



OPEN

Diffusion probabilistic models enhance variational autoencoder for crystal structure generative modeling

Teerachote Pakornchote¹, Natthaphon Choomphon-anomakhun¹, Sorrjit Arrerut¹, Chayanon Atthapak^{1,2}, Sakarn Khamkaeo^{1,2}, Thiparat Chotibut³ & Thiti Bovornratanaraks^{1,2}✉

The crystal diffusion variational autoencoder (CDVAE) is a machine learning model that leverages score matching to generate realistic crystal structures that preserve crystal symmetry. In this study, we leverage novel diffusion probabilistic (DP) models to denoise atomic coordinates rather than adopting the standard score matching approach in CDVAE. Our proposed DP-CDVAE model can reconstruct and generate crystal structures whose qualities are statistically comparable to those of the original CDVAE. Furthermore, notably, when comparing the carbon structures generated by the DP-CDVAE model with relaxed structures obtained from density functional theory calculations, we find that the DP-CDVAE generated structures are remarkably closer to their respective ground states. The energy differences between these structures and the true ground states are, on average, 68.1 meV/atom lower than those generated by the original CDVAE. This significant improvement in the energy accuracy highlights the effectiveness of the DP-CDVAE model in generating crystal structures that better represent their ground-state configurations.

Advances in computational materials science have enabled the accurate prediction of novel materials possessing exceptional properties. Remarkably, these computational advancements have facilitated the successful experimental synthesis of materials that exhibit the anticipated properties. Some predicted materials, such as near-room-temperature superconductors, have been successfully synthesized under high-pressure conditions, with their superconducting temperatures in accordance with density functional theory (DFT) calculations^{1,2}. To achieve accurate predictions, *a priori* knowledge of plausible molecular and crystal structures play a vital role in both theoretical and experimental studies. Several algorithms, such as evolutionary algorithms, swarm particle optimization, random sampling method, and etc., have been employed for structure prediction^{3–5}. These algorithms rely on identifying local minima on the potential energy landscape obtained from, for example, DFT calculations and machine learning-driven methods^{6–8}. In the case of crystal structures, where atoms are arranged in a three-dimensional space with periodic boundaries, additional criteria are necessary to enforce crystal symmetry constraints⁵.

Recent approach for structure prediction employs denoising diffusion models to perform probabilistic inference. These models sample molecular and crystal structures from a probability distribution of atomic coordinates and types^{9–12}, bypassing the computationally intensive DFT calculation that tediously determines the potential energy landscape. By leveraging sufficiently large datasets containing various compounds, this method enables the generation of diverse compositions and combinations of elements simultaneously. Furthermore, the models allow for the control of desired physical properties of the generated structures through conditional probability sampling^{13–15}. These machine learning-based algorithms also hold promise for solving inverse problem efficiently, resolving structures from experimental characterizations, e.g., x-ray absorption spectroscopy and other techniques, a challenging problem in materials science^{16–18}.

¹Extreme Conditions Physics Research Laboratory and Center of Excellence in Physics of Energy Materials (CE:PEM), Department of Physics, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand. ²Thailand Center of Excellence in Physics, Ministry of Higher Education, Science, Research and Innovation, 328 Si Ayutthaya Road, Bangkok 10400, Thailand. ³Chula Intelligent and Complex Systems, Department of Physics, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand. ✉email: thiti.b@chula.ac.th

There are two primary types of denoising diffusion models: score matching approach and denoising diffusion probabilistic models (DDPM)^{19–21}. These two models can denoise (reverse) a normal distribution such that the distribution gradually transforms into the data distribution of interest. The score matching approach estimates the score function of the perturbed data directing the normal distribution toward the data distribution and employing large step sizes of variance. In contrast, DDPM gradually denoises the random noise through a joint distribution of data perturbed at different scales of variance. Both approaches have been utilized for generating molecular structures^{10–12}. However, models based on DDPM tend to sample molecules with higher diversity and energy closer to the ground truth than models based on the score matching approach¹¹.

Since atomic positions in crystal structures are periodic and can be invariant under some rotation groups depending on their crystal symmetry, the core neural networks should favourably possess roto-translational equivariance^{22–24}. Xie et al.⁹ has proposed a model for crystal prediction by a combination between variational autoencoder (VAE)²⁵ and the denoising diffusion model, called crystal diffusion VAE (CDVAE). The model employs the score matching approach with (annealed) Langevin dynamics to generate new crystal structures from random coordinates¹⁹. The neural networks for an encoder and the diffusion model are roto-translationally equivariant. As a result, CDVAE can generate crystal structures with realistic bond lengths and respect crystal symmetry.

Because of the periodic boundary condition imposed on the unit cell, gradually injecting sufficiently strong noises (in the forward process) to the fractional coordinates can lead to the uniform distribution of atomic positions at late times, the consequence of ergodicity in statistical mechanical sense. Rather than beginning with a Gaussian distribution and denoising it as in the original CDVAE formulation, Jiao et al.²⁶ perturbed and sampled atomic positions beginning with a wrapped normal distribution which satisfies the periodic boundary condition. With this approach, the reconstruction performance has been significantly improved. Other circular (periodic) distributions, e.g., the wrapped normal and von Mises distributions, are not natural for DDPM framework since there is no known analytical method to explicitly incorporate such distributions into the framework. There, one needs to resort to an additional sampling procedure to construct the DDPM²⁷.

In this work, we introduce a crystal generation framework called diffusion probabilistic CDVAE (DP-CDVAE). Similar to the original CDVAE, our model consists of two parts: the VAE part and the diffusion part. The purpose of the VAE part is to predict the lattice parameters and the number of atoms in the unit cell of crystal structures. On the other hand, the diffusion part utilizes the diffusion probabilistic approach to denoise fractional coordinates and predict atomic coordinates. By employing the DDPM instead of the score matching approach, the DP-CDVAE model shows reconstruction and generation task performances that are statistically comparable to those obtained from original CDVAE. Importantly, we demonstrate the significantly higher ground-state generation performance of DP-CDVAE, through the distance comparison between generated structures and those optimized using the DFT method. We also analyze the changes in energy and volume after relaxation to gain further insights into models' capabilities.

Results

The performances of DP-CDVAE models are herein presented. There are four DP-CDVAE models, differed with the choice of the encoder (see Fig. S1 in Supplemental Information (SI)). DimeNet⁺⁺ has been employed for the main encoder for every DP-CDVAE models²⁸. We then modify the encoder of DP-CDVAE to encode the crystal structure by two additional neural networks: a multilayer perceptron that takes the number of atoms in the unit cell (N_a) as an input, and a graph isomorphism network (GINE)²⁹. Their latent features are combined with the latent features from DimeNet⁺⁺ through another multilayer perceptron. The N_a is encoded such that the model can decode the N_a accurately, and GINE encoder is inspired by GeoDiff¹¹ whose model is a combination of SchNet³⁰ and GINE which yields better performance.

