Free energy of $(Co_xMn_{1-x})_3O_4$ mixed phases from machine-learning-enhanced *ab initio* calculations

Suzanne K. Wallace [1,2], Anton S. Bochkarev,[3] Ambroise van Roekeghem [1,2]
Javier Carrasco [4], Alexander Shapeev [5] and Natalio Mingo [1,2,*]

[1]*CEA, LITEN, 17 Rue des Martyrs, 38054 Grenoble, France*
[2]*Université Grenoble Alpes, 621 Avenue Centrale, 38400 Saint-Martin-d'Hères, France*
[3]*Interdisciplinary Centre For Advanced Materials Simulation, Universitätsstraße 150, Ruhr-Universität Bochum, 44801 Bochum, Germany*
[4]*Centre for Cooperative Research on Alternative Energies (CIC energiGUNE), Basque Research and Technology Alliance (BRTA),
Alava Technology Park, Albert Einstein 48, 01510 Vitoria-Gasteiz, Spain*
[5]*Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Nobel St. 3, Moscow 143026, Russia*

$(Co_xMn_{1-x})_3O_4$ is a promising candidate material for solar thermochemical energy storage. A high-temperature model for this system would provide a valuable tool for evaluating its potential. However, predicting phase diagrams of complex systems with *ab initio* calculations is challenging due to the varied sources affecting the free energy, and with the prohibitive amount of configurations needed in the configurational entropy calculation. In this work we compare three different machine-learning (ML) approaches for sampling the configuration space of $(Co_xMn_{1-x})_3O_4$, including a simpler ML approach, which would be suitable for application in high-throughput studies. We use experimental data for a feature of the phase diagram to assess the accuracy of model predictions. We find that with some methods, data pretreatment is needed to obtain accurate predictions due to inherently composition-imbalanced training data for a mixed phase. We highlight that the important entropy contributions depend on the physical regimes of the system under investigation and that energy predictions with ML models are more challenging at compositions where there are energetically competing ground state crystal structures. Similar methods to those outlined here can be used to screen other candidate materials for thermochemical energy storage.

I. INTRODUCTION

The mixed phase system $(Co_xMn_{1-x})_3O_4$ is currently under investigation for applications in next-generation concentrated solar energy storage technologies, which are based on reversible redox reactions of metal oxides [1–3]. The motivation of the mixed phase is to attempt to minimize the shortcomings of the pure end members, such as cost and toxicity for Co_3O_4, and sluggish oxidation rate and poor reversibility for Mn_3O_4 [4], by combining them in $(Co_xMn_{1-x})_3O_4$. Gaining atomic-level insights into the stability of Co_3O_4 and Mn_3O_4 mixtures at high temperatures is therefore an important step towards the rational design of these materials. However, the accurate prediction of phase diagrams with free energies derived from *ab initio* calculations can be particularly challenging for complex crystal structures. For such systems, it may be necessary to sample extremely large numbers of configurations.

Based on density functional theory (DFT) calculations of formation energy at $T = 0$ K, published in Ref. [1], all compositions of $(Co_xMn_{1-x})_3O_4$ should decompose into the pure phases Co_3O_4 and Mn_3O_4. However, the mixed phases have been successfully synthesized in a number of works [1–3]. This implies that entropic contributions to the free energy of

the mixed metal oxide phases must be responsible for the stability at finite temperatures [5,6]. It has been suggested that a large number of metastable configurations close in energy to the $T = 0$ K ground state may provide the explanation here [1]. However, the full configuration space of this mixed phase system would require millions of *ab initio* calculations to study exhaustively, even after discounting symmetrically equivalent structures [1]. Recent advances in the coupling of artificial intelligence (AI) techniques and materials design are creating new opportunities to tackle this challenge [7–9]. Specifically, AI-aided approaches, mainly from the subdomain of machine learning (ML) are helping to significantly reduce the number of required computations, making the materials design process faster and cheaper than conventional high-throughput exploration.

