Efficient Autoencoder Pipeline for Discovering High Entropy Alloys with Molecular Dynamics Data

Amirhossein D. Naghdi*
NOMATEN Centre of Excellence, Poland
Amirhossein.Naghdi@ncbj.gov.pl

Grzegorz Kaszuba*
IDEAS NCBR, Poland
Poznań University of Technology, Poland
grzegorz.kaszuba@ideas-ncbr.pl

Stefanos Papanikolaou
NOMATEN Centre of Excellence, Poland
stefanos.papanikolaou@ncbj.gov.pl

Andrzej Jaszkiewicz
Poznań University of Technology, Poland
andrzej.jaszkiewicz@cs.put.poznan.pl

Piotr Sankowski
University of Warsaw, Poland
MIM Solutions, Poland
piotr.sankowski@mim.ai

Abstract

In this work, we utilize computationally efficient Molecular Dynamics (MD) simulations to create a machine learning pipeline for discovery of crystalline multi-component alloys. We employ high-quality interatomic potentials to create a dataset of NiFeCr structures and apply Crystal Diffusion Variational Autoencoder (CDVAE) to maximize their mechanical properties, i.e. bulk modulus. As part of the experiment, we utilize local search coupled with classical interatomic potentials to explore the local structure space and show that utilization of this procedure greatly improves optimization capability of the neural model. We also expand the model with an extra submodule, which attains 42% improvement on modeling the crystalline phase of the structures. Ultimately, we verify the global stability of the created structures with quantum mechanical calculation methods.

1 Introduction

Discovery of novel materials involves both gathering experimental data and *in silico* modeling. Developments in machine learning largely benefit materials science, as neural networks have been applied to processing structure graphs [Gasteiger et al., 2022b,a, Choudhary and DeCost, 2021]. The amount of publicly available crystal structure data is growing rapidly [Jain et al., 2013, Kirklin et al., 2015]. Material structure databases are of significance for the development of the field, as the high-fidelity ab-initio methods such as Density Functional Theory (DFT) simulations bear large computational cost. To avoid the computational bottleneck, especially in the prototyping stage of design, one might resort to more efficient numerical solutions, such as Molecular Dynamics (MD) simulations. MD employs numerical solutions to Newton's equations, facilitating the analysis of the dynamic evolution of atoms by treating them as interacting particles. MD simulations depend on the precise estimation of forces between atoms, acknowledging the intricate nature of atomic interactions in reality. In MD simulations, interatomic potentials, and mathematical formulations employed to approximate force fields, play a crucial role. Notably, recent advancements have explored the use of

*These authors contributed equally.

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

neural networks to enhance the accuracy of these interatomic potentials [Lot et al., 2020, Batzner et al., 2022]. This work focuses on leveraging MD in conjunction with variational autoencoders to streamline the optimization of material properties, particularly in the case of multi-component alloys.

The experimental discovery of alloys containing multiple elements is inherently challenging due to the vast amount of possible compositional and structural combinations. Only two decades ago, a novel alloying technique emerged, proposing to mix multiple chemical elements in relatively high concentrations - the emergent structures are known as high-entropy alloys (HEAs)[Yeh et al., 2004, Cantor et al., 2004]. This innovation has given rise to a new class of materials with exceptional properties, e.g., corrosion resistance or type-II superconductivity [Wu et al., 2016, Xiao et al., 2023]. Consequently, there exists a multitude of potential HEAs, each differing in the number of unique constituent chemical element types (e.g. binary, ternary, quaternary, and quinary alloys) and their respective weight percentages, making them a fascinating subject for study. The complexity of the problem is further heightened by the potential existence of multiple *crystalline phases* (CP) in a given HEA with a specified composition. In materials science, crystal phase refers to the specific arrangement of atoms in a crystalline material, determining its unique structural properties. Due to the challenging and time-consuming nature of discovering novel HEAs with exceptional properties through experimentation and DFT, an alternative approach could involve utilizing generative deep learning.

Usage of diffusion variational autoencoders for material generation has been proposed by Xie et al. [2022], who introduced the Crystal Graph Diffusion Variational Autoencoder (CDVAE), and later followed by Zeni et al. [2024]. Previously to Xie et al. [2022] variational autoencoders have been successfully used to process molecular graphs Gebauer et al. [2019], which, however, are not periodic and thus allow for a less structured approach. In essence, neural models designed for crystal structure processing must consider the periodic nature of crystals. In practical terms, a segment of the crystal is simulated under the assumption that it repeats indefinitely along three directions in space. The modeled region is commonly referred to as a "supercell". CDVAE has been successfully used to explore the composition space and discover new structures in the context of two-dimensional and superconducting materials [Wines et al., 2023, Lyngby and Thygesen, 2022]. However, searching

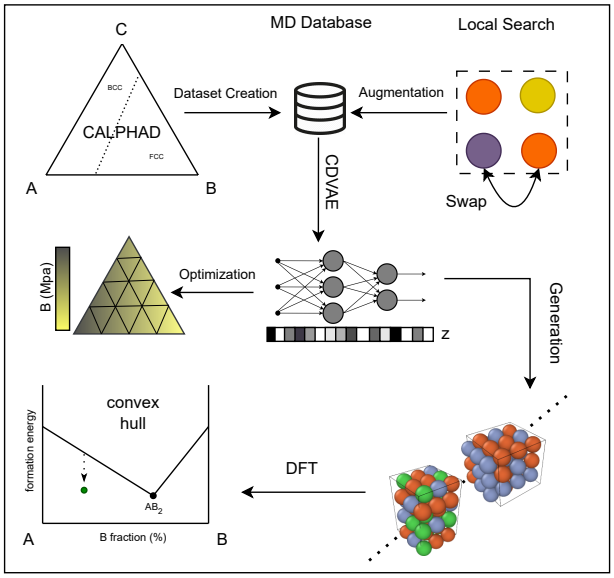


Figure 1: Composition search workflow. The training dataset is created using classical interatomic potentials of NiFeCr supercells initialized with random coordinates. A local search procedure is coupled with MD simulations to additionally include training examples created in an informed way that emphasizes the property of interest. After the model is trained, it's used to optimize the examples from the test set, or, alternatively, create completely new structures when seeded with randomly sampled vectors. Finally, DFT simulations are conducted to verify that the stability of produced structures is in line with the literature.