Three datasets, Perov-5^{31,32}, Carbon-24³³, and MP-20³⁴, were selected to evaluate the performance of the model. The Perov-5 dataset consists of perovskite materials with cubic structures, but with variations in the combinations of elements within the structures. The Carbon-24 dataset comprises carbon materials, where the data consists of carbon element with various crystal systems obtained from *ab initio* random structure searching algorithm at pressure of 10 GPa³³. The MP-20 dataset encompasses a wide range of compounds and structure types.

Reconstruction performance

The reconstruction performance is determined by the similarity between reconstructed and ground-truth structures. The similarity can be evaluated using StructureMatcher algorithm from pymatgen library³⁵. The algorithm takes a pair of crystal structures and performs Niggli reduction to reduce their cells^{36,37}. They are then compared by determining an average displacement between the two structures. If it falls within the error tolerance, the two structures are matched. The reconstructed and ground-truth structures are similar if they pass the criteria of StructureMatcher which are $stol=0.5$, $angle_tol=10$, $ltol=0.3$. Match rate is the percentage of those structures passed the criteria. If the reconstructed and ground-truth structures are similar under the criteria, the root-mean-square distance between their atomic positions is computed and then normalized by $\sqrt[3]{V/N_a}$, where V is the unit-cell volume, and N_a is the number of atoms in the unit cell. An average of the distances of every pair of structures ($\langle\delta_{rms}\rangle$), computed using the StructureMatcher algorithm, is used as the performance metric.

Table 1 presents the reconstruction performance of different models for three different datasets: Perov-5, Carbon-24, and MP-20. Note that Fourier-transformed crystal properties (FTCP) model is presented as a baseline model, which is based on VAE. It encodes and decodes both the real-space and reciprocal-space features of crystal structures³⁸. For the Perov-5 dataset, the DP-CDVAE model achieves a match rate of 90.04%, indicating its ability to reconstruct a significant portion of the ground-truth structures. This performance is 7.48% lower

than the CDVAE model but still demonstrates the effectiveness of our model. In terms of $\langle \delta_{\text{rms}} \rangle$, the DP-CDVAE model achieves a value of 0.0212, comparable to the FTCP model³⁸, but slightly higher than the CDVAE model. Similarly, for the Carbon-24 and MP-20 datasets, the DP-CDVAE model performs well in terms of both match rate and $\langle \delta_{\text{rms}} \rangle$. It achieves match rates of 45.57% and 32.42% for Carbon-24 and MP-20, respectively. The corresponding $\langle \delta_{\text{rms}} \rangle$ values for Carbon-24 and MP-20 are 0.1513 and 0.0383, respectively, comparable to the CDVAE model.

Regarding the DP-CDVAE+ N_a model, the additional encoding of N_a into the model leads to improved match rates for all datasets, with an increase of 2–5%. This enhancement can be attributed to the accurate prediction of N_a . However, in terms of $\langle \delta_{\text{rms}} \rangle$, only the Perov-5 dataset shows an improvement, with a value of 0.0149. On the other hand, for the Carbon-24 and MP-20 datasets, the $\langle \delta_{\text{rms}} \rangle$ values are higher compared to the DP-CDVAE model.

For the DP-CDVAE+GINE and DP-CDVAE+ N_a +GINE models, the additional encoding of GINE into the models leads to a substantial drop in match rates compared to the DP-CDVAE model, particularly for the Perov-5 and Carbon-24 datasets. In contrast, there is a moderate increase in the match rates for the MP-20 dataset. The $\langle \delta_{\text{rms}} \rangle$ values for the Perov-5 and Carbon-24 datasets are comparable to those of the DP-CDVAE model. However, for the MP-20 dataset, the $\langle \delta_{\text{rms}} \rangle$ is noticeably higher in the models with GINE encoder compared to the DP-CDVAE model.

Overall, while the reconstruction performance of the DP-CDVAE model may be lower than the CDVAE model in terms of match rate, it still demonstrates competitive performance with relatively low $\langle \delta_{\text{rms}} \rangle$. The match rate can be enhanced by additionally encoding the N_a , but the performance is traded off by the increase in $\langle \delta_{\text{rms}} \rangle$.

Generation performance

We follow the CDVAE model that used three metrics to determine the generation performance of the models⁹. The first metric is the *Validity* percentage, which encompasses two sub-metrics: *Structural Validity* (Struc.) with a criterion that ensures the distances between every pair of atoms are larger than 0.5 Å, and *Compositional Validity* (Comp.) with a criterion that maintains a neutral total charge in the unit cell. The second metric is called *coverage* (COV), which utilizes structure and composition fingerprints to evaluate the similarity between the generated and ground-truth structures. COV-R (Recall) represents the percentage of ground-truth structures covered by the generated structures. COV-P (Precision) represents the percentage of generated structures that are similar to the ground-truth structures, indicating the quality of the generation. The third metric is the Wasserstein distance between property distributions of generated and ground-truth structures. Three property statistics are density (ρ), which is total atomic mass per volume (unit g/cm³), formation energy (E_{form} , unit eV/atom), and the number of elements in the unit cell (# elem.). A separated and pre-trained neural network is employed to predict E of the structures where the detail of the pre-training can be found in Ref.⁹. The first and second metrics are computed over 10,240 generated structures, and 1000 structures are randomly chosen from the generated structures that pass the validity tests to compute the third metric. The ground-truth structures used to evaluate the generation performance are from the test set.

In Table 2, the DP-CDVAE model achieves a validity rate of 100% for the Perov-5 dataset and close to 100% for the Carbon-24 and MP-20 datasets in terms of structure. The validity rate for composition is comparable to that of the CDVAE model. The DP-CDVAE model also demonstrates comparable COV-R values to the CDVAE model across all three datasets. Furthermore, the DP-CDVAE models with N_a and/or GINE encoders exhibit similar Validity and COV-R metrics to those of the DP-CDVAE model. However, for COV-P, all DP-CDVAE models yield lower values compared to CDVAE.

On the other hand, our models show significant improvements in property statistics. In the case of the MP-20 dataset, the DP-CDVAE models, particularly those with the GINE encoder, yield substantially smaller Wasserstein distances for ρ , E_{form} , and the number of elements compared to other models. For the Carbon-24 dataset, our models also exhibit a smaller Wasserstein distance for ρ compared to the CDVAE model.

