npj | computational materials

Article

Published in partnership with the Shanghai Institute of Ceramics of the Chinese Academy of Sciences

https://doi.org/10.1038/s41524-025-01592-8

Generative deep learning for predicting ultrahigh lattice thermal conductivity materials

Check for updates

Liben Guo[1,4], Yuanbin Liu[2,4], Zekun Chen[3], Hongao Yang[1], Davide Donadio[3] & Bingyang Cao[1]

Developing materials with ultrahigh thermal conductivity is crucial for thermal management and energy conversion. The recent development of generative models and machine learning (ML) holds great promise for predicting new functional materials. However, these data-driven methods are not tailored to identifying energetically stable structures and accurately predicting their thermal properties, as they lack physical constraints and information about the complexity of atomic many-body interactions. Here, we show how combining deep generative models of crystal structures with quantum-accurate, fast ML interatomic potentials can accelerate the prediction of materials with ultrahigh lattice thermal conductivity while ensuring energy optimality. We exploit structural symmetry and similarity metrics derived from atomic coordination environments to enable fast exploration of the structural space produced by the generative model. Additionally, we propose an active-learning-based protocol for the on-the-fly training of ML potentials to achieve high-fidelity predictions of stability and lattice thermal conductivity in prospective materials. Applying this method to carbon materials, we screen 100,000 candidates and identify 34 carbon polymorphs, approximately a quarter of which had not been previously predicted, to have lattice thermal conductivity above 800 W m^{-1} K^{-1}, reaching up to 2,400 W m^{-1} K^{-1} aside from diamond. These findings provide a viable pathway toward the ML-assisted prediction of periodic materials with exceptional thermal properties.

Developing new materials with high lattice thermal conductivity (κ_L) has long attracted academic and industrial interest due to their pivotal role in thermal management techniques for preventing electronic devices and batteries from overheating damage and performance degradation[1–3]. Over the past decades, the discovery and design of high thermal conductivity materials have largely relied on trial-and-error experiments (e.g., random element substitution)[4] within the guidance of empirical rules[5]. More recently, this progress has been significantly accelerated with the rapid development of computational techniques. In this endeavor, quantum-mechanical calculations based on density functional theory (DFT)[6,7] have played a central role. For example, established traditional methods for crystal structure prediction (CSP), such as ab initio random structure searching (AIRSS)[8], CALYPSO[9], and USPEX[10], generally introduce DFT free energy calculations and explicit constraints from prior knowledge to guide the optimization of structures with given chemical compositions. In addition, solving the linearized phonon Boltzmann equation using DFT calculations[11,12] has enabled the accurate prediction of κ_L in crystalline materials[13,14]. However, DFT-driven methods have largely restricted the exploration to a narrow set of candidate materials due to high computational demands[15,16]. So far, only a handful of inorganic crystals have been identified to exhibit ultrahigh κ_L—we here categorize them as having κ_L exceeding 800 W m^{-1} K^{-1} in any direction at room temperature[17,18]—including diamond[19–21], cubic boron arsenide (BAs)[22], cubic boron nitride (BN)[18], and θ-phase tantalum nitride (TaN)[23]. Arguably, several other materials with ultrahigh κ_L likely remain undiscovered.

With the introduction of advanced machine learning (ML) techniques in materials modeling, new approaches have emerged to accelerate the virtual screening of materials[24]. One such approach is the introduction of machine-learned interatomic potentials (MLIPs), which bridge the gap between computational cost and accuracy at the atomic scale. Extensive work has shown that MLIPs can achieve accuracy comparable to DFT while being several orders of magnitude faster[25–27]. Furthermore, MLIPs have proven to be capable of accurately calculating lattice thermal conductivity of both crystalline and disordered solids[28–33]. However, when inferring the

---

[1] Key Laboratory for Thermal Science and Power Engineering of Ministry of Education, Department of Engineering Mechanics, Tsinghua University, 100084 Beijing, China. [2] Inorganic Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford, OX1 3QR, UK. [3] Department of Chemistry, University of California, Davis, Davis, CA, 95616, USA. [4] These authors contributed equally: Liben Guo, Yuanbin Liu. e-mail: ddonadio@ucdavis.edu; caoby@tsinghua.edu.cn

structures and thermal properties of unknown materials, the uncertainty of MLIP models can be high, as their quality depends on the training dataset. Thus, constructing accurate and flexible MLIPs from small datasets remains a challenge.

Another such innovation is the development of generative models for fast CSP. Generative models offer a much more cost-efficient alternative to traditional methods by avoiding the high expense of DFT calculations[34]. Distinct from traditional methods that often need to assume the initial structures and elements, the principle of generative models is to learn the joint probability distribution of materials, from which new composition and structures are sampled and reconstructed. Therefore, generative models enable the rapid exploration of a broader range of potential materials without being limited by predefined assumptions. Furthermore, generative models, such as variational autoencoder (VAE)[35], can map input structures into a continuous vector space, namely latent space in the ML jargon. The differentiable nature of the latent space facilitates the global optimization of material properties through inverse design, typically achieved by training a surrogate model that directly maps the latent space to the desired material properties. Despite the promise of generative models in CSP, they face outstanding challenges in searching for materials with thermodynamical stability due to the lack of rigorous physical constraints[36]. Additionally, while several works have achieved inverse design by mapping properties to structures directly in the latent space or using a conditional approach (e.g., MatterGen)[37–41], building a surrogate model for accurately mapping κ_L from structures remains challenging. This difficulty arises from the complex many-body interactions involved in the structure-κ_L mapping and the temperature dependence of phonon scattering.

In this work, we introduce a unified generative deep learning framework for accelerating the prediction of three-dimensional crystalline materials with ultrahigh κ_L and thermodynamic stability. First, we leverage an SE(3)-equivariant generative model for the fast exploration of materials space. On top of this, we implement an active-learning-based protocol for training MLIPs on the fly, ensuring the effective use of a small amount of training data to achieve robust models that allow for accurate prediction regarding the stability and thermal properties of unseen materials. Furthermore, we propose structural symmetry and similarity metrics to enhance sampling efficiency and screening of the generated chemical space. These components together form an automated loop that accelerates the development of materials. Our method is validated on complex carbon materials, which serve as an ideal test case due to their abundant structural polymorphs and great potential as high κ_L materials (e.g., well-known diamond), resulting from strong covalent bonding and light atomic mass of carbon. We show that our method can not only reproduce experimentally observed carbon materials with high κ_L but also create new, previously unpredicted structures that have great potential for synthesis.

Results

A unified framework for exploring materials space

The overview of our approach is depicted in Fig. 1, which can be divided into three parts. The first part focuses on the data-driven generation of materials with two main objectives: reconstructing the structures in the training set and creating newly valid, diverse structures. We adopt an SE(3)-equivariant crystal diffusion variational autoencoder (CDVAE)[42] as the backbone of our generative model due to its capability of regulating the stability of generated material by incorporating the physical inductive bias—CDVAE uses a noise conditional score network to output gradients that drive atomic coordinates toward lower energy states. Following the framework of CDVAE, initial structures are defined by their elemental compositions $\{A\}$, atomic coordinates $\{x\}$, and lattice parameters $\{L\}$. The encoder maps the initial structures onto the learned joint probability distribution, referred to as the latent space $\{Z\}$. In this space, new structures are synthesized through random sampling. The decoder then handles these samples, determining their number of atoms, lattice parameters, and atomic positions using a diffusion process. These newly generated structures, together with the original structures in the training dataset, constitute the ensembles used for the prediction of ultrahigh κ_L materials. While CDVAE implements the physical stability biases, it remains an approximation and does not strictly adhere to quantum-mechanical theory or calculations. Therefore, this method is not able to guarantee that all generated materials will be thermodynamically stable in a broader chemical space. To address this issue, we introduce data distillation as the second part of our approach. First, we leverage the computational efficiency and accuracy of MLIPs to optimize the structures of all generated materials, ensuring that they are locally stable. To this end, pre-trained models, such as emerging foundation models or general-purpose models for specific elements, can be used out of the box as a starting point due to their robustness across a wide range of structures[43,44]. After optimization, all duplicate structures are removed through structural similarity analysis, leaving only one instance of each. Inspired by empirical rules suggesting that simple unit cells with high symmetry often result in high thermal conductivity[45], we here conducting an initial screening of materials using a structural symmetry metric that is defined based on the number of atoms (N) and symmetry operations (SO) in the unit cell. The selection criterion is that materials with a unit cell of $N \leq 12$ and $\mathrm{SO} \geq 4$ will be retained for the next round of screening, while other structures will be removed.

