

Appendix

A Initialisation of Features

The scalar feature $x_i^{(0)}$ is initialized based on the atomic number, $E(z_i) \in \mathbb{R}^F$, where z_i is the atomic number and E is an embedding layer. The vector feature is initially set to zero, $\vec{x}_i^{(0)} = \vec{0} \in \mathbb{R}^{F \times 3}$. The atomic coordinates, $\vec{r}_i^{(0)}$, are initialized to the coordinates of the initial unrelaxed structure. The lattice node features $y_c^{(0)}$ and $\vec{y}_c^{(0)}$ are initialized analogously to atom features: $y_c^{(0)} = E_{\text{lat}}(c)$, and $\vec{y}_c^{(0)} = \vec{0} \in \mathbb{R}^{F \times 3}$, where E_{lat} is a learnable embedding for each lattice direction. The lattice coordinates, $\vec{l}_c^{(0)}$, are initialized to the lattice vectors of the initial unrelaxed structure.

B Datasets

We evaluate E³Relax on four benchmark datasets: MP (Jain et al. 2013), X–Mn–O (Kim et al. 2020, 2023), C2DB (Gjerding et al. 2021; Lyngby and Thygesen 2022), and layered vdW crystals (Yang et al. 2024). A summary of these datasets is provided in Table 1.

The MP database spans 89 elements and contains 187,687 structure snapshots from 62,783 compounds, each sampled at initial, intermediate, and fully relaxed stages of DFT optimization. After removing entries missing either the initial or final structure, 62,724 usable pairs remain, which we split 8:1:1 into training, validation, and test sets. The hypothetical substitution dataset (X–Mn–O), derived from MP, focuses on ternary oxides for photoanode research. It covers four ‘X’ elements (Mg, Ca, Ba, Sr) substituted into prototype Mn–O frameworks, yielding 28,579 unrelaxed–relaxed pairs. These are partitioned 8:1:1 for training, validation, and testing. The C2DB dataset comprises 11,581 unrelaxed–relaxed 2D structure pairs spanning 62 elements, which we split into training/validation/test sets in a 6:2:2 ratio. The layered–vdW crystal dataset contains 57 2D van der Waals–bonded structures across 29 elements, divided into training, validation, and test subsets in an 8:1:1 ratio.

Figure 1(a)–(c) compares the distributions of atomic coordinate MAE, cell shape deviation, and cell volume MAE between the unrelaxed and DFT-relaxed structures for the four datasets. Three distinct patterns emerge. The MP dataset exhibits narrow distributions in all three metrics, indicating that most unrelaxed structures are already very close to their relaxed counterparts. In contrast, the X–Mn–O and C2DB datasets display long tails, reflecting substantial atomic displacements and lattice deformations during structural optimizations. The layered vdW crystals fall in between: they show modest coordinate MAEs but tightly clustered cell-shape and volume errors near zero, consistent with the minimal lattice adjustments expected from weak van der Waals interlayer interactions.

C Baselines

We compare E³Relax against six SOTA iteration-free ML models: DeepRelax (Yang et al. 2024), PAINN (Schütt, Unke, and Gastegger 2021), EGNN (Satorras, Hoogendoorn, and Welling 2021), EquiformerV2 (Liao et al. 2024), HEGNN (Cen et al. 2024), and GotenNet (Aykent and Xia 2025).

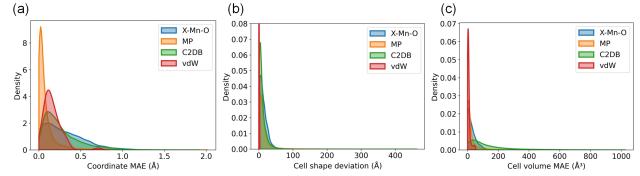


Figure 1: The deviation distribution of the four datasets: (a) coordinate MAE (\AA), (b) cell shape deviation (\AA), and (c) cell volume MAE (\AA^3) between initial and final structures.

HEGNN (Cen et al. 2024), and GotenNet (Aykent and Xia 2025).