Ground-state performance

Another objective of the structure generator is to generate novel structures that also are close to the ground state. To verify that, the generated structures are relaxed using the DFT calculation where the *relaxed structures* exhibit balanced internal stresses with external pressures and reside in local energy minima. These relaxed structures are then compared with the generated structures to evaluate their similarity. In this study, we have chosen a set of 100 generated structures from each of CDVAE, CDVAE+Fourier, and DP-CDVAE models for relaxation where CDVAE+Fourier model is CDVAE model with Fourier embedding features of the perturbed coordinates. However, relaxation procedures for multi-element compounds can be computationally intensive. To address this, we have specifically selected materials composed solely of carbon atoms, using the model trained on Carbon-24 dataset. This selection ensures a convergence of the self-consistent field in DFT calculation. Moreover, in the relaxation, we consider the ground state of the relaxed structures at a temperature of 0 K and a pressure of 10 GPa since the carbon structures in the training set are stable at 10 GPa³³.

We here introduce a ground-state performance presented in Table 3. The StructureMatcher with the same criteria as in the reconstruction performance is used to evaluate the similarity between the generated and relaxed structures. The relaxed structure was used as a based structure to determine if the generated structure can be matched. Four metrics used to determine the similarity are 1) match rate, 2) $\langle \delta_{\text{rms}} \rangle$, 3) ΔV_{rms} and 4) ΔE_{rms} . The ΔV_{rms} and ΔE_{rms} represent the root mean square differences in volume and energy, respectively, between the generated structures and the relaxed structures in the dataset.

In Table 3, the DP-CDVAE model achieves the highest match rate and the lowest $\langle \delta_{\text{rms}} \rangle$ and ΔE_{rms} . Although the CDVAE+Fourier model achieves the lowest ΔV_{rms} , the DP-CDVAE model demonstrates the ΔV_{rms} that is comparable to that of the CDVAE+Fourier model.

Discussion

The DP-CDVAE models significantly enhance the generation performance, particularly in terms of property statistics, while maintaining comparable COVs to those of CDVAE. Specifically, for Carbon-24 and MP-20 datasets, the density distributions between the generated and ground-truth structures from DP-CDVAE models exhibit substantially smaller Wasserstein distance compared those of the CDVAE model (see Table 2). The ΔV_{rms} of the DP-CDVAE model presented in Table 3 is significantly lower than that of the original CDVAE. This is corresponding to smaller Wasserstein distance of ρ shown in Table 2. The DP-CDVAE model also demonstrates significantly smaller $\langle \delta_{\text{rms}} \rangle$ than the original CDVAE. These suggest that our lattice generation closely approximates the relaxed lattice, while also achieving atomic positions that closely resemble the ground-state configuration. This could be an attribute of the DP approach that gradually learns perturbed coordinates, which in turn enhances the quality of sampled coordinates during the reverse process, much like its successful applications in image and molecular structure generation^{11,21,39}. Additionally, the distribution of the number of elements in the unit cells is relatively similar to that of the data in the test set, particularly in the results from the models with GINE encoder. This could be attributed to the capability of GINE to search for graph isomorphism⁴⁰. For the MP-20 dataset, the Wasserstein distances of the E_{form} values generated by our models are notably lower. This suggests that the crystal structures we generate are more likely to have E_{form} values within the specific range we are interested in. Hence, by selecting an appropriate training set, we can concentrate on structures with E_{form} values falling within the synthesizable candidate range.

Moreover, ΔE is the energy difference between the generated structures and their corresponding relaxed structures. The ground-state energy represents a local minimum that the generated structure is relaxed towards. A value of ΔE close to zero indicates that the generated structure is in close proximity to the ground state. In Table 3, it can be observed that our model achieves the ΔE_{rms} value of 400.7 meV/atom which is about 68.1 meV/atom lower than the ΔE_{rms} of CDVAE. The mode of ΔE of our model is 64–128 meV/atom, which is lower than its root-mean-square value (see Fig. S2 in SI). Nevertheless, both the ΔE_{rms} and the mode of ΔE exhibit relatively high values. In many cases, the formation energy of synthesized compounds is reported to be above the convex hull less than 36 meV/atom^{41–43}. To obviate the need for time-consuming DFT relaxation, it is essential for the generated structures to be even closer to the ground state. Therefore, achieving lower ΔE_{rms} values remains a milestone for future work.

Methods

Diffusion probabilistic model

In the diffusion probabilistic model, the data distribution is gradually perturbed by noise in the forward process until it becomes a normal distribution at late times. In this study, the distribution of the fractional coordinate (\mathbf{r}_f) is considered since their values of every crystal structures distribute over the same range, i.e., $\mathbf{r}_f \in [0, 1]^3$. The Markov process is assumed for the forward diffusion such that the joint distribution is a product of the conditional distributions conditioned on the knowledge of the fractional coordinate at the previous time step:

$$q(\mathbf{r}_{1:T}|\mathbf{r}_0) = \prod_{t=1}^T q(\mathbf{r}_t|\mathbf{r}_{t-1}), \quad (1)$$

$$q(\mathbf{r}_t|\mathbf{r}_{t-1}) = \mathcal{N}(\mathbf{r}_t; \sqrt{\bar{\alpha}_t}\mathbf{r}_{t-1}, (1 - \alpha_t)\mathbf{I}),$$

where $\mathbf{r}_0 \sim q(\mathbf{r}_f)$ the data distribution of the fractional coordinate, t is the discretized diffusion time step, T is the final diffusion time, α_t is a noise schedule with a sigmoid scheduler⁴⁴, and the conditional $q(\cdot|\cdot)$ is a Gaussian kernel due to the Markov diffusion process assumption. Then \mathbf{r}_t can be expressed in the Langevin's form through the reparameterization trick as

$$\mathbf{r}_t = \sqrt{\bar{\alpha}_t}\mathbf{r}_0 + \sqrt{1 - \bar{\alpha}_t}\boldsymbol{\epsilon}, \quad (2)$$

where $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I})$, and $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$. This update rule does not necessitate \mathbf{r}_t to remain in $[0, 1]^3$; however, we can impose the periodic boundary condition for the fractional coordinate so that

$$\mathbf{r}_f = \boldsymbol{\pi}(\mathbf{r}_t) := \mathbf{r}_t - \lfloor \mathbf{r}_t \rfloor. \quad (3)$$

Then, $\mathbf{r}_f \in [0, 1]^3$.

