Accelerating crystal structure search through active learning with neural networks for rapid relaxations

Stefaan S. P. Hessmann[1, 2], Kristof T. Schütt[1, 2], Niklas W. A. Gebauer[1, 2], Michael Gastegger[1], Tamio Oguchi[3], Tomoki Yamashita[4]

[1]Machine Learning Group, Technische Universität Berlin, Straße des 17. Juni 135, 10587 Berlin, Germany

[2]Berlin Institute for the Foundations of Learning and Data, Technische Universität Berlin, Marchstraße 23, 10587 Berlin, Germany

[3]Center for Spintronics Research Network, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka, Japan

[4]Department of Electrical, Electronics and Information Engineering, Nagaoka University of Technology, 1603-1 Kamitomioka-machi, Nagaoka, Niigata, Japan

stefaan.hessmann@tu-berlin.de

Novel crystal structures have allowed significant improvements across various research fields, for example, in the discovery of solar cells [1], catalysts [2], superconductors [3], hardware components [4] and batteries [5]. For a single crystal composition, a vast number of stable structures can be found with each having a unique set of physical properties, such as electrical conductivity, thermal conductivity, magnetism, and optical behavior. Stable crystal structures are local minima on the potential energy surface (PES) of the respective composition, and the number of possible stable structures increases exponentially with the number of atoms per cell [6]. As a result of the wide variety of possible stable structures and their physical properties, crystal structure search is an important challenge, with global energy optimization being the most fundamental task.

Many computational methods exist for crystal structure search, typically involving the relaxation of candidate structures to local minima on the PES. These relaxations require evaluating energies, forces, and stress at each step, often relying on computationally expensive density functional theory (DFT). The simplest crystal structure search method is ab initio random structure search (AIRSS) [7], where the structures of the candidate pool are randomly selected for evaluation with structure relaxation. Alternative approaches, such as Bayesian optimization [8], evolutionary algorithms [9], and particle swarm optimization [10] improve efficiency but still depend on full DFT structure relaxations and do not use generated labels for speeding up the relaxation trajectories itself. Machine learning force fields (MLFFs) on the other hand, including neural networks and kernel-based models, can learn the PES of target systems and therefore dramatically reduce computational costs. When trained on a sufficiently large dataset, machine learning models have proven to be useful tools for structure relaxation [11]. However, in many applications, including crystal structure search, suitable datasets for training the MLFFs are often not available and need to be created. To assemble such data sets efficiently while avoiding faulty MLFF predictions for structures that are very different from the training data, active learning has been successfully applied to MLFF training [12]. There also exist some generative approaches that do not rely on the evaluation of candidate structures but instead use diffusion models [13] that are trained on large databases of stable materials such as MaterialsProject [14].

In this work, we propose an iterative high-throughput virtual screening approach to global optimization in crystal structure search based on active learning and structure relaxations of large candidate pools with neural network MLFF ensembles. The key component of our method is a neural network ensemble that accelerates structure relaxations, selects new training data points towards the region of interest, finds low-energy clusters in the candidate pool, and finally provides a stopping criterion to measure convergence. Our process is initialized with unlabeled data, sampling a fixed number of structures for DFT labeling in each active learning cycle and we apply random perturbations to training structures to circumvent bad training performance on symmetric structures. Using ensembles of MLFFs to measure the uncertainty of predictions, which is needed for our active learning algorithm, does not pose any further limitations to the choice of model architecture and therefore allows for any model to be used, as long as it is able to accurately learn the PES of interest. Especially with the increasing availability of pre-trained models and datasets [15], a flexible choice of model architectures allows one to start from trained models and, for this reason, possibly further reduce the need for expensive DFT calculations during active learning. Furthermore, our method enables straight-forward parallelization of computationally demanding tasks, making it efficient for use on high-performance computing.

We evaluate our method for global optimization of Si_{16}, Na_8Cl_8, Ga_8As_8, and Al_4O_6. Here, we reduce the computational effort by up to two orders of magnitude compared to approaches without active learning. We show that the method is capable of finding multiple relevant local minima in the low-energy region of Si_{16} without a large increase in computational effort, compared to only finding the global energy minimum. Finally, we demonstrate the transferability of the neural network ensembles to relatively larger systems by training the models only on the smaller structures of Si_{16} and Al_4O_6 and using them for structure search of Si_{46} and $Al_{16}O_{24}$, respectively. This additionally diminishes the computational expense of DFT calculations throughout the active learning phase.

