nature communications

Article

https://doi.org/10.1038/s41467-024-52378-3

Scalable crystal structure relaxation using an iteration-free deep generative model with uncertainty quantification

Received: 14 April 2024

Accepted: 2 September 2024

Published online: 17 September 2024

Check for updates

Ziduo Yang[1,2,3,10], Yi-Ming Zhao[1,10], Xian Wang[4], Xiaoqing Liu[1], Xiuying Zhang[1], Yifan Li[1], Qiujie Lv[1,2], Calvin Yu-Chian Chen[3,5,6,7,8] ✉ & Lei Shen[1,9] ✉

In computational molecular and materials science, determining equilibrium structures is the crucial first step for accurate subsequent property calculations. However, the recent discovery of millions of new crystals and super large twisted structures has challenged traditional computational methods, both ab initio and machine-learning-based, due to their computationally intensive iterative processes. To address these scalability issues, here we introduce DeepRelax, a deep generative model capable of performing geometric crystal structure relaxation rapidly and without iterations. DeepRelax learns the equilibrium structural distribution, enabling it to predict relaxed structures directly from their unrelaxed ones. The ability to perform structural relaxation at the millisecond level per structure, combined with the scalability of parallel processing, makes DeepRelax particularly useful for large-scale virtual screening. We demonstrate DeepRelax's reliability and robustness by applying it to five diverse databases, including oxides, Materials Project, two-dimensional materials, van der Waals crystals, and crystals with point defects. DeepRelax consistently shows high accuracy and efficiency, validated by density functional theory calculations. Finally, we enhance its trustworthiness by integrating uncertainty quantification. This work significantly accelerates computational workflows, offering a robust and trustworthy machine-learning method for material discovery and advancing the application of AI for science.

Atomic structure relaxation is usually the first step and foundation for further computational analysis of properties in computational chemistry, physics, materials science, and medicine. This includes applications such as chemical reactions on surfaces, complex defects in semiconductor heterostructures, and drug design. To date, computational relaxation algorithms have typically been achieved using iterative optimization, such as traditional ab initio methods, as shown in Fig. 1a. For example, each iteration in density functional theory (DFT) calculations involves solving the Schrödinger equation to determine the electronic density distribution, from which the total

[1]Department of Mechanical Engineering, National University of Singapore, Singapore, Singapore. [2]Artificial Intelligence Medical Research Center, School of Intelligent Systems Engineering, Shenzhen Campus of Sun Yat-sen University, Shenzhen, China. [3]AI for Science (AI4S)-Preferred Program, School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, Shenzhen, China. [4]Department of Physics, National University of Singapore, Singapore, Singapore. [5]State Key Laboratory of Chemical Oncogenomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen, China. [6]Department of Medical Research, China Medical University Hospital, Taichung, Taiwan. [7]Department of Bioinformatics and Medical Engineering, Asia University, Taichung, Taiwan. [8]Guangdong L-Med Biotechnology Co., Ltd., Meizhou, Guangdong, China. [9]National University of Singapore (Chongqing) Research Institute, Chongqing, China. [10]These authors contributed equally: Ziduo Yang, Yi-Ming Zhao. ✉e-mail: cy@pku.edu.cn; shenlei@nus.edu.sg

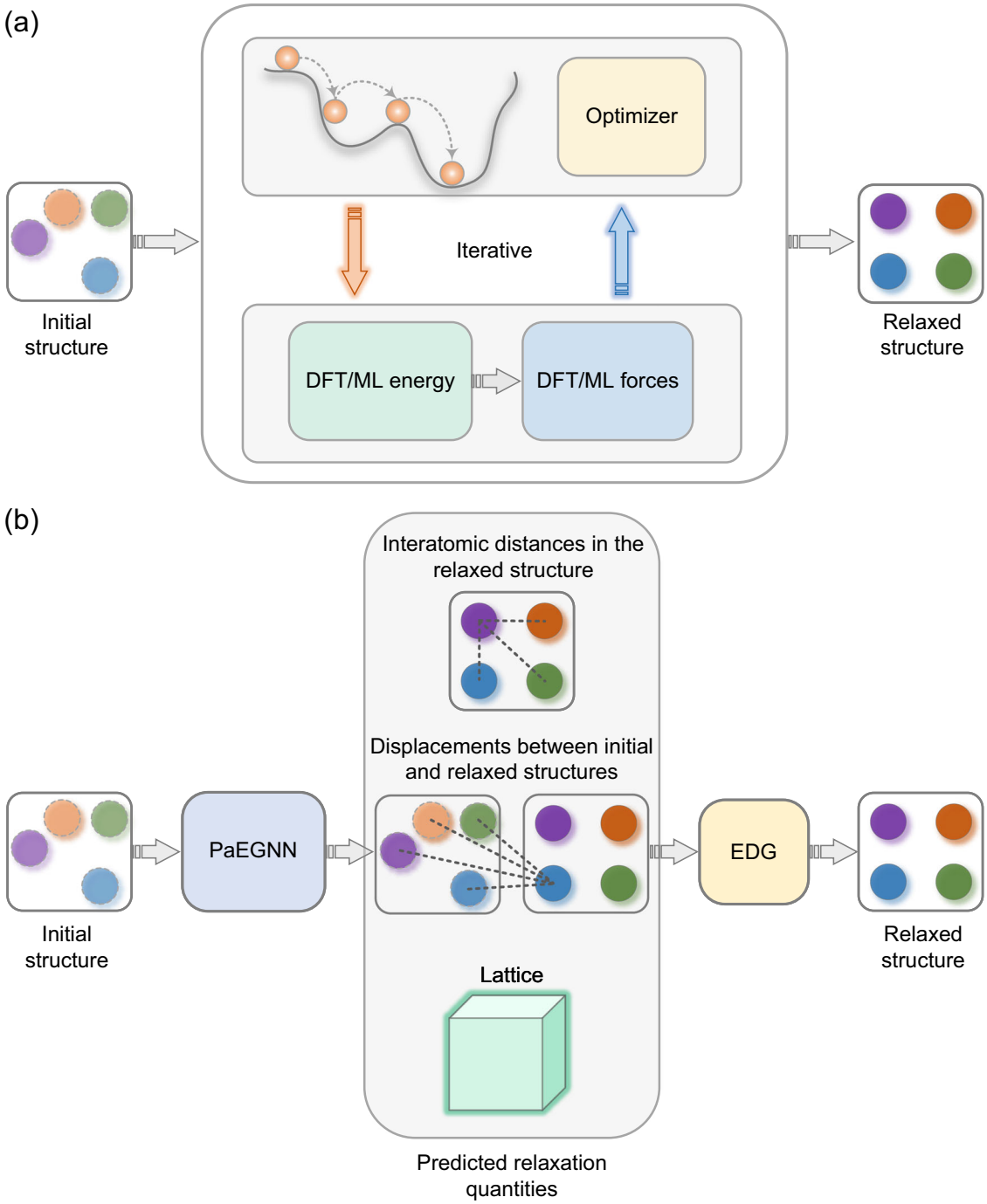

Fig. 1 | An overview of ML methods for crystal structure relaxation. a Iterative ML methods that iteratively estimate energy and force to determine the equilibrium structure. **b** Illustration of our proposed DeepRelax method, which employs a periodicity-aware equivariant graph neural network (PaEGNN) to directly predict the relaxation quantities. Euclidean distance geometry (EDG) is then used to determine the final relaxed structure that satisfies the predicted relaxation quantities.

energy of the system can be calculated. The forces on each atom, derived from differentiating this energy with respect to atomic positions, guide atomic movements to lower the system's energy, typically using optimization algorithms. Despite its effectiveness, the high computational demands and poor scalability of DFT limit its applications across high-dimensional chemical and structural spaces[1], such as the complex chemical reaction surfaces, doped semiconductor interfaces, or in the structural relaxation of the 2.2 million new crystals recently identified by DeepMind[2]. It is worth noting that the discovery of huge new materials has been significantly accelerated by high-throughput DFT calculations[3–6] and advanced machine learning (ML) algorithms[7–10], which is promoting the development of more efficient relaxation algorithms.

ML has emerged as a promising alternative for predicting relaxed structures[1,11–18]. As conventional iterative optimization, iterative ML approaches[1,11–14,17,18] utilize surrogate ML models to approximate energy and forces at each iteration, as shown in Fig. 1a, thereby circumventing the need to solve the computationally intensive Schrödinger equation. A typical example is the defect engineering in crystalline materials[19–21]. Mosquera-Lois et al.[13] and Jiang et al.[22] demonstrated that ML surrogate models could accelerate the optimization of crystals with defects. These ML models can retain DFT-level

accuracy by training on extensive databases containing detailed information on structural relaxations, including energy, forces, and stress.

However, there are two primary challenges in current iterative ML structural optimizers: training data limitations and non-scalability. Their training dataset must include full or partial intermediate steps of DFT relaxation. However, almost all publicized material databases, such as ICSD[23] and 2DMatPedia[6], do not provide such structural information, potentially limiting the application of iterative ML methods. The other challenge is that the large-scale parallel processing capability of iterative ML methods is limited due to their step-by-step nature. To address this, Yoon et al.[16] developed a model called DOGSS, and Kim et al.[15] proposed a model named Cryslator. Both conceptually introduce direct ML approaches to predict the final relaxed structures from their unrelaxed counterparts. However, these approaches have only been validated on specific datasets or systems, and their universal applicability to diverse datasets or systems remains unproven.

In this work, we introduce DeepRelax, a scalable, universal, and trustworthy deep generative model designed for direct structural relaxation. DeepRelax requires only the initial crystal structures to predict equilibrium structures in just a few hundred milliseconds on a single GPU. Furthermore, DeepRelax can efficiently handle multiple crystal structures in parallel by organizing them into mini-batches for simultaneous processing. This capability is especially advantageous in large-scale virtual screening, where rapid assessment of numerous unknown crystal configurations is essential. To demonstrate the reliability and robustness, we evaluate DeepRelax across five different datasets, including diverse 3D and 2D materials: the Materials Project (MP)[24], X-Mn-O oxides[15,25], the Computational 2D Materials Database (C2DB)[26–28], layered van der Waals crystals, and 2D structures with point defects[19,21]. DeepRelax not only demonstrates superior performance compared to other direct ML methods but also exhibits competitive accuracy to the leading iterative ML model, M3GNet[11], while being ~100 times faster in terms of speed. Moreover, we conduct DFT calculations to assess the energy of DeepRelax's predicted structures, confirming our model's ability to identify energetically favorable configurations. In addition, DeepRelax employs an uncertainty quantification method to assess the trustworthiness of the model. Finally, we would like to highlight that the aim of using DeepRelax is not to replace DFT relaxation but pre-relaxation, making the predicted structures very close to the DFT-relaxed configuration. Thus, the DFT method can rapidly complete the residual relaxation steps, significantly speeding up the traditional ab initio relaxation process, especially for complex structures.

Results

DeepRelax architecture

DeepRelax emerges as a solution to the computational bottlenecks faced in DFT methods for crystal structure relaxation. Figure 1b shows the workflow of DeepRelax, which takes an unrelaxed crystal structure as input and uses a periodicity-aware equivariant graph neural network (PaEGNN) to predict the relaxation quantities, including interatomic distances in the relaxed structure, displacements between the initial and relaxed structures, and the lattice matrix of the relaxed structure. DeepRelax then employs a numerical Euclidean distance geometry (EDG) solver to determine the relaxed structure that satisfies the predicted relaxation quantities. In addition, DeepRelax also quantifies bond-level uncertainty for each predicted interatomic distance and displacement. Aggregating these bond-level uncertainties allows for the computation of the system-level uncertainty, offering valuable insights into the trustworthiness of the model.

