Machine-Learning-Assisted Construction of Ternary Convex Hull Diagrams

Hugo Rossignol,[1] Michail Minotakis,[1] Matteo Cobelli,[1] and Stefano Sanvito[1]

[1] *School of Physics and CRANN Institute, Trinity College Dublin, College Green, Dublin 2, Ireland*

(Dated: April 21, 2024)

In the search for novel intermetallic ternary alloys, much of the effort goes into performing a large number of *ab-initio* calculations covering a wide range of compositions and structures. These are essential to build a reliable convex hull diagram. While density functional theory (DFT) provides accurate predictions for many systems, its computational overheads set a throughput limit on the number of hypothetical phases that can be probed. Here, we demonstrate how an ensemble of machine-learning spectral neighbor-analysis potentials (SNAPs) can be integrated into a workflow for the construction of accurate ternary convex hull diagrams, highlighting regions fertile for materials discovery. Our workflow relies on using available binary-alloy data both to train the SNAP models and to create prototypes for ternary phases. From the prototype structures, all unique ternary decorations are created and used to form a pool of candidate compounds. The SNAPs ensemble is then used to pre-relax the structures and screen the most favourable prototypes, before using DFT to build the final phase diagram. As constructed, the proposed workflow relies on no extra first-principles data to train the machine-learning surrogate model and yields a DFT-level accurate convex hull. We demonstrate its efficacy by investigating the Cu-Ag-Au and Mo-Ta-W ternary systems.

I. INTRODUCTION

Systematic materials design aims to develop methods that can help to accelerate the discovery of compounds with tailor-made properties, fit for certain applications. The large investment in the area, not least through the materials genome initiative,[1] underpins the importance of searching for novel compounds to bolster technological progress. Atomistic simulations provide a suitable pathway to achieve this goal, since the search can be performed systematically, at low cost and with a complete control over structure and composition. Density functional theory (DFT) calculations are notably used to predict material properties *in silico*, such as material stability or elastic responses. By performing property predictions across a large range of prototype structures, in the form of high-throughput studies,[2] novel magnetic,[3] high-hardness,[4] and battery materials[5] have been discovered. Extensive databases, grouping large numbers of such calculations, have been created and are open to the community. These include AFLOWlib,[6] Materials Project,[7] OQMD,[8] and NOMAD.[9] While such studies remain faster than experimental investigations, the composition and structural spaces to be searched are incredibly large, limiting the scope of application of pure DFT workflows. Importantly, such a limitation in sampling capacity becomes increasingly critical as the number of elements per compound grows, despite the anticipation that a majority of future compound discoveries would be highly multi-elemental.[10] In order to address this issue and to harness the data available from existing *ab-initio* calculations, machine learning (ML) has proved to be a very powerful tool, as it typically comes at a fraction of the DFT computational cost.

The first step in high-throughput computational studies consists in identifying stable compounds by finding a stoichiometry and an associated structure that can be formed. In order to assess the stability of a given structure, the appropriate convex hull diagram needs to be calculated. The proximity between a compound's enthalpy of formation, ΔH_f, and the closest tie-line on the convex hull serves as a criterion for evaluating its stability. Lower values indicate a higher likelihood of stability. Threshold values, typically up to $\sim$100 meV/atom, are used as stability cut-offs.[11] In order to speed up electronic structure methods such as DFT, one possibility is to predict this quantity directly by using ML models, where compounds' compositional and structural information is encoded and mapped directly onto ΔH_f. This is otherwise known as composition prediction, as it is used to identify which stoichiometries are stable by fixing structural variations. The ML models typically used include neural networks, kernel ridge regression and random trees, while the training data are often taken from OQMD, AFLOWlib or Materials Project. For instance, models where the feature vector is only based on compositional information have been used to predict the stability of compounds forming a set prototype structures (elpasolites, perovskites, heuslers, etc.), which is fixed for the compounds in the training set.[11–14] Including structural information in the definition of a model mainly improves the predictions, if large training datasets ($>10^5$ data points) are used.[15] Graph convolutional neural networks[16–18] have notably been used to predict convex hull distances accurately and benefit greatly from structural features.[19,20] Note that these can also be constructed with compositional information only.[21,22] One downside to the inclusion of structural information in the models is that the optimised structure is not known prior to the search, so that data for unrelaxed structures has to be used. This can notably be corrected by using machine-learning interatomic potentials (MLIAPs), that are capable of performing relaxations.

MLIAPs combine atomic fingerprints, representing in-

dividual atomic environments in the form of feature vectors, with ML algorithms, and effectively map the potential energy surface of a collection of atoms.[23] The last decade has seen an immense expansion of the development and application of such potentials.[24–29] When trained using active learning, MLIAPs have most notably been able to extend the length and timescales of *ab-initio* molecular dynamic simulations by several orders of magnitude.[30–34] Such potentials have been successfully applied to predict the energy and forces of alloys,[35] and have been used to accelerate and assist the construction or further exploration of binary and ternary convex hulls. Workflows built on these potentials use MLIAPs as surrogate models to first relax and then make energy predictions on a large library of prototype structures. The lowest energy structures are then compared to a reference convex hull obtained from DFT calculations. This process allows one to improve the reference convex hull diagram by identifying structures lying below it. The training of such potentials is crucial for adequate performance, and studies insist on using high-energy structures for the relaxations to be reliable.

Work in this area has broadly been split into two categories. In the first, specific MLIAPs are trained for a given system,[36–42] typically using active learning. In the other, MLIAPs are trained on large generic databases, and are used to scan over many phases.[43,44] The former is more accurate than the latter, but it is not transferable to other phases. Due to their higher accuracy, phase-specific MLIAPs can also be regarded as global structure optimisers, in that not only they can be used to identify specific stable compositions, but they can accurately predict their structure as well. Many other ML global structure optimisers exist, either in the form of novel workflows[45–48] or by inserting MLIAPs into the pre-existing state-of-the-art global structure optimisers.[49–51]

In this work, we demonstrate how a MLIAP can be trained on data *readily available* on a mainstream repository, such as AFLOWlib,[6] and used to screen a library of ternary-alloy prototypes constructed from their associated binary systems. Recently, we have shown[52] that an ensemble of spectral neighbor-analysis potential (SNAPs)[26] models, trained on the energy data of the three binary subsystems associated with a ternary one, was able to predict the energies of ternary compounds with a mean absolute error (MAE) of ∼30 meV/atom, as long as the structures were fully relaxed. This, not only provides a fast energy-screening tool for ternary compounds, which only requires existing *ab-initio* data on binary structures, but it also gives the valuable insight that chemical environments within binary and ternary transition-metal alloys are similar. Such observation is at the heart of the workflow introduced here. A selection of binary structures, those close to their respective convex hull tie-plane, are selected as templates for ternary alloys. In a high-throughput set up, these are screened using an ensemble of SNAP models, trained on binaries. The lowest-energy compounds are then selected as the most promising candidates, and their energies are calculated using high-fidelity DFT. The ternary convex hull is thus updated.

What makes this workflow different to tailor-made MLIAPs used for convex hull construction is that all the data, both for the prototypes generation and for training the SNAPs, are taken from the relevant binary phases of the AFLOWlib database.[6] In other words, there is no need to generate any new data for the purpose of training the MLIAPs. Despite its training database not being specifically made, either by including important configurations through physical intuition or through active learning, it still has a low enough error on energy predictions to enable a high-throughput search of novel alloys. This is because stable binary and ternary phases, at least for the materials class of transition-metal intermetallics investigated here, share similar local atomic environments. In some sense, the workflow enables an interpolation of the data already available in AFLOWlib, to scan ternary convex hulls and identify stable compositions. Since only a few high-energy structures and no out-of-equilibrium configurations are included in the SNAP training dataset, additional features are introduced in the workflow to increase the robustness of the predictions. These include constraints on the SNAP-driven relaxation (constant volume and the inclusion of a maximum number of steps), as well as using an ensemble of models.

In this paper, the workflow used to generate novel ternary compounds is presented. The methodology Section II details how ternary prototype structures are built from their binary counterparts, and how binary compound data from AFLOWlib[6] is used to train an ensemble of SNAP[26] models. Such SNAPs are used to relax and screen the ternary prototypes. Then, the results Section III presents how the workflow is used to find stable phases for the Cu-Ag-Au and Mo-Ta-W ternary systems. The so-constructed convex hull diagrams are subsequently compared with their available AFLOWlib counterparts, and conclusions are drawn in Section IV. Finally, Section V presents the computational methods employed.

II. METHODS

The general methodology of our workflow, schematically introduced in Figure 1, is here described in detail. From the AFLOWlib database of binary compounds and their associated DFT-computed energies, an ensemble of SNAP models is trained. A subsection of these structures, the ones with the lowest enthalpy of formation, are also used as parent structures and form a library of prototypes. Note that here, as is standard practice in many DFT-based convex hull constructions, a compound's enthalpy of formation is approximated solely by its ground-state DFT energy. As such, the terms enthalpy and energy will be considered equivalent throughout the rest of the manuscript. Candidate ternaries are then created, by

generating all possible and unique derivative structures of such prototypes[53] at a fixed composition,[54] up to a maximum total number of atoms. These two parts are then combined, as the ensemble of SNAP models is used to relax the candidate ternaries. The final energies are predicted through cross-validation within the ensemble of models, while the standard deviation of the predictions is also used to detect and remove geometries for which the relaxation has failed. The resulting structures with the lowest energy and standard deviation are selected as the best candidates (closest to the convex hull). Finally, full *ab-initio* relaxation is performed for these. The ternary system Cu-Ag-Au is used to develop the methodology and is employed here as an example in each subsection.