for HEAs requires us to consider the CP of the material in hand. A computational approach for identifying the CP of a HEA across various compositions is known as CALPHAD (Calculation of Phase Diagrams). For instance, in a HEA such as NiFeCr with three chemical elements, CALPHAD determines the stable crystalline phase as the composition varies in terms of the percentage of each constituent element. In the case of NiFeCr, the material benchmarked in this paper, CALPHAD indicates that it should exhibit either a Face-Centred Cubic (FCC) or Body-Centred Cubic (BCC) CP [Wu et al., 2017, Lee et al., 2001]. Hence, we must take a distinct approach to composition search within a HEA with specific constituent element types.

In this paper, we apply CDVAE in a workflow that involves obtaining data from classical interatomic potentials, elaborated in Fig.1. In the specific case of NiFeCr, we generated an MD-based dataset comprising various compositions, including binary alloys, with CP determined using the CALPHAD method [Wu et al., 2017, Lee et al., 2001]. This dataset was subsequently utilized to train the CDVAE model.

Next, we enhanced the CDVAE model by incorporating a fully connected neural network for CP classification - the resulting model is called P-CDVAE standing for Phase-CDVAE. This adjustment improved considerably the denoising task's performance and, as a result, better modeling of structures' CP. Afterwards, the model was used to explore the ternary composition space and generate new materials. The stability of the newly generated materials was assessed by calculating their formation energies using DFT. The results were in great agreement with other DFT-based methods aimed at the same study [Wróbel et al., 2015].

Ultimately, we developed a local search framework that, in essence, seeks the optimal value of property within a given configuration of specific composition by manipulating atomic ordering and utilizing feedback from MD simulations. This method was employed to provide training data of the initial training dataset by incorporating the optimized mechanical properties, i.e. bulk modulus, for each data point in the test set identifying the ultimate optimized configuration regarding stability and mechanical properties. In summary, our primary contributions to this study are as follows:

- We applied CDVAE to a machine-learning workflow for generating crystal structures and optimizing their properties. Our approach utilizes efficient classical interatomic potentials to produce training data and assess the adequacy of sampled structures.
- We generated a dataset for ternary NiFeCr alloys, encompassing various atomic orderings of given compositions, and used CDVAE to discover new structures. Ultimately, we assessed the global stability of the generated structures with DFT simulations.
- We enhanced the exploration of the optimization space by crafting improved structures by local search coupled with MD. This facilitated optimization and enabled the model to enhance the bulk modulus of 59% examples to be higher than the 90th percentile of the features among those crafted structures.
- We improved the reconstruction performance of the autoencoder by training it to predict the crystal phase of structures before sampling. This reduced the distance function that evaluates the modeling of crystalline phase geometry by up to 42%.

2 Related Work

Geometric Graph Neural Networks for molecular graphs. In recent years, several Graph Neural Networks (GNNs) have been developed with a specific emphasis on processing molecular graphs. These models employ graph geometry, representing atoms as nodes and their interatomic distances as edges, to generate various types of embeddings. For instance, some models focus on edge lengths, such as Schnet [Schütt et al., 2017] and Physnet Unke and Meuwly [2019], while others consider both edge lengths and angles between them, e.g. DimeNet [Gasteiger et al., 2022b], GemNet-T [Gasteiger et al., 2022a], and AliGNN [Choudhary and DeCost, 2021]. In its advanced version, GemNet-Q [Gasteiger et al., 2022a] goes a step further by incorporating torsion angles, which correspond to triplets of consecutive edges, to create positional embeddings. The autoencoder used in this work utilizes DimeNet++ as the encoder and GemNet-T as the decoder Xie et al. [2022].

Diffusion models for chemical structures Generative models have been applied to the materials science context extensively [Hoffmann et al., 2019, Long et al., 2021, Ren et al., 2022, Zhao et al., 2021]. Diffusion autoencoders are a class of generative models that have shown great performance at

generating structured data by iteratively refining them [Dhariwal and Nichol, 2021, Cai et al., 2020, Shi et al., 2021]. They manage various degrees of noise by optimizing a score function that responds to the distortion of data, e.g. through the physics-inspired Langevin dynamics [Song and Ermon, 2019]. The interpretations of score functions may vary, depending on the specifics of the problem and experimental decisions: a likelihood-based approach utilizes the probability distribution modeled by the network to maximize the likelihood of denoised data. CDVAE utilizes direct prediction of forces acting. There are various approaches to modeling chemical entities with diffusion models: E(3) Equivariant Diffusion Model, for example, uses the notion of likelihood [Hoogeboom et al., 2022].

Quantum mechanical material discovery. The quest for stable materials with exceptional properties poses a significant challenge when exclusively relying on quantum mechanical methods [Oganov et al., 2019]. Efforts have been made to tackle this challenge through various approaches, including evolutionary algorithms[Wang et al., 2012, Glass et al., 2006], random sampling[Pickard and Needs, 2011], and the exploration of substitutional alloying within already established stable materials [Hautier et al., 2011]. Among these methods, cluster expansion (CE) [Wu et al., 2016] techniques utilize quantum mechanical calculations to identify ground-state stable structures in medium-entropy alloys (MEAs), and they are typically applied to alloys containing up to three components. However, it's worth noting that these methods may struggle with multi-component complex materials and the presence of multiple crystalline phases and effect of configurational entropy [Wróbel et al., 2015, Anand et al., 2016] – which is negligible below the room temperature. In contrast, generative models offer a faster and more efficient approach when exploring the vast compositional space of materials.