The final part involves a multistep identification and validation of ultrahigh κ_L materials. Although the symmetry constraints used in the second part can help reduce the search space of material phases, the remaining exploration space is still vast. To further narrow down this exploration space, we propose a two-step sampling method. First, we use the farthest points sampling (FPS) algorithm[46] to select m structures with the greatest diversity as benchmarks (Methods). We then apply a computational protocol for the accurate evaluation of lattice thermal conductivity (ETC) on these benchmarks to identify potential ultrahigh κ_L materials. Second, we cluster m distinct material groups based on their structural similarities with benchmarks using the k-nearest neighbors (KNN) algorithm (Methods). Groups containing benchmarks with $\kappa_L \geq 800$ W m^{-1} K^{-1} are identified as potential ultrahigh κ_L candidates, and all materials within these groups will be screened through the same ETC protocol to identify other materials with ultrahigh κ_L. During the ETC process, the MLIP model will be improved on the fly using a well-established active learning strategy known as "Query by Committee" (QBC)[47,48] before making the final κ_L prediction (Methods). This approach selectively expands the dataset when the uncertainty of the MLIP model surpasses a certain threshold, greatly reduces the cost of curating training data while preserving high-fidelity predictions, and proves to be effective in various physical and chemical applications[49–51]. Finally, the identification of collective ultrahigh κ_L materials is validated through quantum-mechanical calculations or experimental methods.

Material generation performance

We now move to showcase the performance of our method in the practical application to carbon materials. Carbon is known as one of the most challenging and intriguing targets for crystal structure prediction because its tendency to undergo *sp*, *sp*^{2}, and *sp*^{3} hybridization leads to structural complexity and diversity[52]. Recently, a dataset of carbon periodic structures, including 101,529 allotropes whose unit cells have 6–24 atoms, was generated with AIRSS at 10 GPa by Pickard[53]. All structures in this set reside at local energy minima, but no κ_L was provided for any of them. Here, we train the CDVAE model using the lowest-energy 10% of the original dataset, employing a random split of 60% for training, 20% for validation, and 20% for testing, consistent with the approach used in the original literature[42] and subsequently generated 100,000 initial structures at a speed of 0.48 seconds per structure on a single RTX 2080 Ti GPU, several orders of magnitude faster than the DFT-guided methods. To reveal how close the generated structures are to their low-energy states, they are all thoroughly optimized using a pre-trained equivariant message passing neural network potential (Methods), namely Allegro[27], chosen for its high computational efficiency and accuracy. The performance of the CDVAE model is quantified by analysing the structural changes before and after optimization.

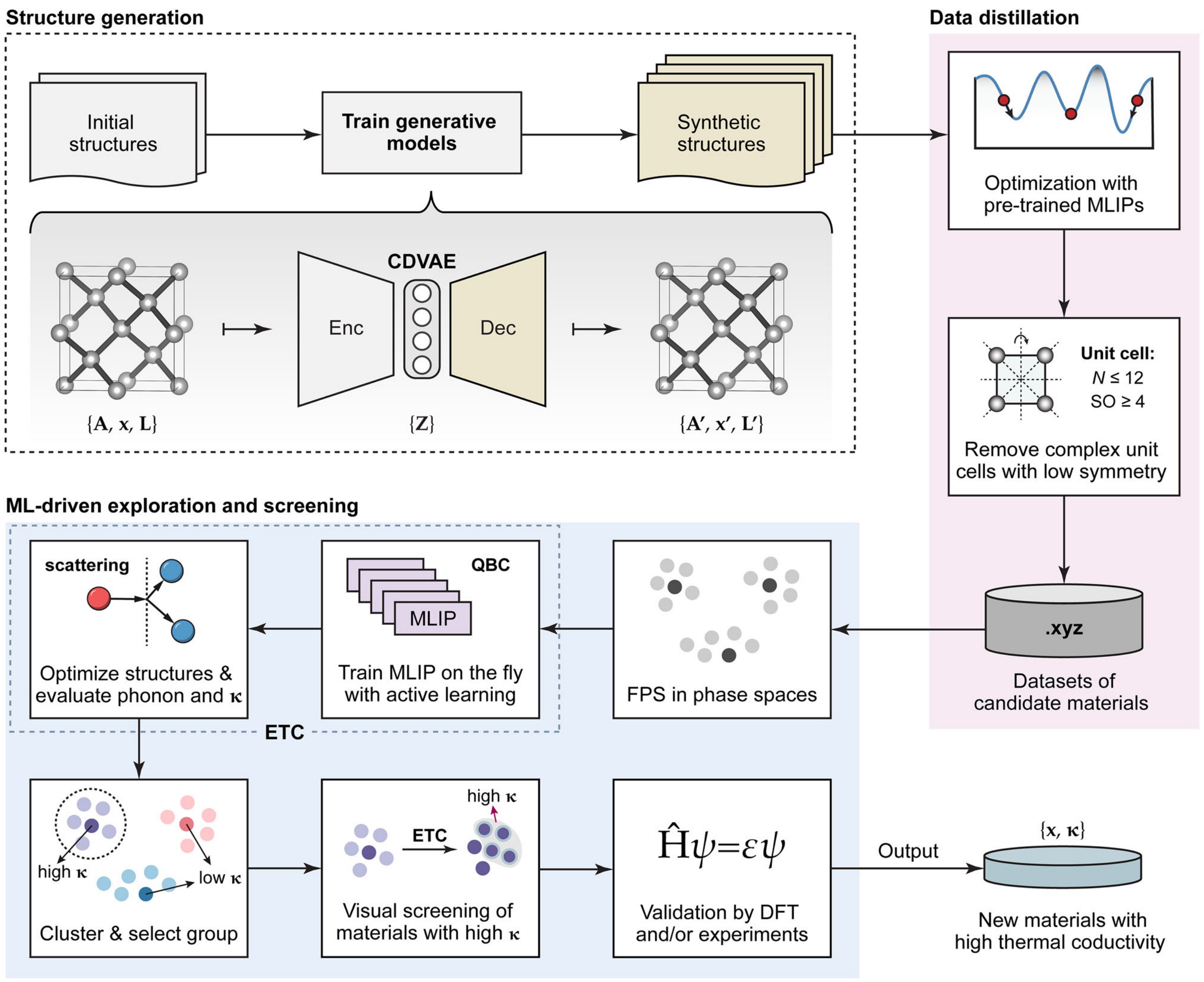

Fig. 1 | The overview of our approach. Starting with the initial crystalline structures (*gray*), an SE(3)-equivariant generative model, namely CDVAE, is trained to produce synthetic structures. The synthetic structures then undergo optimization with pre-trained MLIP models (*pink*). Complex unit cells with low symmetry are filtered out based on the empirical rule, leaving a refined dataset of candidate materials. Subsequently, we select the most diverse *m* structures using FPS algorithm as benchmarks (*blue*) and assess their κ_L via an ETC protocol (*dashed box*). Next, we cluster the candidate materials into *m* groups based on their structural similarity with benchmarks. All materials in those groups containing benchmarks with $\kappa_L \geq 800$ W m^{-1} K^{-1} are further screened with the same ETC protocol. During the ETC process, MLIP models will be improved on the fly using an active learning strategy, ensuring high-fidelity κ_L predictions. Finally, the identified potential materials with ultrahigh κ_L are validated by either DFT calculations or experimental measurements.