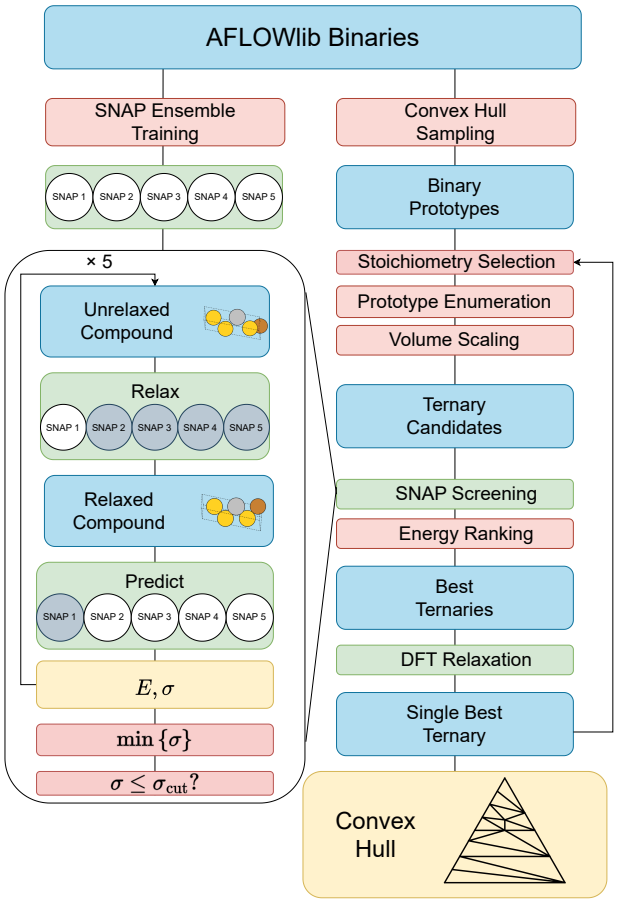


FIG. 1. Diagram of the full stable ternary compounds search workflow implemented in this work. Data available from the three binary subsystems associated with a ternary one (box at the top) is used for two tasks: i) the training of an ensemble of SNAP models, and ii) the construction of a library of parent prototype structures. Derivative structures of the latter are created and all possible ternary decorations of these are produced. Each of them is then relaxed with the SNAPs model and the lowest-enthalpy structures are screened.

A. Generating Prototypes

The first step of the workflow consists in creating a suitable library of ternary prototype materials. The driving idea of this work is that the local atomic environments seen in binary intermetallic alloys are similar to those in the associated ternaries, especially for structures close to equilibrium.[52] In the context of the current work, this insight leads to choosing the binary structures as prototypes for the ternaries. More specifically, the bottom of the three binary convex hulls associated with a ternary system (in our example Ag-Au, Cu-Ag and Cu-Au for Cu-Ag-Au) are scanned to select the lowest-enthalpy compounds. Those within a certain energy range from the convex hull are then selected. All binary structures considered here are taken from the AFLOWlib database.[6] The threshold energy selected differs depending on the system at hand, such as to ensure that roughly the same number of structures are taken from each binary diagram. For instance, in our test system, Cu and Ag are immiscible.[55] Therefore, all binaries have a positive enthalpy of formation and lie far from the Cu-Ag tie line between the two elementary phases (*fcc* Cu and Ag). As a consequence, the energy window above the hull for this binary is larger than that of the other two. The convex hulls of Ag-Au and Cu-Ag are compared in Figure 2 and Table I gives the energy window used as well as the number of structures selected for each binary system.

TABLE I. Number of structures, N_{struct}, selected from each binary system, $X-Y$, to construct the ternary prototypes. Here, we also report the energy window, ΔE, above the convex hull used for the selection.

$X-Y$	N_{struct}	ΔE (meV)
Ag - Au	24	1.7
Cu - Ag	25	65.4
Cu - Au	25	6.2

Once the prototypes are selected, the constituent atoms are stripped of their chemical identity and all structures are compared using the AFLOWlib symmetry tool[56] in order to curate redundancies. This is necessary, since certain structure types (such as *fcc* or *bcc*) may be present several times in the collected database, but may be "decorated" in several different ways for different stoichiometries and binaries. All structures are reduced to their primitive cell at this stage. It is also important to note that single-element structures are also included in this analysis. This leads to a library of unique, undecorated prototypes, taken from the binary convex hulls. For the Cu-Ag-Au system, this results in 40 prototypes. Information on the prototype structures is provided in Supplementary Table S1 in the Supporting Information.

From this set of prototypes, ternary alloys are generated. This task is performed at a fixed stoichiometry and for cells up to a maximum number of atoms, N_{max}. For all the prototypes with a number of atoms compat-

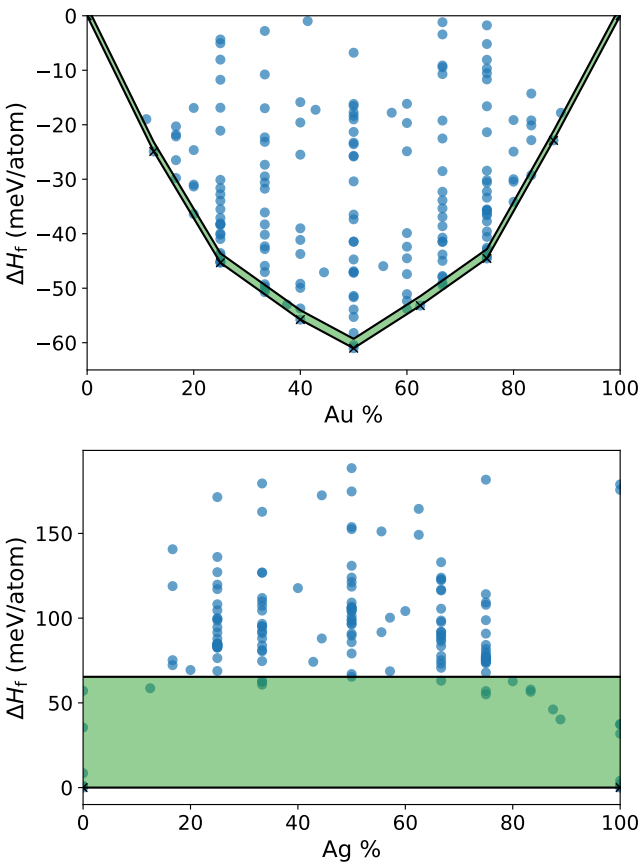


FIG. 2. Convex hull diagrams of the binary Ag-Au (top) and Cu-Ag (bottom) binary systems. Each convex hull is defined by the lower black tie-lines, while the green shadowed regions up to the higher full lines show the energy windows chosen to select the binary structures.

ible with the fixed stoichiometry, the set of all unique derivative structures are created, by following the procedure introduced in references[53,54,57] and by using the associated open source ENUMLIB code. The initial implementation of the algorithm begins from a parent lattice and uses group theory to efficiently enumerate all the unique ways to occupy the sites of supercells constructed from that lattice.[53] Further modifications of the scheme allow for the starting structures to be defined by a lattice and an atomic basis (multi-lattices)[57] and for the generation of derivative geometries at a fixed stoichiometry.[54] This completes the first step of the workflow (blocks in the top right corner of Figure 1) and leads to a set of unique ternary compounds, inspired by the structures of the binaries. The energy of these is then screened using a MLIAP.

B. Ensemble of SNAP Models

MLIAPs typically assume that the total energy, E, of a N-atom system defined by the coordinates $\mathbf{r}^N$ can be written as a sum of atomic energies E_i,

$$E = \sum_{i=1}^{N} E_i \,. \tag{1}$$

Such a partition, first proposed by Behler and Parinello,[25] is based on the principle of nearsightedness.[58,59] The MLIAP of choice for this work is SNAP,[26] which has proved to perform well regardless of the nature of the chemical bond.[60] In this model, the total energy of a compound is written as a sum of linear combinations of the feature vectors describing the chemical environments of each atom i of type α_i in the system. SNAP then takes the bispectrum components, $\mathbf{B}_i^{\alpha_i}$, as feature vectors. The function, E_{SNAP}, that returns the SNAP-predicted total energy is thus defined as

$$E_{\mathrm{SNAP}}\left(\mathbf{r}^N\right) = \sum_{i=1}^{N} \beta_0^{\alpha_i} + \boldsymbol{\beta}^{\alpha_i} \cdot \mathbf{B}_i^{\alpha_i} \,, \tag{2}$$

where $\beta_0^{\alpha_i}$ and $\boldsymbol{\beta}^{\alpha_i}$ are the species-dependent linear coefficients of the ML model. Further details on this potential can be found in Section V. SNAP's linear form allows one to obtain good performance with a small number of features, 56 per species in our case, and when trained on small datasets ($\leq 10^3$ data points).[61] Furthermore, the SNAP hyperparameters are easy to optimise, since the range of optimal values for J_{max} (the maximum angular momentum of the bispectrum) and r_{cut} (the cut-off radius) is wide and consistent for accurate performance.[26,35,62] In our experience, the optimisation of the atomic weights, although generally useful, only leads to modest improvements.[52]

As for our previous study, an ensemble of SNAP models is used to increase the robustness of the predictions. Furthermore, this provides a means of estimating the prediction uncertainty.[52] The ensemble is defined as a set of K functions, $\left\{E_{\mathrm{SNAP}}^k\right\}_{k=1}^K$, where each SNAP model, E_{SNAP}^k, is trained differently, and hence has different linear coefficients. The predicted energy of a new system with atomic positions, $\mathbf{r}^N$, is defined as the mean prediction of the models, $\bar{E}$, and its uncertainty is estimated from the standard deviation, σ, of $\bar{E}$, namely,

$$\bar{E}\left(\mathbf{r}^N\right) = \frac{\sum_{k=1}^{K} E_{\mathrm{SNAP}}^k\left(\mathbf{r}^N\right)}{K} \,, \tag{3}$$

$$\sigma\left(\mathbf{r}^N\right) = \sqrt{\frac{\sum_{k=1}^{K}\left[E_{\mathrm{SNAP}}^k\left(\mathbf{r}^N\right) - \bar{E}\left(\mathbf{r}^N\right)\right]^2}{K}} \,. \tag{4}$$

The training data only consists of binary alloys obtained from the AFLOWlib database. In the case of the Cu-Ag-Au system, the total energies have been recomputed for consistency by single-point DFT calculation (no further relaxation is performed). Differently from what was done when generating the prototypes,