A notable feature of PaEGNN, distinguishing it from previous graph neural networks (GNNs)[29,30] is the explicit differentiation of atoms in various translated cells to encode periodic boundary conditions (PBCs) using a unit cell offset encoding (UCOE). In addition, its design ensures equivariance, facilitating active exploration of crystal symmetries and thus providing a richer geometric representation of crystal structures.

Table 1 | Comparative results of DeepRelax and baseline models on the X-Mn-O dataset, evaluated based on MAE of coordinates (Å), bond length (Å), lattice (Å), cell volume (Å^3), and match rate (%) between the predicted and DFT-relaxed structures

Model	Coordinates	Bond length	Lattice	Cell volume	Match rate
Dummy	0.314	0.429	0.221	32.8	64.8
PAINN	0.159	0.175	0.066	3.8	81.2
EGNN	0.166	0.189	0.066	4.2	77.5
Cryslator*	0.127	–	–	6.2	83.7
DeepRelax	**0.116**	**0.136**	**0.063**	**3.4**	**84.7**

*The results of Cryslator are taken from[15]. DeepRelax is evaluated on the same training, validation, and testing sets as Cryslator for a fair comparison.
The best performance in each metric is highlighted in bold.

Benchmark on X-Mn-O dataset

For our initial benchmarking, we utilize the X-Mn-O dataset, a hypothetical elemental substitution database previously employed for photoanode application studies[25,31]. This dataset derives from the MP database, featuring prototype ternary structures that undergo elemental substitution with X elements (Mg, Ca, Ba, and Sr). It consists of 28,579 data pairs, with each comprising an unrelaxed structure and its corresponding DFT-relaxed state. The dataset is divided into training (N = 22, 863), validation (N = 2, 857), and test (N = 2, 859) sets, adhering to an 8:1:1 ratio. As illustrated in Supplementary Fig. 1, there are significant structural differences between the unrelaxed and DFT-relaxed structures within this dataset.

We conduct a comparative analysis of DeepRelax against the state-of-the-art (SOTA) benchmark model, Cryslator[15]. In addition, we incorporate two types of equivariant graph neural networks-PAINN[29] and EGNN[32]-into our analysis (see Subsection 4.8 for the details). The choice of equivariant models is informed by recent reports highlighting their accuracy in direct coordinate prediction for structural analysis[32–34]. To ensure a fair comparison, we use the same training, validation, and testing sets across all models. As a baseline measure, we introduce a Dummy model, which simply returns the input initial structure as its output. This serves as a control reference in our evaluation process.

To evaluate model performance, we use the mean absolute error (MAE) of Cartesian coordinates, bond lengths, lattice matrix, and cell volume to measure the consistency between predicted and DFT-relaxed structures. In addition, we calculate the match rate–a measure of how closely predicted relaxed structures align with their ground truth counterparts within a defined tolerance, as determined by Pymatgen[3]. Detailed descriptions of these metrics are provided in Subsection 4.10.

Table 1 presents the comparative results, showing that DeepRelax greatly outperforms other baselines. Notably, DeepRelax shows a remarkable improvement in prediction accuracy over the Dummy model, with enhancements of 63.06%, 68.30%, 71.49%, 89.63%, and 30.71% across coordinates, bond lengths, lattice, cell volumes, and match rate, respectively. Moreover, DeepRelax surpasses the previous SOTA model, Cryslator, by 8.66% in coordinate prediction, and 45.16% in cell volume estimation. Figure 2a shows the distribution of MAE for coordinates, lattice matrix, and cell volumes as predicted by the Dummy model and DeepRelax. DeepRelax demonstrates a notable leftward skewness in its distribution, signifying a tendency to predict structures that closely approach the DFT-relaxed state. To visualize the performance of DeepRelax, we take two typical structures, $Sr_4Mn_2O_6$

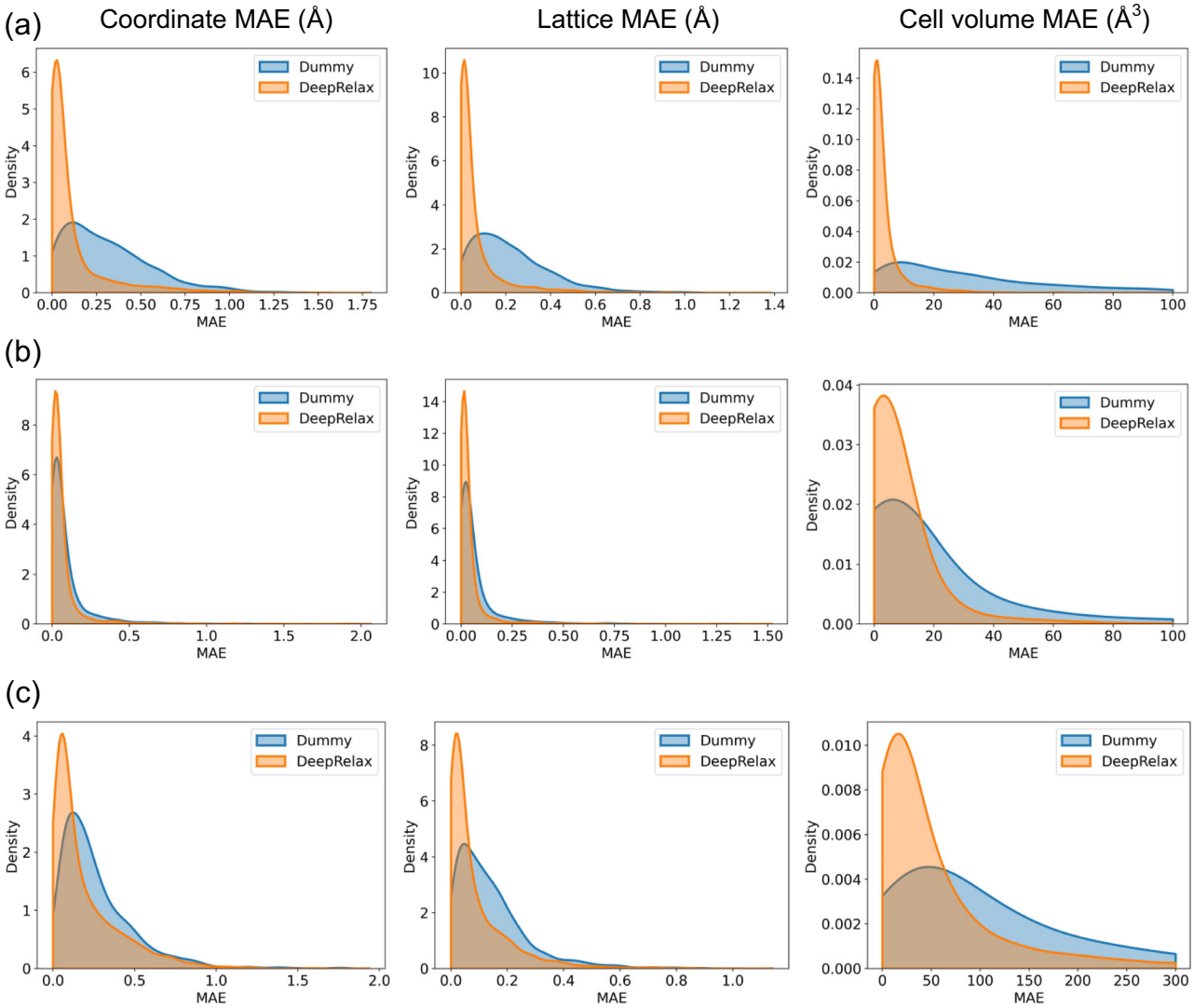

Fig. 2 | Distribution of MAE for predicted structures by the Dummy model and DeepRelax. a X-Mn-O dataset, (**b**) MP dataset, and (**c**) C2DB dataset. Each subfigure, from left to right, displays the MAE for coordinates (Å), lattice matrices (Å), and cell volumes (Å^3), respectively. Source data are provided in this paper.

and $Ba_1Mn_4O_8$, from the X-Mn-O database (see Fig. 3), and relax them using DeepRelax. As can be seen, the DeepRelax-predicted structures are highly consistent with the DFT-relaxed ones. The results demonstrate close agreement with DFT-relaxed structures. More DFT validations are in Subsection 2.7.

Benchmark on Materials Project

To demonstrate DeepRelax's universal applicability across various elements of the periodic table and diverse crystal types, we conduct further evaluations using the Materials Project dataset[11]. This dataset spans 89 elements and comprises 187,687 snapshots from 62,783 compounds captured during their structural relaxation processes. By excluding compounds missing either initial or DFT-relaxed structures, we refined the dataset to 62,724 pairs. Each pair consists of an initial and a corresponding DFT-relaxed structure, providing a comprehensive basis for assessing the performance of DeepRelax. This dataset is then split into training, validation, and test data in the ratio of 90%, 5%, and 5%, respectively. As illustrated in Supplementary Fig. 1, the structural differences for each pair tend toward an MAE of zero, indicating that many initial structures are closely aligned with their DFT-relaxed counterparts.

Training a direct ML model for datasets with varied compositions poses significant challenges, as evidenced in Crystalator[15]. This model shows reduced prediction performance when trained on the diverse MP database. Despite these challenges, DeepRelax demonstrates its robustness and universality. As indicated in Table 2, DeepRelax significantly surpasses the three baseline models in coordinate prediction, highlighting its effectiveness even in diverse and complex datasets. Figure 2b shows the MAE distribution for predicted structures compared to the DFT-relaxed ones for the MP dataset, which is less significant compared to the results for the X-Mn-O dataset shown in Fig. 2a. This is because many initial structures closely resemble their DFT-relaxed structures in the MP database as evidenced by Supplementary Fig. 1. Consequently, the MP dataset presents a more complex learning challenge for structural relaxation models.