3 Methods

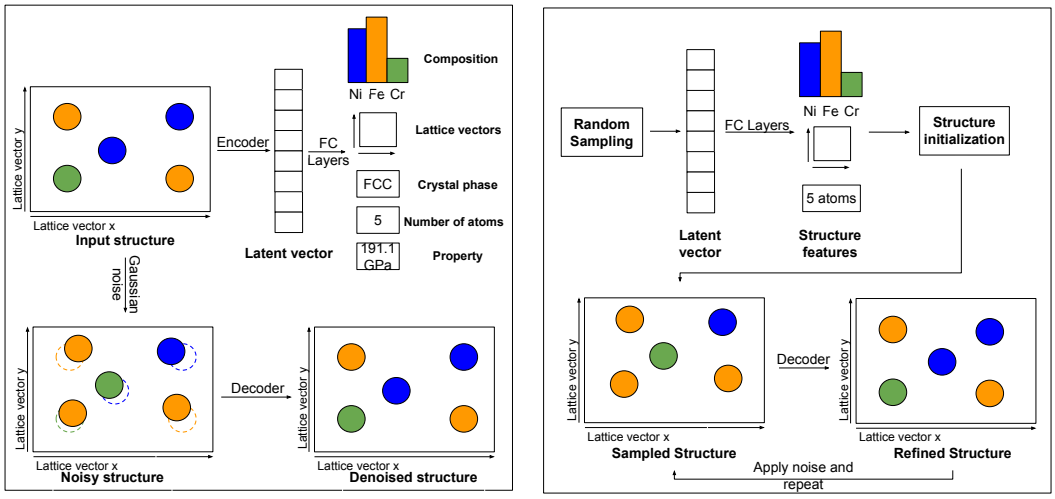
3.1 CDVAE model

CDVAE. Our approach to searching for new NiFeCr alloys revolves around the use of CDVAE [Xie et al., 2022]. The model is composed of DimeNet++ [Gasteiger et al., 2022b] as the encoder and GemNet-dT [Gasteiger et al., 2022a] as the decoder. As shown in Fig. 2**(a)**, the model is trained on denoising task - the input structure is first passed through the encoder, then it undergoes a random perturbation in terms of node coordinates and atom types. The loss function includes the distance the nodes have moved from their assigned position, the discrepancy in atom types between the original structure and the reconstructed one, and Kullback-Leibler divergence. The decoder produces two outputs for each element. Firstly, the score function that governs the denoising is computed. This output is treated as a set of forces that act upon the atoms, and their coordinates are adjusted accordingly. Secondly, the atom types are predicted - this way, the model can correct atom-type errors created by the noise. Creation of new structures (sampling), involves choosing a random feature vector and using the decoder to iteratively refine the atoms' positions and types with annealed Langevin Dynamics [Song and Ermon, 2019]. The key feature of this approach is that, apart from the informed changes made by the model, Gaussian noise is added with every step. As the magnitude of the injected noise is gradually decreased, the process approaches convergence.

Apart from the two GNNs, the model is equipped with a couple of fully connected (FC) submodules that all take the latent vector as the input. As they act on the structure's scalar representation, they predict different structure-level features before the decoder resolves the atom-level features - coordinates and atom types. The remaining components of the loss function are associated with errors made by individual FC submodules. Those serve several purposes:

Initialization of sampled structures. Langevin dynamics is conducted within a supercell of a fixed size and on a fixed number of atoms. The sampling step of the model is depicted in Fig.2**(b)**. To initialize the structure for denoising, three FC modules are used: *FC-n-atom* predicts the number of atoms in a structure, *FC-lattice* predicts the lattice lengths of lattice vectors and the angles between them, while *FC-composition* provides a prior estimation of the content of each element in the structure. With the aforementioned estimates available, initialization is a straightforward process of sampling a known number of atoms: each of them is assigned a random position within the supercell space; atom types are chosen randomly from the distribution provided by *FC-composition*.

Property optimization. *FC-property* predicts the target property from the feature vector. Having trained it, one can use a gradient descent algorithm, e.g. Adam [Kingma and Ba, 2015] to modify the



(a) (b)

Figure 2: **(a)** The training step of CDVAE is based on single-step denoising of randomly perturbed structures. The loss function is applied to the outputs of the decoder: coordinates and atom types. The loss function also includes all outputs of the fully connected layers and Kullback-Leiber divergence. **(b)** Sampling by CDVAE with Langevin dynamics. The structure initialization procedure has been shown.

feature vector with the target of optimizing a property of choice, while the model's parameters are frozen.

Improving the modeling. We propose to equip CDVAE with an additional layer: *FC-phase*. The purpose of this layer is to provide additional knowledge about the crystalline phase of the structure. As was shown in Fig. SM. 2(b), NiFeCr structures form purely single-phase (either FCC or BCC) metallic crystals, based solely on the elementary composition. We train the model to explicitly predict whether the structure is FCC or BCC. We add this information to the decoder's input: the predicted phase is combined with types of individual atoms when passed to the decoder's embedding layer, e.g. we allow the decoder to produce a different embedding for "nickel atom in a BCC structure" and "nickel atom in an FCC structure". This addition is aimed at helping the model assess the correct relative positions of atoms that depend predominantly on the crystal phase. From a graph modeling perspective, it's a high-level description of edge lengths and angles expected to be seen in the decoded structure. The variant of CDVAE equipped with *FC-phase* will be referred to as P-CDVAE.

3.2 Dataset Creation

The dataset consists of NiFeCr crystal structures, as well as binary combinations: NiFe, FeCr, FeCr. All of them are either Face-Centered Cubic (FCC) or Body-Centered Cubic (BCC) structures. To represent a structure of either kind, one needs a minimum of a cell of 2-atoms, repeated periodically and creating the final "supercell". To capture a larger region within the structure (a large enough supercell), one can include multiple unit cells within - we opt to use $3 \times 3 \times 3$ unit cells within a supercell, which amounts to 54 atoms. The benefit of modeling larger supercells is twofold - it can express more complex structures, while also allowing to search for different elementary compositions at a finer resolution. We use classical interatomic potentials to create the dataset: first, we choose a range of elementary compositions. Next, for each of them, we create 20 structures with random orders of atoms within. The ordering of atoms in HEAs plays a crucial role in the properties of those materials. The dataset created this way will be referred to as "NiFeCr-Random". We also create an analogous dataset of $2 \times 2 \times 2$, named "NiFeCr-Small" for analysis. Details of MD and DFT calculations are presented in the appendix section.A.6. The dataset generation process using MD simulations is notably efficient regarding runtime. Although it involves numerous calls to the MD package, it can still produce a dataset of 1,200 configurations in approximately two hours. In contrast, generating the same dataset with DFT would require 3 to 4 hours per configuration on an A100 80GB

NVIDIA GPU for a self-consistent calculation only. For calculating mechanical properties, such as elastic constants, DFT would be nearly ten times more time-consuming.