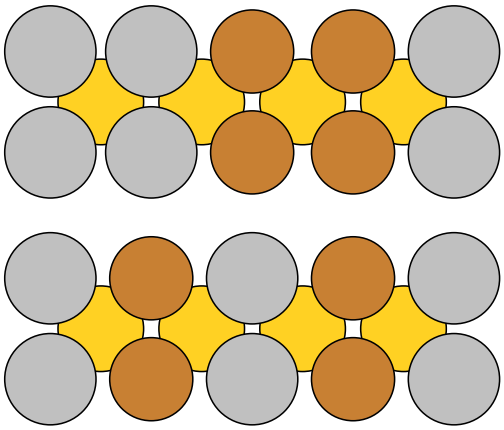
here all binaries, no matter their distance from the convex hull, are included in the training dataset. The same workflow is also used for Mo-Ta-W (the results are described in Section III), for which we demonstrate that the energy values taken directly from AFLOWlib are suitable for training the SNAP models. The full details on the Cu-Ag-Au binary subsystems can be found in reference.[52]

Previously, 10 different SNAP models within the ensemble were obtained by training on different subsets of the same size of the binary-alloy database.[52] For this work, 5 models are trained on the full database, but for each one, a different set of atomic weights for Cu, Ag and Au are used to compute the bispectrum components. This difference is motivated by the need to distinguish compounds with identical site positions in their structure (e.g. the sites of a *bcc* supercell), but different atomic site occupations. If the atomic weights for all species are identical, for some of these structures, notably the high symmetry ones, SNAP will predict identical energies for different compounds. This is illustrated in Figure 3 for two distinct compounds obtained as *bcc* derivative structures with composition $Cu_1Ag_1Au_2$. Prototypes A and B only differ by a permutation of the Ag and Cu atoms. Hence, SNAP models using identical weights for these two atomic types will fail to predict different energies for the compounds. Therefore, by construction, in the ensemble created, the two elements in each pair of atomic types (e.g. Cu and Ag in Cu-Ag) have different weights in at least one model.

Before selecting the values for the atomic weights, r_{cut} and J_{max} are optimised manually and independently by using 10-fold Monte-Carlo cross-validation for fixed identical weights, {1, 1, 1} and thus find: $r_{\text{cut}} = 3.5$Å and $J_{\text{max}} = 4$. For these values, the optimal atomic weights are set by performing a grid search, with the same cross-validation method, and where all three atomic weights are varied from -5 to 5 in steps of 1 (omitting 0). Within this search space, the sets of weights used for the SNAP models in the ensemble are chosen to minimise the cross-validation root-mean squared error (RMSE). The training and cross-validation errors for each model of the ensemble are given in Table II. The ensemble is then used to predict which of the prototype structures have the lowest enthalpy.

C. Energy Screening

The final aim of the workflow is to suggest low-energy ternary structures at a given stoichiometry. Since many compounds with a large energy spread are screened, the suggestions made need to be accurate (must include low-energy structures) and reliable (must not include high-energy and unphysical systems). While the energy error of the SNAP surrogate model is low, it is still prone to making poor predictions on new systems that do not resemble the structures seen in training. As a result, the construction of the workflow highly focuses on the robustness of the final predictions made. Note that choosing parent prototypes from binary compounds already increases the reliability of the predictions.



w_{Cu}, w_{Ag}, w_{Au}			E^{Top}(eV)	E^{Bottom}(eV)
1	1	1	-25.640	-25.640
1	1	2	-25.879	-25.879
1	2	1	-25.717	-25.676
2	1	1	-26.162	-26.047
3	1	2	-25.640	-25.742

FIG. 3. SNAP performance for two structurally identical prototypes. The upper two panels show two different possible site occupations for a $3 \times 1 \times 1$ *bcc* derivative structure with $Cu_1Ag_1Au_2$ stoichiometry. Here we show the z-axis view, with bronze, silver, and gold spheres representing Cu, Ag, and Au respectively. The table shows the SNAP-predicted energies for the two prototypes, when the SNAP is trained with different atomic weights, $\{w_\alpha\}$, as indicated in the first column. Note that when the Cu and Ag weights are identical the two energies coincide by construction. All the crystal structure visualizations are generated with the use of the Atomic Simulation Environment (ASE).[63]

TABLE II. Training and cross-validation (CV) errors for the 5 models, defined by different atomic weights $\{w_\alpha\}$, of the ensemble. Here we report the mean absolute error (MAE) and the root-mean squared error (RMSE). All values are given in meV/atom.

w_{Cu}, w_{Ag}, w_{Au}			Training MAE	Training RMSE	CV MAE	CV RMSE
1	1	2	8.0	13.4	27.1	83.5
1	2	2	8.7	13.5	24.8	64.7
-1	-2	-1	9.7	16.4	30.6	86.4
-1	-2	-2	8.5	13.2	23.5	64.3
-1	-1	-2	7.7	13.1	25.6	75.0

The first step in the energy-screening process consists in setting the compounds' unit-cell volume. This is chosen by taking the weighted average of the elemental volumes of the constituent atoms, an approximation that reproduces the results from *ab-initio*-relaxed compounds

quite well, as is illustrated in Figure 4. Then, the volume and all lattice parameters are kept fixed during any relaxation driven by the SNAP models. This is because, while the training database includes a diverse set of structures, they are all at equilibrium, namely their forces and stress-tensor elements are close to zero. Therefore, no configurations are strongly compressed or expanded, a fact that causes the SNAP models to perform poorly on the prediction of equilibrium volumes and lattice parameters. The volume is only allowed to change for the final, most promising, structures selected for the DFT relaxation.

Each of the K SNAP models available is used to drive an ionic relaxation, with a maximum of N_s steps, for all of the prototypes, leading to K differently relaxed structures per prototype. A "cross-validated" energy prediction is given for each relaxed structure. Given a candidate obtained by relaxing a prototype with the k-th SNAP model, the energy prediction is made with K-1 models, namely all SNAPs bar the one used for the relaxation of the candidate at hand. The mean and standard deviation of the energy predictions of the K-1 models are then saved. For every prototype, one of the K relaxed structures obtained is selected, namely the one with lowest "cross-validated" standard deviation. This is the structure whose final total energy has received the largest consensus among the SNAP models. Therefore, there is only one relaxed structure per prototype. This procedure is illustrated in the flowchart in Figure 1.

The reason why this process is not a single ionic relaxation stems from the drive towards robustness of the predictions. Without the inclusion of the N_s iteration cutoff, some of the relaxations would lead to structures that are trapped in unphysical local minima of the potential energy surface (PES) of the driving SNAP model. By stopping the relaxation process at a low number of steps ($N_\mathrm{s} = 10$ in this study), this effect is mostly avoided, as the structures cannot change too drastically. For the relaxations that are accurately driven by SNAP, the largest drop in energy typically occurs during the first few steps of the relaxation process. While accurate relaxations are also cut before convergence, as they are not distinguished from the inaccurate ones, the final structures are still lower in energy than the initial prototypes. This reduces the likelihood of obtaining high-energy structures and the total run time of the workflow remains modest.

Using "cross-validated" energy predictions of the relaxed structures helps to remove the bias of specific SNAP models. The SNAP driving the relaxation typically predicts the final structure to have a lower energy than the initial one, since it moves the geometric configuration towards, an (at least) local minimum of *its* particular PES. If the relaxation is inaccurate, the resulting structure will be, in fact, high in energy (as predicted by DFT). The relaxation-driving SNAP model is, therefore, not used for energy predictions. Instead, the other models of the ensemble are, since they are less likely to present a bias towards that relaxed structure and therefore to predict it being low in energy. Note that since

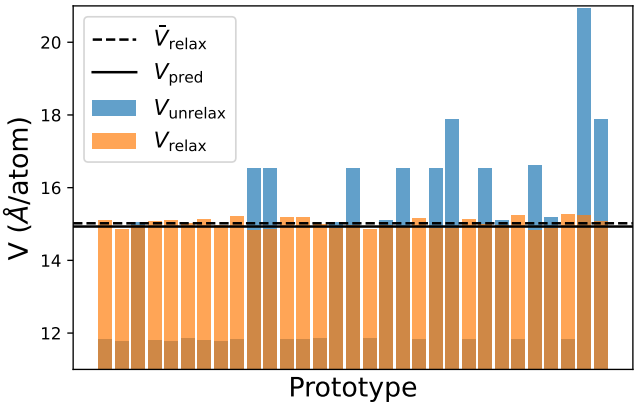


FIG. 4. Plot showing the initial unrelaxed volumes, V_unrelax, and relaxed equilibrium ones, V_relax, of a set of ternary prototypes at the stoichiometry $Cu_2Ag_1Au_1$. The unrelaxed volumes are chosen from the volume of the binary associated with each ternary compound's structure. The dashed line indicates the mean equilibrium volumes for these compounds, $\bar{V}_\mathrm{relax}$, while the full line shows the volume predicted by the weighted average of the elemental volumes, V_pred.

the compounds used as training data are the same for all the K models, they could also lead to biases for the same structure. This is accounted for by using the "cross-validated" standard deviation, rather than the mean energy, to select the "best" structure out of the K relaxed ones. Indeed, even if several/all SNAP models are biased towards a particular structure and collectively predict it to have a low energy, the inaccurate predictions of each model, will be different. This is because they are inaccurate, extrapolated predictions. It has indeed been shown that the standard deviation of SNAPs correlates with the error of the ensemble's prediction.[52] Hence, while the mean prediction of the ensemble may give a low energy value, the standard deviation will be large.

Finally the "cross-validated" standard deviation prediction for all structures must be lower than a cut-off value, σ_cut, to be considered for the final energy screening. This typically excludes structures with low SNAP-predicted energy that DFT returns to be high-energy, as well as structures with high SNAP-predicted energy. From the sample of structures selected, the ones with the lowest "cross-validated" mean energy are chosen and relaxed with DFT. In this study, 15 structures per stoichiometry are selected through such a process.

In summary, the workflow described creates a set of prototypes and uses an ensemble of ML potentials to relax and screen the structures, which are most likely to have low energy. This is done iteratively at fixed stoichiometries. The final selected compounds are then recomputed with full DFT relaxation. The workflow, therefore, allows one to perform all the computationally intensive DFT calculations only on the most promising candidates. In the following section, this workflow will be used to reconstruct the ternary-alloy convex hulls of Cu-

Ag-Au and Mo-Ta-W.

III. RESULTS

This section, which is separated into two subsections, presents the key outcomes of our method. Firstly, we examine the performance of our workflow against the well-established and extensively studied Cu-Ag-Au phase diagram.[64] Then, we provide a comparison between our results and those of one of the better-characterised phase diagrams available in AFLOWlib, specifically Mo-Ta-W. By benchmarking the workflow phase diagram predictions to the DFT created ones available in AFLOWlib, we gain valuable insights into the effectiveness of our approach.