Transfer learning on 2D materials database

Given that most materials databases do not provide the energy and force information of unrelaxed structures, it is difficult for conventional iterative ML models to transfer the trained model from Materials Project to other materials databases. This difficulty arises because transfer learning typically depends on the availability of energy and force information to fine-tune the model. DeepRelax, with its direct structural prediction feature, is more compatible with transfer learning, making it a flexible tool even when only structural data are available.

https://doi.org/10.1038/s41467-024-52378-3

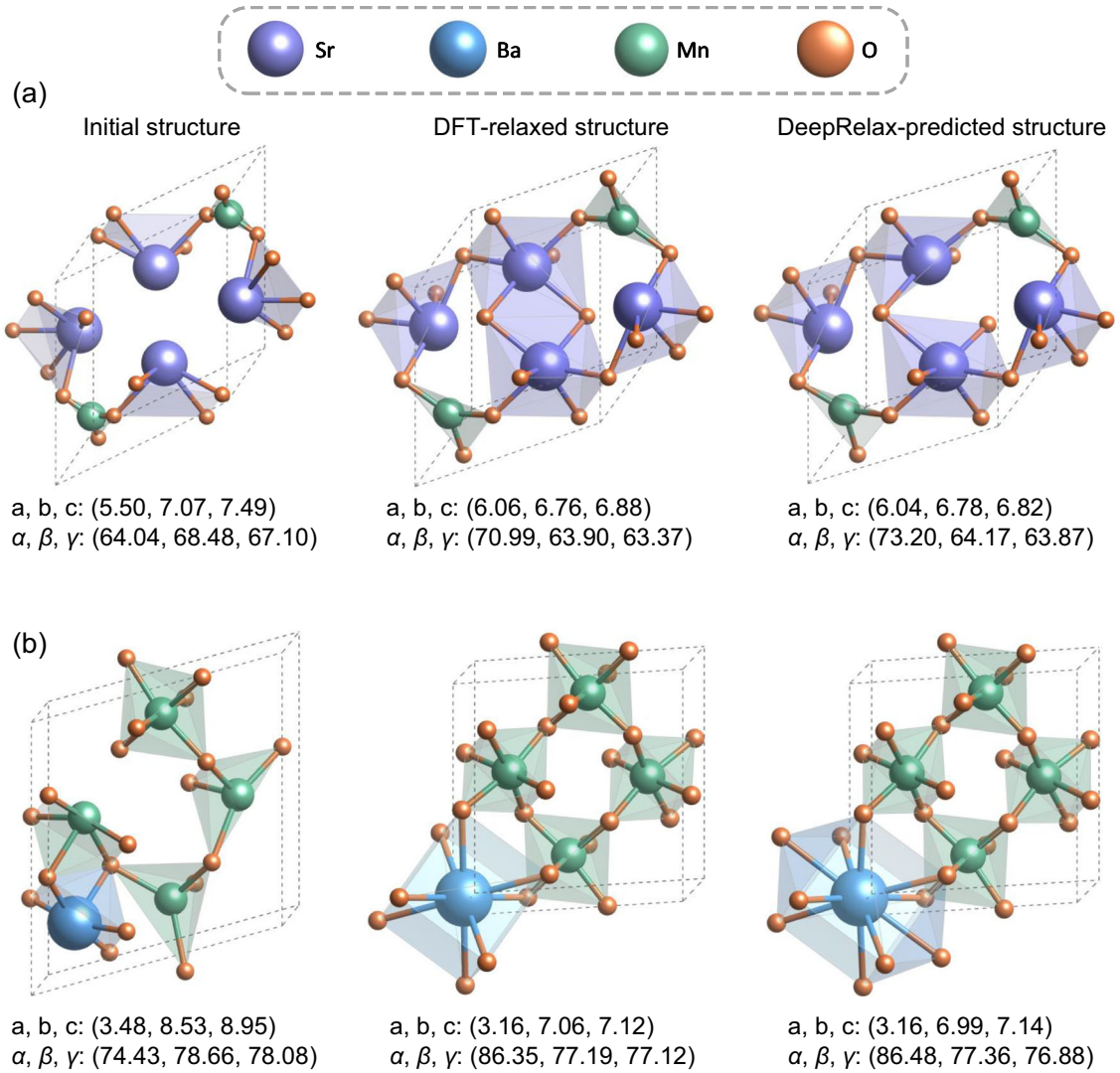

Fig. 3 | Visualization of two crystal structures relaxed by DeepRelax. a $Sr_4Mn_2O_6$ and (**b**) $Ba_1Mn_4O_8$, where *a*, *b*, and *c* are lattice constants in angstroms (Å), and *α*, *β*, and *γ* are angles in degrees (°). The results demonstrate close agreement between DeepRelax-predicted structures and DFT-relaxed structures.

To demonstrate the reliable application of DeepRelax, we extend the application of DeepRelax, initially pre-trained on 3D materials from the MP dataset, to 2D materials through transfer learning. We take the C2DB dataset[26–28] as an example, which covers 62 elements and comprises 11,581 pairs of 2D crystal structures, each consisting of an initial and a DFT-relaxed structure. The dataset is divided into training, validation, and testing subsets, maintaining a ratio of 6:2:2. The structural differences for each pair in this dataset fall within the range observed for the X-Mn-O and MP datasets, as shown in Supplementary Fig. 1.

Table 2 | Comparison of results between the proposed DeepRelax and other models on the MP dataset. The performances are evaluated by the MAE of coordinates (Å), bond length (Å), lattice (Å), and cell volume (Å^3) between the predicted and DFT-relaxed structures

Model	Coordinates	Bond length	Lattice	Cell volume
Dummy	0.095	0.112	0.072	27.0
PAINN	0.088	**0.082**	0.043	9.3
EGNN	0.086	0.086	0.043	**9.3**
DeepRelax	**0.066**	0.094	**0.041**	9.6
Improvement	30.53%	16.07%	43.06%	64.44%

The improvement is calculated by comparing DeepRelax with the Dummy model. The best performance in each metric is highlighted in bold.

In this application, DeepRelax trained via transfer learning is denoted as DeepRelaxT to differentiate it from DeepRelax. Table 3 illustrates our key findings: Firstly, both DeepRelax and DeepRelaxT outperform the other three baselines in the C2DB dataset, proving the applicability of our direct ML model to 2D materials. Figure 2c presents the MAE distribution for predicted structures by the Dummy model and DeepRelax on the C2DB dataset. These results suggest a modest improvement over the Dummy model. Notably, this improvement surpasses those observed for the MP dataset, as depicted in Fig. 2b. Secondly, DeepRelaxT demonstrates notable improvements over DeepRelax, with enhancements of 5.61% in coordinates, 38.43% in bond length, 3.53% in lattice, and 5.81% in cell volume in terms of MAE. Finally, DeepRelaxT shows a faster convergence rate than DeepRelax, as detailed in Supplementary Fig. 2. These results underline the benefits of large-scale pretraining and the efficacy of transfer learning.

Application in layered vdW crystals

Layered vdW crystals are of significant interest in the field of materials science and nanotechnology because of their unique tunable

Table 3 | Comparison of results among DeepRelax, DeepRelaxT (transfer learning version), and other models on the C2DB dataset. The performances are evaluated by the MAE of coordinates (Å), bond length (Å), lattice (Å), and cell volume (Å^3) between the predicted and DFT-relaxed structures

Model	Coordinates	Bond length	Lattice	Cell volume
Dummy	0.268	0.400	0.142	149.6
PAINN	0.226	0.283	0.086	61.9
EGNN	0.232	0.311	0.089	67.9
DeepRelax	0.196	0.268	0.085	60.2
DeepRelaxT	**0.185**	**0.165**	**0.082**	**56.7**

The best performance in each metric is highlighted in bold.

structures, such as twisting and sliding configurations[35]. One notable characteristic of these crystals is that the weak inter-layer vdW force may significantly change upon full relaxation, while the strong intra-layer chemical bonds undergo relatively small changes.

To demonstrate the reliable performance of our DeepRelax model on this type of crystal, we performed DFT relaxation of 58 layered vdW crystals covering 29 elements using van der Waals corrections, parameterized within the DFT-D3 Grimme method. Given the small sample size, we employ transfer learning, utilizing a model pre-trained on the Materials Project dataset.

Supplementary. Table 1 shows the inter-layer distances for the unrelaxed, DFT-D3-relaxed, and DeepRelax-predicted structures of six vdW layered crystals. The inter-layer distances of the predicted structures closely match those of the relaxed structures, highlighting the effectiveness of transferred DeepRelax on layered vdW crystals. Furthermore, an analysis of the MAE in bond length for representative bonding pairs, detailed in Supplementary Table 2, further demonstrates DeepRelax's precision in predicting structural changes in layered vdW crystals.

Application in crystals with defects

Most crystals have intrinsic defects. To demonstrate the robustness of DeepRelax to crystal structures with neutral point defects, we employ MoS_2 structures with a low defect concentration, including 5933 different defect configurations within an 8×8 supercell, as cataloged by Huang et al.[21], to evaluate DeepRelax. Supplementary Fig. 3 demonstrates a notably lower MAE in both atom coordinates and bond lengths for DeepRelax compared to the Dummy model, thereby underscoring DeepRelax's robustness and efficacy in defect structure calculations, which is further validated by DFT calculations in the next chapter.

DFT validations

Usually, the initial crystal structure may deviate from or be close to the final relaxed structure. To demonstrate the efficacy and robustness of DeepRelax, we perform DFT validations on two types of initial structures: those from the X-Mn-O dataset, which exhibit large deviations from the DFT-relaxed state, and those from the MP dataset, which are generally closer to their DFT-relaxed structures, as illustrated in Supplementary Fig. 1. The detailed settings for the DFT calculations are provided in Subsection 4.9.

In the first experiment, we evaluated our model's predictive capability under challenging conditions using the X-Mn-O dataset. We filtered out unrelaxed structures from the X-Mn-O test set that are structurally similar to their DFT-relaxed counterparts using Pymatgen's "Structure_matcher" function. From the remaining test set ($N = 1007$), we randomly selected 100 samples. Figure 4a shows the deviation distribution for the selected unrelaxed structures, which closely aligns with that of the complete test set, thus confirming the representativeness of the selected subset. Subsequently, we employed DeepRelax to predict the relaxed structures for these samples. Figure 4b shows box plots of the energy distributions for the unrelaxed, DFT-relaxed, and DeepRelax-predicted structures. The energy distributions of the DeepRelax-predicted and DFT-relaxed structures show similar medians and interquartile ranges, validating the model's accuracy in predicting energetically favorable structures. The MAE in energy is significantly reduced by an order of magnitude from 32.51 to 5.97.

In the second experiment, we tested whether DeepRelax remains effective with structures starting from a relatively rational initial unrelaxed state using the Materials Project dataset. Here, we again randomly selected 100 samples from the test set. Figure 4c shows the deviation distribution for these samples. The energies of the unrelaxed, DFT-relaxed, and DeepRelax-predicted structures were calculated using DFT. Figure 4d shows that the predicted structures feature an energy distribution nearly identical to that of the DFT-relaxed structures, demonstrating the model's effectiveness in handling relatively rational initial unrelaxed structures.

Besides the energy indicator, we further demonstrated our model's effectiveness using the number of residual optimizing (ionic) steps required for DFT relaxation. Specifically, we randomly selected 20 structures from the test set of the point-defect dataset, with their deviation distribution as shown in Fig. 4e. We then conducted DFT calculations starting from the unrelaxed and DeepRelax-predicted structures, respectively. As shown in Fig. 4f, starting DFT relaxation from the DeepRelax-predicted structures significantly reduces the number of required ionic steps, which is also robust.

Analysis of uncertainty

A critical challenge in integrating artificial intelligence (AI) into material discovery is establishing trustworthy AI models. Current deep learning models typically offer accurate predictions only within the chemical space covered by their training datasets, known as the applicability domain[36]. Predictions for samples outside this domain can be questionable. Thus, uncertainty quantification has become critical for AI models by quantifying prediction confidence levels, thereby aiding researchers in decision-making and experimental planning.

To validate the efficacy of our proposed uncertainty quantification in reflecting the confidence level of model predictions, we compute Spearman's rank correlation coefficient between the total predicted distance error and its associated system-level uncertainty. Figure 5a–c shows the hexagonal binning plots of system-level uncertainty against total distance MAE for the X-Mn-O, MP, and C2DB datasets, respectively. Correlation coefficients of 0.95, 0.83, and 0.88 for these datasets demonstrate a strong correlation between predicted error and predicted system-level uncertainty. Figure 5d, e presents the bond-level uncertainty visualization for two predicted structures, illustrating the correlation between predicted bond length error and associated bond-level uncertainty. These results indicate that the model's predicted uncertainty is a good indicator of the predicted structure's accuracy.