Here, in order to investigate possible entropic contributions to the reduction of the free energy of $(Co_xMn_{1-x})_3O_4$, and to develop a fuller theoretical understanding of the stability of these mixed phases, we use and compare different ML methods to augment the DFT calculations performed in Ref. [1]. To evaluate the accuracy of the sampling of the configuration space with these different ML approaches, we use the methods described in Ref. [10] to calculate the phase coexistence region between the tetragonal hausmannite (H) and spinel (S) phases of $(Co_xMn_{1-x})_3O_4$ from the predicted free energy curves. In each case, the configurational entropy contribution to the free energy is calculated using the different
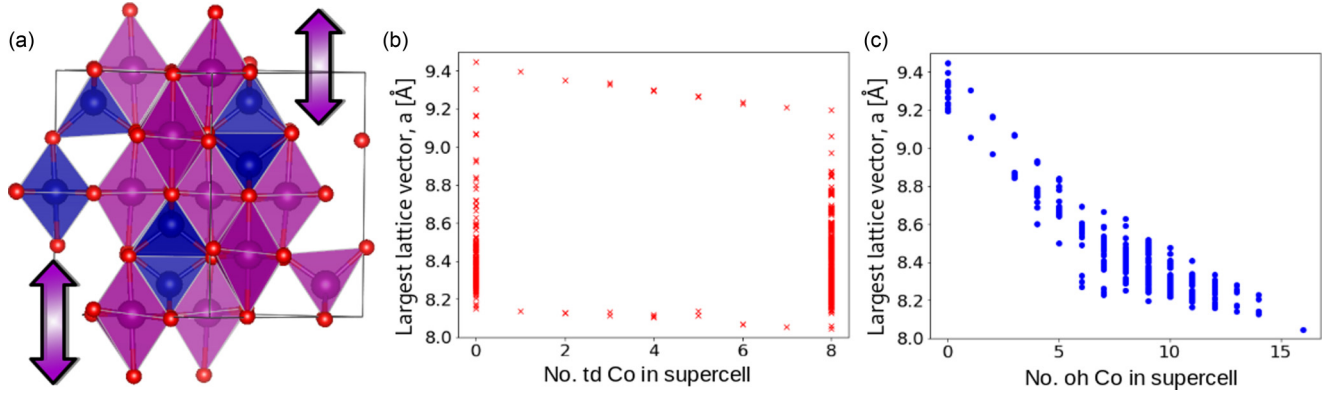
*Corresponding author: natalio.mingo@cea.fr

FIG. 1. (a) $(Co_xMn_{1-x})_3O_4$ crystal structure indicating tetragonal distortion introduced by Mn ions on octahedral lattice sites. Produced with VESTA [11]. Largest lattice parameter a in $(Co_xMn_{1-x})_3O_4$ supercells as a function of (b) number of Co on tetrahedral (td) lattice sites and (c) number of Co on octahedral (oh) lattice sites, data from Ref. [1] for set A- and B-type structures.

ML approaches and compared to experimental data for this feature of the phase diagram. We discuss the relative strengths of the different methods employed in this work and show that an accurate prediction of different parts of the phase diagram of this system depends the most strongly on accurately accounting for different contributions to the free energy.

II. MIXED METAL OXIDE PHASES

The pure Co_3O_4 phase has a cubic spinel structure (S) with space group 227. Below $1170\,°C$, the pure Mn_3O_4 is a tetragonally distorted spinel structure (hausmannite, H) with space group 141 and distortion along one crystallographic axis relative to the S structure. Above $1170\,°C$, Mn_3O_4 transitions to a S structure [12]. The transition metal (TM) sites in these structures are either tetrahedrally (td) or octahedrally (oh) coordinated to the oxygen atoms. Mn ions substituted on oh Co sites in the S structure are Jahn-Teller active, inducing a tetragonal distortion of the lattice [13]. These Jahn-Teller active oh sites are largely, but not solely, responsible for a composition-dependent tetragonal distortion of the mixed phases as shown in Fig. 1, where small changes in lattice parameters are also correlated to a number of td Co. Plots of all DFT lattice parameters as a function of Co composition are shown in the Supplemental Material (SM) [14] (Fig. 1).