In order to accurately evaluate the stability of our predicted prototypes and ensure consistency in our analysis, we have used the QHull[65] library to calculate the convex hulls presented in this work. The stable, ground-state compounds used to construct the reference convex hull (a subset of the full database) were downloaded from AFLOWlib. For both ternary systems studied, in order to guarantee consistency, we have re-calculated the energies of these compounds with the Vienna Ab initio Simulation Package (VASP).[66] Throughout the entire process, we have strictly followed the AFLOWlib standards as outlined in reference,[67] with an energy cut-off of 600 eV to ensure tight convergence. More information regarding the DFT calculations can be found in Section V A. It should be noted that, for the Mo-Ta-W ternary system, we directly use the AFLOWlib pre-computed energies for the training of the SNAPs ensemble. For the Cu-Ag-Au system, energy data are taken from a previous project,[52] where they were re-calculated with VASP.

A. Cu-Ag-Au Ternary Convex Hull

In order to evaluate the performance of our workflow, it is essential to select a well-studied phase diagram that meets specific requirements. Another key consideration is the availability of sufficient data to train an accurate MLIAP. To facilitate the identification and correction of any errors during the development of the workflow, it is also beneficial to choose a relatively simple phase diagram. With these criteria in mind, we chose the Cu-Ag-Au ternary system, a choice further supported by the fact that the MLIAPs for this phase diagram have already been optimized and trained in our previous work.[52]

As a proof of concept, we have focused on the equiatomic $Cu_1Ag_1Au_1$ ternary phase as well as phases with stoichiometric ratios of 2-1-1[a] and 2-2-1. The reason for this choice is that data at these stoichiometries are available in AFLOWlib for comparison. The results of the workflow are presented in Figure 5 and Table III. In order to quantitatively assess the stability of the structures proposed by the workflow, we use, δ, the distance from the reference convex hull (AFLOWlib). A negative value indicates that the predicted structure lies below the calculated convex hull, establishing its stability as an intermetallic compound. Then, the convex hull needs to be recalculated and corrected by taking into account the newly predicted stable structure. In contrast, a positive distance from the convex hull provides a criterion for assessing whether the structure is metastable or unstable. In Table III values predicted by the workflow (AFLOWlib) are labelled as δ^{WP} (δ^{AFLOW}).

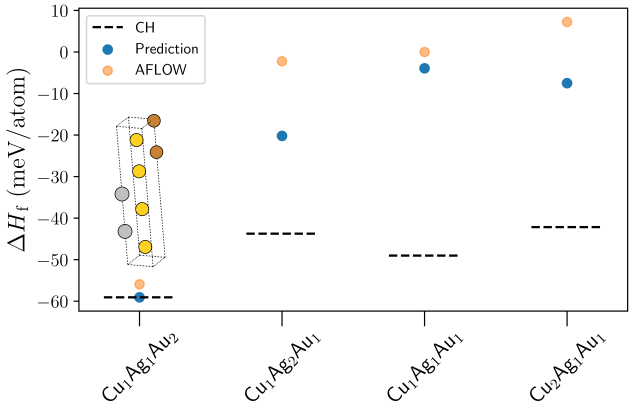


FIG. 5. Workflow predictions for the Cu-Ag-Au ternary system across different stoichiometries. The graph presents the different compositions and their corresponding enthalpy of formation, ΔH_f. The blue points are associated with the predictions from the proposed workflow, whereas the orange ones represent the lowest-energy AFLOWlib points. The dashed line (CH) marks the tie-plane position of the convex hull. The proposed workflow manages to identify one stable intermetallic phase among these, namely $Cu_1Ag_1Au_2$. Furthermore, it manages to outperform the AFLOW dictionary method in all of the presented cases. The unit cell of the newly discovered crystal structure on the convex hull is presented as well. Here, Au atoms are in gold, Ag in silver, and Cu in bronze.

The scalability and speed of the algorithm allow us, in principle, to investigate more regions of the phase diagram, in a single study, than a pure DFT phase diagram construction scheme. This is exemplified by using the proposed workflow to predict structures that are not in AFLOWlib's database, namely compounds with 3-1-1 stoichiometry. The results of the benchmark are presented in Table III, alongside the crystal structure of the new stable phase, $Cu_1Ag_1Au_3$, in Figure 6.

Our approach outperforms the AFLOW dictionary method in all cases, demonstrating a better predictive capability, which arises from the exploration of a larger pool of prototypes. Interestingly, the structures predicted by the proposed workflow are consistently closer to the con-

[a] By this we mean all permutations of a stoichiometric ratio. For a X-Y-Z ternary system, 2-1-1 refers to three compositions: $X_2Y_1Z_1$, $X_1Y_2Z_1$, and $X_1Y_1Z_2$.

TABLE III. Workflow predictions for the Cu-Ag-Au ternary system with 2-2-1 and 3-1-1 compositions. The stoichiometries and their corresponding distance from the convex hull, δ^{WP}, are presented. For the 2-2-1 compounds, the distance from the convex hull of the phases available in the AFLOWlib database, δ^{AFLOW}, are also given. Note that for all materials, the distance from the AFLOWlib convex hull tie-plane is used as reference. A new gold-heavy intermetallic, namely $Cu_1Ag_1Au_3$ is predicted as stable (see Fig. 6).

Stoichiometry	δ^{AFLOW} (meV/atom)	δ^{WP} (meV/atom)
$Cu_2Ag_2Au_1$	208.95	25.99
$Cu_2Ag_1Au_2$	205.69	37.45
$Cu_1Ag_2Au_2$	90.27	17.21
$Cu_3Ag_1Au_1$	–	20.35
$Cu_1Ag_3Au_1$	–	31.05
$Cu_1Ag_1Au_3$	–	-0.02

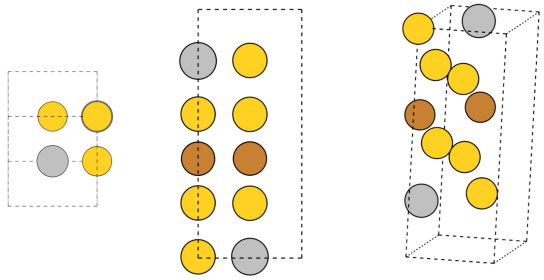


FIG. 6. The unit cell of the crystal structure found on the convex hull at the composition $Cu_1Ag_1Au_3$ is presented in both a top-view with respect to the z-axis (left), a side-view along the x-axis (middle) and a tilted view (right). In this structure, Au atoms are colored in gold, Ag atoms in silver, and Cu atoms in bronze.

vex hull than those predicted by the AFLOW dictionary method. This is to be expected, since the workflow effectively selects the relevant structures for creating the pool of ternary candidates. Furthermore, our model consistently predicts structures with negative or almost negative (< 10 meV/atom) enthalpy of formation, a fact that gives us confidence in the reliability of the predicted structures. Notably, we have been able to identify two new gold-heavy stable phases, namely $Cu_1Ag_1Au_2$ and $Cu_1Ag_1Au_3$. This indicates that stable intermetallic phases may exist on the gold side of the phase diagram. We have confidence in our prediction, given the fact that the dictionary method structure for $Cu_1Ag_1Au_2$ is within 3 meV/atom of the convex hull, suggesting the possibility of the existence of a stable phase. This is consistent with the formation of the solid solutions in the gold-rich region of the experimental phase diagram.[64] The rest of the structures are considered to be potentially metastable, with an average distance from the convex hull of around 30 meV/atom.[68] Overall, our analysis demonstrates the ability of the workflow introduced here to predict structures closer to the convex hull than those from the state-of-the-art dictionary method and possibly uncover novel phases, should these exist.

B. Mo-Ta-W ternary convex hull

As a second benchmark, we wish to explore a phase diagram that exhibits a variety of stable phases. Thus, the main criterion for our selection, among all the possible transition-metal ternary combinations, is the total number of stable compounds. The Mo-Ta-W ternary system emerged as a good candidate, based on a search run with the AFLOW REST-API.[69] In fact, it exhibits the highest number of stable ternary phases of the entire database of transition metal alloys. In order to compare our proposed workflow with the dictionary method, we have made predictions corresponding to the same stoichiometries presented in the previous section. Furthermore, we have used our method to explore areas of the phase diagram poorly covered in AFLOWlib.

We now perform a similar analysis as that described in the previous section. The structure prototypes used for the element decoration are extracted from those of the binaries closest to their respective convex hulls. Information on the prototype structures is provided in Supplementary Table S2 in the Supporting Information. Then, an ensemble of ML models relax the created structures and order them based on their predicted energy. A set of 15 structures for each stoichiometry, corresponding to those with the lowest predicted energies, is sampled and proceeds to the next stage. The latter consists in performing a DFT relaxation and a static calculation for each one of these predictions. A significant difference with respect to the Cu-Ag-Au system is that we now use AFLOWlib's database to train the models without further re-calculation. The AFLOW REST-API is used to download the energies and the crystal structures for the three binary convex hulls (Mo-W, Ta-W, and Mo-Ta). The models are trained as explained in the Methods section (see Section II). Recycling data already available on AFLOWlib allows us to avoid about 1,500 DFT relaxation calculations, some of them for cells up to 46 atoms, just for the training of the model.

The results for the 1-1-1 and 2-1-1 compositions, those with stable phases in AFLOWlib, are presented first in Figure 7. In this case as well we predict a new stable intermetallic phase, $Mo_1Ta_2W_1$. However, this time our workflow does not consistently outperform the dictionary method. In fact, for two out of the four stoichiometries investigated in Figure 7, we obtain compounds with energies similar to the ones already present in AFLOWlib, while for one, $Mo_2Ta_1W_1$, our search delivers a compound with a higher energy. Interestingly, in this last case our newly found structure and the original one, contained in AFLOWlib, belong to different space groups.