Ablation study

DeepRelax's technical contributions are twofold: it utilizes UCOE for handling PBCs explicitly, and it employs a method for estimating bond-level data uncertainty to encourage the model to capture a more comprehensive representation of the underlying data distribution.

To validate the effectiveness of these two strategies, we introduce three additional baseline models for comparison:

- Vanilla: Excludes both UCOE and data uncertainty estimation.
- DeepRelax (UCOE): Integrates UCOE but omits data uncertainty estimation.

https://doi.org/10.1038/s41467-024-52378-3

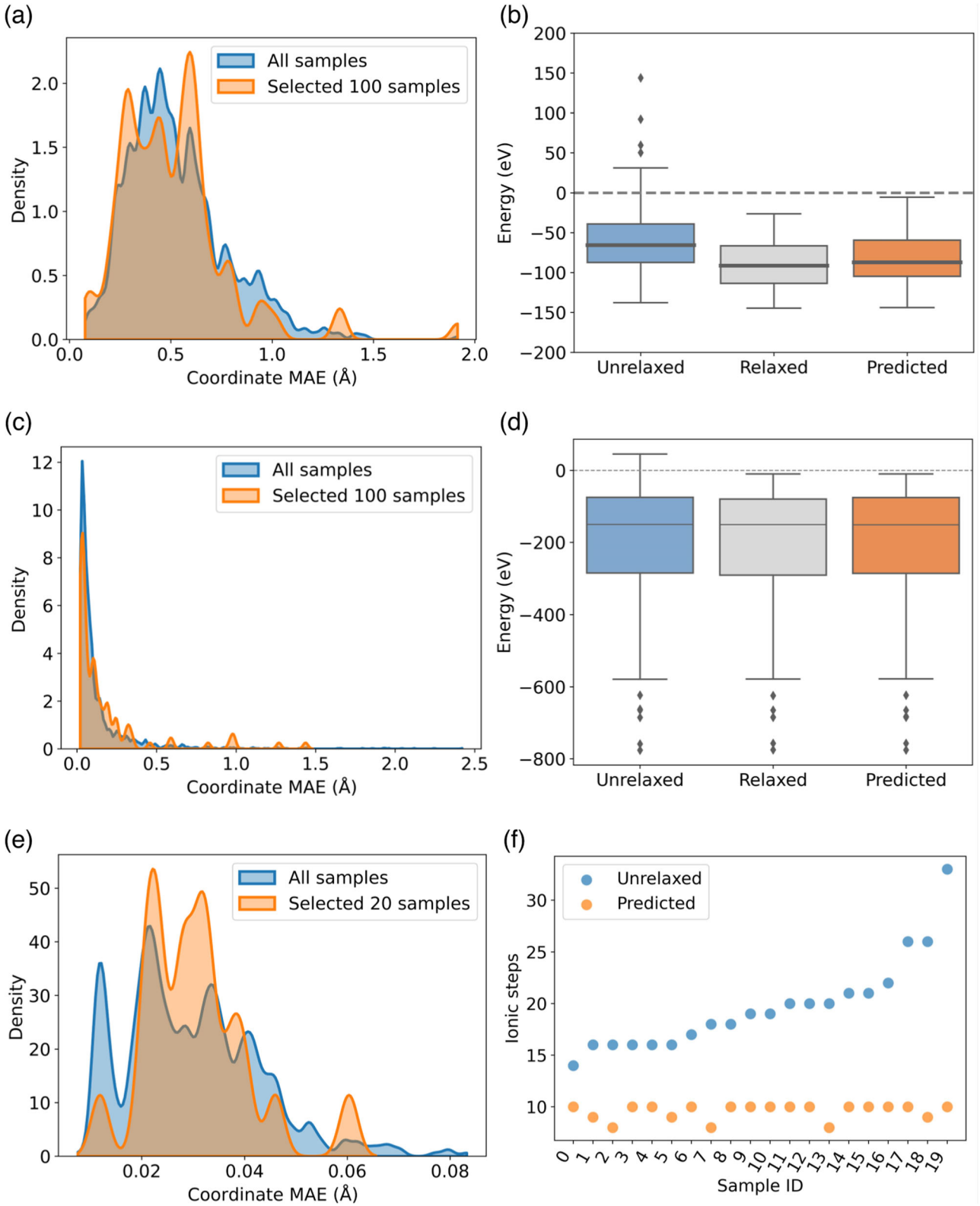

Fig. 4 | DFT validations. a Distributions of deviations for 100 random samples from the X-Mn-O dataset, measured using MAE in coordinates (Å) between the unrelaxed and DFT-relaxed structures. **b** Energy distribution for the three types of structures among the 100 random samples from the X-Mn-O dataset. The boxplots show the median (black line inside the box), interquartile range (box), and whiskers extending to 1.5 times the interquartile range, with outliers plotted as individual points. **c** Distribution of deviations for 100 random samples from the Materials Project dataset with relatively rational initial structures. **d** Energy distribution for the three types of structures among the 100 random samples from the Materials Project dataset. **e** Distributions of deviations for 20 random samples from the 2D materials defect dataset. **f** The number of DFT ionic steps required to complete DFT structure relaxation, starting from the initial unrelaxed structures and the DeepRelax-predicted structures, respectively. The samples are sorted based on the number of ionic steps required by the unrelaxed structures for better observation. Source data are provided in this paper.

https://doi.org/10.1038/s41467-024-52378-3

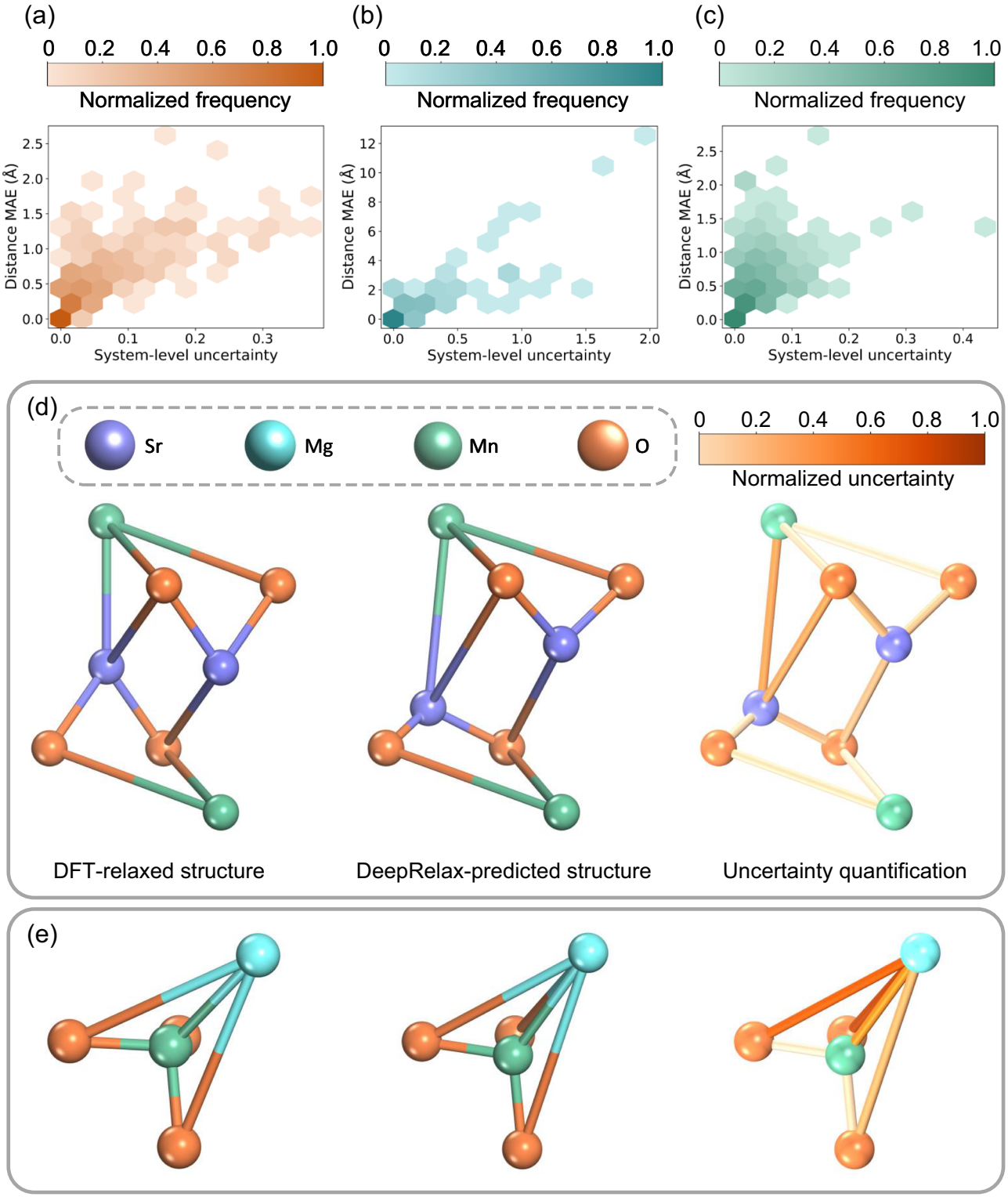

Fig. 5 | Uncertainty quantification. Hexagonal binning plots comparing system-level uncertainty with distance MAE (Å) for the (**a**) X-Mn-O, (**b**) MP, and (**c**) C2DB datasets. **d**, **e** illustrate the bond-level uncertainty for each predicted pairwise distance in $Sr_2Mn_2O_4$ and $Mg_1Mn_1O_3$, respectively, demonstrating the correlation between distance prediction errors and their associated bond-level uncertainties. Source data are provided in this paper.

- DeepRelax (BLDU): Implements bond-level data uncertainty estimation but not UCOE.

Table 4 demonstrates that DeepRelax (UCOE) attains a significant performance enhancement over the Vanilla model, suggesting the UCOE contributes greatly to model performance. On the other hand, DeepRelax (BLDU) shows a more modest improvement, which indicates the added value of data uncertainty estimation. Overall, DeepRelax shows a 25.16% improvement in coordinate MAE and a 20.00% advancement in bond length MAE over the Vanilla model. These

Table 4 | Ablation study to investigate the impact of unit cell offset encoding (UCOE) and bond-level data uncertainty (BLDU) estimation on model performance. The performances are evaluated by the MAE of coordinates (Å), bond length (Å), lattice (Å), and cell volume (Å^3) between the predicted and DFT-relaxed structures

Model	Coordinates	Bond length	Lattice	Cell volume
Dummy	0.314	0.429	0.221	32.839
Vanilla	0.155	0.170	0.063	3.478
DeepRelax (UCOE)	0.121	0.147	0.063	3.563
DeepRelax (BLDU)	0.142	0.171	0.064	3.539
DeepRelax	**0.116**	**0.136**	**0.063**	**3.442**

The best performance in each metric is highlighted in bold.

comparative results underscore the combined effectiveness of UCOE and data uncertainty estimation in our final DeepRelax model.