References

[1] Green, M. A., Ho-Baillie, A., & Snaith, H. J., Nature Photonics, 8 (2014) 506–514.
[2] Nørskov, J. K., Bligaard, T., Rossmeisl, J., & Christensen, C. H., Nature Chemistry, 1 (2009) 37–46.
[3] Bednorz, J. G., & Müller, K. A., Zeitschrift für Physik B Condensed Matter, 64 (1986) 189–193.
[4] Liu, C. et al., Nature Nanotechnology, 15 (2020) 545–557.
[5] Mizushima, K., Jones, P., Wiseman, P., & Goodenough, J. B., Materials Research Bulletin, 15 (1980) 783–789.
[6] Stillinger, F. H., Physical Review E, 59 (1999) 48.
[7] Pickard, C. J., & Needs, R., Physical Review Letters, 97 (2006) 045504.
[8] Yamashita, T. et al., Physical Review Materials, 2 (2018) 013803.
[9] Oganov, A. R., & Glass, C. W., The Journal of Chemical Physics, 124 (2006).
[10] Wang, Y., Lv, J., Zhu, L., & Ma, Y., Physical Review B, 82 (2010) 094116.
[11] Unke, O. T. et al., Nature Communications, 12 (2021) 7273.
[12] Vandermause, J. et al., npj Computational Materials, 6 (2020) 1–11.
[13] Zeni, C. et al., ArXiv preprint arXiv:2312.03687 (2023).
[14] Jain, A. et al., APL Materials, 1 (2013) 011002.
[15] Merchant, A. et al., Nature, (2023) 1–6.

Figures

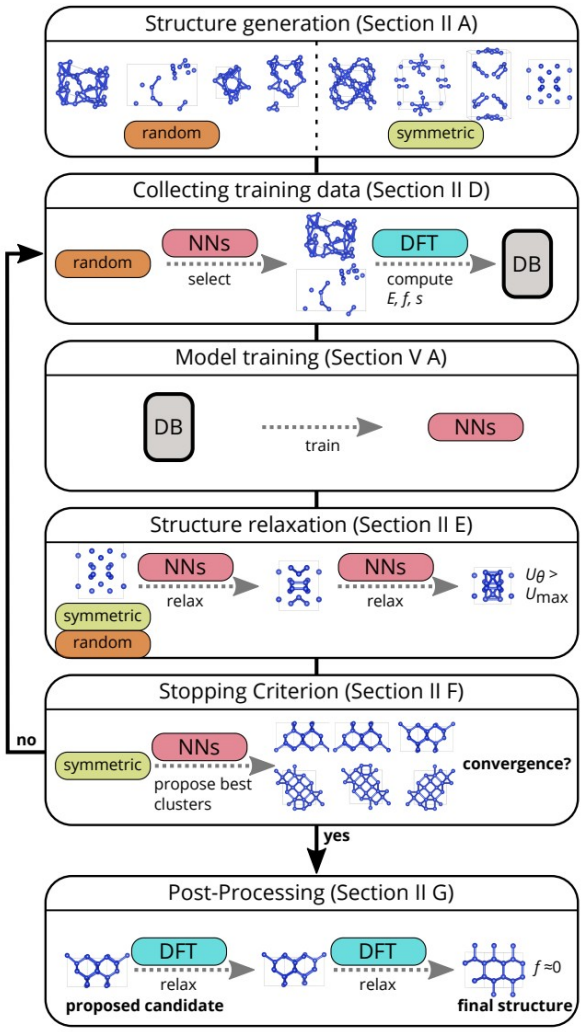

Figure 1. Schematic overview of the method. Initially, a pool of candidate structures is generated based on random or symmetric generation. Next, the active learning cycle is started by the selection of promising datapoints and by training the neural network ensembles. These are then used to accelerate structure relaxations for all structures in the candidate pool. At the end of the active learning cycle, the neural network models propose the most promising clusters of low-energy structures and a stopping criterion is applied. In the case of convergence, the most promising structures are further validated through structure relaxation with DFT.