Figure 2a-c shows the distribution of relative deviations in lattice lengths, lattice angles, and atomic fractional coordinates between the generated and optimized structures, respectively. As indicated in Fig. 2d, a great number of the generated structures undergo evident changes after optimization, such as lattice twisting or tensile and compressive deformations. More specifically, the maximum relative deviation in lattice lengths can exceed 20%, with approximately half of them showing a relative error over 10%, while lattice angles exhibit a relative average deviation of around 5.6%. Such deviations arise from the lack of rigorous physical constraints in the CDVAE model. In terms of material property prediction, lattice strain has a large impact on the elastic moduli and acoustic group velocities, which in turn significantly influence κ_L. For instance, a uniaxial strain of +1% in diamond can result in an approximate −12% decrease in κ_L[54]. In this context, accurately predicting the equilibrium geometry and lattice parameters becomes particularly important, underscoring the necessity of using either DFT or an accurate MLIP model to optimize the generated structures. Interestingly, we observe higher symmetry in the optimized structures than in the generated ones.

Figure 2e further explores the uniqueness and novelty of generated structures after optimization. Uniqueness reflects the diversity of the generated structures, while novelty indicates that these structures are not present in the training dataset (refer to metrics for determining uniqueness and novelty in "Methods" section)[37,42]. As the number of the generated structures increases, the total number of unique and novel materials also grows, although their proportion decreases and approaches saturation. Specifically, a total of 7213 structures are found to be unique, accounting for approximately 7.2% of the total generated structures. We then refine the dataset using the established structural symmetry metric, reducing the pool to 1361 candidate structures, from which ultrahigh κ_L materials will be identified.

Exploration of ultrahigh thermal conductivity materials

The ML-driven exploration of ultrahigh κ_L materials, as outlined in Fig. 1, is conducted on the 1361 candidates. The visualization of the entire candidate set is illustrated in Fig. 3a mapped by Smooth Overlap of Atomic Positions (SOAP) descriptors[55]. It highlights some well-known carbon allotropes, such as diamond, graphite, and Bct-C12[56], which have been reproduced through the generative process. We first use the FPS algorithm to select the

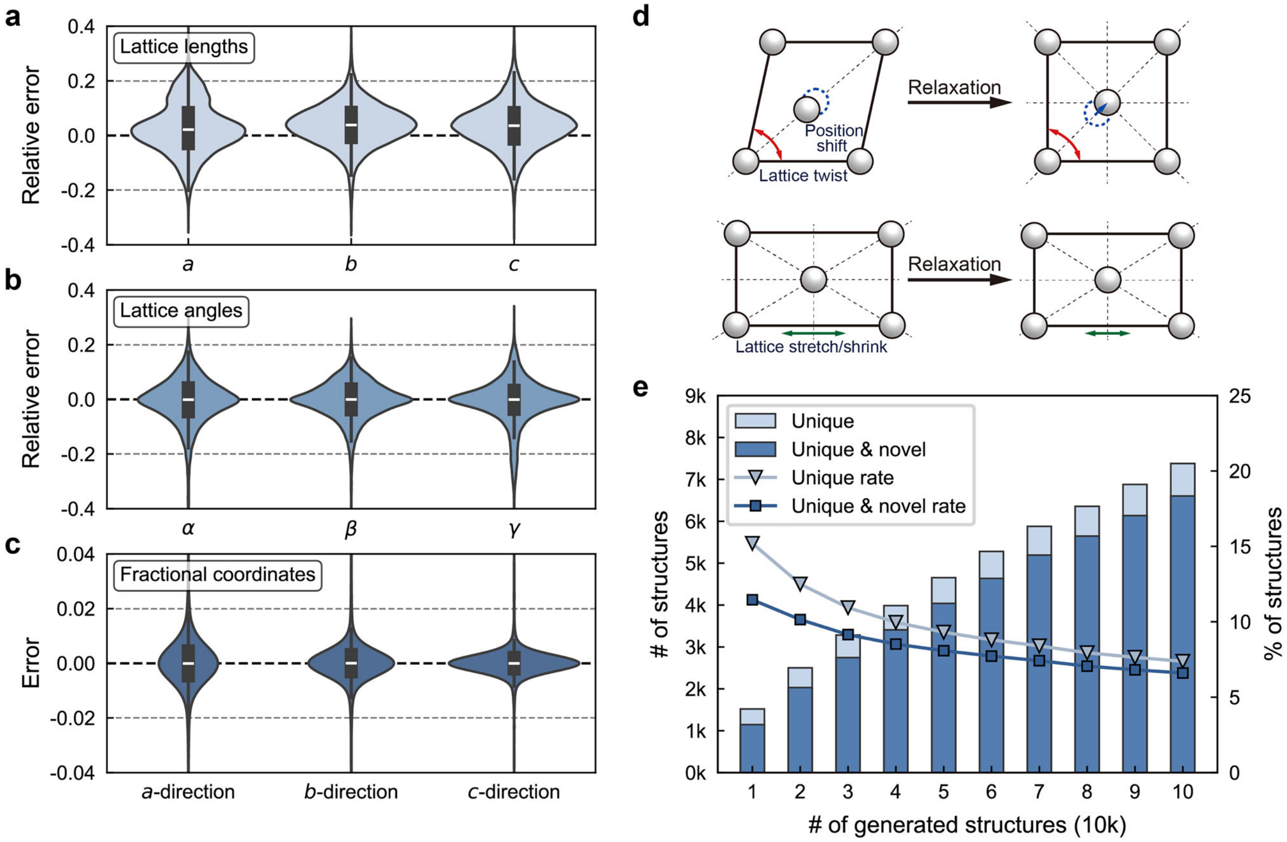

Fig. 2 | Performance evaluation of the generative model. Violin plots illustrating the distribution of relative deviations between the generated and optimized structures for (**a**) lattice lengths, **b** lattice angles, and (**c**) atomic fractional coordinates, respectively. These deviations demonstrate that the structural changes occur during the optimization process due to the lack of rigorous physical constraints in the generative model. **d** Schematic representation of structural relaxation, showing how optimization can lead to position shifts, lattice twists, and stretch/shrink deformations. **e** The generative performance of CDVAE quantified in terms of the number and rate of unique and novel structures. As the number of generated structures increases, the total number of unique and novel structures grows, but their proportion decreases and approaches saturation.

50 most diverse structures as benchmarks, as indicated by the green squares in Fig. 3a. In addition, five reported ultrahigh κ_L materials (listed in Table 1) are included as benchmarks to further enhance the search process. Note that by executing FPS, we identify 53 benchmarks, among which diamond and AA T12-carbon[57] are included.

The κ_L evaluation of the benchmarks is carried out using the high-fidelity Allegro MLIPs that have been evolved and trained through the active learning protocol based on the uncertainty estimation outlined in the Methods section. Figure 3b depicts the evolution of energy uncertainty for each benchmark configuration. In the beginning, ten configurations exceed the empirical uncertainty threshold of 15 meV atom^{-1}, with the highest reaching 35.6 meV atom^{-1}. As DFT reference configurations with the highest uncertainties are added iteratively to the training set, the overall uncertainty decreases. By the final iteration, all uncertainties have converged below 15 meV atom^{-1}, with their distribution shifting toward lower uncertainty. Lastly, nine benchmark structures are confirmed to exhibit a κ_L exceeding 800 W m^{-1} K^{-1}, and these are highlighted with red borders, among which diamond and AA T12-carbon are included.