Discussion

The rapid advancement of generative models like CDVAE[8], PGCGM[10], and MatterGen[9], has opened avenues for the prolific generation of hypothetical materials with potentially desirable properties, such as 2.2 million new materials recently discovered by Google DeepMind. Clearly, it is impossible to relax such a huge number of structures using the traditional ab initio method, and it is also very difficult using the iterative ML relax models. For example, we further compare the efficiency between DeepRelax and M3GNet, a leading iterative ML relaxation model. DeepRelax offers a substantial speed advantage, being ~100 times faster than M3GNet (see Supplementary Table 3). Based on this estimation, to relax the 2.2 million new materials, our DeepRelax model only needs around 100 hours or 4 days, while M3GNet will take around 400 days. Moreover, our DeepRelax model supports parallel GPU processing, which can further significantly reduce computer time. While there are other direct structure-prediction ML methods, such as DOGSS[16] and Cryslator[15], detailed comparisons with these methods are provided in Supplementary Note 6. Overall, we introduce a fast, scalable, and trustworthy deep generative model, DeepRelax, for direct structural relaxation. Despite its advancements, opportunities for further improvement remain, which we explore in subsequent discussions.

Firstly, DeepRelax primarily focuses on predicting interatomic distances, which are quantities fundamentally involving two-body interactions. Incorporating the prediction of higher-order many-body quantities could further enhance the accuracy of structural predictions. Removing the PaEGNN and EDG modules to an end-to-end model could further increase the speed of process of structure relaxation.

Secondly, implementing active learning strategies[37,38] may further enhance DeepRelax's performance, particularly in underexplored chemical spaces. Active learning efficiently reduces the need for extensive training data by strategically choosing the most informative samples. DeepRelax's capability to assess prediction uncertainty aligns well with the principles of active learning, suggesting its feasibility as a future enhancement method.

Thirdly, DeepRelax is not designed to replace DFT but to significantly speed up the traditional ab initio relaxation process, especially for complex structures, such as complex chemical reaction surfaces or doped/disordered/twisted crystal interfaces.

In conclusion, DeepRelax represents a significant advancement in crystal structure prediction, offering efficient, scalable, universal, and trusted structural relaxation capabilities. It excels at direct predictions from initial configurations and effectively handles periodic boundary conditions, incorporating uncertainty quantification. Thus, DeepRelax stands as a powerful tool in advancing material science research.

Methods

Periodicity in crystals

A crystal can be conceptualized as a periodic arrangement of atoms in 3D space[8]. This periodicity is typically captured by a unit cell, that effectively represents the crystal structure. Such a unit cell, containing N atoms, can be fully characterized by three components:

- Atom Types: Represented by $\boldsymbol{A} = (a_0,...,a_N) \in \mathbb{A}^N$, where $\mathbb{A}$ denotes the set of all chemical elements.
- Atom Coordinates: Denoted by $\boldsymbol{R} = (\vec{\boldsymbol{r}}_0,..., \vec{\boldsymbol{r}}_N) \in \mathbb{R}^{N\times 3}$.
- Lattice Vectors: Expressed as $\boldsymbol{L} = (\vec{\boldsymbol{l}}_1, \vec{\boldsymbol{l}}_2, \vec{\boldsymbol{l}}_3) \in \mathbb{R}^{3\times 3}$.

Given $\boldsymbol{M} = (\boldsymbol{A}, \boldsymbol{R}, \boldsymbol{L})$, we can model the infinite periodic structure as:

$$\{(a_i', \vec{\boldsymbol{r}}_i') | a_i' = a_i, \vec{\boldsymbol{r}}_i' = \vec{\boldsymbol{r}}_i + k_1\vec{\boldsymbol{l}}_1 + k_2\vec{\boldsymbol{l}}_2 + k_3\vec{\boldsymbol{l}}_3, k_1, k_2, k_3 \in \mathbb{Z}\}, \quad (1)$$

where (k_1, k_2, k_3) are the unit cell offsets used to replicate the unit cell across the 3D space.

Multi-graph representation for crystal structures

Multi-graphs offer an intuitive way to represent crystal structures under periodic boundary conditions (PBCs)[8], as depicted in Fig. 6(a). These graphs can be effectively processed by GNNs through graph convolutions or message passing, which simulate many-body interactions[11,12,29,30,32,35,39–57]. Formally, we define a multi-graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ to encode these periodic structures. Here, $\mathcal{V} = \{v_1,...,v_N\}$ represents the set of nodes (atoms), and $\mathcal{E} = \{e_{ij,(k_1,k_2,k_3)} | i,j \in \{1,...,N\}, k_1, k_2, k_3 \in \mathbb{Z}\}$ signifies the set of edges (bonds). The edge $e_{ij,(k_1,k_2,k_3)}$ denotes a directed connection from node v_i in the original unit cell to node v_j in the unit cell translated by $k_1\vec{\boldsymbol{l}}_1 + k_2\vec{\boldsymbol{l}}_2 + k_3\vec{\boldsymbol{l}}_3$. Nodes are interconnected with their nearest neighbors within a cutoff distance D (set to 6 Å in our study).

To actively explore the crystal symmetry, each node $v_i \in \mathcal{V}$ is assigned both a scalar feature $\boldsymbol{x}_i \in \mathbb{R}^F$ and a vector feature $\vec{\boldsymbol{x}}_i \in \mathbb{R}^{F\times 3}$, i.e., retaining F scalars and F vectors for each node. These features are updated in a way that preserves symmetry during training. The scalar feature $\boldsymbol{x}_i^{(0)}$ is initialized as an embedding dependent on the atomic number, $E(z_i) \in \mathbb{R}^F$, where z_i is the atomic number, and E is an embedding layer mapping z_i to a F-dimensional feature vector. This embedding is similar to the one-hot vector but is trainable. The vector feature is initially set to zero, $\vec{\boldsymbol{x}}_i^{(0)} = \vec{\boldsymbol{0}} \in \mathbb{R}^{F\times 3}$. In addition, we define $\vec{\boldsymbol{r}}_{ij} = \vec{\boldsymbol{r}}_j - \vec{\boldsymbol{r}}_i$ as the vector from node v_i to v_j.

Periodicity-aware equivariant graph neural network

PaEGNN iteratively updates node representations in two phases: message passing and updating. These phases are illustrated in Fig. 6b and further detailed in Fig. 6c–e. During message passing, nodes receive information from neighboring nodes, expanding their accessible radius. In the updating phase, PaEGNN utilizes the node's internal messages (composed of F scalars and F vectors) to update its features. To prevent over-smoothing[58,59], skip connections are added to each layer.

In subsequent sections, we define the norm $\|\cdot\|$ and dot product $\langle\cdot,\cdot\rangle$ as operations along the spatial dimension, while concatenation $\oplus$ and the element-wise product $\circ$ are performed along the feature dimension.

Unit cell offset encoding. A notable feature of PaEGNN, distinguishing it from previous models[29,30], is the explicit differentiation of atoms in various translated unit cells to encode PBCs. To achieve this, we define the set $\mathcal{C} = \{-2, -1, 0, 1, 2\}$. We then use this set to generate translated unit cells with offsets $(k_1, k_2, k_3) \in \mathcal{C}\times\mathcal{C}\times\mathcal{C}$. The translated unit cells, resulting from the offsets $(k_1, k_2, k_3) \in \mathcal{C}\times\mathcal{C}\times\mathcal{C}$, are generally sufficient

https://doi.org/10.1038/s41467-024-52378-3

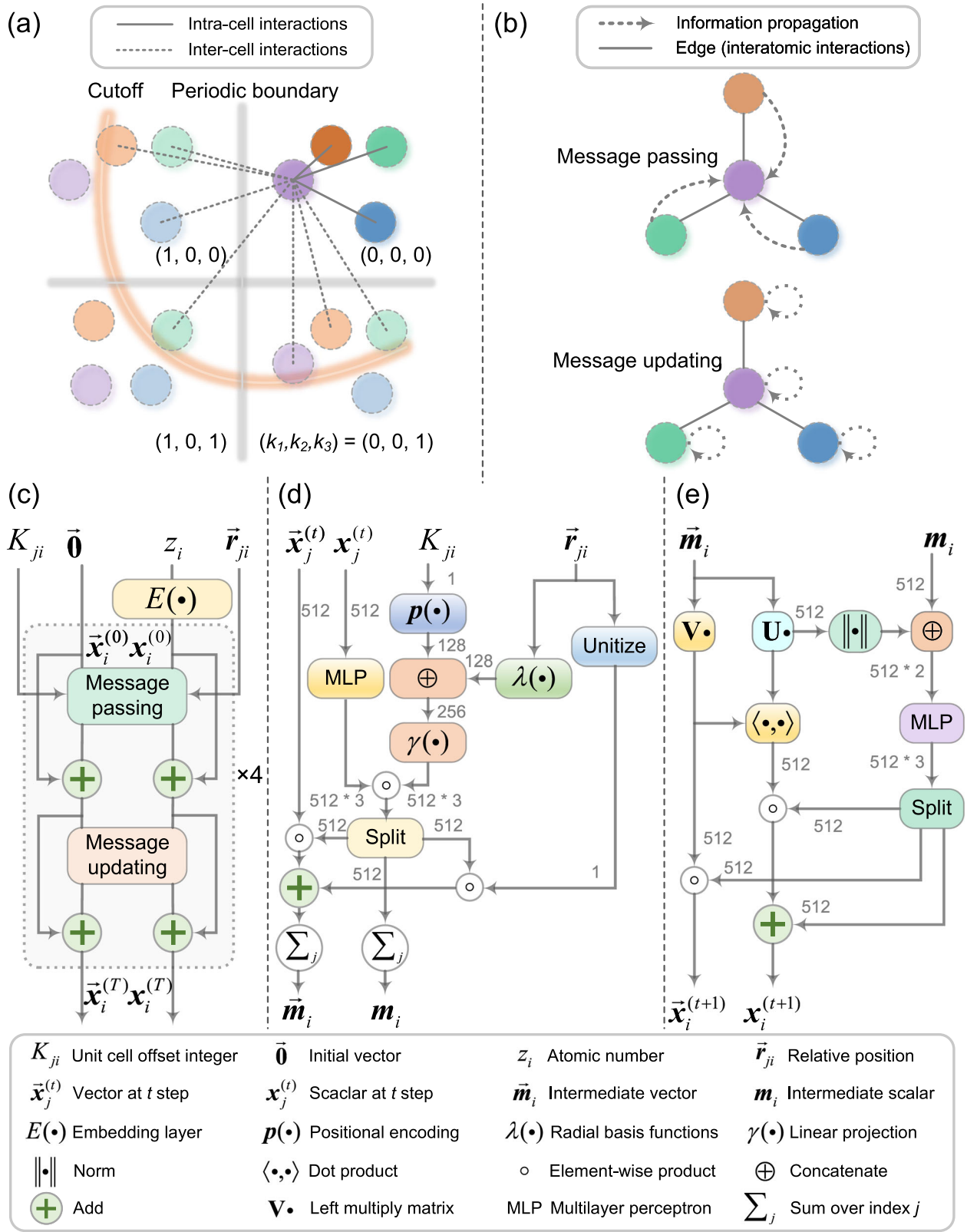

Fig. 6 | The architecture of PaEGNN. a Illustration of the multi-graph representation designed to capture atomic interactions across cell boundaries in periodic structures. **b** Message passing that collects messages from a node's neighbors and message updating that updates node representations using a node's internal states. **c** Overview of PaEGNN, comprising four layers, each with message passing and message updating phases, taking unit cell offset integer K_{ji}, initial vector $\vec{\boldsymbol{x}}_i^{(0)} = \vec{\mathbf{0}}$, initial scalar $\boldsymbol{x}_i^{(0)} = E(z_i)$, and relative position $\vec{\boldsymbol{r}}_{ji}$ as inputs and outputting the final vector $\vec{\boldsymbol{x}}_i^{(T)}$ and scalar $\boldsymbol{x}_i^{(T)}$ representations. **d** During the message passing phase, a node v_i aggregates messages from neighboring vectors $\vec{\boldsymbol{x}}_j^{(t)}$ and scalars $\boldsymbol{x}_j^{(t)}$, forming intermediate vector and scalar variables $\vec{\boldsymbol{m}}_i$ and $\boldsymbol{m}_i$. **e** The message updating phase integrates the F vectors and F scalars within $\vec{\boldsymbol{m}}_i$ and $\boldsymbol{m}_i$ to generate updated vector $\vec{\boldsymbol{x}}_i^{(t+1)}$ and scalar $\boldsymbol{x}_i^{(t+1)}$.