Three distinct Co-Mn substitution schemes were considered in Ref. [1] and form the basis of the training data used for the ML methods in this study. Taking as a base the pure Mn_3O_4 with the cubic structure of Co_3O_4, the data sets referred to as "set A" ("set B") in this study correspond to starting to substitute only td (oh) sites with Co until all td (oh) sites are occupied by Co, and only then beginning to substitute onto the oh (td) sites. Set A and B, as referred to in this work, correspond to "scenario 1" and "scenario 2", respectively, in Ref. [1]. In the data set referred to as "set C" in this work (or "scenario 3" in Ref. [1]), Co ions are allowed to substitute freely on all TM sites in the crystal. The DFT training data are composed of such structures after atomic positions and cell volumes have been relaxed.

It is expected that the set C-type structures should be higher in energy than those of either set A or set B types [1], whichever is lower at the concentration of interest. However,

without the constraint of preferential occupation, the total possible combination space is dramatically larger, opening up the possibility of entropy stablization of the mixed phases [6].

To compare the sampling size required in each of the three substitution scenarios, we take the 56 atom supercell used in the DFT calculations in Ref. [1]. Each supercell contains 24 TM sites, 8 of which are td coordinated and 16 are oh coordinated. In the set C scenario for N Co ions in the supercell, the total number of possible combinations on the 24 TM sites is given by

$$C_N = \frac{24!}{N!(24-N)!}. \tag{1}$$

Whereas in set B, the total number of combinations for $N < 16$ would be

$$C_N = \frac{16!}{N!(16-N)!}. \tag{2}$$

Then for $N > 16$, this would be

$$C_N = \frac{8!}{(N-16)![8-(N-16)]!}. \tag{3}$$

Similarly for set A, the constraint of preferential filling gives a total number of possible combinations for $N < 8$ as

$$C_N = \frac{8!}{N!(8-N)!}. \tag{4}$$

Then for $N > 8$, this would be

$$C_N = \frac{16!}{(N-8)![16-(N-8)]!}. \tag{5}$$

Due to the availability of all 24 sites for substitutions in set C-type structures, at intermediate compositions the total number of possible configurations C_N is enormous. For example, for $x = 0.5$ in $(Co_xMn_{1-x})_3O_4$, C_N for set C is 2.70×10^6. In comparison, for the same x for sets A and B, C_N is just 1820. For this reason, the data set in Ref. [1] sampled sets A and B thoroughly, but was unable to substantially sample set C, even after eliminating symmetrically equivalent structures. Therefore, it was not possible in this study to consider configurational entropy contributions from this large portion of the total configuration space. For this reason, we have used

the DFT data set of Ref. [1] as a starting point for training different ML models to predict the energies of the set C-type structures.

III. MACHINE-LEARNING METHODS

Recent years have seen many successes in the use of ML methods for various applications in materials science [15–17], where data sets are used to train algorithms to predict properties of interest. Here, to approximate the formation energy in complex $(Co_xMn_{1-x})_3O_4$ mixed phases, we apply supervised machine learning [18]. For this, various learning algorithms and various representations for the atomic configurations (or "descriptors") exist which take into account rotational and translational invariance [19–23]. Comparisons of some of these methods can be found, highlighting their differing predictive capabilities and computational cost [24,25].

Common to all of these methods is the need to optimize the model. During training it is important to consider overfitting and underfitting, or "bias-variance" trade-off [26]. To avoid overfitting or underfitting, it is common to trial various "hyperparameters" associated with the model complexity for the particular ML method and to then compare the root mean square error (RMSE) between reference outputs and model predictions at the end of the training process. This is done with both the training and validation data sets, to optimize the hyperparameters, and a "hold-out" set of data, which was not seen by the algorithm during training, to assess model accuracy. Note that some differing terminologies are used in the literature where the term "validation set" is sometimes used for the "hold-out set," or is sometimes replaced by "test set." In the case of underfitting, the RMSE on the training data is large. While for overfitting the RMSE on the training data is typically small, but coincides with either a plateau or increase in the RMSE for the validation set with increasing model complexity.

