|} \right) \right) \quad (9)$$

where \mathbf{TP}_0 is a tensor product operation yielding outputs with rotation order $l = 0$. We further add a skip connection and gate mechanism to output updated node features:

$$\mathbf{h}'_i = \psi(\mathbf{W}_{\text{skip}} \mathbf{h}_i + \mathbf{W}_E \mathbf{f}_i) \quad (10)$$

where ψ is an equivariant gate activation function. In practice, we sequentially apply several of our equivariant message passing layers.

B.5. Model Optimization

We optimize the model using Huber loss functions for energy, force, and stress:

$$\mathcal{L} = L_{\text{huber}} \left(\frac{E}{N}, \frac{\hat{E}}{N} \right) + \lambda_F \sum_{i=1}^N \sum_{k=1}^3 L_{\text{huber}} \left(F_{i,k}, \hat{F}_{i,k} \right) + \lambda_S \sum_{l=1}^6 L_{\text{huber}} \left(\sigma_l, \hat{\sigma}_l \right) \quad (11)$$

where λ_F, λ_S control the relative weight of force and stress losses. In practice, we set $\lambda_F = 1$ and $\lambda_S = 0.1$. We compute per-atom energy and decompose the 3×3 stress matrix into virial stress components. We train the model using the AdamW optimizer (Loshchilov & Hutter, 2019) and Cosine Annealing learning rate schedule (Loshchilov & Hutter, 2022).

C. Proofs of Propositions

C.1. Proof of Prop. 3.1

Proposition C.1. *HINet predictions $\hat{\mathbf{F}}_i$ form a conservative vector field.*

By definition, a vector field $\mathbf{v} : \mathbb{R} \rightarrow \mathbb{R}^n$ is conservative if there exists a continuously differentiable scalar field φ such that:

$$\mathbf{v} = \nabla \varphi$$

Importantly, we use the polynomial envelope function of Gasteiger et al. (2021) and continuously differentiable activation functions throughout the model to ensure that the potential energy predicted by our model \hat{E} is continuously differentiable with respect to atom positions.

And because we compute forces using:

$$\hat{\mathbf{F}} = -\nabla_{\mathbf{P}} \hat{E}$$

the resulting vector field $\hat{\mathbf{F}}$ is conservative. \square

C.2. Proof of Prop. 3.2

Proposition C.2. *HINet predictions satisfy force equilibrium $\sum_{i=1}^N \hat{\mathbf{F}}_i = \mathbf{0}$ when no external influences applied.*

We define edge force $\hat{\mathbf{F}}_{ji}$ as:

$$\hat{\mathbf{F}}_{ji} = -\frac{\partial \hat{E}}{\partial \mathbf{r}_{ji}}$$

We can then decompose the forces acting on each atom as:

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

Summing over all atoms we get:

$$\begin{aligned}
\sum_{i=1}^N \hat{\mathbf{F}}_i &= \sum_{i=1}^N \sum_{j \in \mathcal{N}_i} (\hat{\mathbf{F}}_{ji} - \hat{\mathbf{F}}_{ij}) \\
&= \sum_{(i,j) \in \mathcal{E}} (\hat{\mathbf{F}}_{ij} - \hat{\mathbf{F}}_{ji}) \\
&= 0
\end{aligned}$$

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

D. Additional Ablations

Table 9. Mean absolute errors on MPtrj validation set for HIENet compared without architechtrual changes from ComFormer. HIENet + RBF denotes replacing the HIENet edge embeddings with the RBF embeddings from ComFormer. HIENet + EdgeConv denotes unclude the Edge-wise transformer layer from iComFormer. Models trained for 20 epochs on the MPtrj dataset. Best performing model in **bold**.

Model	Energy ↓ (meV/atom)	Force ↓ (meV/Å)	Stress ↓ (kBar)
HIENet + RBF	18.87	55.58	3.97
HIENet + EdgeConv	17.97	54.27	3.69
HIENet	16.26	49.29	3.48

To justify some of the design choices of HIENet compared to ComFormer, we provide several additional ablation studies in Tables 9 and 10. Using RBF kernels to embed edge vectors negatively impacts performance, and the inclusion of an edge convolution layer both negatively impacts the performance and reduces the model efficiency. As such, both were removed in the final HIENet design.