to encompass all atoms within a 6Å cutoff distance. We use $(k_1,k_2,k_3)_{ij}$ to denote the unit cell offset from node v_i to node v_j, where node v_j is located in a unit cell translated by $k_1\vec{l}_1+k_2\vec{l}_2+k_3\vec{l}_3$ relative to node v_i. Let $K_{ij}=(k_1+2)+(k_2+2)5+(k_3+2)25$ be a positive integer that uniquely indexes the unit cell offset, the sinusoidal positional encoding[60] for K_{ij} is computed as:

$$p(K_{ij},f)=\begin{cases}\sin(K_{ij}/10000^{f/F}), & \text{if } f\in\{0,2,4,\ldots,F-2\}\\ \cos(K_{ij}/10000^{(f-1)/F}), & \text{if } f\in\{1,3,5,\ldots,F-1\}\end{cases} \tag{2}$$

The full positional encoding vector is then

$$\boldsymbol{p}(K_{ij})=\Big(p(K_{ij},0),p(K_{ij},1),\ldots,p(K_{ij},F-1)\Big)\in\mathbb{R}^F \tag{3}$$

The unit cell offset encoding $\boldsymbol{p}(K_{ij})$ explicitly encodes the relative position of the unit cells in which the two nodes, v_i and v_j, are located. This encoding enables the GNN to explicitly recognize the periodic structure, thereby enhancing predictive performance.

Message passing phase. During this phase, a node v_i aggregates messages from neighboring scalars $\boldsymbol{x}_j^{(t)}$ and vectors $\vec{\boldsymbol{x}}_j^{(t)}$, forming intermediate scalar and vector variables $\boldsymbol{m}_i$ and $\vec{\boldsymbol{m}}_i$ as follows:

$$\boldsymbol{m}_i=\sum_{v_j\in\mathcal{N}(v_i)}(\boldsymbol{W}_h\boldsymbol{x}_j^{(t)})\circ\gamma_h\Big(\lambda(\|\vec{\boldsymbol{r}}_{ji}\|)\oplus\boldsymbol{p}(K_{ji})\Big) \tag{4}$$

$$\begin{aligned}\vec{\boldsymbol{m}}_i=\sum_{v_j\in\mathcal{N}(v_i)}&(\boldsymbol{W}_u\boldsymbol{x}_j^{(t)})\circ\gamma_u\Big(\lambda(\|\vec{\boldsymbol{r}}_{ji}\|)\oplus\boldsymbol{p}(K_{ji})\Big)\circ\vec{\boldsymbol{x}}_j^{(t)}\\&+(\boldsymbol{W}_v\boldsymbol{x}_j^{(t)})\circ\gamma_v\Big(\lambda(\|\vec{\boldsymbol{r}}_{ji}\|)\oplus\boldsymbol{p}(K_{ji})\Big)\circ\frac{\vec{\boldsymbol{r}}_{ji}}{\|\vec{\boldsymbol{r}}_{ji}\|}\end{aligned} \tag{5}$$

Here, $\oplus$ denotes concatenation, $\mathcal{N}(v_i)$ represents the neighboring nodes of v_i, $\boldsymbol{W}_h,\boldsymbol{W}_u,\boldsymbol{W}_v\in\mathbb{R}^{F\times F}$ are trainable weight matrices, λ is a set of Gaussian radial basis functions (RBF)[46] that are used to expand bond distances, and γ_h, γ_u, and γ_v are a linear projection mapping the concatenated feature back to F-dimensional space.

Message updating phase. The updating phase concentrates on integrating the F scalars and F vectors within $\boldsymbol{m}_i$ and $\vec{\boldsymbol{m}}_i$ to generate updated scalar $\boldsymbol{x}_i^{(t+1)}$ and vector $\vec{\boldsymbol{x}}_i^{(t+1)}$:

$$\boldsymbol{x}_i^{(t+1)}=\boldsymbol{W}_{s1}(\boldsymbol{m}_i\oplus\|\boldsymbol{U}\vec{\boldsymbol{m}}_i\|)+\boldsymbol{W}_{s2}(\boldsymbol{m}_i\oplus\|\boldsymbol{U}\vec{\boldsymbol{m}}_i\|)\langle\boldsymbol{V}\vec{\boldsymbol{m}}_i,\boldsymbol{U}\vec{\boldsymbol{m}}_i\rangle \tag{6}$$

$$\vec{\boldsymbol{x}}_i^{(t+1)}=\boldsymbol{W}_v(\boldsymbol{m}_i\oplus\|\boldsymbol{U}\vec{\boldsymbol{m}}_i\|)\circ(\boldsymbol{V}\vec{\boldsymbol{m}}_i) \tag{7}$$

where $\boldsymbol{W}_{s1},\boldsymbol{W}_{s2},\boldsymbol{W}_v\in\mathbb{R}^{F\times 2F}$ and $\boldsymbol{U},\boldsymbol{V}\in\mathbb{R}^{F\times F}$.

Predicting relaxation quantities

Assuming PaEGNN comprises T layers, we define the bond feature $\boldsymbol{h}_{ij}=\gamma\Big(\lambda(\|\vec{\boldsymbol{r}}_{ij}\|)\oplus\boldsymbol{p}(k_1,k_2,k_3)\Big)$, where γ is a linear projection mapping the concatenated feature back to F-dimensional space. The prediction of a pairwise distance $\hat{d}_{ij}$ for the edge $e_{ij,(k_1,k_2,k_3)}$ is formulated as:

$$\hat{d}_{ij}=|f_d(\boldsymbol{W}_d\boldsymbol{x}_i^{(T)}\oplus\boldsymbol{W}_d\boldsymbol{x}_j^{(T)}\oplus\boldsymbol{h}_{ij})| \tag{8}$$

where $\boldsymbol{W}_d\in\mathbb{R}^{F\times F}$ is a learnable matrix and $f_d:\mathbb{R}^{3F}\to\mathbb{R}$ is a linear map. Using Eqn. (8), we can predict both the interatomic distances in the relaxed structure and the displacements between the initial and relaxed structures. In addition, DeepRelax predicts the lattice matrix of the relaxed structure as follows:

$$\hat{\boldsymbol{L}}=r_L\left(f_L\left(\boldsymbol{W}_l\big(\vec{\boldsymbol{l}}_1\oplus\vec{\boldsymbol{l}}_2\oplus\vec{\boldsymbol{l}}_3\big)\oplus\left(\sum_{v_i\in\mathcal{G}}\boldsymbol{W}_x\boldsymbol{x}_i\right)\right)\right) \tag{9}$$

Here, $\boldsymbol{W}_x\in\mathbb{R}^{F\times F}$, $\boldsymbol{W}_l\in\mathbb{R}^{9\times F}$, and $f_L:\mathbb{R}^{2F}\to\mathbb{R}^9$ is a linear mapping yielding a 9-dimensional vector $\boldsymbol{L}_v$. The operation r_L reshapes $\boldsymbol{L}_v$ into a 3×3 matrix $\hat{\boldsymbol{L}}$ to reflect the lattice vectors.

Uncertainty-aware loss function

In real scenarios, each predicted distance is subject to inherent noise (e.g., measurement errors or human labeling errors). To capture this uncertainty, we can model the pairwise distances as random variables following a Laplace distribution, i.e., $d_{ij}\sim\text{Laplace}(\hat{d}_{ij},\hat{b}_{ij})$. Here, $\hat{d}_{ij}$ and $\hat{b}_{ij}$ are the location parameter and scale parameter, respectively. In our application, $\hat{d}_{ij}$ represents the predicted distance and $\hat{b}_{ij}$ represents the associated bond-level data uncertainty. The scale parameter $\hat{b}_{ij}$ is predicted as follows:

$$\hat{b}_{ij}=f_b(\boldsymbol{W}_b\boldsymbol{x}_i^{(T)}\oplus\boldsymbol{W}_b\boldsymbol{x}_j^{(T)}\oplus\boldsymbol{h}_{ij}) \tag{10}$$

where $\boldsymbol{W}_b\in\mathbb{R}^{F\times F}$ is a learnable matrix, and $f_b:\mathbb{R}^{3F}\to\mathbb{R}$ is a linear map.

To train DeepRelax such that its output follows the assumed Laplace distribution, we propose an uncertainty-aware loss $\mathcal{L}_u$, which comprises interatomic distance loss $\mathcal{L}_i$ and displacement loss $\mathcal{L}_d$:

$$\mathcal{L}_i=\sum_{e_{ij,(k_1,k_2,k_3)}\in\mathcal{E}}\log(2\hat{b}_{ij})+\frac{|d_{ij}-\hat{d}_{ij}|}{\hat{b}_{ij}} \tag{11}$$

$$\mathcal{L}_d=\sum_{e_{ij,(0,0,0)}\in\mathcal{E}}\log(2\hat{b}_{ij})+\frac{|d_{ij}-\hat{d}_{ij}|}{\hat{b}_{ij}} \tag{12}$$

$$\mathcal{L}_u=\mathcal{L}_i+\mathcal{L}_d \tag{13}$$

In these expressions, d_{ij} represents the ground truth distance. The edges $e_{ij,(k_1,k_2,k_3)}\in\mathcal{E}$ pertain to interatomic distance predictions, whereas $e_{ij,(0,0,0)}\in\mathcal{E}$ denotes edges used for displacement predictions within the unit cell, discounting PBCs. In essence, $\mathcal{L}_i$ and $\mathcal{L}_d$ represents the negative log-likelihood of the Laplace distribution, thereby capturing the data uncertainty. Consequently, a larger $\hat{b}_{ij}$ indicates greater bond-level data uncertainty in the prediction, and vice versa. The total loss $\mathcal{L}$ is consist of $\mathcal{L}_u$ and a lattice loss $\mathcal{L}_l$:

$$\mathcal{L}_l=\sum|\hat{\boldsymbol{L}}-\boldsymbol{L}| \tag{14}$$

$$\mathcal{L}=\mathcal{L}_u+\mathcal{L}_l \tag{15}$$

where $\boldsymbol{L}$ represents the ground lattice matrix of the relaxed structure.

Numerical Euclidean distance geometry solver

We propose a numerical EDG solver to determine the relaxed structure that aligns with the predicted relaxation quantities. Specifically, for a given graph $\mathcal{G}=(\mathcal{V},\mathcal{E},d)$ and a dimension K, the EDG problem[61–63] seeks a realization-specifically, a coordinate matrix $\hat{\boldsymbol{R}}\in\mathbb{R}^{N\times K}:\mathcal{V}\to\mathbb{R}^K$ in K-dimensional space that satisfies the distance constraint d as follows:

$$\forall(u,v)\in\mathcal{E},\ \|\hat{\boldsymbol{R}}(u)-\hat{\boldsymbol{R}}(v)\|=d_{uv} \tag{16}$$

For simplicity in notation, $\hat{\boldsymbol{R}}(u)$ and $\hat{\boldsymbol{R}}(v)$ are typically written as $\hat{\boldsymbol{R}}_u$ and $\hat{\boldsymbol{R}}_v$.