We then use KNN classification to select 10 nearest neighbors for each ultrahigh κ_L benchmark. However, to minimize the overlap among the selected clusters, a total of 64 unique structures are ultimately chosen, represented as purple circles in Fig. 3a. A similar ETC process based on active learning is executed again for the KNN samples (Fig. 3c). It is noted that only for one structure the MLIPs have an uncertainty slightly exceeding the threshold of 15 meV atom^{-1}. Figure 3d shows the distribution of κ_L values for the entire set of candidate materials identified through our workflow. In the end, 34 structures (see supplementary materials) are found to have a κ_L greater than 800 W m^{-1} K^{-1}, among which five structures are identified as reproduction of known benchmarks with ultrahigh κ_L listed in Table 1. Over 50% of the total KNN candidate materials exhibit ultrahigh κ_L, suggesting that a greater structural similarity to high-κ_L materials increases the likelihood of exhibiting high κ_L. For example, structures in the vicinity of diamond mostly show ultrahigh κ_L, while those near BCO-C16[21,58] with a κ_L of 462 W m^{-1} K^{-1}, consistently have κ_L below 800 W m^{-1} K^{-1} (see Table S2). Therefore, screening high κ_L materials with structural similarity metrics can be considered as an effective approach.

DFT validation and phonon properties

To validate our findings, we select six representative carbon allotropes and conduct DFT calculations (see Methods) to determine their respective ground-state energy and κ_L (Fig. 4 and Table S3). These materials include three diamond polytypes[59,60] and three other candidates that are structurally distinct from diamond. The three diamond polytypes consist of alternating layers of cubic diamond and hexagonal diamond, while the other three candidates are tubulanes with 4- and 8-membered rings based on cross-linked carbon nanotubes[61,62]. The results show that the κ_L predicted by our MLIP are in close agreement with those calculated by DFT, confirming the reliability and accuracy of our method. To gain insights into the thermodynamic stability of the six carbon allotropes, we calculated their positions on the energy convex hull leveraging the active-learned Allegro potential introduced before. It indicates that all six candidates lie within 0.5 eV atom^{-1}—a value below the thermodynamic upper limit of 0.933 eV atom^{-1}, which is used to identify metastable materials as detailed in ref. 63—or less from the convex hull. Additionally, 21 other structures, not shown here, have an energy-above-hull of less than 0.25 eV atom^{-1}, and 8 are within 0.1 eV atom^{-1} (Fig. S2). These two thresholds, commonly referenced in the

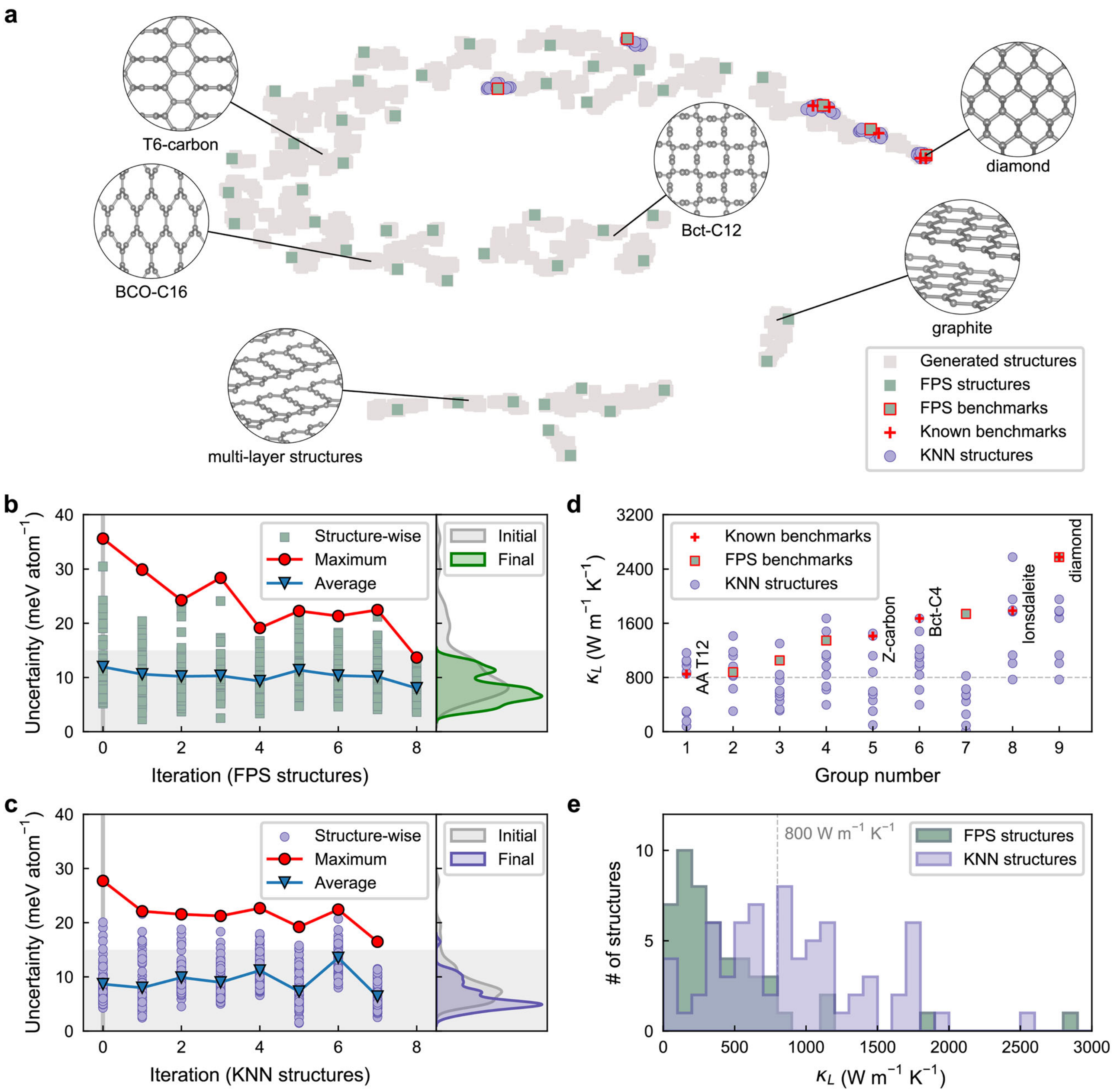

Fig. 3 | Machine learning-driven exploration of ultrahigh κ_L materials. **a** Visualization of the 1361 candidate materials by t-SNE, highlighting well-known carbon allotropes such as diamond, graphite, and Bct-C12 reproduced through the generative process. Energy uncertainty evolution of (**b**) FPS- and (**c**) KNN-selected benchmarks over multiple active learning iterations. **d** Distribution of κ_L values for the entire set of selected candidate materials. **e** Histogram of κ_L values for FPS- and KNN-selected structures. The fact that over 50% of KNN-selected materials exhibit κ_L greater than 800 W m^{-1} K^{-1} underscores that structural similarity to known ultrahigh κ_L materials being a strong indicator of ultrahigh κ_L potential.

literature[36,37,64,65], suggest that these structures have promising potential for experimental synthesis.

Figure 5 shows the phonon dispersion curves of the six representative carbon allotropes, comparing the results from DFT with those predicted by our MLIP model. The absence of imaginary frequencies in the phonon dispersion curves indicates the dynamical stability of these structures. Notably, the phonon branches in these predicted materials are more complex than those of diamond. This increased complexity enhances phonon-phonon scattering, contributing to the fact that, although these materials exhibit high lattice thermal conductivities, they are still lower than that of diamond. Additionally, the bandwidths of the three carbon tubulanes shown in Fig. 5d-f are significantly wider compared with those of diamond polytypes (a to c), which results in a higher rate of phonon scattering especially within a range of 15–40 THz (Fig. S3 in supplementary materials). This further explains that phonons above 15 THz in these three tubulanes contribute negligible thermal transport, leading their κ_L being obviously lower than those of diamond polytypes. Nonetheless, their κ_L remain exceeding 800 W m^{-1} K^{-1}.