Crafting examples with local search. The dataset comprises structures obtained in MD simulations with randomly chosen ordering for different atom types. Mechanical properties of alloys are largely affected by chemical short-range ordering of atoms [Zhang et al., 2020, Naghdi et al., 2023]. We hypothesize that the short-range patterns that contribute to the highest bulk modulus consist of local structures that can't be easily obtained at random. To provide the model with such examples, we extended the dataset with structures crafted in an informed way. For that, we implemented a local search that involves swapping positions of pairs of atoms to optimize the bulk modulus with feedback from MD for each considered configuration. During each iteration, 20 random atom transpositions are tried, each followed by relaxation of the structure and computation of the bulk modulus.

For the next iteration, the structure with the highest value of this property is chosen, among the ground structure and the created ones. Experiments have been conducted on 3 datasets. NiFeCr-LS is a dataset which, apart from all structures from NiFeCr-Random, includes structures that have undergone 20 local search iterations as described: 10% of examples from NiFeCr-Random's training, validation, and test examples were randomly chosen, optimized this way, and added to their respective partitions of the data. The changes in the mean property during the local search are shown in Fig. SM. 3.

4 Experiments

4.1 Reconstruction

We evaluate the model's capability to encode and reconstruct the structures. This task provides valuable insights into the deterministic nature of the sampling process. We use *pymatgen*'s *StructureMatcher* class. We use error tolerances of *StructureMatcher* originally used in Xie et al. [2022]. We also test Cartesian distance between structure fingerprints produced by *pymatgen*'s *CrystalNN* tool, which we deem useful as they give a measure of similarity without the need to correctly match the structure. First, we compare the reconstruction performance of CDVAE to the Carbon dataset, investigated in Xie et al. [2022]. The match rate is higher than in the original experiment, which can be attributed to the double number of steps we use in Langevin dynamics. We also asses reconstruction performance of NiFeCr-Small, which is a variant with a similar number of atoms in a supercell (up to 24 for carbon and 16 for NiFeCr-Small), and NiFeCr-Random, which possesses significantly larger structures. The results can be seen in Tab.1. Committing to larger structures makes the match unlikely and thus, the sampling procedure is not deterministic. The results also point to a qualitative difference between the NiFeCr datasets and Carbon. Fingerprint distances suggest that, on average, the reconstructions of NiFeCr structures are composed of sites with similar local neighbourhoods. This, however, does not translate to high match rates. We hypothesize that the relative position and orientation of those local clusters are different, which inevitably compromises the global match, especially for NiFeCr-Random. Nonetheless, even for NiFeCr-Small, it could be attributed to permutational freedom of NiFeCr structures. The dichotomy between the local action of variational autoencoders and the global integrity of reconstruction has been pinpointed byXie et al. [2022].

Table 1: Reconstruction scores of CDVAE on Carbon, NiFeCr-Small and NiFeCr-Random datasets.

DATASET	MATCH RATE	FP DIST
CARBON	66.2%	0.2603
NIFECR-SMALL	30.0%	0.1758
NIFECR-RANDOM	6.4%	0.1358

Next, we compare the models trained on datasets composed of 54-atom structures: either NiFeCr-Random or NiFeCr-LS. For brevity, we refer to the experiments by the type of the model and data - e.g. P-CDVAE-LS is a P-CDVAE trained on NiFeCr-LS. The match rate is disregarded in comparison. Instead, we aim to evaluate the reconstruction of the geometry under which atoms are distributed, aside from the elements that those atoms represent. For that, we use G-vectors. G-vectors are atomic fingerprints that encode the local environment with regard to both the types of nearest atoms and

geometric features - lengths of the edges and angles between them. Based on G-vectors, we compute G-distances between two structures, which is explained in detail in the appendix section.A.5. We also introduce CP G-distances by disregarding the atomic types during the computation of G-vectors - those descriptors match the geometric environment of individual atoms. The reconstruction of the models trained on 54-atom NiFeCr structures has been compared with regard to *CrystalNN* fingerprint distance and both G-vector distances (Tab.2). Notably, both P-CDVAE models showed significant improvement in CP G-distance compared to CDVAE: 29% and 42%, respectively. This suggests an improvement in the modeling of the crystal phase.

Table 2: Reconstruction scores of CDVAE trained on NiFeCr-Random and instances of P-CDVAE trained on NiFeCr-Random or NiFeCr-LS. Reconstruction is evaluated against the common part of the test set, which is the test set of NiFeCr-Random. The metrics used were Fingerprint distance, G-distance, and CP G-distance.

MODEL	FP DIST	G-DIST	CP G-DIST
CDVAE-RANDOM	0.1358	0.3618	0.0984
P-CDVAE-RANDOM	0.1039	0.3513	0.0691
P-CDVAE-LS	0.0823	0.3521	0.0560

4.2 Material Generation

The P-CDVAE model succeeded in achieving similar results compared to pure quantum-based calculations [Wróbel et al., 2015] for the generation task. The success of this task depends on various factors, including the performance of the generative model framework and the reconstruction task, among others. As a result, we conducted an evaluation of both CDVAE and P-CDVAE in the reconstruction and generation task to determine their respective performance, ultimately choosing generated materials from P-CDVAE for subsequent DFT validation due to its enhanced capability in denoising the structures within their ground truth CP (more details in appendix section. A.3). In order to achieve this, latent vectors from a multidimensional Gaussian distribution were utilized to query the CDVAE and P-CDVAE models. This process resulted in the generation of 500 structures, encompassing a significant portion of both binary and ternary alloy composition spaces.