- **DeepRelax** first predicts interatomic distances, atomic displacements, and lattice parameters, then solves a Euclidean-distance-geometry optimization to reconstruct Cartesian coordinates of relaxed structures. To predict relaxed lattice vectors, it pools the final GNN node embeddings into a single global vector, concatenates this with a linear embedding of the initial lattice vectors, passes the result through a small feed-forward network that outputs nine values, and reshapes them into a 3×3 matrix.
- **PAINN** is an equivariant graph neural network originally developed for interatomic potential and force prediction. Although not designed for iteration-free optimization, we adapt its force-prediction branch to output atomic displacements, and we use the same pooling-and-feedforward strategy as DeepRelax to predict the relaxed lattice.
- **EGNN** is a lightweight SE(3)-equivariant architecture that updates atomic coordinates at each layer, enabling direct prediction of atomic coordinates of relaxed structures. We employ the identical lattice-prediction pipeline used by DeepRelax.
- **EquiformerV2** is an equivariant Transformer that incorporates efficient eSCN convolutions, attention renormalization, separable S^2 activations, and separable layer normalization to model higher-order geometric interactions. We adapt its force-prediction head to output atomic displacements and apply the DeepRelax lattice-prediction procedure.
- **HEGNN** generalizes EGNN by adding high-degree steerable representations through efficient inner-product scalarization, preserving equivariance while improving expressivity.
- **GotenNet** rethinks efficient 3D equivariant GNNs by combining geometric tensor attention and hierarchical refinement—achieving high expressiveness without the computational burden of Clebsch–Gordan tensor products.

D Performance indicators

MAE of Coordinate The MAE of coordinates quantifies the positional discrepancies between ML-predicted and

Table 1: Summary of the four benchmark datasets used in our experiments.

Dataset	Dimensionality	Sample size	Elements covered	Train : Val : Test
Materials Project (MP)	3D	62 724	89	8 : 1 : 1
X-Mn-O	3D	28 579	6	8 : 1 : 1
C2DB	2D	11 581	62	6 : 2 : 2
Layered vdW Crystals	2D	57	29	8 : 1 : 1

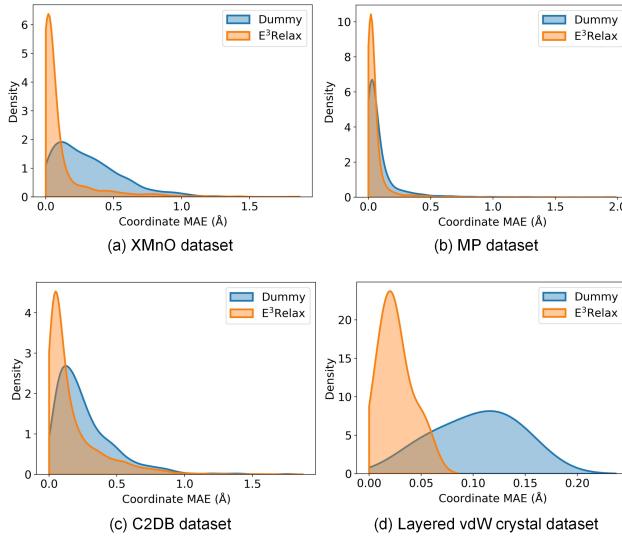


Figure 2: Distribution for coordinates MAE (\AA) for structures predicted by the Dummy model and $E^3\text{Relax}$ across various datasets. (a) X-Mn-O, (b) MP, (c) C2DB, and (d) layered vdW crystals.

DFT-relaxed structures. It is defined as follows:

$$\Delta_{\text{coord}} = \frac{1}{3N} \sum_{v_i \in \mathcal{V}^a} |\vec{r}_i^{(T)} - \tilde{r}_i| \quad (1)$$

where N is the total number of nodes in graph \mathcal{G} , T is the total number of layers, and $\vec{r}_i^{(T)}$ and \tilde{r}_i represent the predicted and ground truth Cartesian coordinates, respectively.

Cell Shape Deviation We measure cell shape deviation by the Frobenius norm of the difference between the ML-predicted and DFT-relaxed lattice metric tensors. Given the lattice matrix $\mathbf{L} = [\vec{l}_1, \vec{l}_2, \vec{l}_3]$, the lattice metric tensor is $\mathbf{G} = \mathbf{L}^T \mathbf{L}$. The cell shape deviation is then

$$\Delta_{\text{shape}} = \sqrt{\sum_{i=1}^3 \sum_{j=1}^3 (\mathbf{G}_{\text{pred}}^{ij} - \mathbf{G}_{\text{true}}^{ij})^2} \quad (2)$$

where $\mathbf{G}_{\text{pred}}^{ij}$ and $\mathbf{G}_{\text{true}}^{ij}$ represent the components of the predicted and truth lattice metric tensors, respectively.