In the reverse diffusion process, if the consecutive discretized time step is small compared to the diffusion timescale, the reverse coordinate trajectories can be approximately sampled also from the product of Gaussian diffusion kernels as

$$p_\theta(\mathbf{r}_{0:T}) = p(\mathbf{r}_T) \prod_{t=1}^T p_\theta(\mathbf{r}_{t-1}|\mathbf{r}_t), \quad (4)$$

$$p_\theta(\mathbf{r}_{t-1}|\mathbf{r}_t) = \mathcal{N}(\mathbf{r}_{t-1}; \boldsymbol{\mu}_\theta, \sigma_t^2\mathbf{I}),$$

where

$$\begin{aligned}\boldsymbol{\mu}_\theta &= \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{r}_t - \frac{1 - \alpha_t}{\sqrt{1 - \alpha_t}} \boldsymbol{\epsilon}_\theta \right), \\ \sigma_t^2 &= \frac{(1 - \bar{\alpha}_{t-1})(1 - \alpha_t)}{1 - \bar{\alpha}_t}.\end{aligned}\quad (5)$$

The reverse conditional distribution can be trained by minimizing the Kullback-Leibler divergence between $p_\theta(\mathbf{r}_{t-1}|\mathbf{r}_t)$ and $q(\mathbf{r}_{t-1}|\mathbf{r}_t, \mathbf{r}_0)$, the posterior of the corresponding forward process²¹. We use GemNetT for the diffusion network to train the parametrized noise $\boldsymbol{\epsilon}_\theta$ ⁴⁵. Then, the coordinate in the earlier time can be sampled from $\mathbf{r}_{t-1} \sim p_\theta(\mathbf{r}_{t-1}|\mathbf{r}_t)$, whose corresponding reverse Langevin's dynamics reads

$$\mathbf{r}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{r}_t - \frac{1 - \alpha_t}{\sqrt{1 - \alpha_t}} \boldsymbol{\epsilon}_\theta \right) + \sigma_t \boldsymbol{\epsilon}', \quad (6)$$

where $\boldsymbol{\epsilon}' \sim \mathcal{N}(0, \mathbf{I})$. Crucially, we empirically found that the final reconstruction performance is considerably improved when we impose the periodic boundary condition on the fractional coordinate at every time step such that $\mathbf{r}_{t-1} \sim p_\theta(\mathbf{r}_{t-1}|\mathbf{r}_{f_t})$ and α_t in the first term of Eq. (6) is replaced by $\bar{\alpha}_t$. Namely, in our modified reverse process, the coordinate is sampled from

$$\begin{aligned}\mathbf{r}_{t-1} &= \frac{1}{\sqrt{\bar{\alpha}_t}} \left(\mathbf{r}_{f_t} - \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}_\theta \right) + \sigma_t \boldsymbol{\epsilon}', \\ \mathbf{r}_{f_t} &= \boldsymbol{\pi}(\mathbf{r}_t).\end{aligned}\quad (7)$$

An illustration of denoising atomic coordinates with Eq. (7) is demonstrated in Fig. 2. The model performance using Eq. (6) is shown in Table S1 in SI, whereas the performance using Eq. (7) is shown in Table 1.

Graph neural networks

Graph neural networks architecture facilitate machine learning of crystal graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, graph representations of crystal structures. \mathcal{V} and \mathcal{E} are sets of nodes and edges, respectively, defined as

$$\begin{aligned}\mathcal{V} &= \{(\mathbf{f}_n, \mathbf{r}_{c_n}) \mid \mathbf{f}_n \in \mathbb{R}^M, \mathbf{r}_{c_n} = \mathbf{L}\mathbf{r}_{f_n} \in \mathbb{R}^3\}, \\ \mathcal{E} &= \{\Delta\mathbf{r}_{c_{mn}}^{(\mathbf{T})} \mid \Delta\mathbf{r}_{c_{mn}}^{(\mathbf{T})} = \mathbf{r}_{c_m} - \mathbf{r}_{c_n} + \mathbf{T}; \mathbf{r}_{c_m}, \mathbf{r}_{c_n} \in \mathbb{R}^3\},\end{aligned}$$

where n and m are indices of atoms in a crystal structure, \mathbf{f}_n is a vector of M features of an atom in the unit cell, \mathbf{T} is a translation vector, and \mathbf{L} is a lattice matrix that converts a fractional coordinate \mathbf{r}_{f_n} into its atomic Cartesian coordinate \mathbf{r}_{c_n} . The atomic features, fractional coordinates, and atomic Cartesian coordinates of the crystal structure are vectorized (concatenated) as $\mathbf{f} = (\mathbf{f}_1, \dots, \mathbf{f}_{N_a}) \in \mathbb{R}^{N_a \times M}$, $\mathbf{r}_f = (\mathbf{r}_{f_1}, \dots, \mathbf{r}_{f_{N_a}}) \in \mathbb{R}^{N_a \times 3}$, and $\mathbf{r}_c = (\mathbf{r}_{c_1}, \dots, \mathbf{r}_{c_{N_a}}) \in \mathbb{R}^{N_a \times 3}$. Three graph neural networks implemented in this work are DimeNet⁺⁺²⁸, GINE²⁹, and GemNetT⁴⁵. DimeNet⁺⁺ and GINE are employed for encoders, and GemNetT is used for a diffusion network. DimeNet⁺⁺ and GemNetT, whose based architecture concerns geometry of the graphs, are rotationally equivariant. GemNetT has been devised by incorporating the polar angles between four atoms into DimeNet⁺⁺. This development grants GemNetT a higher degree of expressive power compared to DimeNet⁺⁺⁴⁶. Furthermore, GINE has been developed to distinguish a graph isomorphism, but not graph geometry nor the distance between nodes, which is important for our study. Thus we supplement the edge attributes into GINE with the distances between atoms, i.e. $\mathcal{E} = \{|\Delta\mathbf{r}_{c_{mn}}^{(\mathbf{T})}|\}$.