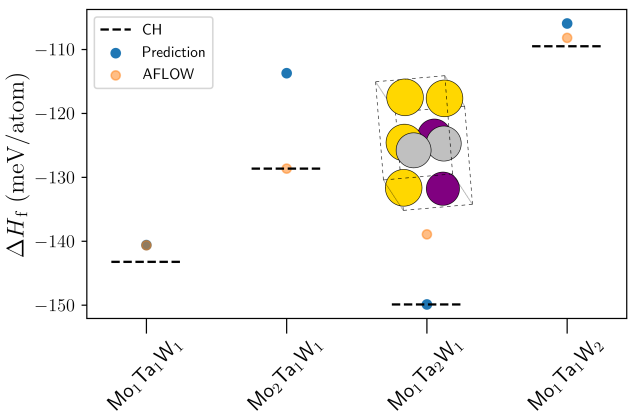


FIG. 7. Workflow predictions for the Mo-Ta-W ternary system across different stoichiometries, 1-1-1 and 2-1-1. The graph presents the different compositions and their corresponding enthalpy of formation, ΔH_f. The blue points are associated with the predictions from the proposed workflow, whereas the orange ones represent the lowest-energy AFLOWlib data. The dashed line marks the tie-plane position of the convex hull (CH). The proposed workflow has managed to identify one previously unknown intermetallic phase, namely $Mo_1Ta_2W_1$, whose unit cell is shown as an inset. Here Mo atoms are in purple, Ta in gold, and W in silver.

The AFLOW-predicted one has space group 107 (tetragonal), while our scheme finds a low-symmetry monoclinic crystal structure with space group 9. The final geometries are not equivalent as determined by the AFLOW-SYM tool.[56] Nevertheless, the compound discovered with the workflow only has an enthalpy of formation 14.91 meV/atom higher than the AFLOWlib compound.

As a force-field based approach, our workflow gets better when the MLIAP improves. In this case, we have extracted the data used to train the SNAPs from the AFLOWlib repository, a detail that led to a less accurate force-field than the one used for the Cu-Ag-Au system. In fact, minor inconsistencies in the energy data, due to unconverged results, may generate errors in the force-field.[70,71] That being understood, we have still demonstrated that new phases can be predicted by an almost DFT-free workflow, since our initial data for model training are readily available in the AFLOWlib database. The workflow systematically assesses a wide range of compositions and potential compounds. Specifically, it involves the evaluation of 331,734 ternaries based on their calculated SNAPs energies. Following this, the 15 lowest-enthalpy structures, for each stoichiometry, undergo relaxation with DFT. Interestingly, the DFT analysis reveals that, on average, the most stable compound ranks 7^{th} among the suggested options. Additionally, the *ab-initio* computations are shortened since all compounds move closer to their equilibrium geometry after the SNAP-guided relaxation, in contrast to their fully unrelaxed counterparts.

Perhaps a more accurate force-field would also be able to find the AFLOWlib minimum for $Mo_2Ta_1W_1$ (see Fig. 7). Nevertheless, our workflow is already able to identify the majority of the structures close to the convex hull. It should also be noted that this is the phase diagram for which AFLOWlib's dictionary method works best, as it is able to detect four intermetallic phases, more than for any other transition metal alloy phase diagram.

Then, we move to analyse stoichiometries poorly explored in AFLOWlib, namely 2-2-1 and 4-1-1. In Table IV we provide a comparison of the distance from the convex hull for the structures predicted with our method, δ^{WP}, and the ones from AFLOWlib, δ^{AFLOW}. For these compositions, the AFLOWlib compounds are unstable as they all have a positive enthalpy of formation. In contrast, those found by our workflow all have a negative enthalpy of formation and are found near or at the convex hull. These results provide a comparison between our method and AFLOWlib for structures predicted as unstable by the latter.

TABLE IV. Workflow predictions for the Mo-Ta-W ternary system with 2-2-1 and 4-1-1 compositions. The stoichiometries and their corresponding distance from the convex hull, δ, are presented (δ^{WP} is for compounds generated by our workflow, while δ^{AFLOW} is for the AFLOWlib compounds). Three intermetallic phases are predicted as stable and two others metastable. Surprisingly, our algorithm is able to find structures with an energy of up to 1 eV/atom lower than those identified by the dictionary method of AFLOWlib.

Stoichiometry	δ^{AFLOW}(meV/atom)	δ^{WP}(meV/atom)
$Mo_2Ta_2W_1$	880.90	0.00
$Mo_1Ta_2W_2$	962.84	0.00
$Mo_2Ta_1W_2$	1032.50	8.50
$Mo_4Ta_1W_1$	320.95	46.56
$Mo_1Ta_4W_1$	516.30	3.25
$Mo_1Ta_1W_4$	334.16	0.00

The ability of our workflow to consistently predict structures that (i) are close to the convex hull and (ii) have a negative enthalpy of formation is thus demonstrated. The former point means that we have an effective algorithm to use for structure search in regions of interest. The latter validates our physical intuition behind the assumption that the crystal structures of the binary alloys close to the convex hull can be used as a template for atomic decoration in the search for ternary phases. This approach has allowed us to identify three new intermetallic compounds (see Table IV), namely $Mo_2Ta_2W_1$, $Mo_1Ta_2W_2$, and $Mo_1Ta_1W_4$. Such positive results demonstrate the value of the enhanced freedom in the structure search provided by our algorithm with respect to dictionary methods.

Finally, following the same spirit as for the analysis of the Cu-Ag-Au system, we now turn our attention to previously unexplored areas of the ternary convex hull. Our results for the $Mo_1Ta_2W_3$ and 3-1-1 compositions are shown in Figure 8. As one can observe, together with

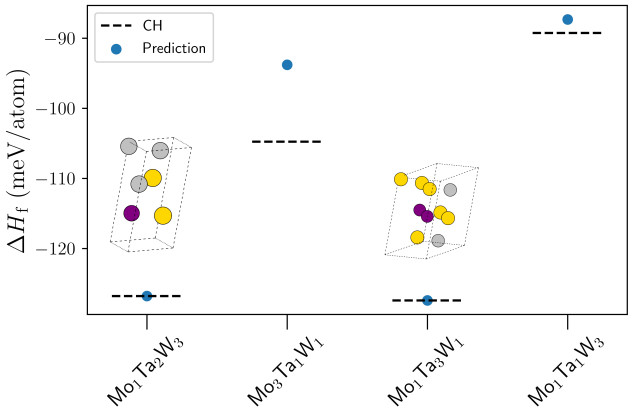


FIG. 8. Workflow predictions (blue points) of the enthalpy of formation for the Mo-Ta-W ternary system across the $Mo_1Ta_2W_3$ and 3-1-1 compositions. The enthalpy of formation for each composition at the appropriate convex hull tie-plane (CH) is shown as a dashed line. The unit cell of the crystal structures found on the convex hull are presented as well. Here, Mo atoms are in purple, Ta in gold, and W in silver. Two new intermetallics alloys have been identified, namely $Mo_1Ta_2W_3$ and $Mo_1Ta_3W_1$.

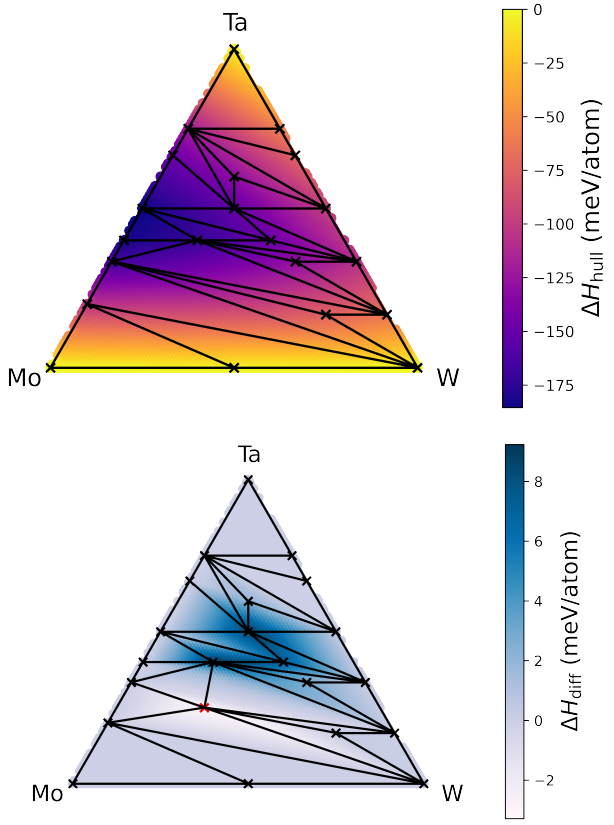


FIG. 9. Workflow-computed convex hull for the Mo-Ta-W system (upper panel). The color heat map corresponds to the calculated enthalpy of formation at a given stoichiometry. In the lower panel, we present the difference between the convex hull of AFLOWlib (reference) and that computed by our workflow. Black crosses for the ternary region symbolize the newly predicted intermetallic phases, and the red cross denotes the only stable intermetallic originally predicted by the AFLOW dictionary method.

structures away from the tie-plane, we also find two new stable compounds, namely $Mo_1Ta_2W_3$ and $Mo_1Ta_3W_1$. Such new phases, together with the low-energy ones previously discussed, call for a modification of the ternary convex hull that exists in AFLOWlib. The new diagram is presented in the top panel of Figure 9. In order to facilitate the comparison, the lower panel of the same figure shows the difference between the AFLOW- and our workflow-predicted convex hulls (positive values mean that our predicted convex hull is lower in energy than the original AFLOWlib one).

The new convex hull returns a picture, where most of the stable ternary structures identified belong to the Ta-W heavy area, and only one intermetallic alloy exists in the Mo-rich region of the compositional space. The latter is the compound found on AFLOWlib. Interestingly, the new phases predicted by our workflow undercut $Mo_1Ta_1W_1$, $Mo_1Ta_2W_1$ and $Mo_1Ta_1W_2$, the other intermetallic alloys initially predicted as stable by AFLOWlib. These are now respectively, 2.59, 10.95 and 1.31 meV/atom, above their associated tie-planes and have to be considered metastable. Experimentally, there is evidence that the Mo-Ta-W system forms a ternary solid solution[72] at finite temperature across the entire phase diagram. It should be noted that the Mo-Ta binary space is far better sampled by AFLOWlib than the Ta-W and Mo-W ones. This could imply that it is more difficult to reach the convex hull close to such a facet of the diagram. In contrast, the Mo-W system only displays small enthalpies of formation for the stable binary phases, implying that both Mo and W form more stable phases with Ta than amongst themselves. These two reasons could explain why it is more difficult to find stable intermetallic phases in the Mo-rich part of the composition space.