MAE of Cell Volume The cell volume error measures the model’s ability to preserve the overall scale of the unit cell. It is defined as the absolute difference between the predicted and reference volumes:

$$\Delta_{\text{volume}} = \left| |\vec{l}_1^{(T)} \cdot (\vec{l}_2^{(T)} \times \vec{l}_3^{(T)})| - |\tilde{l}_1 \cdot (\tilde{l}_2 \times \tilde{l}_3)| \right| \quad (3)$$

where $\vec{l}_1^{(T)}, \vec{l}_2^{(T)}, \text{ and } \vec{l}_3^{(T)}$ are the predicted lattice vectors, and $\tilde{l}_1, \tilde{l}_2, \text{ and } \tilde{l}_3$ are their corresponding DFT-relaxed reference lattice vectors.

E Distribution of Predicted Structures

Figure 2 compares the coordinate MAE distributions obtained with the Dummy model (blue) and $E^3\text{Relax}$ (orange) across the four datasets. In every case the $E^3\text{Relax}$ curve is shifted markedly toward lower errors, confirming that the model recovers geometries much closer to the DFT-relaxed ground truth. The gain is most significant for the X-Mn-O, C2DB and layered-vdW sets, where large structural optimizations occur. For the MP dataset, however, the improvement over the Dummy model is less visually pronounced. This is attributed to the relatively minor structural differences between the unrelaxed and relaxed structures in the MP dataset, as shown in Figure 1.

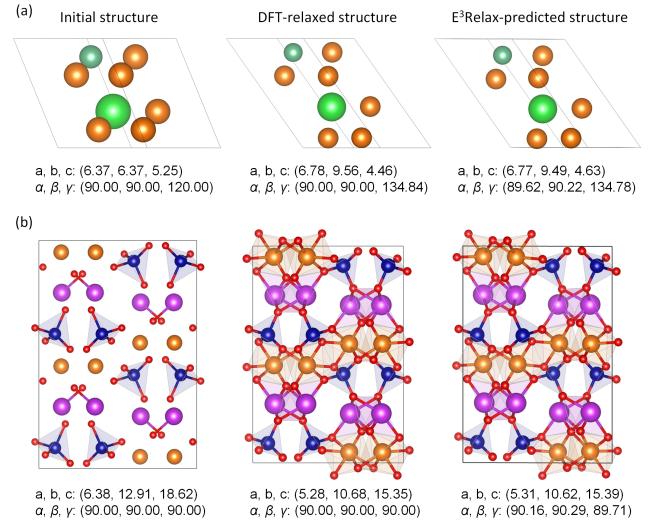


Figure 3: Two crystal structures from the MP database after optimization with $E^3\text{Relax}$: (a) BaMg₆Nb and (b) Bi₈Cr₈Mg₈O₄₀. a , b , and c are lattice constants in angstroms (\AA), and α , β , and γ are angles in degrees ($^\circ$).

F Visualization of $E^3\text{Relax}$ -predicted Structures

To further qualitatively evaluate $E^3\text{Relax}$, we selected representative test-set examples from MP, C2DB, and layered vdW crystals dataset and relaxed them using our model (Figures 3–5). In these examples, the $E^3\text{Relax}$ -predicted geometries closely match the corresponding DFT-relaxed structures, particularly in terms of atomic arrangements.

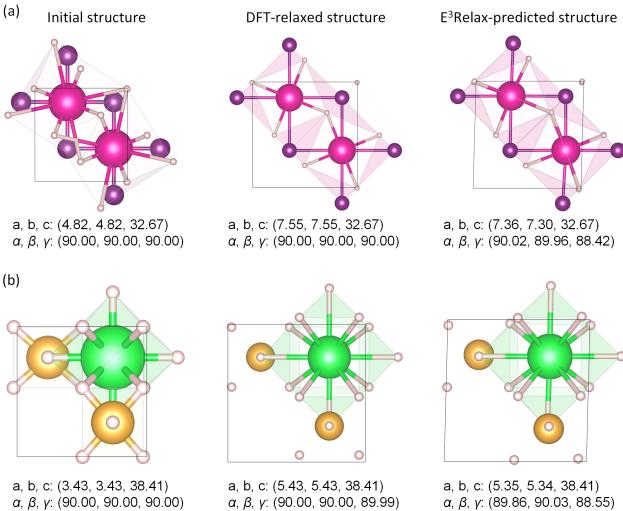


Figure 4: Visualization of two crystal structures from the C2DB dataset relaxed by E³Relax: (a) H₈I₂Rb₂ and (b) H₁₂Au₂Sr, where a , b , and c are lattice constants in angstroms (\AA), and α , β , and γ are angles in degrees ($^\circ$).