Discussion

Generative models are becoming increasingly popular in the field of materials prediction due to their high efficiency in exploring vast material spaces, yet they face significant challenges in identifying local energy minima. As shown in our analysis, nearly half of the generated structures experience large lattice strain (over 10%) and distortion, highlighting the necessity of further structural optimization to ensure local structural stability. Here, we propose a unified framework to improve the accuracy and speed of identifying potential ultrahigh κ_L materials, by integrating deep generative models with active-learned MLIPs, ensuring the thermodynamic stability of generated materials. The use of structural symmetry metrics streamlines the identification of promising materials by filtering out low-symmetry or complex structures. In addition, the introduction of structural similarity metrics allows for a more targeted selection process, reducing the computational load by focusing on materials that share structural characteristics with known high-performance materials, thereby increasing the likelihood of predicting ultrahigh κ_L candidates. This prioritizes inverse design methods when targeting novel materials with no or little prior property knowledge. Our method is applied to carbon materials, leading to the prediction of 34 carbon polymorphs with ultrahigh κ_L (exceeding 800 W m^{-1} K^{-1} in any direction), with 33 out of 34 carbon polymorphs are confirmed to be insulators or semiconductors (see Fig. S4). Among these, five materials with known ultrahigh κ_L (such as diamond and lonsdaleite) are reproduced, and eight materials are entirely new in both structure and thermal properties. The remaining 21 materials have had their structures previously reported[66], but their ultrahigh κ_L properties had not been studied or documented before. This demonstrates the dual capability of our approach in validating existing materials and predicting entirely new candidates with ultrahigh κ_L. Remarkably, the identified materials displayed lattice thermal conductivity values reaching up to 2400 W m^{-1} K^{-1}, making them the highest known thermal conductors aside from diamond.

Our proposed framework offers an efficient solution not limited to carbon materials but fully scalable to multi-component systems. With the continuous advancement of high-accuracy, universal MLIP foundation models[43,44,67,68], this approach holds great promise for unlocking new high-performance materials across a broad range of applications.

Table 1 | Five reported carbon materials with ultrahigh κ_L

Benchmark name	κ_L (reported)	κ_L (this work)
Diamond	2000–3500 (Expt.)[19,20]	2573
Lonsdaleite	1687 (BTE)[21]	1785
Bct-C4	1411 (BTE)[21]	1672
Z-carbon	1262 (BTE)[21]	1412
AA T12-carbon	1049 (BTE)[57]	852

The entries are obtained either from experimental measurements or theoretical computation by BTE. (Unit of κ_L: W m^{-1} K^{-1}).

Methods

Structural analysis on synthetic materials

The metrics of uniqueness and novelty are defined based on the *StructureMatcher* utility from the Python library *pymatgen*[69] which can distinguish whether two crystals have the same structure. When synthesizing structures using generative model, some of the them may have the same structure. We remove all the duplicates and leave only one instance of each. These leaving structures are defined to be unique from each other. As for novelty, a synthesized structure is considered novel if it cannot be matched with any structures from the training dataset of the generative model. The *StructureMatcher* utility follows the default parameters: *ltol* = 0.2, *stol* = 0.3, *angle_tol* = 5, which collectively give the tolerances on cell and site positions when comparing two structures.

The determination of the primitive cell and space group are based on the *SpacegroupAnalyzer* utility from *pymatgen*, using parameters of

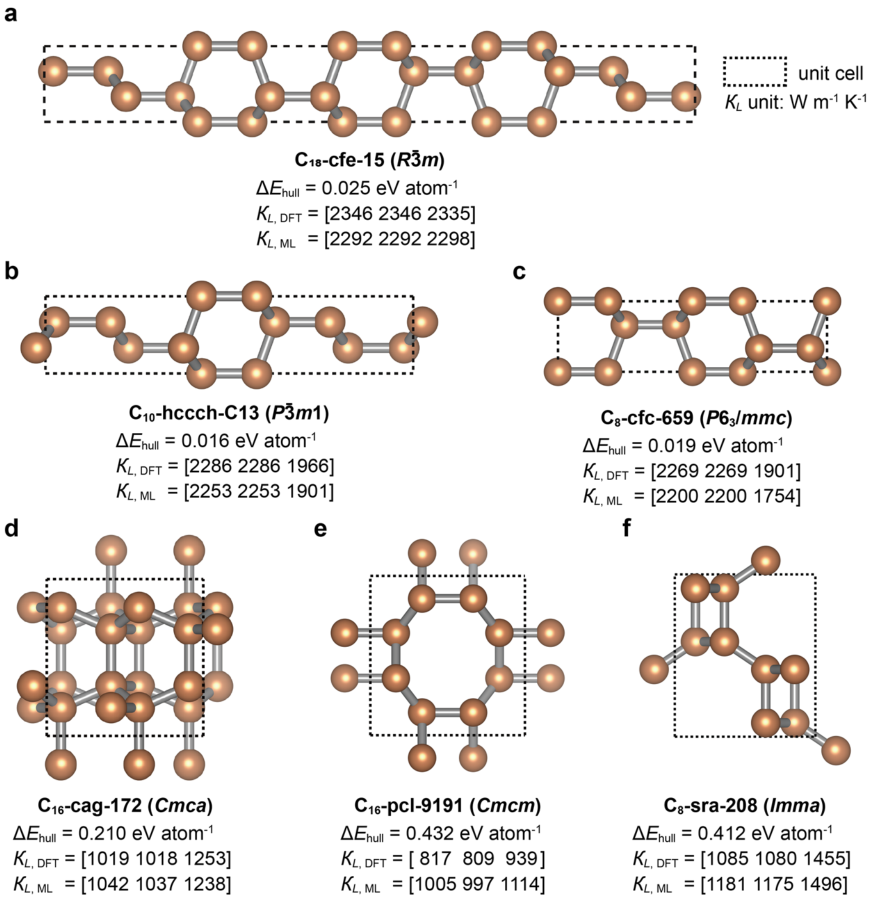

Fig. 4 | Crystal structures and thermodynamic properties of the six predicted carbon allotropes with ultrahigh κ_L. **a** 18-atom trigonal structure (6 atoms in primitive cell). **b** 10-atom trigonal structure. **c** 8-atom hexagonal structure. **d** 16-atom orthorhombic carbon tubulanes (8 atoms in primitive cell), having space group *Cmca* with a glide plane *a* (the last letter in *Cmca*) comparing with **e** 16-atom orthorhombic carbon tubulanes (8 atoms in primitive cell), having space group *Cmcm*. **f** 8-atom orthorhombic carbon tubulanes (4 atoms in primitive cell). The top three structures (**a**–**c**) are diamond polytypes, while the next three (**d**–**f**) represent tubulanes with 4 + 8 membered rings.