In this work we investigate the use of three different supervised ML methods. The models are trained with a data set of hundreds of final, relaxed structures of $(Co_xMn_{1-x})_3O_4$ and their corresponding total energies. In the next sections we provide a brief outline of each method and a description of their corresponding training hyperparameters.

A. Artificial neural networks (ANNs)

Our ANN methodology is that of Ref. [27], which makes use of symmetry functions for representing chemical environments, similar to those developed by Behler and Parrinello [19,28,29]. During the construction of a descriptor for the chemical environment, it is important to ensure that it is invariant with respect to rotations and translations of the system and also to permutations in the ordering of the atomic inputs. The design of the descriptor can also allow for more automatic optimization of the model, and hence fewer hyperparameters that need to be tuned manually [27].

The representation of input data to the ANN in this study is an n-dimensional vector for each atom i in the system,

$$D_i^l = \sum c^l(z_i)c^l(z_j)e^{-\sigma^l(r_{ij}-\eta^l)^2}f_{cut}(r_{ij}, R_{cut}), \quad (6)$$

which contains information about the atoms surrounding atom i out to a certain cutoff distance R_{cut} by using a Gaussian expansion of interatomic distances r_{ij} to probe the surrounding space where l labels each Gaussian in the expansion. Parameters η and σ are related to the probing of this space [27]. In this particular implementation, these two parameters are automatically optimized during training so that regions of space which are more important for distinguishing different atomic configurations are selected. f_{cut} is the cutoff function which ensures that the contributions from neighboring atoms smoothly goes to zero as r_{ij} approaches R_{cut}. $c^l(z_i)$ and $c^l(z_j)$ are the parts of the vector descriptor that are related to the chemical identity of atom i and its neighbor j, defined by their atomic numbers z_i and z_j. These vectors are initialized randomly and are also optimized automatically during training.

The output of this ANN is the energy of atom i. The total energy of a system is then the summation of all of the atomic energies,

$$E_{tot}^{ANN} = \sum_i E_i. \quad (7)$$

The ANN is trained by minimizing the loss function. The original implementation in Ref. [27] and also in Ref. [30] trains simultaneously with forces and energy, whereas in this study we train only with energies due to the lack of force information in the initial data set. Without including the forces, the loss function in Ref. [27] is reduced to

$$\left\| E_{tot}^{ANN} - E_{tot}^{DFT} \right\|^2, \quad (8)$$

where E_{DFT} are the reference total energies of the training data and E_{ANN} are the predictions by the ANN model for the same atomic configurations. Essentially, the ANN model is fit to minimize the RMSE between DFT energies and predicted energies for the training data.

Despite the automatic optimization of some of the parameters in Eq. (6), there are still various hyperparameters that need to be tuned for the ANN, and carefully chosen to avoid overfitting or underfitting. First, there is the cutoff distance R_{cut} used when creating the descriptor for each atom in the system. Then for model complexity, the architecture of the neural network can be tuned by varying the layer size and depth. A larger network allows for more model flexibility, but too large can result in overfitting to the training data. Lastly, the number of epochs can also be tuned. This is the number of training cycles the model is optimized over. Again, here if the model is trained over too many epochs this can result in overfitting.