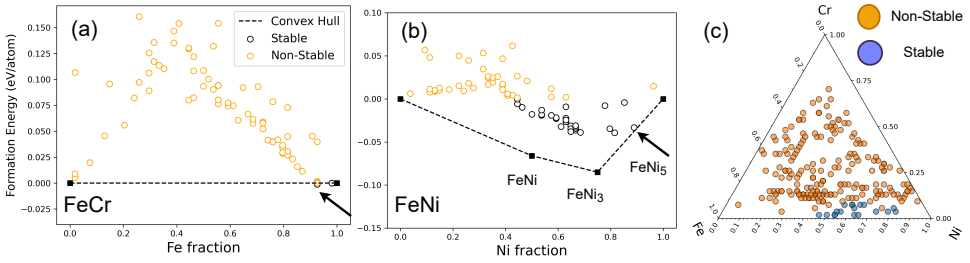


Figure 3: Stability of generated materials validated through DFT. (a-b) Show the generated materials for different FeCr/FeNi alloys and their stability determined with their formation energy relative to the convex hull. Arrows indicate the new stable materials on/below the convex hull which were not in the dataset before. (c) Similarly illustrates the stable ternary alloys within the composition space. The results are similar to the pure quantum-mechanical calculations [Wróbel et al., 2015].

The distribution of generated structures in terms of their composition is shown in Fig. SM.1(c). Afterwards, we did spin-polarized DFT calculations, and based on the formation energy results, we found great agreement with the cluster expansion method (CE), which is purely based on quantum mechanical calculations, in terms of the stability of the materials for NiFeCr structures [Wróbel et al., 2015]. Both P-CDVAE and CE agree on the stable compositions of ternary NiFeCr and their binary alloys. Fig. 3(a-b) shows the convex hull plot for binary FeCr/NiFe alloys in the dashed line. In materials science, a convex hull (of stable structures) is a representation of the most thermodynamically stable compositions within an alloy system. It is determined by plotting the formation energies, calculated using high-fidelity methods like DFT, of known stable materials

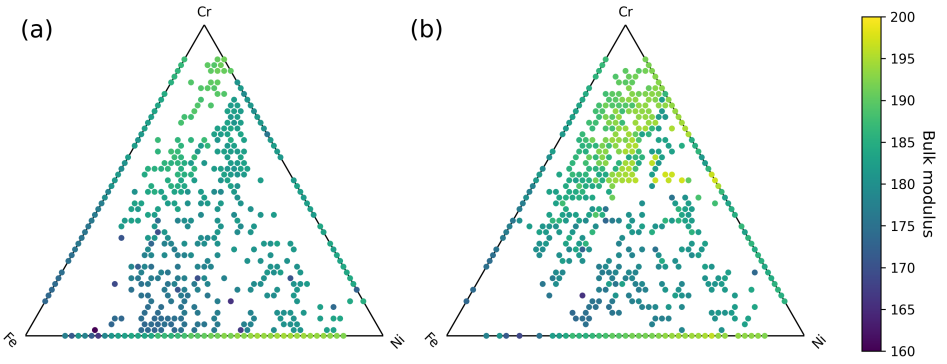


Figure 4: The highest values of bulk modulus of structures generated during optimization, by composition. (a) Shows the results for the model trained on NiFeCr-Random, (b) the results for the model trained on NiFeCr-LS.

against the alloy's composition, providing insights into optimal alloy compositions for stability. For example, in the case of NiFe, P-CDVAE generated a new stable material ($FeNi_5$) with formation energy laying on the convex hull and many meta-stable materials (which could be stable under special thermodynamical situations such as high pressure and temperature) with a negative formation energy. all materials with a positive formation energy are considered unstable. The same goes with ternary NiFeCr (Fig. 3(c)), where our model found stable materials (blue dots) in the ternary composition space. This benchmarks the method designed in this work promising for the study of higher than three component alloys.

4.3 Optimization

We employed the models we obtained to optimize the bulk modulus. Two previously tested instances of P-CDVAE were used - one trained on NiFeCr-Random and one trained on NiFeCr-LS. To provide the same experimental conditions for both models, we used the test partition of NiFeCr-Random (composed of 327 structures) as the initial optimization points. The structures from the test set were encoded and the obtained latent vectors were optimized with the Adam algorithm (with a learning rate 0.001). Each step of optimizing a latent vector involves querying the *FC-property* network and using back-propagation to maximize the bulk modulus. The encodings were used to sample new structures at several optimization stages: after 200, 500, 1000, 1500, and 2500 steps. The bulk moduli of sampled structures were assessed with Molecular Dynamics simulations, as described in section 3.2. MD simulations were also employed to rule out the incorrectly denoised structures - we require that, after the simulation converges, per-atom energy is not higher than the threshold set by the highest per-atom energy found in the training data. Additionally, we require that the output of the neural model is already close to convergence - the structure's total energy can be at most 10 eV above the final threshold. This is to ensure that structures passed to MD are already close to the relaxed state - otherwise, passing structures that do not meet these criteria may cause MD simulations that do not converge, or, having reached the maximum number of iterations, yield unrelaxed structures and return anomalous values of the target property. As encodings are optimized with a simple goal to maximize the prediction of a fully connected network, the moment when most of them cease to translate into structures that fit these criteria marks the end of the experiment.

Results. Tab. 3 shows the amount of correctly decoded structures, as well as the values of elastic moduli present in the population of optimized structures at consecutive stages - those include the mean bulk modulus among all correct structures and the mean bulk modulus of 10 best structures present in the population. P-CDVAE, trained on NiFeCr-Random showed little capability to optimize the bulk elastic modulus (Tab.3). The inclusion of examples crafted with local search during training allowed for more effective optimization by the model trained on NiFeCr-LS. Tab.A1 shows the percentage of initial structures whose optimization achieved bulk modulus of value thresholds during one of the sampling steps. For the model trained on NiFeCr-Random, the thresholds were based on the distribution of the training part of NiFeCr-Random. For the model trained on NiFeCr-LS, the thresholds were set only by those training examples, which were the product of local search, making

Table 3: Results of property optimization with **(a)** the model trained on the dataset of randomly initialized NiFeCr structures and **(b)** dataset with addition of structures. Latent vectors that describe structures have been optimized with the Adam algorithm. At the given steps of this process, they were utilized to produce crystal structures. For each of those steps, the number of correctly decoded structures is shown, along with the statistics of their properties. The mean value of the property is shown for all correctly decoded structures and for 10 structures with the highest value of the property.