Table 10. 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 ↓ (meV/atom)	Force ↓ (meV/Å)	Stress ↓ (kBar)
SO(3)	19.13	56.12	3.98
O(3)	16.26	49.29	3.48

To empirically justify why we use O(3) equivariant crystal graph representations instead of the geometrically complete but SO(3) equivariant crystal graphs from ComFormer, we provide an ablation study where we include the additional periodic encodings of Yan et al. (2024). As expected, we observe that SO(3) equivariant HIENet does not perform as well.

E. Phonon and Bulk Modulus Workflows

E.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) (Perdew et al., 1996) 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 (Togo et al., 2023; Togo, 2023) where foundation models 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 (Hinuma et al., 2017; Atsushi Togo & Tanaka, 2024). Using this workflow, the high-symmetry k-paths and the sampling grids were identical between the reference phonon band structure from PhononDB and the predicted band structure from our model.

E.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 different models. 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) (Langreth & Mehl, 1983) as implemented the Vienna Ab-initio Simulation Package (VASP) (Kresse & Furthmüller, 1996b). 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 (De Jong et al., 2015). We then calculate elastic tensors for all materials in the validation set using our HIENet model as well as MACE-MP-0, CHGNet SevenNet-l3i5, SevenNet-0, and EquiformerV2 (Batatia et al., 2023; Deng et al., 2023; Park et al., 2024; Barroso-Luque et al., 2024). Of the 2,000 structures computed, 1,763 were ultimately plotted after filtering the structures where K_{VRH} was not reported or where reported K_{VRH} values were extreme. Following Batatia et al. (2023), K_{VRH} values between -50 GPa to 600 GPa were considered.

F. Model Settings and Experimental Details

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 ($l = 0$), 128 vector channels ($l = 1$), 64 tensor channels ($l = 2$), and 32 higher order-tensor channels ($l = 3$). We use 8 radial Bessel basis functions for distance encoding and a polynomial envelope (Gasteiger et al., 2020) with $p = 6$. We use SiLU activation functions (Elfwing et al., 2018) throughout the network to ensure smooth and continuously differentiable gradients. Additionally, we scale the input energies by the root mean square (RMS) of forces from the training dataset and shifted by element-wise reference energies from the same dataset.

Following Batatia et al. (2023), we split the Materials Project Trajectory (MPtrj) Dataset (Deng et al., 2023) into training (95%) and validation (5%) sets. The model is trained for 250 epochs using AdamW optimizer (Loshchilov & Hutter, 2019) with a Cosine Annealing learning rate scheduler (Loshchilov & Hutter, 2022). We use Huber loss functions for each of the targets with $\delta = 0.01$. To prevent overfitting, we use model checkpoint from epoch 221 for materials discovery applications. For the Matbench Discovery benchmark, we additionally fine-tune the model on energy for 50 epochs. Training is performed

on a platform with ten CPUs, Intel Xeon 6248R (Cascade Lake), 3.0GHz, 24-core, 384GB DDR4 memory, and twenty GPUs, NVIDIA A100 40GB GPU accelerator. The total batch size is 440 (22 per GPU) and requires 55 GPU minutes per training epoch and 30 GPU seconds per validation epoch. The training hyperparameters for both the training and fine-tuning stages are summarized in Table 11.

Table 11. Hyperparameters for training and fine-tuning stages.

Hyper-parameters	Training	Fine-tuning
Optimizer	AdamW	AdamW
Learning rate scheduling	Cosine	Cosine
Maximum learning rate	0.01	0.0005
Minimum learning rate	0.0001	0.0
Warmup epochs	0.1	0.1
Warmup factor	0.2	0.1
Number of epochs	250	50
Batch size	22	64
Weight decay	0.001	0.001
Energy loss weight, λ_E	1.0	1.0
Force loss weight, λ_F	1.0	0.1
Stress loss weight, λ_σ	0.01	0.001
Model EMA Decay	0.999	0.999