DP-CDVAE's architecture

The forward process of DP-CDVAE model is illustrated in Fig. 1. The model is a combination of two generative models, which are VAE and diffusion probabilistic model. The pristine crystal structures consist of the fractional coordinate (\mathbf{r}_f), the lattice matrix (\mathbf{L}), ground-truth atomic type (Z), and the number of atoms in a unit cell (N_a). For crystal graphs of the encoders, their node features are $\mathbf{f} = Z$. The number of atoms in a unit cell N_a is encoded through multilayer perceptron before concatenated with the latent features from other graph encoders. They are encoded to train $\boldsymbol{\mu}_\phi$ and $\log\text{var}_\phi$ where ϕ is a learnable parameter of the encoders. The latent variables (\mathbf{z}) can be obtained by

$$\mathbf{z} = \boldsymbol{\mu}_\phi + e^{\log\text{var}_\phi} \boldsymbol{\epsilon}'', \quad (8)$$

where $\boldsymbol{\epsilon}'' \sim \mathcal{N}(0, \mathbf{I})$. Then, \mathbf{z} will be decoded to compute the lattice lengths and angles, which then yield the lattice matrix (\mathbf{L}_z), N_a , and \mathbf{A}_z . In the original CDVAE, \mathbf{A}_z is the probability vector indicating the fraction of each atomic type in the compound and is used to perturb Z by

$$Z_t \sim \mathcal{M}(\text{softmax}(\mathbf{A} + \sigma'_t \mathbf{A}_z)) \quad (9)$$

where \mathcal{M} is a multinomial distribution, \mathbf{A} is a one-hot vector of ground-truth atomic type Z , and σ'_t is the variance for perturbing atomic types at time t , which is distinct from σ_t used for perturbing the atomic coordinates. Similar to the original CDVAE, σ'_t is selected from the range of [0.01, 5].

For the diffusion network, the input structures are constructed from \mathbf{r}_{f_t} , Z_t , and \mathbf{L}_z where the (Cartesian) atomic coordinates at time t are computed by $\mathbf{r}_{c_t} = \mathbf{L}_z \mathbf{r}_{f_t}$. These are then utilized by the crystal graphs for the diffusion network, whose node features are $\mathbf{f}_t = (Z_t, \mathbf{F}_t, \mathbf{z}, t)$ where \mathbf{F}_t is a Fourier embedding feature of \mathbf{r}_t (see SI). As proposed by Ho et al.²¹, we use the simple loss to train the model such that

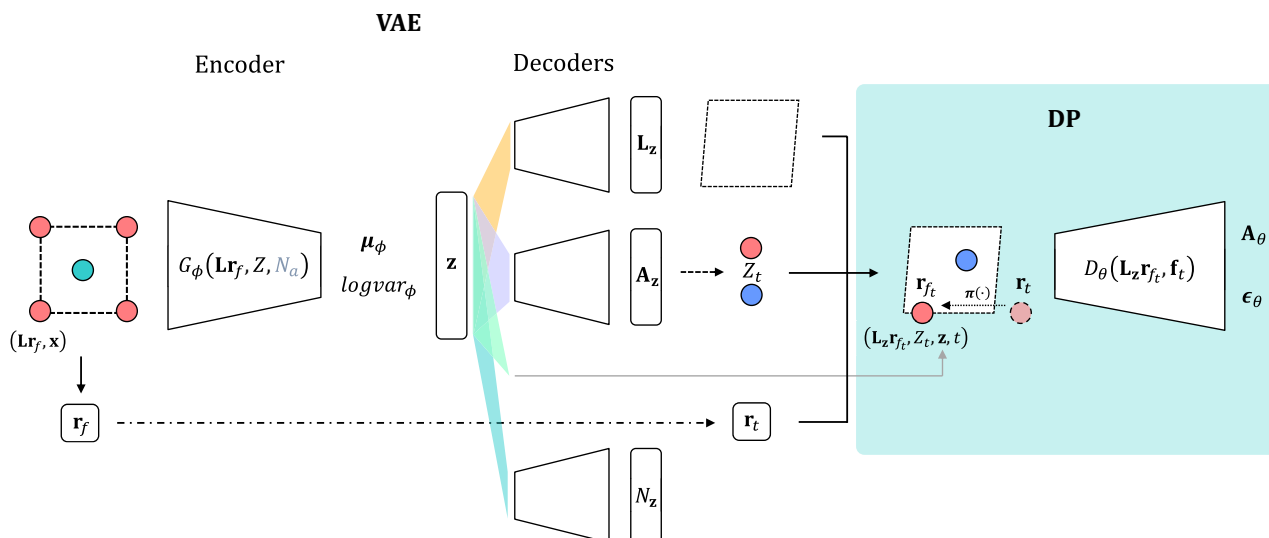


Figure 1. The schematic summarizing the architecture for training the DP-CDVAE model. Multiple sub-networks are trained to minimize the total loss function. The encoder ($G_\phi(\mathbf{L}\mathbf{r}_f, Z, N_a)$) compresses input pristine crystal structures into the latent feature (\mathbf{z}). The predicted lattice parameters (\mathbf{L}_z), the predicted number of atoms (N_z), and all the decoded features enable the reconstruction of crystal structures. Here, \mathbf{A}_z enables the sampling of atomic types (Z_t), and all the decoded features enable the reconstruction of crystal structures. The input fractional coordinates \mathbf{r}_f undergo perturbation (dash-dotted line) at time step t and then are transformed by $\pi(\cdot)$ to satisfy the periodic boundary condition (dotted line), serving as the coordinates for the reconstructed crystal structures. These reconstructed structures, $(\mathbf{L}_z\mathbf{r}_{f_t}, Z_t, \mathbf{z}, t)$, are subsequently fed into the diffusion network ($D_\theta(\mathbf{L}_z\mathbf{r}_{f_t}, \mathbf{f}_t)$), where \mathbf{f}_t is a node feature composing of Z_t, \mathbf{z} , and t . The diffusion network predicts the noise added to the fractional coordinates (ϵ_θ) as well as the one-hot vector of atomic types (\mathbf{A}_θ). Dashed-line boxes represent the unit cells of the crystal structures.

Models	Match rate (%) \uparrow			$\langle \delta_{rms} \rangle \downarrow$		
	Perov-5	Carbon-24	MP-20	Perov-5	Carbon-24	MP-20
FTCP ⁹	99.34	62.28	69.89	0.0259	0.2563	0.1593
CDVAE ⁹	97.52	55.22	45.43	0.0156	0.1251	0.0356
DP-CDVAE	90.04	45.57	32.42	0.0212	0.1513	0.0383
DP-CDVAE+ N_a	91.86	50.99	36.17	0.0149	0.1612	0.0560
DP-CDVAE+GINE	80.50	49.02	34.08	0.0214	0.1599	0.0455
DP-CDVAE+ N_a +GINE	88.30	38.28	37.44	0.0180	0.1921	0.0525

Table 1. Reconstruction performance. Significant values are in [bold].

$$\mathcal{L}_{simple} = \|\epsilon - \epsilon_\theta(\mathbf{r}_{c_t}, \mathbf{f}_t)\|^2. \quad (10)$$

Since the diffusion model is trained to predict both ϵ and \mathbf{A} , so the loss of the diffusion network is

$$\mathcal{L}_{diff} = \mathcal{L}_{simple} + \lambda \mathcal{L}_{CE}(\mathbf{A}, \mathbf{A}_\theta(\mathbf{r}_{c_t}, \mathbf{f}_t)), \quad (11)$$

where \mathcal{L}_{CE} is the cross entropy loss, λ is a loss scaling factor, and $t \in \{1, \dots, T\}$ where $T = 1000$. In this work, t is randomly chosen for each crystal graph and randomly reinitialized for each epoch in the training process. The total loss in the training process is shown in Eq. S1 in SI.