In summary, the workflow developed provides a comprehensive scan across the ternary composition space, enabling the construction of a convex hull with DFT-accuracy and identifying areas prone to alloy stability through the discovery of new hull points. This is accomplished by screening of the order of 10^5 candidate compounds with the SNAPs ensemble. In the case of the Cu-Ag-Au test system, two stable Au-rich compounds are found, which are not present on the AFLOWlib convex hull. The new $Cu_1Ag_1Au_2$ (Figure 5) and $Cu_1Ag_1Au_3$ (Figure 6) ternaries' structures have space groups 123 and 63 respectively. They respectively resemble distorted *bcc* and *hcp* structures. The workflow hence highlights the region in which Cu-Ag-Au alloys are likely to form. The experimental structures for these phases are *fcc* solid-state solutions.[64,73,74] For the Mo-Ta-W system, six novel compounds are identified, undercutting

three of the AFLOWlib compounds, leaving only the $Mo_2Ta_1W_1$ intermetallic to lie on the convex hull. Experimentally, *bcc* solid-state solutions form across the full compositional range.[72] One of the phases discovered, $Mo_1Ta_1W_4$, possesses a tetragonally distorted *bcc* structure, while the others have space groups 74 ($Mo_1Ta_2W_1$), 12 ($Mo_1Ta_2W_2$ and $Mo_1Ta_3W_1$), and 38 ($Mo_1Ta_2Ta_3$ and $Mo_2Ta_2W_1$). The results found here suggest that at low temperatures the Mo-rich corner is dominated by an intermetallic phase, $Mo_2Ta_1W_1$. However, the regions with lower Mo concentrations form many more different intermetallics that lie close to each other in energy. This region is thus more susceptible to forming solid-state phases at finite temperatures. The convex hulls obtained with this workflow suggest promising regions of material stability, notably in the form solid-state solutions for the two systems studied. This could help guide experimental studies in the synthesis of stable ternary alloys. Our strategy resembles the MatLearn[75] approach in spirit, but is a more accurate method as it provides DFT-level convex hulls. This is illustrated by the fact that no novel ternary phases are predicted by MatLearn for Cu-Ag-Au and Mo-Ta-W, and for the former, the only 'known' phase (from DFT) is in the Cu-rich region, which differs from the experimental phase diagram.

IV. CONCLUSION

We have developed a workflow that predicts the crystal structure and assesses the stability of ternary compounds of a particular stoichiometry. A library of prototype structures is formed from the lowest-enthalpy alloys of the associated binary subsystems. From this database, derivative ternary structures are generated by site decoration. Then, an ensemble of SNAP force-fields is used to select the most promising structures among them, bypassing the majority of the *ab-initio* calculations. Therefore, the proposed workflow highly increases the throughput in the search for stable ternary compounds without compromising the quality of the predictions. This is used here to map the ternary convex hull of transition-metal alloys. The crucial aspect of the proposed scheme is that both the training of the force-fields and the creation of the prototype ternary structures are based solely on the knowledge of the binary phases. As such, no additional DFT calculations are required, since both the structures and their corresponding energies are readily available on the AFLOWlib database. Employing *ab-initio* calculations solely in the final stage of the workflow and focusing them on the most promising candidates, allows us to perform a comprehensive exploration of the phase diagram of a ternary system with only a few hundred DFT calculations. This enables us to map previously unexplored portions of the ternary space and to identify regions of interest, thus driving the discovery of novel compounds.

We have demonstrated that the proposed workflow is able to predict crystal structures with negative enthalpy of formation and effectively identify the stable intermetallics, should they exist. In particular, we have used the Cu-Ag-Au and Mo-Ta-W ternary systems as an example. In the first case, we have predicted several new phases that, although not all thermodynamically stable, have an enthalpy of formation lower than those found by the AFLOW dictionary method. In addition, we have identified an Au-rich composition region, where stable intermetallic phases are expected, in accordance with the location of solid solutions in the experimental phase diagram.[64] Interestingly, in the case of Mo-Ta-W, one of the ternary systems with the largest number of stable intermetallics in AFLOWlib, our method is capable of identifying a plethora of new phases, resulting in the correction of the original DFT-calculated convex hull proposed by AFLOW.

In summary, we have developed a novel way to integrate machine learning to accelerate a DFT workflow. Although the ML model introduced here does not perform as well as force-fields with tailor-made databases, its construction requires no new DFT calculations and simply recycles pre-existing results, already present on large-scale databases. This represents an example of how machine-learning interatomic potentials can be seamlessly integrated into a materials design pipeline without the need to generate *ad hoc* large training sets.

V. COMPUTATIONAL METHODS

The details of the computational methods are presented in this section. The parameters used for the DFT calculations run with VASP[66] are first discussed. A brief presentation of the SNAP[26] is then given, along with details of the implementations used for the current work.

A. DFT Calculations

All DFT calculations are performed using the Vienna Ab initio Simulation Package (VASP),[66] version 5.4.4. Projector augmented wave (PAW) pseudopotentials are used for each element together with the Perdew-Burke-Ernzerhof (PBE) functional.[76] A plane wave cutoff of 600 eV is used for all calculations. The energy convergence criterion for each self-consistent cycle is of 10^{-4} eV. Full atomic relaxations are performed (update of atomic positions, cell volume and lattice parameters) with a stopping criterion on the forces of 10^{-3} eV/Å. A Fermi-Dirac smearing of 0.2 eV is chosen for all calculations.

For the k-point sampling, a Gamma-centered mesh is employed for all calculations. The density of the mesh and the spacing between k-points is chosen based on AFLOWlib's convergence criteria.[67] The mesh is system-specific and determined from the N_{KPPRA} (number of k-points per reciprocal atom). The number of sampling points along each direction is proportional to the norm of the corresponding reciprocal lattice vector. The total

number of sampling points per reciprocal atom is then minimised and N_{KPPRA} is used as a lower bound. Values of 10×10^3 and 6×10^3 are used for static calculations and relaxations respectively.

B. Spectral Analysis Neighbor Potential

The Spectral Analysis Neighbor Potential (SNAP)[26] is used as an energy predictor. As described in Section II B, an ensemble of models is employed for predictions. Equation (2) defines the expression of the function, E_{SNAP}, and combined with Equation (1), gives the energy of a system with N atoms. The atomic fingerprints that define the chemical environments of each atom i in the system, belonging to species α_i, are the bispectrum components.[23] These are used to represent configurations instead of seemingly more obvious choices (e.g. atomic Cartesian coordinates), as they are invariant upon rotation and permutations of identical atoms. Note that invariance with respect to translations is guaranteed by Eq. (1). For each atom, the vector $\mathbf{B}_i^{\alpha_i}$, which collects the first components up to a maximum index, is taken as a feature for the machine-learning model (ridge regression in the case of SNAP). A short description of the bispectrum components is given below.

The neighborhood of an atom i atom can be described by a density function, ρ_i, centered at that atom with delta functions at the sites of surrounding atoms, within a sphere of radius r_{cut}. It is defined in three dimensions as

$$\rho_i(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_i) + \sum_j w^{\alpha_j} \delta(\mathbf{r} - \mathbf{r}_j) f_c(r_{ij}), \qquad (5)$$

where the sum is over all atoms within r_{cut} from the central atom. Here, $\mathbf{r}_i$ is the position of atom i, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, w^{α_j} is the specie-specific weight of atom j and f_c is a cut-off function that smoothly runs to zero as r_{ij} approaches r_{cut}, as defined in reference.[25] In order to represent this density distribution as a vector, it is expanded in a suitable basis. Atomic positions are first mapped onto the 4D sphere, by switching to polar coordinates (θ, ϕ, r) and by defining a third polar angle, θ_0, from the radial coordinate (see reference[23] for details). The density function is then expanded in terms of hyperspherical harmonics $U^J_{m',m}$, the natural basis for expansion on the 4D sphere. Dropping the atomic index, ρ is written as

$$\rho(\mathbf{r}) = \sum_{J=0}^{\infty} \sum_{m,m'=-J}^{J} c^J_{m',m} U^J_{m',m}(\theta, \phi, \theta_0) . \qquad (6)$$

The hyperspherical harmonic index, J, runs in half-integer steps, while m and m' run between $-J$ and J in integer steps. The outer sum is truncated in practice at a value J_{max}, treated as a hyperparameter. The expansion coefficients, $c^J_{m',m}$, cannot be used as descriptors, since they are complex and are not invariant under system rotation. From them, however, the rotationally-invariant and real-valued bispectrum components B_{J,J_1,J_2} are constructed

$$B_{J,J_1,J_2} = \sum_{m'_1,m_1=-J_1}^{J_1} c^{J_1}_{m'_1,m_1} \sum_{m'_2,m_2=-J_2}^{J_2} c^{J_2}_{m'_2,m_2} \times \sum_{m',m=-J}^{J} C^{J,J_1,J_2}_{mm_1m_2} C^{J,J_1,J_2}_{m'm'_1m'_2} \left(c^J_{m',m}\right)^* . \qquad (7)$$

Here, $C^{J,J_1,J_2}_{mm_1m_2}$ and $C^{J,J_1,J_2}_{m'm'_1m'_2}$ are the Clebsch-Gordan coefficients, which possess the same symmetry invariances as the system. After taking the non-zero and unique distinct components, the bispectrum vector is formed, denoted $\mathbf{B}_i^{\alpha_i}$, with atomic and specie indices. The bispectrum components are a highly non-linear representation of the local atomic coordinates and account for up to four-body interactions. Their complexity is what makes it possible for them to be effectively used together with a simple regressor in SNAP to accurately map structures to energies.

The fitting, testing and predictions of the SNAP models used are performed using an in-house python library built with SCIKIT-LEARN[77] and the Atomic Simulation Environment (ASE[63]) python libraries. The bispectrum components are computed using LAMMPS.[78] The pipeline is built in python to perform the API download of binary structures and energies from the AFLOWlib[79] database and to generate derivative structures from the prototypes using ENUMLIB.[53] DFT calculations are managed by using a combination of ASE[63] and PYMATGEN.[80]

DATA AND SOFTWARE AVAILABILITY

Data associated with this project, including the AFLOWlib AUIDs for the training data, the AFLOWlib labels of the structures used as prototypes, the parameters of the SNAP models and the lowest values of the enthalpy of formation found at each composition are available on the Github repository (https://github.com/HugoRossignol/Workflow_Ternary_ConvexHull).

