Exploring 2D Materials: Discovery and Characterization via Generative Models and DFT

Peder Lyngby[1], Kartikeya Sharma[1], Kristian Thygesen[1]
[1]Computational Atomic-scale Materials Design (CAMD), Department of Physics, Technical University of Denmark, Denmark

pmely@dtu.dk

Efficient algorithms to generate candidate crystal structures with good stability properties can play a key role in data-driven materials discovery. Here, we show that a crystal diffusion variational autoencoder (CDVAE) is capable of generating two-dimensional (2D) materials of high chemical and structural diversity and formation energies mirroring the training structures. Specifically, we train the CDVAE on 2615 2D materials with energy above the convex hull ΔH_{hull} < 0.3 eV/atom, and generate 5003 materials that we relax using density functional theory (DFT). We also generate 14192 new crystals by systematic element substitution of the training structures. We find that the generative model and lattice decoration approach are complementary and yield materials with similar stability properties but very different crystal structures and chemical compositions. In total we find 11630 predicted new 2D materials, where 8599 of these have ΔH_{hull}< 0.3 eV/atom as the seed structures, while 3331 are within 100 meV of the convex hull and could potentially be synthesised.

However, for practical purposes knowledge of the structure and composition of stable materials is not very useful in itself because the decision to synthesise and deploy a given material usually requires some presumptions about the material's properties. Therefore, to make computational materials discovery relevant for experiments and applications, the determination of stable crystal structures must be complemented by a characterisation of the most basic materials properties.

Subsequently, we calculate the elementary physical properties of the 3331 most stable monolayers within 100 meV/atom of the convex hull. As a testimony to the good thermodynamic stability and chemical validity of the structures, we find that the majority of them (about 70%) are dynamically stable, i.e. stable against small perturbations of the atom positions and unit cell shape. For this subset of 2759 materials we employ the computational workflow behind the C2DB to compute a wide variety of properties, such as the stiffness tensor, piezoelectric tensor, deformation potentials, Born and Bader charges, electronic band structure, effective masses, plasma frequency, Fermi surface, projected density of states, magnetic moments, magnetic exchange couplings, magnetic anisotropy, topological indices, optical- and infrared polarisability.

Prior to the present work, the C2DB (which is currently the largest 2D materials database) contained 1345 monolayers with convex hull energy below 0.1 eV/atom. Thus the new set of monolayers characterised in this work, triples the number of (theoretically) known stable 2D materials.

Our work exposes previously unknown parts of the 2D chemical space and provides a basis for the discovery of 2D materials with specific properties. All data is available in the C2DB.

Moreover, this substantial collection of newly characterized 2D materials serves as an extensive training set for the development of a novel C2DB-ML database. Leveraging the CDVAE and lattice decoration protocol, we generate thousands of additional 2D structures, which are then relaxed using machine learning (ML) universal interatomic potentials, fine-tuned based on nearly 12,000 DFT-relaxation trajectories from our initial generated materials. Subsequently, employing graph neural networks, we predict – instead of calculate - the properties in the C2DB workflow. We envision this comprehensive ML database to act as a foundation for high-throughput screening studies, facilitating the identification of promising candidate materials for further exploration using the ab initio DFT workflow on a select subset of intriguing candidates.

References

[1] P Lyngby, KS Thygesen, npj Computational Materials 8 (2022) 232
[2] P Lyngby, KS Thygesen, (2024) arXiv:2402.02783
[3] T Xie, (2021) arXiv:2110.06197