In the reverse diffusion process, we measure the model performance of two tasks: reconstruction and generation tasks. For the former task, \mathbf{z} is obtained from Eq. (8) by using the ground-truth structure as an input of the encoders. For the latter task, $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I})$, which is then used to predict $N_a, \mathbf{L}_z, \mathbf{A}_z$, and concatenate with the node feature of the crystal graph in the diffusion network. At the initial step, $t = T$, Z_T is sampled from the highest probability of \mathbf{A}_z , and the final-time coordinate is obtained from sampling a Gaussian distribution, i.e. $\mathbf{r}_T \sim \mathcal{N}(0, \mathbf{I})$. The coordinates can be denoised using Eq. (7), and the predicted atomic types are updated in each reversed time step by $\text{argmax}(\mathbf{A}_\theta)$.

Datasets	Models	Validity (%) ↑		COV (%) ↑		Property statistics ↓		
		Struc.	Comp.	R.	P.	ρ	E_{form}	# elem.
Perov-5	G-SchNet ⁹	99.92	98.79	0.18	0.23	1.625	4.746	0.0368
	P-G-SchNet ⁹	79.63	99.13	0.37	0.25	0.2755	1.388	0.4552
	CDVAE ⁹	100	98.59	99.45	98.46	0.1258	0.0264	0.0628
	DP-CDVAE	100	98.07	99.52	98.39	0.1807	0.0713	0.0767
	DP-CDVAE+N _a	99.99	97.34	99.55	97.22	0.1027	0.0287	0.0437
	DP-CDVAE+GINE	100	96.11	98.94	95.63	0.2114	0.0832	0.0498
	DP-CDVAE+N _a +GINE	100	97.09	99.52	96.73	0.1368	0.0425	0.0210
Carbon-24	G-SchNet ⁹	99.94	–	0.00	0.00	0.9427	1.320	–
	P-G-SchNet ⁹	48.39	–	0.00	0.00	1.533	134.7	–
	CDVAE ⁹	100	–	99.80	83.08	0.1407	0.2850	–
	DP-CDVAE	99.92	–	99.56	77.98	0.1109	0.2596	–
	DP-CDVAE+N _a	99.73	–	99.61	72.29	0.1080	0.3030	–
	DP-CDVAE+GINE	99.50	–	100	68.13	0.0977	0.3623	–
	DP-CDVAE+N _a +GINE	98.61	–	99.21	65.13	0.1267	0.4136	–
MP-20	G-SchNet ⁹	99.65	75.96	38.33	99.57	3.034	42.09	0.6411
	P-G-SchNet ⁹	77.51	76.40	41.93	99.74	4.04	2.448	0.6234
	CDVAE ⁹	100	86.70	99.15	99.49	0.6875	0.2778	1.432
	DP-CDVAE	99.59	85.44	98.93	98.96	0.4037	0.1547	0.9179
	DP-CDVAE+N _a	99.81	84.95	99.36	99.33	0.4889	0.1800	1.053
	DP-CDVAE+GINE	99.82	81.92	99.48	99.00	0.2785	0.0603	0.5679
	DP-CDVAE+N _a +GINE	99.90	83.89	95.51	99.27	0.1790	0.0522	0.6909

Table 2. Generation performance. Significant values are in [bold].

Model	Match rate (%) ↑	$\langle \delta_{rms} \rangle$ ↓	ΔV_{rms} (Å ³ /atom) ↓	ΔE_{rms} (meV/atom) ↓
CDVAE	63	0.0321	0.0227	468.8
CDVAE+Fourier	62	0.0216	0.0157	494.4
DP-CDVAE	64	0.0141	0.0158	400.7

Table 3. Ground-state performance. Significant values are in [bold].

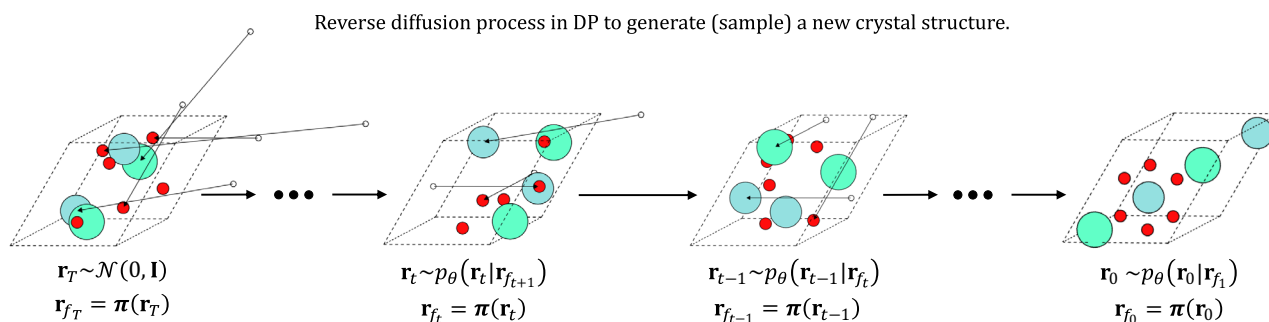


Figure 2. The schematic depicting the reverse diffusion process of the DP-CDVAE model. Initially, atomic coordinates are sampled from a normal distribution and subsequently mapped into the unit cell (dashed-line box) using the periodic boundary-imposing function $\pi(\cdot)$. White circles outside the unit cell depict the atomic coordinates prior to the periodic boundary condition is imposed, while colored circles represent atoms that are inside the unit cell of interest. The action of $\pi(\cdot)$ on the atoms outside the unit cell is represented by an arrow that translates the white circles into colored circles in the unit cell. Left to right show the reverse direction of the arrow of time, depicting the reverse diffusion process.

DFT calculations

The Vienna *ab initio* Simulation Package (VASP) was employed for structural relaxations and energy calculations based on DFT^{47,48}. The calculations were conducted under the generalized gradient approximation (GGA), which is Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional, and the project augmented wave (PAW) method^{49,50}. The thresholds for energy and force convergence were set to 10^{-5} eV and 10^{-5} eV/Å, respectively. The plane-wave energy cutoff was set to 800 eV, and the Brillouin zone integration was carried out on a k-point mesh of $5 \times 5 \times 5$ created by the Monkhorst-Pack method^{51,52}.