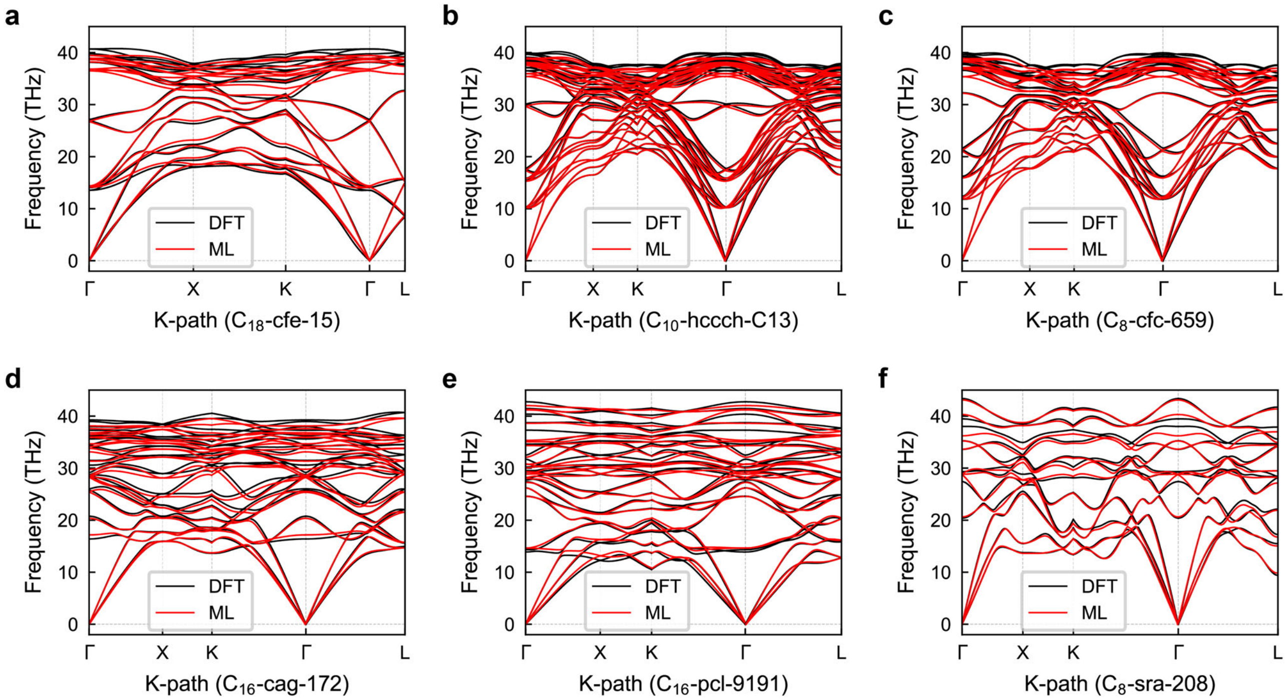

Fig. 5 | Phonon dispersion curves for six representative carbon allotropes, comparing the results from DFT (black lines) and our "Allegro" MLIP (red lines). **a–c** Three diamond polytypes with highest κ_L. **d–f** Other three representative carbon tubulanes with high κ_L. The close alignment between the DFT- and MLIP-predicted curves indicates the accuracy of our model in reproducing the phonon dispersion properties of these materials.

symprec = 0.1 and *angle_tolerance* = 1.0 which give the tolerances of symmetry coordinates and angle respectively.

Global visualization of synthetic materials

The global overview of the structural distribution is mapped by SOAP descriptor, a high-dimensional vector that captures local structural information around each atom in a material. It is constructed by expanding the atomic density around each atom using a set of orthogonal basis functions, typically spherical harmonics and radial basis functions. This results in a rotationally and translationally invariant representation of the local atomic environment, which is crucial for accurately describing the structural properties of materials. We use an atomic environment cutoff radius of 5.0 Å, Gaussian expansion smoothness of 0.5 Å and spherical harmonic functions expanded to $n_{\max} = 6$, $l_{\max} = 6$. While SOAP is generally employed to distinguish the local environments of individual atoms, here we use an averaged SOAP descriptor over all atoms in the cells[70,71]. These structures descriptors are then decomposed to two-dimension for similarity analysis and visualization by t-SNE algorithm[72], where the perplexity parameter is set to 10.

The FPS algorithm is then used to sample diverse points in this embedding space. It works by iteratively selecting points that are farthest from the already chosen points, ensuring that the sampled points are spread out across the entire space. This diversity in sampling helps in exploring the search space more effectively, leading to the identification of a wide range of potential candidates with varying properties.

Pre-trained Allegro potential

The aim of the pre-trained MLIP is to model a generalized PES covering enormous variety of structures with diverse environments of *sp*, sp^2, and sp^3. To achieve this, its training dataset ought to include representative yet diverse configurations of possible structures. Here we curate the dataset designated as GAP-2020 which is a subset derived from ref. 73. It encompasses a variety of configuration types including RSS crystalline, normal bulk crystalline, Samara Carbon Allotrope Database (SACADA)[66] crystal, diamond, graphite, nanotubes, fullerenes and amorphous forms. Here, we select structures from GAP-2020 where the energy is less than −4.5 eV atom^{-1} and the magnitude of force components is below 30 eV Å^{-1} to form our training dataset. The dataset consists of 5445 configurations, which have been stratified based on formation energy and allocated into training, validation, and testing sets in a default 60-20-20 ratio based on energy.

The pre-trained Allegro model has three layers and $l_{\max} = 3$ with 16 features for both even and odd irreps, summing up to 232,776 parameters in total. The 2-body latent multi-layer perceptron (MLP) consists of 3 layers of dimensions [128, 128, 128], using SiLU nonlinearities[74]. The later latent MLPs consist of a single layer of dimensionality 128, also using SiLU nonlinearities. The embedding weight projection was implemented as a single matrix multiplication without a hidden layer or a nonlinearity. The final edge energy multi-layer perceptron has 3 hidden layers of dimensions [256, 128, 64] and again SiLU nonlinearities. The model is initialized according to a uniform distribution. We use a radial cutoff of 5.5 Å. We use a basis of 8 non-trainable Bessel functions for the basic encoding with the polynomial envelope function using $p = 48$. Models are trained using a mean-square-error (MSE) loss on the energy with the Adam optimizer[75,76] in *PyTorch*[77], with default parameters of $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-8}$ without weight decay.

The accuracy of the pre-trained Allegro potential has a mean-average-error (MAE) of 24.3 meV atom^{-1} of energy on testing set and a root-mean-square-error (RMSE) of 273 meV Å^{-1} of force on testing set. The specific prediction errors of each configuration types are detailed in Fig. S1 and Table S1 in supplementary materials.

QbC based active learning protocol for high-fidelity MLIPs

In order to train high-fidelity MLIPs for stability determination and κ_L prediction, the training datasets are built through iterative processes based on Query by Committee technique[47,48]. Since small ensembles may under-represent the hypothesis space, and the required memory and computational effort increase linearly with the number of ensemble members[78], we chose an ensemble size of 8 to balance this trade-off relation, which is also a recommended size by Shang and Liu[79]. Starting with an initial dataset, a committee of 8 MLIPs is trained on 8 respective subsets slightly different from each other, while in each case a small fraction of data is intentionally left out. Due to the difference of training datasets and stochastic initiation of model weights, a diverse committee of MILPs is provided. Given a structure to have its κ_L evaluated, this committee give different predictions of its potential energy, and the standard deviation of them is defined as the structural uncertainty.

For each iteration, the structure with highest uncertainty will be rattled 8 times and have their DFT reference data calculated to be added to dataset. This protocol will stop if all structural uncertainties are below the threshold of 15 meV atom^{-1}. Noted that the configurations of each structure will only be added to the dataset once, which may also call the stop of the active learning progress. Ultimately, all configurations are used to train a final MLIP.

To be more specific, the initial dataset encompasses 1000 configurations chosen from GAP-2020 utilizing FPS on the feature space defined by their respective SOAP descriptors. This dataset is progressively enriched with configurations of structures identified to have high uncertainty marked by DFT calculations during the ETC process. The committee and final MLIPs are all Allegro potentials with the same hyperparameters as the pretrained Allegro model. As for configurations of DFT calculation, they are detailed in the follow part.

Configurations of DFT calculation

The configurations of DFT calculation for MLIPs dataset construction and κ_L validation are the same, both implemented in the Vienna Ab initio Simulation Package (VASP). We perform spin-polarized calculations with the optB88-vdW dispersion inclusive exchange–correlation functional[80,81], with an energy cutoff of 600 eV and a dense k-spacing of 0.15 Å^{-1}.

Lattice thermal conductivity calculations

In this work the intrinsic κ_L is calculated by solving the phonon Boltzmann transport equation (BTE) in the framework of anharmonic lattice dynamics[82,83], with interatomic force constants (IFCs) derived from either MLIPs or DFT. Although the relaxation time approximation (RTA) method is known to underestimate κ_L compared to the direct inversion solution of BTE[84–86], when exploring ultrahigh κ_L materials using MLIPs, the RTA method is used for computational efficiency. The difference between the two methods turns out to have minimal impact on our final outcomes (Fig. S5). For DFT validation, κ_L are obtained by full converged solution of BTE. Only the lowest order of anharmonic scattering, namely three-phonon scattering, is considered here for computational efficiency. The harmonic and anharmonic IFCs are calculated under $3 \times 3 \times 3$ supercells by finite displacement difference approach with a cutoff radius of 3 Å utilizing the *Phonopy* and *Phono3py* software[87,88] and the mesh numbers multiply the respective lattice lengths are set to exceed 25. We conduct two convergence tests: one for the cutoff radius (Fig. S6) and one for the mesh sizes (Fig. S7). The good convergence of both tests verifies that our choice of these two parameters is reasonable.