(a) NiFeCr-Random

STEP	MEAN B	MEAN B (TOP 10)	CORRECT
INITIAL	181.7	192.0	327
200	181.4	193.1	237
500	181.7	193.3	225
1000	181.1	193.8	178
1500	182.8	193.8	129
2500	185.8	193.3	45

(b) NiFeCr-20

STEP	MEAN B	MEAN B (TOP 10)	CORRECT
INITIAL	181.7	192.0	327
200	181.7	192.9	265
500	182.4	193.0	247
1000	185.4	194.4	199
1500	189.8	195.0	186
2500	193.0	196.5	124

for a more demanding baseline. Nevertheless, the percentage of structures optimized above the thresholds was higher for this model. The model trained on NiFeCr-LS also found the new maximum value of the bulk modulus, 197.26 GPa, compared to the highest one obtained during local search: 196.58 GPa. Fig. 4 shows the highest property values of property for each composition obtained during optimization for both models. The model trained on NiFeCr-LS generated structures with high bulk modulus across various compositions.

5 Conclusions

In this work, a computationally efficient pipeline for generating crystal structures has been proposed. The experimental procedure employed classical interatomic potentials to generate the data and assess the quality of the outputs yielded by the autoencoder. MD simulations were used to relax the created structures further and predict the bulk modulus. The feasibility of the proposed experimental procedure depends on the availability of interatomic potentials for the structures considered. The development of new interatomic potentials is a trend that benefits the applicability of MD. As part of the data creation procedure, local search was used to craft examples in an informed way. This was done by changing the order of atoms to maximize the target property. The addition of 10% of structures crafted with local search to the base dataset proved crucial for the optimization efficiency. This speaks to the importance of conditioning data to convey knowledge relevant to the problem at hand.

NiFeCr structures form two distinct crystalline phases. Apart from that, the template of the structure doesn't vary significantly. The prior estimation of crystal phase and its use in diffusion, allowed for a better reconstruction of this template. The reconstruction of 54-atom structures was not faithful enough to match the reconstructed examples to their ground truth counterparts with the *StructureMatcher* Ong et al. [2013]. Even for 16-atom structures, faithful reconstruction posed a significant challenge. Therefore, it must be noted that the sampling performed by the model was not a deterministic process. Nonetheless, neural optimization was successful. The reconstruction scores achieved by the model suggest that the discrepancies between the encoded and decoded structures stem mostly from the permutation of atoms inside, while the spatial distribution of nodes is accurate. It can be hypothesized that the results could be improved by employing changes that facilitate faithful, global reconstruction of the atomic order. It is especially challenging for a diffusion model, whose principle is to model the structure bottom-up from local interactions Xie et al. [2022]. To better preserve the input structure, one could experiment with the part of CDVAE responsible for the assignment of atom types to every node or ensure that the encoding is descriptive enough to reconstruct it by increasing the size of the dataset, the length of the latent vector, and the length of interactions modeled by the employed GNNs. It is possible that for optimal reconstruction of alloy supercells with the proposed size, a hierarchical graph autoencoder would be more appropriate than a diffusion model.

Code Availability

The code for training the P-CDVAE model and performing local search is available at https://github.com/grzegorzkaszuba/allogen_cdvae. Additionally, the code for generating datasets based on molecular dynamics can be found here: https://github.com/Amirhossein4131/AlloGen.git.

Acknowledgements

N.D.A and P.St were supported by the European Regional Development Fund through the Foundation for Polish Science (grant No. MAB PLUS/2018/8) and the Ministry of Science and Higher Education (agreement No. MEiN/2023/DIR/3795), with additional support from the European Union Horizon 2020 program (Grant Agreement No. 857470). K.G was supported by IDEAS NCBR. J.A was supported by the Polish Ministry of Science and Higher Education and S.Pi was partially supported by ERC consolidator grant TUgbOAT no 772346 and NCN no 2020/37/B/ST6/04179.

References

G. Anand, R. Goodall, and Colin L. Freeman. Role of configurational entropy in body-centred cubic or face-centred cubic phase formation in high entropy alloys. *Scripta Materialia*, 124: 90–94, 2016. ISSN 1359-6462. doi: https://doi.org/10.1016/j.scriptamat.2016.07.001. URL https://www.sciencedirect.com/science/article/pii/S1359646216303025.

Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P. Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E. Smidt, and Boris Kozinsky. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13(1):2453, May 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-29939-5. URL https://doi.org/10.1038/s41467-022-29939-5.

Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.*, 98:146401, Apr 2007. doi: 10.1103/PhysRevLett.98.146401. URL https://link.aps.org/doi/10.1103/PhysRevLett.98.146401.

P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50:17953–17979, Dec 1994. doi: 10.1103/PhysRevB.50.17953. URL https://link.aps.org/doi/10.1103/PhysRevB.50.17953.

Ruojin Cai, Guandao Yang, Hadar Averbuch-Elor, Zekun Hao, Serge Belongie, Noah Snavely, and Bharath Hariharan. Learning gradient fields for shape generation, 2020.

B. Cantor, I.T.H. Chang, P. Knight, and A.J.B. Vincent. Microstructural development in equiatomic multicomponent alloys. *Materials Science and Engineering: A*, 375-377:213–218, 2004. ISSN 0921-5093. doi: https://doi.org/10.1016/j.msea.2003.10.257. URL https://www.sciencedirect.com/science/article/pii/S0921509303009936.