B. Moment tensor potentials (MTPs)

MTPs are a class of ML potentials first proposed in Ref. [20] and later extended for multicomponent systems [31]. The latest release of the software is described in Ref. [32]. Similarly to the ANN methodology outlined above, MTPs represent the energy of an atomic configuration as a sum of contributions of local atomic environments of each atom i,

$$E_{tot}^{MTP} = \sum_{i=1}^n E_i. \quad (9)$$

Each contribution E_i is linearly expanded via a set of basis functions B_α:

$$E_i = \sum_\alpha \xi_\alpha B_i^\alpha, \tag{10}$$

where the set of parameters $\boldsymbol{\xi} = \{\xi_\alpha\}$ are obtained during training. Unlike ANNs, the atomic environments are instead represented by the moments of inertia of the neighboring atoms. The moment tensor descriptors of the ith atom consist of radial and angular parts and are given by

$$M_i^{\mu,\nu} = \sum_j f^\mu(r_{ij}, z_i, z_j) \underbrace{\mathbf{r}_{ij} \otimes \cdots \otimes \mathbf{r}_{ij}}_{\nu \text{ times}}. \tag{11}$$

Here $\mathbf{r}_{ij}$ is the position of the jth atom relative to the ith atom and r_{ij} is, as before, its length. The radial part is further expanded as

$$f^\mu(r_{ij}, z_i, z_j) = \sum_{\beta=1}^{N_Q} c_{\mu, z_i, z_j}^{(\beta)} Q^{(\beta)}(r_{ij}) \tag{12}$$

and defines different shells around atom i and contains a set of radial parameters $\mathbf{c} = \{c_{\mu, z_i, z_j}^{(\beta)}\}$ which are also obtained during training, and radial basis functions $Q^{(\beta)}(r_{ij})$ based on polynomials. A cutoff radius R_{cut} is also used here to ensure smooth behavior at the edges of the atomic environments. The angular part $\underbrace{\mathbf{r}_{ij} \otimes \cdots \otimes \mathbf{r}_{ij}}_{\nu \text{ times}}$ contains the angular information about the atomic environment and is a rank ν tensor. When the radial and angular components are combined to form $M^{\mu,\nu}$, for $\mu = 0$, the increasing ranks of the tensor can be interpreted mechanically as the number of atoms within R_{cut} of atom i (or as the "mass" of these atoms) for $M^{0,0}$, the center of mass scaled by the mass for $M^{0,1}$ and the tensor of the second moments of inertia for $M^{0,2}$, etc. For $\mu > 0$, this can then be interpreted as weighted moments of inertia [20].

The basis functions B_α in Eq. (10) are constructed by defining the "level of moments" via

$$\text{lev} M^{\mu,\nu} = 2 + 4\mu + \nu, \tag{13}$$

where the coefficients in Eq. (13) were found to be optimal in Ref. [33]. The tensor contractions of a number of moments are defined by adding together such levels. All such contractions of one or more moments form the basis functions B_α. As for the symmetry functions representation with the ANN, these basis functions are invariant to atomic permutations, rotations, and reflections. The functional form of the MTP is defined first by choosing a maximum level for the basis set lev_{max}, and then including all basis functions whose level is less than or equal to that maximum. And second, by the size of the radial basis N_Q in Eq. (12).

The parameters $\boldsymbol{\xi}$ and $\mathbf{c}$ make up the total set of parameters that are found during training, $\boldsymbol{\theta} = \{\boldsymbol{\xi}, \mathbf{c}\}$. The total number of basis functions (and hence the number of the corresponding parameters $\boldsymbol{\xi}$) grows exponentially with lev_{max}, but the number of radial functions increases linearly with lev_{max} and N_Q. lev_{max} and N_Q are therefore the hyperparameters which define the total set of parameters to be found during training and hence define the model complexity. The computational expense increases with the total number of free parameters;

and the optimal number of such parameters depends on the total training set size, with the possibility of overfitting with too many parameters for a small data set. During the training process, fitting with MTP is performed using output variables for each configuration in the training set for quantities from *ab initio* calculations: total energy, forces, and the stress tensor. Weights can be set to express the importance of each of these quantities during the optimization. In the case of this work, as the initial data set did not contain sufficient force information, weights for forces were set to zero.