We reformulate the conventional EDG problem into a global optimization task:

$$\mathcal{L}_g = \sum_{(u,v)\in\mathcal{E}} |\,\|\hat{\boldsymbol{R}}_u - \hat{\boldsymbol{R}}_v\| - d_{uv}| \tag{17}$$

This is a non-convex optimization problem, and minimizing $\mathcal{L}_g$ gives an approximation solution of $\hat{\boldsymbol{R}}$.

In our specific application, we aim to find a coordinate matrix $\hat{\boldsymbol{R}} \in \mathbb{R}^{N\times 3}$ for a system of N atoms in three-dimensional space, meeting the constraints imposed by $\hat{d}_{ij}$, $\hat{b}_{ij}$, and $\hat{\boldsymbol{L}}$. Specifically, we first define an upper bound and a lower bound using $\hat{d}_{ij}$, $\hat{b}_{ij}$ as following:

$$\hat{d}^u_{ij} = \hat{d}_{ij} + \exp(\hat{b}_{ij}) \tag{18}$$

$$\hat{d}^l_{ij} = \hat{d}_{ij} - \exp(\hat{b}_{ij}) \tag{19}$$

Subsequently, we propose minimizing a bounded Euclidean distance (BED) loss:

$$\begin{aligned}\mathcal{L}_g = &\sum_{\substack{e_{ij,(k_1,k_2,k_3)}\in\mathcal{E}\\ e_{ij,(0,0,0)}\in\mathcal{E}}} \max(0, \|\hat{\boldsymbol{R}}_u - \hat{\boldsymbol{R}}_v\| - \hat{d}^u_{ij})\\ &+ \sum_{\substack{e_{ij,(k_1,k_2,k_3)}\in\mathcal{E}\\ e_{ij,(0,0,0)}\in\mathcal{E}}} \max(0, \hat{d}^l_{ij} - \|\hat{\boldsymbol{R}}_u - \hat{\boldsymbol{R}}_v\|)\end{aligned} \tag{20}$$

For each edge $e_{ij,(k_1,k_2,k_3)}$, the location of node v_j is dictated by $k_1\hat{\boldsymbol{l}}_1 + k_2\hat{\boldsymbol{l}}_2 + k_3\hat{\boldsymbol{l}}_3$, where $\hat{\boldsymbol{l}}_1, \hat{\boldsymbol{l}}_2, \hat{\boldsymbol{l}}_3$ are predicted lattice vectors. The BED loss only penalizes coordinate pairs whose distances fall outside the lower and upper bounds, thus mitigating the impact of less accurate predictions. In our work, we use Adam optimizer to minimize $\mathcal{L}_g$.

Uncertainty quantification

We initially quantify bond-level uncertainties and subsequently aggregate these to determine the system-level uncertainty of the predicted structure. The bond-level uncertainty can be further decomposed into data uncertainty and model uncertainty. Data uncertainty arises from the inherent randomness in the data, while model uncertainty arises from a lack of knowledge about the best model to describe the data[64].

We employ ensemble-based uncertainty techniques[36,65], which involve training an ensemble of T independent model replicates, with $T = 5$ used in this study. The T model replicates have the same neural network architectures and hyperparameters, but the learnable parameters are initialized with different random seeds. For the t-th model replicate, let $\hat{d}_{ij}(t)$ denote the predicted distance, $\hat{b}_{ij}(t)$ the associated bond-level data uncertainty, and $\hat{w}_{ij}(t)$ the associated bond-level model uncertainty. Model uncertainty for each pair is calculated as the deviation from the mean predicted distance $\bar{d}_{ij}$:

$$\hat{w}_{ij}(t) = |\,\hat{d}_{ij}(t) - \bar{d}_{ij}| \tag{21}$$

where the mean predicted distance $\bar{d}_{ij}$ is given by:

$$\bar{d}_{ij} = \frac{1}{T}\sum_{t=1}^{T} \hat{d}_{ij}(t) \tag{22}$$

The total bond-level uncertainty $\hat{U}_{ij}$ is the sum of the exponential of the data uncertainties and the model uncertainties across T models:

$$\hat{U}_{ij} = \frac{1}{T}\sum_{t=1}^{T}\left(\exp(\hat{b}_{ij}(t)) + \hat{w}_{ij}(t)\right) \tag{23}$$

Finally, the system-level uncertainty $\hat{U}$ is computed as the average of all bond-level uncertainties:

$$\hat{U} = \frac{1}{N}\sum_{\substack{e_{ij,(k_1,k_2,k_3)}\in\mathcal{E}\\ e_{ij,(0,0,0)}\in\mathcal{E}}} \hat{U}_{ij} \tag{24}$$

where N represents the total number of evaluated pairs.

Implementation details

The DeepRelax model is implemented using PyTorch. Experiments are conducted on an NVIDIA RTX A6000 with 48 GB of memory. The training objective is to minimize Eqn. (15). We use the AdamW optimizer with a learning rate of 0.0001 to update the model's parameters. In addition, we implement a learning rate decay strategy, reducing the learning rate if there is no improvement in a specified metric for a duration of 5 epochs.

We implement PAINN[29] and EGNN[32] models, utilizing the source code available at https://github.com/Open-Catalyst-Project/ocp and https://github.com/vgsatorras/egnn, respectively. These equivariant models are adept at directly predicting the coordinates of a relaxed structure from its unrelaxed counterpart, leveraging the intrinsic property that coordinates are equivariant quantities.

DFT calculations

In our study, DFT calculations are performed using the Vienna Ab initio Simulation Package (VASP)[66], employing the generalized gradient approximation (GGA) with the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional. All VASP calculations are performed using the electronic minimization algorithm "all band simultaneous update of orbitals" (ALGO = All), with a cut-off energy of 550 eV, an energy convergence criterion of 1.0×10^{-5} eV, and a Gaussian smearing width of 0.02 eV. For the X-Mn-O dataset, we run the self-consistent calculation to obtain the total energy without spin polarization. The K-point mesh is a $9\times9\times9$ grid, ensuring precise total energy calculations. The effective on-site Coulomb interactions (U value) of Mn 3d orbital is chosen as 3.9 eV, aligning with that used in Cryslator[15]. For the MP dataset, the self-consistent is running with a $5\times5\times5$ K-point mesh for structures containing fewer than 60 atoms and $3\times3\times3$ for those with more than 60 atoms. Spin polarization is applied to structures exhibiting magnetism to enhance the convergence of total energy calculations. For layered vdW crystals, we performed DFT calculations with van der Waals corrections (DFT-D3 Grimme method). For MoS_2 structures with defects, the structure is relaxed until the interatomic force is smaller than 0.05 eV/Å. Spin polarization is included following previous studies[19,21]. These high-throughput self-consistent and structural relaxation calculations are implemented utilizing the AiiDA computational framework[67].

Performance indicators

MAE of coordinate. The MAE of coordinates assesses the structural difference between the predicted and DFT-relaxed structures. It is defined as:

$$\Delta_{\text{coord}} = \frac{1}{N}\sum_{v_i\in\mathcal{G}} |\hat{\boldsymbol{r}}_i - \vec{\boldsymbol{r}}_i| \tag{25}$$

where N represents the total number of nodes in $\mathcal{G}$, $\hat{\boldsymbol{r}}_i$ and $\vec{\boldsymbol{r}}_i$ represent the predicted and ground truth Cartesian coordinates, respectively.

MAE of bond length. The MAE of bond length measures the error in predicting interatomic distances:

$$\Delta_{\text{bond}} = \frac{1}{M}\sum_{e_{ij,(k_1,k_2,k_3)}\in\mathcal{E}} |\hat{d}_{ij} - d_{ij}| \tag{26}$$

where M is the total number of chemical bonds, $\hat{d}_{ij}$ and d_{ij} are the predicted and ground interatomic distances.

MAE of lattice. This metric calculates the error in predicting the lattice matrices:

$$\Delta_{\text{lattice}} = \frac{1}{9}\sum |\hat{\boldsymbol{L}} - \boldsymbol{L}| \tag{27}$$

where $\hat{\boldsymbol{L}}$ and $\boldsymbol{L}$ are the predicted and ground lattice matrices.

MAE of cell volume. The error in predicting the cell volume is given by:

$$\Delta_{\text{volume}} = \left| |\hat{\boldsymbol{l}}_1 \cdot (\hat{\boldsymbol{l}}_2 \times \hat{\boldsymbol{l}}_3)| - |\vec{\boldsymbol{l}}_1 \cdot (\vec{\boldsymbol{l}}_2 \times \vec{\boldsymbol{l}}_3)| \right| \tag{28}$$

where $\times$ is the cross product and $\hat{\boldsymbol{l}}_i$ and $\vec{\boldsymbol{l}}_i$ are the predicted and ground truth lattice vectors.

Match rate. We utilize the "Structure_matcher" function from the Pymatgen package[3] to compare the predicted structure with the DFT-relaxed structure. Default parameters are used for this function (ltol = 0.2, stol = 0.3) to ensure consistent and objective comparisons.

Data availability

The dataset for X-Mn-O is available at https://zenodo.org/records/8081655(ref. 68). The dataset for Materials Project is available at https://figshare.com/articles/dataset/MPF_2021_2_8/19470599 (ref. 69). The dataset for C2DB is available at https://cmr.fysik.dtu.dk/c2db/c2db.html. The dataset for MoS2 structures with defects is available at https://research.constructor.tech/p/2d-defects-prediction. The layered vdW crystals dataset is in-house collected and currently unpublished; access can be obtained by contacting Dr. Lei Shen with reasonable requests. Source data and a Python script to reproduce the figures in this paper are provided. Source data are provided in this paper.

Code availability

Code for DeepRelax is available at https://github.com/Shen-Group/DeepRelaxand https://zenodo.org/records/13160937 (ref. 70).