Data availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Received: 17 December 2024; Accepted: 28 March 2025;

References

1. Zhao, Q., Stalin, S., Zhao, C.-Z. & Archer, L. A. Designing solid-state electrolytes for safe, energy-dense batteries. *Nat. Rev. Mater.* **5**, 229–252 (2020).
2. Yu, R. et al. Ultrawide-bandgap semiconductor AlN crystals: Growth and applications. *J. Mater. Chem. C.* **9**, 1852–1873 (2021).
3. Black, J. R. Electromigration—A brief survey and some recent results. *IEEE Trans. Electron Devices* **16**, 338–347 (1969).
4. Wu, H., Fan, H. & Hu, Y. Ab initio determination of ultrahigh thermal conductivity in ternary compounds. *Phys. Rev. B* **103**, L041203 (2021).
5. Slack, G. A. Nonmetallic crystals with high thermal conductivity. *J. Phys. Chem. Solids* **34**, 321–335 (1973).
6. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
7. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
8. Pickard, C. J. & Needs, R. J. Ab initio random structure searching. *J. Phys.: Condens. Matter* **23**, 053201 (2011).
9. Wang, Y., Lv, J., Zhu, L. & Ma, Y. CALYPSO: A method for crystal structure prediction. *Comput. Phys. Comm.* **183**, 2063–2070 (2012).
10. Glass, C. W., Oganov, A. R. & Hansen, N. USPEX—Evolutionary crystal structure prediction. *Comput. Phys. Comm.* **175**, 713–720 (2006).
11. Baroni, S., Giannozzi, P. & Testa, A. Green's-function approach to linear response in solids. *Phys. Rev. Lett.* **58**, 1861–1864 (1987).
12. Gonze, X. Perturbation expansion of variational principles at arbitrary order. *Phys. Rev. A* **52**, 1086–1095 (1995).
13. Turney, J. E., Landry, E. S., McGaughey, A. J. H. & Amon, C. H. Predicting phonon properties and thermal conductivity from anharmonic lattice dynamics calculations and molecular dynamics simulations. *Phys. Rev. B* **79**, 064301 (2009).
14. Broido, D. A., Malorny, M., Birner, G., Mingo, N. & Stewart, D. A. Intrinsic lattice thermal conductivity of semiconductors from first principles. *Appl. Phys. Lett.* **91**, 231922 (2007).
15. Zhang, L., Wang, Y., Lv, J. & Ma, Y. Materials discovery at high pressures. *Nat. Rev. Mater.* **2**, 17005 (2017).
16. Oganov, A. R., Pickard, C. J., Zhu, Q. & Needs, R. J. Structure prediction drives materials discovery. *Nat. Rev. Mater.* **4**, 331–348 (2019).
17. Dames, C. Ultrahigh thermal conductivity confirmed in boron arsenide. *Science* **361**, 549–550 (2018).
18. Chen, K. et al. Ultrahigh thermal conductivity in isotope-enriched cubic boron nitride. *Science* **367**, 555–559 (2020).
19. Wei, L., Kuo, P. K., Thomas, R. L., Anthony, T. R. & Banholzer, W. F. Thermal conductivity of isotopically modified single crystal diamond. *Phys. Rev. Lett.* **70**, 3764–3767 (1993).
20. Inyushkin, A. V. et al. Thermal conductivity of high purity synthetic single crystal diamonds. *Phys. Rev. B* **97**, 144305 (2018).
21. Hu, M. & Wang, Q. Ultrahigh thermal conductivity of carbon allotropes with correlations with the scaled Pugh ratio. *J. Mater. Chem. A* **7**, 6259–6266 (2019).
22. Lindsay, L., Broido, D. A. & Reinecke, T. L. First-principles determination of ultrahigh thermal conductivity of boron arsenide: A competitor for diamond? *Phys. Rev. Lett.* **111**, 025901 (2013).
23. Kundu, A. et al. Ultrahigh thermal conductivity of θ-phase tantalum nitride. *Phys. Rev. Lett.* **126**, 115901 (2021).
24. Zuo, Y. et al. Accelerating materials discovery with Bayesian optimization and graph deep learning. *Mater. Today* **51**, 126–135 (2021).
25. Friederich, P., Häse, F., Proppe, J. & Aspuru-Guzik, A. Machine-learned potentials for next-generation matter simulations. *Nat. Mater.* **20**, 750–761 (2021).
26. Lysogorskiy, Y. et al. Performant implementation of the atomic cluster expansion (PACE) and application to copper and silicon. *npj Comput. Mater.* **7**, 97 (2021).
27. Musaelian, A. et al. Learning local equivariant representations for large-scale atomistic dynamics. *Nat. Commun.* **14**, 579 (2023).
28. Sosso, G. C., Donadio, D., Caravati, S., Behler, J. & Bernasconi, M. Thermal transport in phase-change materials from atomistic simulations. *Phys. Rev. B* **86**, 104301 (2012).
29. Campi, D., Donadio, D., Sosso, G. C., Behler, J. & Bernasconi, M. Electron-phonon interaction and thermal boundary resistance at the crystal-amorphous interface of the phase change compound GeTe. *J. Appl. Phys.* **117**, 015304 (2015).