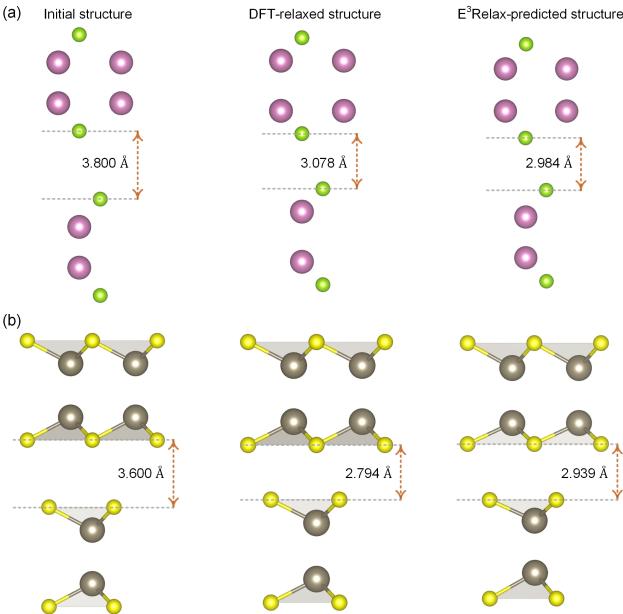


Figure 5: Visualization of two layered vdW crystal structures relaxed by E³Relax: (a) In₄Se₄. (b) S₄Tl₄.

G Hyperparameter Sensitivity Analysis

We further analyze the sensitivity of E³Relax to key hyperparameters, including the number of layers (T), hidden dimension size (F), and layer weights (α_t) (Figure 6). Overall, E³Relax demonstrates strong robustness to hyperparameter variations—its performance remains stable across a broad range of network depths, hidden dimensions, and layer-weight configurations.

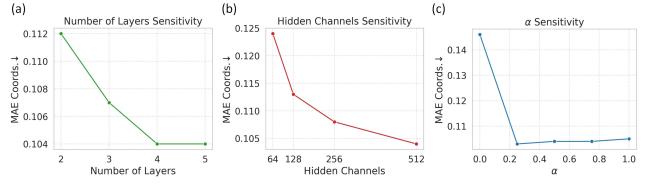


Figure 6: Sensitivity analysis of E³Relax with respect to network depth (T), hidden dimension (F), and layer weights (α_t).

H Training M3GNet and CHGNet

In addition to iteration-free models, another popular approach for accelerating structural optimization is iteration-based methods. These methods train GNN-based surrogate models, such as M3GNet (Chen and Ong 2022), CHGNet (Deng et al. 2023), and MACE (Batazia et al. 2022), to approximate energy forces, and stress. These predictions then drive an outer optimization loop, in which both atomic positions and lattice parameters are updated repeatedly until convergence.

In this study, we selected two well-known and straightforward-to-implement iterative models, M3GNet (Chen and Ong 2022) and CHGNet (Deng et al. 2023), as baseline models. These models were implemented directly from the authors’ public repositories using default settings. We follow the original M3GNet protocol (Chen and Ong 2022) by extracting three snapshots per trajectory—the initial, an intermediate, and the final relaxed structures—yielding three structures per sample. The two models were then trained to predict energy, forces, and stress. Once trained, the models were used to relax crystal structures by iteratively guiding atoms toward lower-energy regions.

I DFT Parameters

In this study, DFT calculations are conducted using the Vienna Ab initio Simulation Package (VASP) (Kresse and Furthmüller 1996) and Atomic Simulation Environment (ASE) Python library (Larsen et al. 2017). These calculations employ the Generalized Gradient Approximation (GGA) with the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional (Perdew, Burke, and Ernzerhof 1996). For the XMnO dataset, we set the energy cut-off at 500 eV, the energy convergence criterion at 1.0×10^{-5} eV, and use a K-point mesh of $6 \times 6 \times 6$. Gaussian smearing with a width of 0.05 eV is used for the electronic occupations. For the C2DB dataset, the energy convergence criterion is set to 1.0×10^{-4} eV, and the K-point mesh is set to $6 \times 6 \times 1$. For layered vdW crystals, DFT calculations were conducted with van der Waals corrections using the DFT-D3 Grimme method. For MoS₂ structure with point defects, optimization was performed until interatomic forces were below 0.05 eV/ \AA . Spin polarization was included, following the approaches of previous studies (Huang et al. 2023; Kazeev et al. 2023).