ACKNOWLEDGMENTS

This work has been supported by the Irish Research Council Advanced Laureate Award (IRCLA/2019/127), and by the Irish Research Council postgraduate program (MC). We acknowledge the DJEI/DES/SFI/HEA Irish Centre for High-End Computing (ICHEC) and Trinity Centre for High Performance Computing (TCHPC) for the provision of computational resources. We would like to acknowledge Eve Gilligan for creating the cover illustration.

AUTHOR CONTRIBUTIONS

This section is written according to the CRediT system. H.R. and M.M. contributed equally to this work. H.R. and M.M. contributed to conceptualisation, methodology, software, data curation, formal analysis, investigation, validation, visualisation, writing the original draft, as well as reviewing and editing the manuscript. M.C. contributed to conceptualisation, software as well as reviewing and editing the manuscript. S.S. contributed to conceptualisation, funding acquisition, project administration, resources, supervision, as well as reviewing and editing the manuscript.

SUPPORTING INFORMATION AVAILABLE

Supporting Information is available to provide additional data showing the diversity of the structures used as ternary prototypes. Information on the prototype structures (AFLOWlib auid, space group of the parent binary compound, and space group of the prototype structure) is provided in Supplementary Tables S1 and S2 for the Cu-Ag-Au and Mo-Ta-W systems, respectively.

REFERENCES

[1] Materials Genome Initiative, *https://www.mgi.gov/*.

[2] Curtarolo, S., Hart, G. L. W., Nardelli, M. B., Mingo, N., Sanvito, S. and Levy, O., The high-throughput highway to computational materials design, *Nature Mater. 12*, pp. 191–201 **2013**.

[3] Sanvito, S., Oses, C., Xue, J., Tiwari, A., Zic, M., Archer, T., Tozman, P., Venkatesan, M., Coey, M. and Curtarolo, S., Accelerated discovery of new magnets in the Heusler alloy family, *Sci. Adv. 3*(4) , pp. e1602241 **2017**.

[4] Sarker, P., Harrington, T., Toher, C., Oses, C., Samiee, M., Maria, J.-P., Brenner, D. W., Vecchio, K. S. and Curtarolo, S., High-entropy high-hardness metal carbides discovered by entropy descriptors, *Nat. Commun. 9*(1) , pp. 4980 **2018**.

[5] Kim, J. C., Li, X., Moore, C. J., Bo, S.-H., Khalifah, P. G., Grey, C. P. and Ceder, G., Analysis of charged state stability for monoclinic $LiMnBO_3$ cathode, *Chem. Mater. 26*(14) , pp. 4200–4206 **2014**.

[6] Curtarolo, S., Setyawan, W., Wang, S., Xue, J., Yang, K., Taylor, R. H., Nelson, L. J., Hart, G. L., Sanvito, S., Buongiorno-Nardelli, M., Mingo, N. and Levy, O., AFLOWLIB.org: A distributed materials properties repository from high-throughput ab initio calculations, *Comput. Mater. Sci. 58*, pp. 227–235 **2012**.

[7] Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G. and Persson, K. A., Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, *APL Mater. 1*(1) , pp. 011002 **2013**.

[8] Saal, J. E., Kirklin, S., Aykol, M., Meredig, B. and Wolverton, C., Materials design and discovery with high-throughput density functional theory: the open quantum materials database (OQMD), *JOM 65*, pp. 1501–1509 **2013**.

[9] Draxl, C. and Scheffler, M., The NOMAD laboratory: from data sharing to artificial intelligence, *J. Phys. Mater. 2*(3) , pp. 036001 **2019**.

[10] Toher, C., Oses, C., Hicks, D. and Curtarolo, S., Unavoidable disorder and entropy in multi-component systems, *npj Comput. Mater. 5*, pp. 69 **2019**.

[11] Kim, K., Ward, L., He, J., Krishna, A., Agrawal, A. and Wolverton, C., Machine-learning-accelerated high-throughput materials screening: Discovery of novel quaternary Heusler compounds, *Phys. Rev. Mater. 2*(12) , pp. 123801 **2018**.

[12] Faber, F., Lindmaa, A., Von Lilienfeld, O. A. and Armiento, R., Crystal structure representations for machine learning models of formation energies, *Int. J. Quantum Chem. 115*(16) , pp. 1094–1101 **2015**.

[13] Faber, F. A., Lindmaa, A., Von Lilienfeld, O. A. and Armiento, R., Machine learning energies of 2 million elpasolite (ABC_2D_6) crystals, *Phys. Rev. Lett. 117*(13) , pp. 135502 **2016**.

[14] Schmidt, J., Shi, J., Borlido, P., Chen, L., Botti, S. and Marques, M. A., Predicting the thermodynamic stability of solids combining density functional theory and machine learning, *Chem. Mater. 29*(12) , pp. 5090–5103 **2017**.

[15] Ward, L., Liu, R., Krishna, A., Hegde, V. I., Agrawal, A., Choudhary, A. and Wolverton, C., Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations, *Phys. Rev. B 96*(2) , pp. 024104 **2017**.

[16] Xie, T. and Grossman, J. C., Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, *Phys. Rev. Lett. 120*(14) , pp. 145301 **2018**.

[17] Park, C. W. and Wolverton, C., Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery, *Phys. Rev. Mater. 4*(6) , pp. 063801 **2020**.

[18] Schmidt, J., Hoffmann, N., Wang, H.-C., Borlido, P., Carriço, P. J., Cerqueira, T. F., Botti, S. and Marques, M. A., Machine-Learning-Assisted Determination of the Global Zero-Temperature Phase Diagram of Materials, *Adv. Mater. 35*(22) , pp. 2210788 **2023**.

[19] Bartel, C. J., Trewartha, A., Wang, Q., Dunn, A., Jain, A. and Ceder, G., A critical examination of compound stability predictions from machine-learned formation energies, *npj Comput. Mater. 6*(1) , pp. 97 **2020**.

[20] Pandey, S., Qu, J., Stevanović, V., John, P. S. and Gorai, P., Predicting energy and stability of known and hypothetical crystals using graph neural network, *Patterns 2*(11) , pp. 100361 **2021**.

[21] Goodall, R. E. A. and Lee, A. A., Predicting materials properties without crystal structure: Deep representation learning from stoichiometry, *Nat. Commun. 11*(1) , pp. 6280 **2020**.

[22] Jha, D., Ward, L., Paul, A., Liao, W.-k., Choudhary, A., Wolverton, C. and Agrawal, A., Elemnet: Deep learning the chemistry of materials from only elemental composition, *Sci. Rep. 8*(1) , pp. 1–13 **2018**.

[23] Bartók, A. P., Kondor, R. and Csányi, G., On representing chemical environments, *Phys. Rev. B 87*(18) , pp. 184115 **2013**.

[24] Bartók, A. P., Payne, M. C., Kondor, R. and Csányi, G., Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons, *Phys. Rev. Lett. 104*(13) , pp. 136403 **2010**.

[25] Behler, J. and Parrinello, M., Generalized neural-network representation of high-dimensional potential-energy surfaces, *Phys. Rev. Lett. 98*(14) , pp. 146401 **2007**.

[26] Thompson, A. P., Swiler, L. P., Trott, C. R., Foiles, S. M. and Tucker, G. J., Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials, *J. Comput. Phys. 285*, pp. 316–330 **2015**.

[27] Shapeev, A. V., Moment tensor potentials: A class of systematically improvable interatomic potentials, *Multiscale Modeling & Simulation 14*(3) , pp. 1153–1173 **2016**.

[28] Drautz, R., Atomic cluster expansion for accurate and transferable interatomic potentials, *Phys. Rev. B 99*(1) , pp. 014104 **2019**.

[29] Domina, M., Patil, U., Cobelli, M. and Sanvito, S., Cluster expansion constructed over Jacobi-Legendre polynomials for accurate force fields, *Phys. Rev. B 108*(9) , pp. 094102 **2023**.

[30] Deringer, V. L. and Csányi, G., Machine learning based interatomic potential for amorphous carbon, *Phys. Rev. B 95*(9) , pp. 094203 **2017**.

[31] Caro, M. A., Csányi, G., Laurila, T. and Deringer, V. L., Machine learning driven simulated deposition of carbon films: From low-density to diamondlike amorphous carbon, *Phys. Rev. B 102*(17) , pp. 174201 **2020**.

[32] Jinnouchi, R., Karsai, F. and Kresse, G., On-the-fly machine learning force field generation: Application to melting points, *Phys. Rev. B 100*(1) , pp. 014105 **2019**.

[33] Mortazavi, B., Javvaji, B., Shojaei, F., Rabczuk, T., Shapeev, A. V. and Zhuang, X., Exceptional piezoelectricity, high thermal conductivity and stiffness and promising photocatalysis in two-dimensional MoSi2N4 family confirmed by first-principles, *Nano Energy 82*, pp. 105716 **2021**.

[34] Mortazavi, B., Silani, M., Podryabinkin, E. V., Rabczuk, T., Zhuang, X. and Shapeev, A. V., First-principles multiscale modeling of mechanical properties in graphene/borophene heterostructures empowered by machine-learning interatomic potentials, *Advanced Materials 33*(35) , pp. 2102807 **2021**.

[35] Chen, C., Deng, Z., Tran, R., Tang, H., Chu, I.-H. and Ong, S. P., Accurate force field for molybdenum by machine learning large materials data, *Phys. Rev. Mater. 1*(4) , pp. 043603 **2017**.

[36] Gubaev, K., Podryabinkin, E. V., Hart, G. L. and Shapeev, A. V., Accelerating high-throughput searches for new alloys with active learning of interatomic potentials, *Comput. Mater. Sci. 156*, pp. 148–156 **2019**.

[37] Bernstein, N., Csányi, G. and Deringer, V. L., De novo exploration and self-guided learning of potential-energy surfaces, *npj Comput. Mater. 5*(1) , pp. 99 **2019**.