References

1. Zuo, Y. et al. Accelerating materials discovery with bayesian optimization and graph deep learning. *Materials Today* **51**, 126–135 (2021).
2. Merchant, A. et al. Scaling deep learning for materials discovery. *Nature* **624**, 80–85 (2023).
3. Ong, S. P. et al. Python materials genomics (pymatgen): A robust, open-source Python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
4. Saal, J. E., Kirklin, S., Aykol, M., Meredig, B. & Wolverton, C. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (oqmd). *Jom* **65**, 1501–1509 (2013).
5. Curtarolo, S. et al. Aflow: An automatic framework for high-throughput materials discovery. *Comput. Mater. Sci.* **58**, 218–226 (2012).
6. Zhou, J. et al. 2dmatpedia, an open computational database of two-dimensional materials from top-down and bottom-up approaches. *Sci. Data* **6**, 86 (2019).
7. Chen, B. et al. Phase stability and superconductivity of lead hydrides at high pressure. *Phys. Rev. B* **103**, 035131 (2021).
8. Xie, T., Fu, X., Ganea, O.-E., Barzilay, R., Jaakkola, T. Crystal diffusion variational autoencoder for periodic material generation. In *International Conference on Learning Representations (ICLR)* (2022).
9. Zeni, C. et al. Mattergen: a generative model for inorganic materials design. Preprint at https://doi.org/10.48550/arXiv.2312.03687 (2023).
10. Zhao, Y. et al. Physics guided deep learning for generative design of crystal materials with symmetry constraints. *Npj Comput. Mater.* **9**, 38 (2023).
11. Chen, C. & Ong, S. P. A universal graph deep learning interatomic potential for the periodic table. *Nat. Comput. Sci.* **2**, 718–728 (2022).
12. Deng, B. et al. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nat. Mach. Intell.* **5**, 1031–1041 (2023).
13. Mosquera-Lois, I., Kavanagh, S. R., Ganose, A. M. & Walsh, A. Machine-learning structural reconstructions for accelerated point defect calculations. *Npj Comput. Mater.* **10**, 121 (2024).
14. Kolluru, A. et al. Open challenges in developing generalizable large-scale machine-learning models for catalyst discovery. *ACS Catal.* **12**, 8572–8581 (2022).
15. Kim, S., Noh, J., Jin, T., Lee, J. & Jung, Y. A structure translation model for crystal compounds. *Npj Comput. Mater.* **9**, 142 (2023).
16. Yoon, J. & Ulissi, Z. W. Differentiable optimization for the prediction of ground state structures (dogss). *Phys. Rev. Lett.* **125**, 173001 (2020).
17. Wang, Z. et al. Concurrent learning scheme for crystal structure prediction. *Phys. Rev. B* **109**, 094117 (2024).
18. Omee, S. S., Wei, L., Hu, M. & Hu, J. Crystal structure prediction using neural network potential and age-fitness pareto genetic algorithm. *J. Mater. Inform.* **4**, 2 (2024).
19. Kazeev, N. et al. Sparse representation for machine learning the properties of defects in 2d materials. *Npj Comput. Mater.* **9**, 113 (2023).
20. Mosquera-Lois, I., Kavanagh, S. R., Walsh, A. & Scanlon, D. O. Identifying the ground state structures of point defects in solids. *Npj Comput. Mater.* **9**, 25 (2023).
21. Huang, P. et al. Unveiling the complex structure-property correlation of defects in 2d materials based on high throughput datasets. *Npj 2D Mater. Appl.* **7**, 6 (2023).
22. Jiang, C., Marianetti, C. A., Khafizov, M. & Hurley, D. H. Machine learning potential assisted exploration of complex defect potential energy surfaces. *Npj Comput. Mater.* **10**, 21 (2024).
23. Belsky, A., Hellenbrandt, M., Karen, V. L. & Luksch, P. New developments in the inorganic crystal structure database (icsd): accessibility in support of materials research and design. *Acta Crystallogr. Sec. B Struct. Sci.* **58**, 364–369 (2002).
24. Jain, A. et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
25. Kim, S., Noh, J., Gu, G. H., Aspuru-Guzik, A. & Jung, Y. Generative adversarial networks for crystal structure prediction. *ACS Cent. Sci.* **6**, 1412–1420 (2020).
26. Haastrup, S. et al. The computational 2d materials database: high-throughput modeling and discovery of atomically thin crystals. *2D Mater.* **5**, 042002 (2018).
27. Gjerding, M. N. et al. Recent progress of the computational 2d materials database (c2db). *2D Mater.* **8**, 044002 (2021).
28. Lyngby, P. & Thygesen, K. S. Data-driven discovery of 2d materials by deep generative models. *Npj Comput. Mater.* **8**, 232 (2022).
29. Schütt, K., Unke, O., Gastegger, M. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International Conference on Machine Learning*, pp. 9377–9388 PMLR (2021).
30. Xie, T. & Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).

31. Noh, J. et al. Unveiling new stable manganese based photoanode materials via theoretical high-throughput screening and experiments. *Chem. Commun.* **55**, 13418–13421 (2019).
32. Satorras, V.G., Hoogeboom, E., Welling, M. E (n) equivariant graph neural networks. In *International Conference on Machine Learning*, pp. 9323–9332 PMLR (2021).
33. Zhang, X. et al. Efficient and accurate large library ligand docking with karmadock. *Nat. Comput. Sci.* **3**, 789–804 (2023).
34. Dong, T., Yang, Z., Zhou, J., Chen, C.Y.-C. Equivariant flexible modeling of the protein–ligand binding pose with geometric deep learning. *J. Chem. Theory Comput.* **19**, 8446–8459 (2023).
35. Li, H. et al. Deep-learning density functional theory hamiltonian for efficient ab initio electronic-structure calculation. *Nat. Comput. Sci.* **2**, 367–377 (2022).
36. Yu, J., Wang, D., Zheng, M. Uncertainty quantification: Can we trust artificial intelligence in drug discovery? *Iscience* **25**, 104814 (2022).
37. Tran, K. & Ulissi, Z. W. Active learning across intermetallics to guide discovery of electrocatalysts for co2 reduction and h2 evolution. *Nat. Catal.* **1**, 696–703 (2018).
38. Szymanski, N. J. et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* **624**, 86–91 (2023).
39. Li, Y. et al. Local environment interaction-based machine learning framework for predicting molecular adsorption energy. *J. Mater. Inform.* **4**, 4 (2024).
40. Musaelian, A. et al. Learning local equivariant representations for large-scale atomistic dynamics. *Nat. Commun.* **14**, 579 (2023).
41. Pablo-García, S. et al. Fast evaluation of the adsorption energy of organic molecules on metals via graph neural networks. *Nat. Comput. Sci.* **3**, 433–442 (2023).
42. Gong, X. et al. General framework for e (3)-equivariant neural network representation of density functional theory hamiltonian. *Nat. Commun.* **14**, 2848 (2023).
43. Zhong, Y., Yu, H., Su, M., Gong, X. & Xiang, H. Transferable equivariant graph neural networks for the hamiltonians of molecules and solids. *Npj Comput. Mater.* **9**, 182 (2023).
44. Zhong, Y. et al. Universal machine learning kohn-sham hamiltonian for materials. *Chinese Phys. Lett.* **41**, 077103 (2024).
45. Park, C. W. & Wolverton, C. Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery. *Phys. Rev. Mater.* **4**, 063801 (2020).
46. Schütt, K. et al. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in Neural Information Processing Systems* **30**, (2017).
47. Chen, C., Ye, W., Zuo, Y., Zheng, C. & Ong, S. P. Graph networks as a universal machine learning framework for molecules and crystals. *Chem. Mater.* **31**, 3564–3572 (2019).
48. Gasteiger, J., Groß, J., Günnemann, S.: Directional message passing for molecular graphs. In *International Conference on Learning Representations (ICLR)* (2020).
49. Choudhary, K. & DeCost, B. Atomistic line graph neural network for improved materials property predictions. *Npj Comput. Mater.* **7**, 185 (2021).
50. Unke, O. T. et al. Spookynet: Learning force fields with electronic degrees of freedom and nonlocal effects. *Nat. Commun.* **12**, 7273 (2021).
51. Batzner, S. et al. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453 (2022).
52. Banik, S. et al. Cegann: Crystal edge graph attention neural network for multiscale classification of materials environment. *Npj Comput. Mater.* **9**, 23 (2023).
53. Unke, O. T. & Meuwly, M. Physnet: A neural network for predicting energies, forces, dipole moments, and partial charges. *J. Chem. Theory Comput.* **15**, 3678–3693 (2019).
54. Zhang, X., Zhou, J., Lu, J. & Shen, L. Interpretable learning of voltage for electrode design of multivalent metal-ion batteries. *Npj Comput. Mater.* **8**, 175 (2022).
55. Omee, S.S. et al. Scalable deeper graph neural networks for high-performance materials property prediction. *Patterns* **3**, 100491 (2022).
56. Haghighatlari, M. et al. Newtonnet: A newtonian message passing network for deep learning of interatomic potentials and forces. *Digit. Discov.* **1**, 333–343 (2022).
57. Han, J. et al. A survey of geometric graph neural networks: Data structures, models and applications. Preprint at https://doi.org/10.48550/arXiv.2403.00485 (2024).
58. Yang, Z., Zhong, W., Lv, Q. & Chen, C. Y.-C. Learning size-adaptive molecular substructures for explainable drug–drug interaction prediction by substructure-aware graph neural network. *Chem. Sci.* **13**, 8693–8703 (2022).
59. Yang, Z., Zhong, W., Zhao, L. & Chen, C. Y.-C. Mgraphdta: deep multiscale graph neural network for explainable drug–target binding affinity prediction. *Chem. Sci.* **13**, 816–833 (2022).
60. Vaswani, A. et al. Attention is all you need. *Advances in Neural Information Processing Systems* **30**, (2017).
61. Liberti, L., Lavor, C., Maculan, N. & Mucherino, A. Euclidean distance geometry and applications. *SIAM Rev.* **56**, 3–69 (2014).
62. Lu, W. et al. Tankbind: Trigonometry-aware neural networks for drug-protein binding structure prediction. *Adv. Neural Inform. Process. Syst.* **35**, 7236–7249 (2022).
63. Masters, M. R., Mahmoud, A. H., Wei, Y. & Lill, M. A. Deep learning model for efficient protein–ligand docking with implicit side-chain flexibility. *J. Chem. Inform. Model.* **63**, 1695–1707 (2023).
64. Gawlikowski, J. et al. A survey of uncertainty in deep neural networks. *Artif. Intell. Rev.* **56**, 1513–1589 (2023).
65. Luo, Y., Liu, Y., Peng, J. Calibrated geometric deep learning improves kinase–drug binding predictions. *Nat. Mach. Intell.* **5**, 1390–1401 (2023).
66. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
67. Uhrin, M., Huber, S. P., Yu, J., Marzari, N. & Pizzi, G. Workflows in aiida: Engineering a high-throughput, event-based engine for robust and modular computational workflows. *Comput. Mater. Sci.* **187**, 110086 (2021).
68. Kim, S. A Structure Translation Model for Crystal Compounds Release for manuscript acceptance. Npj. Comput. Mater. **9**, 142 (2023).
69. Chen, C., Ong, S.P. MPF.2021.2.8. https://doi.org/10.6084/m9.figshare.19470599.v3 (2022).
70. Yang, Z. et al. Source Code for "Scalable Crystal Structure Relaxation Using an Iteration-Free Deep Generative Model with Uncertainty Quantification". https://doi.org/10.5281/zenodo.13160937 (2024).

Acknowledgements

We acknowledge Prof. Kristian Sommer Thygesen and Peder Lyngby for their generous provision of the C2DB database, complete with both initial and DFT-relaxed structures. This work was supported by the National Natural Science Foundation of China (Grant No. 62176272, to C.Y.-C.C.), Research and Development Program of Guangzhou Science and Technology Bureau (No. 2023B01J1016, to C.Y.-C.C.), Key-Area Research and Development Program of Guangdong Province (No. 2020B1111100001, to C.Y.-C.C.), Singapore MOE Tier 1 (No. A-8001194-00-00, to L.S.), and Singapore MOE Tier 2 (No. A-8001872-00-00, to L.S.).

Author contributions

Conceived and designed the research: L.S. and Z.Y. Conducted the experiments: Z.Y., Y.-M.Z., and X.W. Analyzed the data and interpreted the results: Z.Y., Y.-M.Z., X.W., X.L., X.Z., Y.L., Q.L., C.Y.-C.C., and L.S. Wrote the manuscript: Z.Y. and L.S. Edited and revised the manuscript: Z.Y., L.S., and C.Y.-C.C.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-52378-3.

Correspondence and requests for materials should be addressed to Calvin Yu-Chian Chen or Lei Shen.

Peer review information *Nature Communications* thanks Zhimei Sun, and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.