Data availability

The code and datasets generated and/or analysed during the current study are available at <https://github.com/trachote/dp-cdvae>.

Received: 27 August 2023; Accepted: 4 January 2024

Published online: 13 January 2024

References

- Needs, R. J. & Pickard, C. J. Perspective: Role of structure prediction in materials discovery and design. *APL Mater.* **4**, 053210. <https://doi.org/10.1063/1.4949361> (2016).
- Kohn, W. & Sham, L. J. *Phys. Rev.* **140**, A1133 (1965).
- Oganov, A. R. & Glass, C. W. Crystal structure prediction using *ab initio* evolutionary techniques: Principles and applications. *J. Chem. Phys.* **124**, 244704. <https://doi.org/10.1063/1.2210932> (2006).
- Wang, Y., Lv, J., Zhu, L. & Ma, Y. Crystal structure prediction via particle-swarm optimization. *Phys. Rev. B* **82**, 094116. <https://doi.org/10.1103/PhysRevB.82.094116> (2010).
- Pickard, C. J. & Needs, R. J. *Ab initio* random structure searching. *J. Phys. Condens. Matter* **23**, 053201. <https://doi.org/10.1088/0953-8984/23/5/053201> (2011).
- Oganov, A. R., Pickard, C. J., Zhu, Q. & Needs, R. J. Structure prediction drives materials discovery. *Nat. Rev. Mater.* **4**, 331 (2019).
- Schön, J. C., Doll, K. & Jansen, M. Predicting solid compounds via global exploration of the energy landscape of solids on the *ab initio* level without recourse to experimental information. *Physica Status Solidi (b)* **247**, 23. <https://doi.org/10.1002/pssb.200945246> (2010).
- Podryabinkin, E. V., Tikhonov, E. V., Shapeev, A. V. & Oganov, A. R. Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. *Phys. Rev. B* **99**, 064114. <https://doi.org/10.1103/PhysRevB.99.064114> (2019).
- Xie, T., Fu, X., Ganea, O.-E., Barzilay, R. & Jaakkola, T. S. Crystal diffusion variational autoencoder for periodic material generation. In *International Conference on Learning Representations*. https://openreview.net/forum?id=03RLpj-tc_ (2022).
- Shi, C., Luo, S., Xu, M. & Tang, J. Learning gradient fields for molecular conformation generation. In *Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research*, Vol. 139, 9558–9568 (eds Meila, M. & Zhang, T.) (PMLR, 2021).
- Xu, M., Yu, L., Song, Y., Shi, C., Ermon, S. & Tang, J. Geodiff: A geometric diffusion model for molecular conformation generation. In *International Conference on Learning Representations*. <https://openreview.net/forum?id=PzcvxEMzvQC> (2022).
- Guan, J., Qian, W. W., Peng, X., Su, Y., Peng, J., & Ma, J. 3d equivariant diffusion for target-aware molecule generation and affinity prediction. In *The Eleventh International Conference on Learning Representations*. <https://openreview.net/forum?id=kJqXEPXMsE0> (2023).
- Kang, S. & Cho, K. Conditional molecular design with deep generative models. *J. Chem. Inf. Model.* **59**, 43. <https://doi.org/10.1021/acs.jcim.8b00263> (2019).
- Lim, J., Ryu, S., Kim, J. W. & Kim, W. Y. Molecular generative model based on conditional variational autoencoder for *de novo* molecular design. *J. Cheminform.* **10**, 31 (2018).
- Song, Y., Shen, L., Xing, L. & Ermon, S. Solving inverse problems in medical imaging with score-based generative models. In *International Conference on Learning Representations*. <https://openreview.net/forum?id=vaRCHVj0uGI> (2022).
- Cui, A. *et al.* Decoding phases of matter by machine-learning Raman spectroscopy. *Phys. Rev. Appl.* **12**, 054049. <https://doi.org/10.1103/PhysRevApplied.12.054049> (2019).
- Carbone, M. R., Topsakal, M., Lu, D. & Yoo, S. Machine-learning x-ray absorption spectra to quantitative accuracy. *Phys. Rev. Lett.* **124**, 156401. <https://doi.org/10.1103/PhysRevLett.124.156401> (2020).
- Liang, Z. *et al.* Decoding structure-spectrum relationships with physically organized latent spaces. *Phys. Rev. Mater.* **7**, 053802. <https://doi.org/10.1103/PhysRevMaterials.7.053802> (2023).
- Song, Y. & Ermon, S. Generative modeling by estimating gradients of the data distribution. In *Advances in Neural Information Processing Systems*, Vol. 32, (eds Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E. & Garnett, R.) (Curran Associates, Inc., 2019). https://proceedings.neurips.cc/paper_files/paper/2019/file/3001ef257407d5a371a96dcd947c7d93-Paper.pdf.
- Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S. & Poole, B. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*. <https://openreview.net/forum?id=PxTIG12RRHS> (2021).
- Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, Vol. 33, 6840–6851 (eds Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M. & Lin, H.) (Curran Associates, Inc., 2020). https://proceedings.neurips.cc/paper_files/paper/2020/file/4c5bfcfec8584af0d967f1ab10179ca4b-Paper.pdf.
- Bronstein, M. M., Bruna, J., Cohen, T. & Velickovic, P. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *CoRR arXiv:2104.13478* (2021).
- Cohen, T. S., Geiger, M., Köhler, J. & Welling, M. Spherical CNNs. In *International Conference on Learning Representations*. <https://openreview.net/forum?id=Hkdb5xZRb> (2018).
- Thomas, N., Smidt, T., Kearnes, S., Yang, L., Li, L., Kohlhoff, K. & Riley, P. Tensor field networks: Rotation- and translation-equivariant neural networks for 3d point clouds. <https://doi.org/10.48550/ARXIV.1802.08219> (2018).
- Kingma, D. P. & Welling, M. Auto-encoding variational bayes. In *International Conference on Learning Representations* (2014).
- Jiao, R., Huang, W., Lin, P., Han, J., Chen, P., Lu, Y. & Liu, Y. Crystal structure prediction by joint equivariant diffusion on lattices and fractional coordinates. In *Workshop on "Machine Learning for Materials" ICLR 2023*. <https://openreview.net/forum?id=VPBypHdu24j> (2023).
- Okhotin, A., Molchanov, D., Arkipkin, V., Bartosh, G., Alanov, A. & Vetrov, D. Star-shaped denoising diffusion probabilistic models. [arXiv:2302.05259](https://arxiv.org/abs/2302.05259) [stat.ML] (2023).
- Gasteiger, J., Giri, S., Margraf, J. T. & Günnemann, S. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. [arXiv:2011.14115](https://arxiv.org/abs/2011.14115) [cs.LG] (2022).