[38] Artrith, N., Urban, A. and Ceder, G., Constructing first-principles phase diagrams of amorphous Li_xSi using machine-learning-assisted sampling with an evolutionary algorithm, *The Journal of chemical physics 148*(24) , pp. 241711 **2018**.

[39] Kharabadze, S., Thorn, A., Koulakova, E. A. and Kolmogorov, A. N., Prediction of stable Li-Sn compounds: boosting ab initio searches with neural network potentials, *npj Comput. Mater. 8*(1) , pp. 136 **2022**.

[40] Seko, A., Machine learning potentials for multicomponent systems: The Ti-Al binary system, *Phys. Rev. B 102*(17) , pp. 174104 **2020**.

[41] Kelvinius, F. E., Armiento, R. and Lindsten, F., Graph-based machine learning beyond stable materials and relaxed crystal structures, *Phys. Rev. Mater. 6*(3) , pp. 033801 **2022**.

[42] Wang, R., Xia, W., Slade, T. J., Fan, X., Dong, H., Ho, K.-M., Canfield, P. C. and Wang, C.-Z., Machine learning guided discovery of ternary compounds involving La and immiscible Co and Pb elements, *npj Computational Materials 8*(1) , pp. 258 **2022**.

[43] Law, J. N., Pandey, S., Gorai, P. and St. John, P. C., Upper-Bound Energy Minimization to Search for Stable Functional Materials with Graph Neural Networks, *JACS Au* **2022**.

[44] Chen, C. and Ong, S. P., A universal graph deep learning interatomic potential for the periodic table, *Nat. Comput. Sci. 2*(11) , pp. 718–728 **2022**.

[45] Bisbo, M. K. and Hammer, B., Global optimization of atomic structure enhanced by machine learning, *Phys. Rev. B 105*(24) , pp. 245404 **2022**.

[46] Paleico, M. L. and Behler, J., Global optimization of copper clusters at the ZnO (10 1 0) surface using a DFT-based neural network potential and genetic algorithms, *J. Chem. Phys. 153*(5) , pp. 054704 **2020**.

[47] Yamashita, T., Sato, N., Kino, H., Miyake, T., Tsuda, K. and Oguchi, T., Crystal structure prediction accelerated by Bayesian optimization, *Phys. Rev. Mater. 2*(1) , pp. 013803 **2018**.

[48] Podryabinkin, E. V., Tikhonov, E. V., Shapeev, A. V. and Oganov, A. R., Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning, *Phys. Rev. B 99*(6) , pp. 064114 **2019**.

[49] Deringer, V. L., Pickard, C. J. and Csányi, G., Data-driven learning of total and local energies in elemental boron, *Phys. Rev. Lett. 120*(15) , pp. 156001 **2018**.

[50] Pickard, C. J., Ephemeral data derived potentials for random structure search, *Phys. Rev. B 106*(1) , pp. 014102 **2022**.

[51] Tong, Q., Xue, L., Lv, J., Wang, Y. and Ma, Y., Accelerating CALYPSO structure prediction by data-driven learning of a potential energy surface, *Faraday Discuss. 211*, pp. 31–43 **2018**.

[52] Minotakis, M., Rossignol, H., Cobelli, M. and Sanvito, S., Machine-learning surrogate model for accelerating the search of stable ternary alloys, *Phys. Rev. Mater. 7*(9) , pp. 093802 **2023**.

[53] Hart, G. L. and Forcade, R. W., Algorithm for generating derivative structures, *Phys. Rev. B 77*(22) , pp. 224115 **2008**.

[54] Hart, G. L., Nelson, L. J. and Forcade, R. W., Generating derivative structures at a fixed concentration, *Comput. Mater. Sci. 59*, pp. 101–107 **2012**.

[55] Kusoffsky, A., Thermodynamic evaluation of the ternary Ag–Au–Cu system—including a short range order description, *Acta Mater. 50*(20) , pp. 5139–5145 **2002**.

[56] Hicks, D., Oses, C., Gossett, E., Gomez, G., Taylor, R. H., Toher, C., Mehl, M. J., Levy, O. and Curtarolo, S., AFLOW-SYM: platform for the complete, automatic and self-consistent symmetry analysis of crystals, *Acta Crystallogr. A 74*(3) , pp. 184–203 **2018**.

[57] Hart, G. L. W. and Forcade, R. W., Generating derivative structures from multilattices: Algorithm and appli-

cation to hcp alloys, *Phys. Rev. B 80*(1) , pp. 014120 **2009**.

[58] Kohn, W., Density functional and density matrix method scaling linearly with the number of atoms, *Phys. Rev. Lett. 76*(17) , pp. 3168 **1996**.

[59] Prodan, E. and Kohn, W., Nearsightedness of electronic matter, *Proceedings of the National Academy of Sciences 102*(33) , pp. 11635–11638 **2005**.

[60] Lunghi, A. and Sanvito, S., A unified picture of the covalent bond within quantum-accurate force fields: From organic molecules to metallic complexes' reactivity, *Sci. Adv. 5*, pp. eaaw221 **2019**.

[61] Zuo, Y., Chen, C., Li, X., Deng, Z., Chen, Y., Behler, J., Csányi, G., Shapeev, A. V., Thompson, A. P., Wood, M. A. and Ong, S. P., Performance and cost assessment of machine learning interatomic potentials, *The Journal of Physical Chemistry A 124*(4) , pp. 731–745 **2020**.

[62] Li, X.-G., Hu, C., Chen, C., Deng, Z., Luo, J. and Ong, S. P., Quantum-accurate spectral neighbor analysis potential models for Ni-Mo binary alloys and fcc metals, *Phys. Rev. B 98*(9) , pp. 094104 **2018**.

[63] Larsen, A. H., Mortensen, J. J., Blomqvist, J., Castelli, I. E., Christensen, R., Dułak, M., Friis, J., Groves, M. N., Hammer, B., Hargus, C., Hermes, E. D., Jennings, P. C., Jensen, P. B., Kermode, J., Kitchin, J. R., Kolsbjerg, E. L., Kubal, J., Kaasbjerg, K., Lysgaard, S., Maronsson, J. B., Maxson, T., Olsen, T., Pastewka, L., Peterson, A., Rostgaard, C., Schiøtz, J., Schütt, O., Strange, M., Thygesen, K. S., Vegge, T., Vilhelmsen, L., Walter, M., Zeng, Z. and Jacobsen, K. W., The atomic simulation environment—a Python library for working with atoms, *J. Phys.: Condens. Matter 29*(27) , pp. 273002 **2017**.

[64] Prince, A., Phase diagrams of ternary gold alloys, *Inst. Metals* , pp. 7–42 **1990**.

[65] Barber, C. B., Dobkin, D. P. and Huhdanpaa, H., The quickhull algorithm for convex hulls, *ACM Trans. Math. Softw. 22*(4) , pp. 469–483 **1996**.

[66] Kresse, G. and Furthmüller, J., Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci. 6*(1) , pp. 15–50 **1996**.

[67] Calderon, C. E., Plata, J. J., Toher, C., Oses, C., Levy, O., Fornari, M., Natan, A., Mehl, M. J., Hart, G., Nardelli, M. B. and Curtarolo, S., The AFLOW standard for high-throughput materials science calculations, *Comput. Mater. Sci. 108*, pp. 233–238 **2015**.

[68] Sun, W., Dacek, S. T., Ong, S. P., Hautier, G., Jain, A., Richards, W. D., Gamst, A. C., Persson, K. A. and Ceder, G., The thermodynamic scale of inorganic crystalline metastability, *Sci. Adv. 2*, pp. e1600225 **2016**.

[69] Taylor, R. H., Rose, F., Toher, C., Levy, O., Yang, K., Buongiorno Nardelli, M. and Curtarolo, S., A RESTful API for exchanging materials data in the AFLOWLIB.org consortium, *Comput. Mater. Sci. 93*, pp. 178–192 **2014**.

[70] Deringer, V. L., Bartók, A. P., Bernstein, N., Wilkins, D. M., Ceriotti, M. and Csányi, G., Gaussian process regression for materials and molecules, *Chemical Reviews 121*(16) , pp. 10073–10141 **2021**.

[71] Bayerl, D., Andolina, C. M., Dwaraknath, S. and Saidi, W. A., Convergence acceleration in machine learning potentials for atomistic simulations, *Digital Discovery 1*(1) , pp. 61–69 **2022**.

[72] Rostoker, W., A Study of Ternary Phase Diagrams of Tungsten and Tantalum, *Battelle Memorial Inst. Defense Metals Information Center, Columbus, Ohio* **1963**.

[73] Kusoffsky, A., Thermodynamic evaluation of the ternary Ag–Au–Cu system—including a short range order description, *Acta materialia 50*(20) , pp. 5139–5145 **2002**.

[74] Cao, W., Chang, Y., Zhu, J., Chen, S. and Oates, W., Thermodynamic modeling of the Cu–Ag–Au system using the cluster/site approximation, *Intermetallics 15*(11) , pp. 1438–1446 **2007**.

[75] Peterson, G. G. and Brgoch, J., Materials discovery through machine learning formation energy, *Journal of Physics: Energy 3*(2) , pp. 022002 **2021**.

[76] Perdew, J. P., Burke, K. and Ernzerhof, M., Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett. 77*, pp. 3865–3868 **1996**.

[77] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E., Scikit-learn: Machine Learning in Python, *J. Mach. Learn. Res. 12*, pp. 2825–2830 **2011**.

[78] Plimpton, S., Fast parallel algorithms for short-range molecular dynamics, *J. Comput. Phys. 117*(1) , pp. 1–19 **1995**.

[79] Oses, C., Gossett, E., Hicks, D., Rose, F., Mehl, M. J., Perim, E., Takeuchi, I., Sanvito, S., Scheffler, M., Lederer, Y., Levy, O., Toher, C. and Curtarolo, S., AFLOW-CHULL: cloud-oriented platform for autonomous phase stability analysis, *J. Chem. Inf. Model 58*(12) , pp. 2477–2490 **2018**.

[80] Ong, S. P., Richards, W. D., Jain, A., Hautier, G., Kocher, M., Cholia, S., Gunter, D., Chevrier, V. L., Persson, K. A. and Ceder, G., Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis, *Comput. Mater. Sci. 68*, pp. 314–319 **2013**.