References

- Aykent, S.; and Xia, T. 2025. Gotennet: Rethinking efficient 3d equivariant graph neural networks. In *The Thirteenth International Conference on Learning Representations*.
- Batatia, I.; Kovacs, D. P.; Simm, G.; Ortner, C.; and Csányi, G. 2022. MACE: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in Neural Information Processing Systems*, 35: 11423–11436.
- Cen, J.; Li, A.; Lin, N.; Ren, Y.; Wang, Z.; and Huang, W. 2024. Are high-degree representations really unnecessary in equivariant graph neural networks? *Advances in Neural Information Processing Systems*, 37: 26238–26266.
- Chen, C.; and Ong, S. P. 2022. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11): 718–728.
- Deng, B.; Zhong, P.; Jun, K.; Riebesell, J.; Han, K.; Bartel, C. J.; and Ceder, G. 2023. CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 1–11.
- Gjerding, M. N.; Taghizadeh, A.; Rasmussen, A.; Ali, S.; Bertoldo, F.; Deilmann, T.; Knøsgaard, N. R.; Kruse, M.; Larsen, A. H.; Manti, S.; et al. 2021. Recent progress of the computational 2D materials database (C2DB). *2D Materials*, 8(4): 044002.
- Huang, P.; Lukin, R.; Faleev, M.; Kazeev, N.; Al-Maeeni, A. R.; Andreeva, D. V.; Ustyuzhanin, A.; Tormasov, A.; Castro Neto, A.; and Novoselov, K. S. 2023. Unveiling the complex structure-property correlation of defects in 2D materials based on high throughput datasets. *npj 2D Materials and Applications*, 7(1): 6.
- Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; et al. 2013. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1).
- Kazeev, N.; Al-Maeeni, A. R.; Romanov, I.; Faleev, M.; Lukin, R.; Tormasov, A.; Castro Neto, A.; Novoselov, K. S.; Huang, P.; and Ustyuzhanin, A. 2023. Sparse representation for machine learning the properties of defects in 2D materials. *npj Computational Materials*, 9(1): 113.
- Kim, S.; Noh, J.; Gu, G. H.; Aspuru-Guzik, A.; and Jung, Y. 2020. Generative adversarial networks for crystal structure prediction. *ACS central science*, 6(8): 1412–1420.
- Kim, S.; Noh, J.; Jin, T.; Lee, J.; and Jung, Y. 2023. A structure translation model for crystal compounds. *npj Computational Materials*, 9(1): 142.
- Kresse, G.; and Furthmüller, J. 1996. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 54(16): 11169.
- Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; et al. 2017. The atomic simulation environment—a Python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27): 273002.
- Liao, Y.-L.; Wood, B.; Das, A.; and Smidt, T. 2024. EquiformerV2: Improved Equivariant Transformer for Scaling to Higher-Degree Representations. In *International Conference on Learning Representations (ICLR)*.
- Lyngby, P.; and Thygesen, K. S. 2022. Data-driven discovery of 2D materials by deep generative models. *npj Computational Materials*, 8(1): 232.
- Perdew, J. P.; Burke, K.; and Ernzerhof, M. 1996. Generalized gradient approximation made simple. *Physical review letters*, 77(18): 3865.
- Satorras, V. G.; Hoogeboom, E.; and Welling, M. 2021. E (n) equivariant graph neural networks. In *International conference on machine learning*, 9323–9332. PMLR.
- Schütt, K.; Unke, O.; and Gastegger, M. 2021. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International Conference on Machine Learning*, 9377–9388. PMLR.
- Yang, Z.; Zhao, Y.-M.; Wang, X.; Liu, X.; Zhang, X.; Li, Y.; Lv, Q.; Chen, C. Y.-C.; and Shen, L. 2024. Scalable crystal structure relaxation using an iteration-free deep generative model with uncertainty quantification. *Nature Communications*, 15(1): 8148.