30. Mangold, C. et al. Transferability of neural network potentials for varying stoichiometry: Phonons and thermal conductivity of Mn_xGe_y compounds. *J. Appl. Phys*. **127**, 244901 (2020).
31. Fan, Z., Ding, J. & Ma, E. Machine learning bridges local static structure with multiple properties in metallic glasses. *Mater. Today* **40**, 48–62 (2020).
32. Liu, Y.-B. et al. Machine learning interatomic potential developed for molecular simulations on thermal properties of β-Ga2O3. *J. Chem. Phys.* **153**, 144501 (2020).
33. Yang, G., Liu, Y.-B., Yang, L. & Cao, B.-Y. Machine-learned atomic cluster expansion potentials for fast and quantum-accurate thermal simulations of wurtzite AlN. *J. Appl. Phys.* **135**, 085105 (2024).
34. 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).
35. Kingma, D. P. & Welling, M. Auto-encoding variational bayes. In *Int. Conf. Learn. Represent*. (ICLR, 2014).
36. Zhao, Y. et al. Physics guided deep learning for generative design of crystal materials with symmetry constraints. *npj Comput. Mater.* **9**, 38 (2023).
37. Zeni, C. et al. A generative model for inorganic materials design. *Nature*. https://doi.org/10.1038/s41586-025-08628-5 (2025)
38. Antunes, L. M., Butler, K. T. & Grau-Crespo, R. Crystal structure generation with autoregressive large language modeling. *Nat. Commun.* **15**, 10570 (2024).
39. Noh, J. et al. Inverse design of solid-state materials via a continuous representation. *Matter* **1**, 1370–1384 (2019).
40. Ren, Z. et al. An invertible crystallographic representation for general inverse design of inorganic crystals with targeted properties. *Matter* **5**, 314–335 (2022).
41. Chen, L., Zhang, W., Nie, Z., Li, S. & Pan, F. Generative models for inverse design of inorganic solid materials. *J. Mater. Inf.* **1**, 4 (2021).
42. Xie, T., Fu, X., Ganea, O.-E., Barzilay, R. & Jaakkola, T. S. Crystal diffusion variational autoencoder for periodic material generation. In *Int. Conf. Learn. Represent*. (ICLR, 2022).
43. Chen, C. & Ong, S. P. A universal graph deep learning interatomic potential for the periodic table. *Nat. Comput. Sci.* **2**, 718–728 (2022).
44. Deng, B. et al. CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nat. Mach. Intell.* **5**, 1031–1041 (2023).
45. Liu, C. et al. An efficient strategy for searching high lattice thermal conductivity materials. *ACS Appl. Energy Mater.* **5**, 15356–15364 (2022).
46. Eldar, Y., Lindenbaum, M., Porat, M. & Zeevi, Y. Y. The farthest point strategy for progressive image sampling. *IEEE Trans. Image Process.* **6**, 1305–1315 (1997).
47. Krogh, A. & Vedelsby, J. Neural network ensembles, cross validation, and active learning. *Adv. Neural Inf. Process. Syst.* **7**, 231–238 (1995).
48. Seung, H. S., Opper, M. & Sompolinsky, H. Query by committee. *Adv. Neural Inf. Process. Syst.* **7**, 287–294 (1992).
49. Schran, C. et al. Machine learning potentials for complex aqueous systems made simple. *Proc. Natl. Acad. Sci. USA* **118**, e2110077118 (2021).
50. Schran, C., Brezina, K. & Marsalek, O. Committee neural network potentials control generalization errors and enable active learning. *J. Chem. Phys.* **153**, 104105 (2020).
51. Sonti, S. et al. Stability and dynamics of zeolite-confined gold nanoclusters. *J. Chem. Theory Comput.* **20**, 8261–8269 (2024).
52. Hirsch, A. The era of carbon allotropes. *Nat. Mater.* **9**, 868–871 (2010).
53. C. J. Pickard, AIRSS data for carbon at 10GPa and the C+N+H+O system at 1GPa, Materials Cloud; https://doi.org/10.24435/MATERIALSCLOUD:2020.0026/V1 (2020)
54. Wang, B., He, Y., Rodionov, N. & Zhu, J. First-principles predictions of thermal conductivity of bulk diamond under isotropic and uniaxial (100) strains. *Mater. Today Phys.* **36**, 101182 (2023).
55. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).
56. Dong, X., Hu, M., He, J., Tian, Y. & Wang, H.-T. A new phase from compression of carbon nanotubes with anisotropic Dirac fermions. *Sci. Rep.* **5**, 10713 (2015).
57. Liu, Z. et al. Bond saturation significantly enhances thermal energy transport in two-dimensional pentagonal materials. *Nano Energy* **45**, 1–9 (2018).
58. Yue, S.-Y. et al. Thermal transport in novel carbon allotropes with sp^2 or sp^3 hybridization: An ab initio study. *Phys. Rev. B* **95**, 085207 (2017).
59. Wen, B., Zhao, J., Bucknum, M. J., Yao, P. & Li, T. First-principles studies of diamond polytypes. *Diam. Relat. Mater.* **17**, 356–364 (2008).
60. Zheng, W., Liu, Q.-J., Liu, Z.-T. & Zhang, Z.-Q. First-principles calculations of structural, electronic and elastic properties of carbon allotropes. *Mater. Sci. Semicond. Process.* **146**, 106692 (2022).
61. Greshnyakov, V. A. & Belenkov, E. A. Structures of diamond-like phases. *J. Exp. Theor. Phys.* **113**, 86–95 (2011).
62. Zhu, Q., Zeng, Q. & Oganov, A. R. Systematic search for low-enthalpy sp^3 carbon allotropes using evolutionary metadynamics. *Phys. Rev. B* **85**, 201407 (2012).
63. Aykol, M., Dwaraknath, S. S., Sun, W. & Persson, K. A. Thermodynamic limit for synthesis of metastable inorganic materials. *Sci. Adv.* **4**, eaaq0148 (2018).
64. Han, S. et al. Design of new inorganic crystals with the desired composition using deep learning. *J. Chem. Inf. Model.* **63**, 5755–5763 (2023).
65. Sun, W. et al. The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2**, e1600225 (2016).
66. Hoffmann, R., Kabanov, A. A., Golov, A. A. & Proserpio, D. M. Homo citans and carbon allotropes: For an ethics of citation. *Angew. Chem. Int. Ed.* **55**, 10962–10976 (2016).
67. Yang, H. et al. MatterSim: A deep learning atomistic model across elements, temperatures and pressures. Preprint at http://arxiv.org/abs/2405.04967 (2024).
68. Batatia, I. et al. A foundation model for atomistic materials chemistry. Preprint at http://arxiv.org/abs/2401.00096 (2024).
69. 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).
70. Cheng, B. et al. Mapping materials and molecules. *Acc. Chem. Res.* **53**, 1981–1991 (2020).
71. Mavračić, J., Mocanu, F. C., Deringer, V. L., Csányi, G. & Elliott, S. R. Similarity between amorphous and crystalline phases: The case of TiO_2. *J. Phys. Chem. Lett.* **9**, 2985–2990 (2018).
72. van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
73. Rowe, P., Deringer, V. L., Gasparotto, P., Csányi, G. & Michaelides, A. An accurate and transferable machine learning potential for carbon. *J. Chem. Phys.* **153**, 034702 (2020).
74. Hendrycks, D. & Gimpel, K. Gaussian Error Linear Units (GELUs). Preprint at https://doi.org/10.48550/arXiv.1606.08415 (2023).
75. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. In *Int. Conf. Learn. Represent*. (ICLR, 2015).
76. Loshchilov, I. & Hutter, F. Decoupled weight decay regularization. In *Int. Conf. Learn. Represent*. (ICLR, 2019).
77. Paszke, A. et al. PyTorch: An imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* **32**, 8024–8035 (2019).
78. Gawlikowski, J. et al. A survey of uncertainty in deep neural networks. *Artif. Intell. Rev.* **56**, 1513–1589 (2023).
79. Shang, C. & Liu, Z.-P. *Quantum Chemistry in the Age of Machine Learning* (Elsevier, 2023)
80. Klimeš, J., Bowler, D. R. & Michaelides, A. Chemical accuracy for the van der Waals density functional. *J. Phys.: Condens. Matter* **22**, 022201 (2010).
81. Klimeš, J., Bowler, D. R. & Michaelides, A. Van der Waals density functionals applied to solids. *Phys. Rev. B* **83**, 195131 (2011).

82. McGaughey, A. J. H., Jain, A., Kim, H.-Y. & Fu, B. Phonon properties and thermal conductivity from first principles, lattice dynamics, and the Boltzmann transport equation. *J. Appl. Phys.* **125**, 011101 (2019).
83. Barbalinardo, G., Chen, Z., Lundgren, N. W. & Donadio, D. Efficient anharmonic lattice dynamics calculations of thermal transport in crystalline and disordered solids. *J. Appl. Phys.* **128**, 135104 (2020).
84. Omini, M. & Sparavigna, A. Beyond the isotropic-model approximation in the theory of thermal conductivity. *Phys. Rev. B* **53**, 9064–9073 (1996).
85. Ward, A., Broido, D. A., Stewart, D. A. & Deinzer, G. Ab initio theory of the lattice thermal conductivity in diamond. *Phys. Rev. B* **80**, 125203 (2009).
86. Fugallo, G. et al. Thermal conductivity of graphene and graphite: Collective excitations and mean free paths. *Nano Lett.* **14**, 6109–6114 (2014).
87. Togo, A., Chaput, L., Tadano, T. & Tanaka, I. Implementation strategies in phonopy and phono3py. *J. Phys.: Condens. Matter* **35**, 353001 (2023).
88. Togo, A. First-principles phonon calculations with Phonopy and Phono3py. *J. Phys. Soc. Jpn.* **92**, 012001 (2023).

Acknowledgements

This study was financially supported by the National Natural Science Foundation of China (Nos. 52425601, 52250273, 52327809, U20A20301, and 82361138571), National Key Research and Development Program of China (No. 2023YFB4404104), and Beijing Natural Science Foundation (No. L233022).

Author contributions

Conceptualization (L.G., Y.L., B.C.); Methodology (L.G., Y.L., Z.C., H.Y.); Investigation (L.G., Y.L., Z.C., H.Y.); Visualization (L.G., Y.L., Z.C., D.D., B.C.); Supervision (D.D., B.C.); Writing—original draft (L.G.); Writing—review and editing (L.G., Y.L., Z.C., H.Y., D.D., B.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/s41524-025-01592-8.

Correspondence and requests for materials should be addressed to Davide Donadio or Bingyang Cao.