29. Hu*, W., Liu*, B., Gomes, J., Zitnik, M., Liang, P., Pande, V. & Leskovec, J. Strategies for pre-training graph neural networks, In *International Conference on Learning Representations*. <https://openreview.net/forum?id=HJlWWSFDH> (2020).
30. Schütt, K., Kindermans, P.-J., Sauceda Felix, H. E., Chmiela, S., Tkatchenko, A. & Müller, K.-R. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. In *Advances in Neural Information Processing Systems* Vol. 30 (eds Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S. & Garnett, R.) (Curran Associates, Inc, 2017). https://proceedings.neurips.cc/paper_files/paper/2017/file/303ed4c69846ab36c2904d3ba8573050-Paper.pdf.
31. Castelli, I. E. *et al.* New cubic perovskites for one-and two-photon water splitting using the computational materials repository. *Energy Environ. Sci.* **5**, 9034 (2012).
32. Castelli, I. E. *et al.* Computational screening of perovskite metal oxides for optimal solar light capture. *Energy Environ. Sci.* **5**, 5814 (2012).
33. Pickard, C. J. <https://doi.org/10.24435/MATERIALSCLOUD:2020.0026/V1>
34. Jain, A. *et al.* Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
35. Grosse-Kunstleve, R. W., Sauter, N. K. & Adams, P. D. Numerically stable algorithms for the computation of reduced unit cells. *Acta Crystallogr. A* **60**, 1. <https://doi.org/10.1107/S010876730302186X> (2004).
36. Gruber, B. The relationship between reduced cells in a general Bravais lattice. *Acta Crystallogr. A* **29**, 433. <https://doi.org/10.1107/S0567739473001063> (1973).
37. Křivý, I. & Gruber, B. A unified algorithm for determining the reduced (Niggli) cell. *Acta Crystallogr. A* **32**, 297. <https://doi.org/10.1107/S0567739476000636> (1976).
38. Ren, Z. *et al.* An invertible crystallographic representation for general inverse design of inorganic crystals with targeted properties. *Matter* **5**, 314. <https://doi.org/10.1016/j.matt.2021.11.032> (2022).
39. Nichol, A. Q. & Dhariwal, P. Improved denoising diffusion probabilistic models, In *Proceedings of the 38th International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 139, 8162–8171 (eds Meila, M. & Zhang, T.) (PMLR, 2021). <https://proceedings.mlr.press/v139/nichol21a.html>.
40. Xu, K., Hu, W., Leskovec, J. & Jegelka, S. How powerful are graph neural networks? In *International Conference on Learning Representations*. <https://openreview.net/forum?id=ryGs6iA5Km> (2019).
41. Wu, Y., Lazić, P., Hautier, G., Persson, K. & Ceder, G. First principles high throughput screening of oxynitrides for water-splitting photocatalysts. *Energy Environ. Sci.* **6**, 157. <https://doi.org/10.1039/C2EE23482C> (2013).
42. Ishikawa, T. & Miyake, T. Evolutionary construction of a formation-energy convex hull: Practical scheme and application to a carbon–hydrogen binary system. *Phys. Rev. B* **101**, 214106. <https://doi.org/10.1103/PhysRevB.101.214106> (2020).
43. Ektarawong, A., Johansson, E., Pakornchote, T., Bovornratanaarak, T. & Alling, B. Boron vacancy-driven thermodynamic stabilization and improved mechanical properties of alB₂-type tantalum diborides as revealed by first-principles calculations. *J. Phys. Mater.* **6**, 025002. <https://doi.org/10.1088/2515-7639/acbe69> (2023).
44. Kingma, D. P., Salimans, T., Poole, B. & Ho, J. On density estimation with diffusion models, In *Advances in Neural Information Processing Systems*, (eds Beygelzimer, A., Dauphin, Y., Liang, P. & Vaughan, J. W.). <https://openreview.net/forum?id=2LdBqxc1Yv> (2021).
45. Klicpera, J., Becker, F. & Günnemann, S. Gnnmet: Universal directional graph neural networks for molecules, In *Advances in Neural Information Processing Systems* (eds Beygelzimer, A., Dauphin, Y., Liang, P. & Vaughan, J. W.). https://openreview.net/forum?id=HS_sOaxS9K- (2021).
46. Joshi, C. K., Bodnar, C., Mathis, S. V., Cohen, T. & Lio, P. On the expressive power of geometric graph neural networks. https://openreview.net/forum?id=Rkxj1GXn9_ (2023).
47. Kresse, G. & Furthmüller, J. *Comput. Mater. Sci.* **6**, 15 (1996).
48. Kresse, G. & Furthmüller, J. *Phys. Rev. B* **54**, 11169 (1996).
49. Perdew, J. P., Burke, K. & Ernzerhof, M. *Phys. Rev. Lett.* **77**, 3865 (1996).
50. Blöchl, P. E. *Phys. Rev. B* **50**, 17953 (1994).
51. Monkhorst, H. J. & Pack, J. D. *Phys. Rev. B* **13**, 5188 (1976).
52. Pack, J. D. & Monkhorst, H. J. *Phys. Rev. B* **16**, 1748 (1977).

Acknowledgements

This research project is supported by the Second Century Fund (C2F), Chulalongkorn University. This Research is funded by Thailand Science Research and Innovation Fund Chulalongkorn University and National Research Council of Thailand (NRCT): (NRCT5-RSA63001-04). T.C. acknowledges funding support from the NSRF via the Program Management Unit for Human Resources and Institutional Development, Research and Innovation [grant number B05F650024]. The authors acknowledge high performance computing resources including NVIDIA A100 GPUs from Chula Intelligent and Complex Systems Lab, Faculty of Science, and from the Center for AI in Medicine (CU-AIM), Faculty of Medicine, Chulalongkorn University, Thailand. We acknowledge the supporting computing infrastructure provided by NSTDA, CU, CUAASC, NSRF via PMUB [B05F650021, B37G660013] (Thailand). URL: www.e-science.in.th. The Computational Materials Physics (CMP) Project, SLRI, Thailand, is acknowledged for providing computational resource.

Author contributions

T.P. and N.C. implemented the methodology into the code. T.P., N.C., R.A., and C.A. designed the project and contributed to discussion of the results. S.K. performed the DFT calculations. T.P., N.C., R.A., and T.C. contributed to discussion of the theory and the algorithm. T.P. and T.C. wrote the paper. T.B. supervised all aspects of the project.

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/s41598-024-51400-4>.

Correspondence and requests for materials should be addressed to T.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024