Kamal Choudhary and Brian DeCost. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7(1):185, Nov 2021. ISSN 2057-3960. doi: 10.1038/s41524-021-00650-1. URL https://doi.org/10.1038/s41524-021-00650-1.

Prafulla Dhariwal and Alex Nichol. Diffusion models beat gans on image synthesis, 2021.

Johannes Gasteiger, Florian Becker, and Stephan Günnemann. Gemnet: Universal directional graph neural networks for molecules, 2022a.

Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs, 2022b.

Niklas W. A. Gebauer, Michael Gastegger, and Kristof T. Schütt. *Symmetry-Adapted Generation of 3d Point Sets for the Targeted Discovery of Molecules*. Curran Associates Inc., Red Hook, NY, USA, 2019.

Colin W. Glass, Artem R. Oganov, and Nikolaus Hansen. Uspex—evolutionary crystal structure prediction. *Computer Physics Communications*, 175(11):713–720, 2006. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2006.07.020. URL https://www.sciencedirect.com/science/article/pii/S0010465506002931.

Geoffroy Hautier, Chris Fischer, Virginie Ehrlacher, Anubhav Jain, and Gerbrand Ceder. Data mined ionic substitutions for the discovery of new compounds. *Inorganic Chemistry*, 50(2):656–663, Jan 2011. ISSN 0020-1669. doi: 10.1021/ic102031h. URL https://doi.org/10.1021/ic102031h.

Jordan Hoffmann, Louis Maestrati, Yoshihide Sawada, Jian Tang, Jean Michel D. Sellier, and Yoshua Bengio. Data-driven approach to encoding and decoding 3-d crystal structures. *CoRR*, abs/1909.00949, 2019. URL http://arxiv.org/abs/1909.00949.

Emiel Hoogeboom, Victor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3d, 2022.

Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A. Persson. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 07 2013. ISSN 2166-532X. doi: 10.1063/1.4812323. URL https://doi.org/10.1063/1.4812323.

Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, San Diega, CA, USA, 2015.

Scott Kirklin, James E. Saal, Bryce Meredig, Alex Thompson, Jeff W. Doak, Muratahan Aykol, Stephan Rühl, and Chris Wolverton. The open quantum materials database (oqmd): assessing the accuracy of dft formation energies. *npj Computational Materials*, 1(1):15010, Dec 2015. ISSN 2057-3960. doi: 10.1038/npjcompumats.2015.10. URL https://doi.org/10.1038/npjcompumats.2015.10.

G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169–11186, Oct 1996. doi: 10.1103/PhysRevB.54.11169. URL https://link.aps.org/doi/10.1103/PhysRevB.54.11169.

G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996. ISSN 0927-0256. doi: https://doi.org/10.1016/0927-0256(96)00008-0. URL https://www.sciencedirect.com/science/article/pii/0927025696000080.

G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B*, 47:558–561, Jan 1993. doi: 10.1103/PhysRevB.47.558. URL https://link.aps.org/doi/10.1103/PhysRevB.47.558.

G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59:1758–1775, Jan 1999. doi: 10.1103/PhysRevB.59.1758. URL https://link.aps.org/doi/10.1103/PhysRevB.59.1758.

Byeong-Joo Lee, Jae-Hyeok Shim, and Hyun Min Park. A semi-empirical atomic potential for the fe-cr binary system. *Calphad*, 25(4):527–534, 2001. ISSN 0364-5916. doi: https://doi.org/10.1016/S0364-5916(02)00005-6. URL https://www.sciencedirect.com/science/article/pii/S0364591602000056.

Teng Long, Nuno M. Fortunato, Ingo Opahle, Yixuan Zhang, Ilias Samathrakis, Chen Shen, Oliver Gutfleisch, and Hongbin Zhang. Constrained crystals deep convolutional generative adversarial network for the inverse design of crystal structures. *npj Computational Materials*, 7(1):66, May 2021. ISSN 2057-3960. doi: 10.1038/s41524-021-00526-4. URL https://doi.org/10.1038/s41524-021-00526-4.

Ruggero Lot, Franco Pellegrini, Yusuf Shaidu, and Emine Küçükbenli. Panna: Properties from artificial neural network architectures. *Computer Physics Communications*, 256:107402, 2020. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2020.107402. URL https://www.sciencedirect.com/science/article/pii/S0010465520301843.

Peder Lyngby and Kristian Sommer Thygesen. Data-driven discovery of 2d materials by deep generative models. *npj Computational Materials*, 8(1):232, Nov 2022. ISSN 2057-3960. doi: 10.1038/s41524-022-00923-3. URL https://doi.org/10.1038/s41524-022-00923-3.

M. Methfessel and A. T. Paxton. High-precision sampling for brillouin-zone integration in metals. *Phys. Rev. B*, 40:3616–3621, Aug 1989. doi: 10.1103/PhysRevB.40.3616. URL https://link.aps.org/doi/10.1103/PhysRevB.40.3616.

Hendrik J. Monkhorst and James D. Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13:5188–5192, Jun 1976. doi: 10.1103/PhysRevB.13.5188. URL https://link.aps.org/doi/10.1103/PhysRevB.13.5188.

Amir H. Naghdi, Kamran Karimi, Axel E. Poisvert, Amin Esfandiarpour, Rene Alvarez, Pawel Sobkowicz, Mikko Alava, and Stefanos Papanikolaou. Dislocation plasticity in equiatomic nicocr alloys: Effect of short-range order. *Phys. Rev. B*, 107:094109, Mar 2023. doi: 10.1103/PhysRevB.107.094109. URL https://link.aps.org/doi/10.1103/PhysRevB.107.094109.

Artem R. Oganov, Chris J. Pickard, Qiang Zhu, and Richard J. Needs. Structure prediction drives materials discovery. *Nature Reviews Materials*, 4(5):331–348, May 2019. ISSN 2058-8437. doi: 10.1038/s41578-019-0101-8. URL https://doi.org/10.1038/s41578-019-0101-8.

Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L. Chevrier, Kristin A. Persson, and Gerbrand Ceder. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013. ISSN 0927-0256. doi: https://doi.org/10.1016/j.commatsci.2012.10.028. URL https://www.sciencedirect.com/science/article/pii/S0927025612006295.

Chris J Pickard and R J Needs. Ab initio random structure searching. *Journal of Physics: Condensed Matter*, 23(5):053201, jan 2011. doi: 10.1088/0953-8984/23/5/053201. URL https://dx.doi.org/10.1088/0953-8984/23/5/053201.

Zekun Ren, Siyu Isaac Parker Tian, Juhwan Noh, Felipe Oviedo, Guangzong Xing, Jiali Li, Qiaohao Liang, Ruiming Zhu, Armin G. Aberle, Shijing Sun, Xiaonan Wang, Yi Liu, Qianxiao Li, Senthilnath Jayavelu, Kedar Hippalgaonkar, Yousung Jung, and Tonio Buonassisi. An invertible crystallographic representation for general inverse design of inorganic crystals with targeted properties. *Matter*, 5(1):314–335, jan 2022. doi: 10.1016/j.matt.2021.11.032. URL https://doi.org/10.1016%2Fj.matt.2021.11.032.

Kristof T. Schütt, Pieter-Jan Kindermans, Huziel E. Sauceda, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions, 2017.

Chence Shi, Shitong Luo, Minkai Xu, and Jian Tang. Learning gradient fields for molecular conformation generation. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 9558–9568. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/shi21b.html.

J. S. Smith, O. Isayev, and A. E. Roitberg. Ani-1: an extensible neural network potential with dft accuracy at force field computational cost. *Chem. Sci.*, 8:3192–3203, 2017. doi: 10.1039/C6SC05720A. URL http://dx.doi.org/10.1039/C6SC05720A.

Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/file/3001ef257407d5a371a96dcd947c7d93-Paper.pdf.

A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comp. Phys. Comm.*, 271:108171, 2022.

Oliver T. Unke and Markus Meuwly. Physnet: A neural network for predicting energies, forces, dipole moments, and partial charges. *Journal of Chemical Theory and Computation*, 15(6):3678–3693, Jun 2019. ISSN 1549-9618. doi: 10.1021/acs.jctc.9b00181. URL https://doi.org/10.1021/acs.jctc.9b00181.

Yanchao Wang, Jian Lv, Li Zhu, and Yanming Ma. Calypso: A method for crystal structure prediction. *Computer Physics Communications*, 183(10):2063–2070, 2012. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2012.05.008. URL https://www.sciencedirect.com/science/article/pii/S0010465512001762.

Daniel Wines, Kevin F Garrity, Tian Xie, and Kamal Choudhary. Forward and inverse design of high t_c superconductors with DFT and deep learning. In *Workshop on "Machine Learning for Materials" ICLR 2023*, 2023. URL https://openreview.net/forum?id=642Z6YL1pk1.

Jan S. Wróbel, Duc Nguyen-Manh, Mikhail Yu. Lavrentiev, Marek Muzyk, and Sergei L. Dudarev. Phase stability of ternary fcc and bcc fe-cr-ni alloys. *Phys. Rev. B*, 91:024108, Jan 2015. doi: 10.1103/PhysRevB.91.024108. URL https://link.aps.org/doi/10.1103/PhysRevB.91.024108.

Changjun Wu, Byeong-Joo Lee, and Xuping Su. Modified embedded-atom interatomic potential for fe-ni, cr-ni and fe-cr-ni systems. *Calphad*, 57:98–106, 2017. ISSN 0364-5916. doi: https://doi.org/10.1016/j.calphad.2017.03.007. URL https://www.sciencedirect.com/science/article/pii/S0364591616302048.

Qu Wu, Bing He, Tao Song, Jian Gao, and Siqi Shi. Cluster expansion method and its application in computational materials science. *Computational Materials Science*, 125:243–254, 2016. ISSN 0927-0256. doi: https://doi.org/10.1016/j.commatsci.2016.08.034. URL https://www.sciencedirect.com/science/article/pii/S0927025616304049.

Guorui Xiao, Wuzhang Yang, Qinqing Zhu, Shijie Song, Guang-Han Cao, and Zhi Ren. Superconductivity with large upper critical field in noncentrosymmetric cr-bearing high-entropy alloys. *Scripta Materialia*, 223:115099, 2023. ISSN 1359-6462. doi: https://doi.org/10.1016/j.scriptamat.2022.115099. URL https://www.sciencedirect.com/science/article/pii/S1359646222005942.

Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi Jaakkola. Crystal diffusion variational autoencoder for periodic material generation, 2022.

J.-W. Yeh, S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, and S.-Y. Chang. Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes. *Advanced Engineering Materials*, 6(5):299–303, 2004. doi: https://doi.org/10.1002/adem.200300567. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/adem.200300567.

Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Sasha Shysheya, Jonathan Crabbé, Lixin Sun, Jake Smith, Bichlien Nguyen, Hannes Schulz, Sarah Lewis, Chin-Wei Huang, Ziheng Lu, Yichi Zhou, Han Yang, Hongxia Hao, Jielan Li, Ryota Tomioka, and Tian Xie. Mattergen: a generative model for inorganic materials design, 2024. URL https://arxiv.org/abs/2312.03687.

Ruopeng Zhang, Shiteng Zhao, Jun Ding, Yan Chong, Tao Jia, Colin Ophus, Mark Asta, Robert O. Ritchie, and Andrew M. Minor. Short-range order and its impact on the crconi medium-entropy alloy. *Nature*, 581(7808):283–287, May 2020. ISSN 1476-4687. doi: 10.1038/s41586-020-2275-z. URL https://doi.org/10.1038/s41586-020-2275-z.

Yong Zhao, Mohammed Al-Fahdi, Ming Hu, Edirisuriya M. D. Siriwardane, Yuqi Song, Alireza Nasiri, and Jianjun Hu. High-throughput discovery of novel cubic crystal materials using deep generative neural networks. *Advanced Science*, 8(20), aug 2021. doi: 10.1002/advs.202100566. URL https://doi.org/10.1002%2Fadvs.202100566.