C. Extended mean field (EMF) model for a two-species, two-site-type compound

An approach commonly known as the Miedema model is often used to describe the mixing enthalpy of alloys [5,34]. The Miedema model was originally developed in the context of liquid alloys, and it was expressed in terms of solution enthalpies and interfacial area between species. Nevertheless, its underlying basic concept is shared with a mean field model in which the average energy of atomic configurations with average species occupations ϕ_i is given by atom-pair contributions $\phi_i \epsilon_{i,j} \phi_j$, where i and j denote the species type [35]. In binary mixtures of species a and b the two concentrations are related by $\phi_b = 1 - \phi_a$. The total energy is

$$\langle E \rangle = \phi_a^2 \epsilon_{aa} + 2\phi_a \phi_b \epsilon_{ab} + \phi_b \epsilon_{bb}. \tag{14}$$

The case of $(Co_x Mn_{1-x})_3 O_4$ is more complex because there are two distinct site types: oh and td, and two species: Co and Mn. We thus extend the model to account for all the possible interactions between pairs of species and site types. There are four types of atoms: Co on a td site, Co on an oh site, Mn on a td site, and Mn on an oh site. This results in a 4×4 matrix of pair interactions ϵ_{ij}. However, the matrix is symmetric because $\epsilon_{ij} = \epsilon_{ji}$, so only the ten elements contained in the diagonal and the upper triangle are independent parameters. We can write the energy as the quadratic form:

$$\langle E \rangle = \underline{x} \epsilon \overline{x} = x_i \epsilon_{ij} x_j, \tag{15}$$

where we assume summation over repeated indices. Here $\underline{x} = \overline{x}^T = \{x_{\text{td}}, 1 - x_{\text{td}}, x_{\text{oh}}, 2 - x_{\text{oh}}\}$ defines the average occupations of the system, where $0 \leqslant x_{\text{td}} \leqslant 1$ ($0 \leqslant x_{\text{oh}} \leqslant 2$) denote the average Co concentration in the td (oh) sites, and $1 - x_{\text{td}}$ and $2 - x_{\text{oh}}$ denote the corresponding Mn concentrations. The occupation range of the oh sites is twice as large as for the td sites, since there are twice as many of the former than of the latter. The ϵ_{ij} are adjustable parameters to be fitted to the DFT energies. To do this, we reshape the ten independent elements of ϵ as a single column array $\overline{M}$ and write N equations for each of the DFT configurations that we want to include in the fit as

$$E^{(n)} = x_i^{(n)} x_j^{(n)} M_u = A_{n,u} M_u, \tag{16}$$

where $u \equiv (i, j)$ labels the ordered list of pairs (i, j), with $1 \leqslant i \leqslant 4, 1 \leqslant j \leqslant i$, containing a total of ten elements. $E^{(n)}$ is the DFT energy corresponding to configuration n, whose concentrations array is $\overline{x}^{(n)} = \{x_{\text{td}}^{(n)}, 1 - x_{\text{td}}^{(n)}, x_{\text{oh}}^{(n)}, 2 - x_{\text{oh}}^{(n)}\}$. Directly substituting this into Eq. (16), the explicit values of $\overline{A_n}$ in

terms of the elements of $\overline{x}^{(n)}$ are

$$\overline{A} = \{x_{td}^2, (1-x_{td})x_{td}, x_{oh}x_{td}, (2-x_{oh})x_{td},$$
$$(1-x_{td})^2, x_{oh}(1-x_{td}), (2-x_{oh})(1-x_{td}), \quad (17)$$
$$x_{oh}^2, (2-x_{oh})x_{oh}, (2-x_{oh})^2\},$$

where the superscript (n) is omitted from the x for simplicity.

Equation (16) is a system of N equations with ten unknowns. It can be written in compact form as

$$\mathbf{A}\overline{M} = \overline{E}. \quad (18)$$

The solution to this (usually overdetermined) system can be obtained as

$$\overline{M} = \mathbf{A}^{-1}\overline{E}, \quad (19)$$

where $\mathbf{A}^{-1}$ is the pseudoinverse of $\mathbf{A}$. We implement this solution via a standard least-squares algorithm. Rather than fit to the total DFT energies, the model is fit with the formation energies calculated via

$$E_{form}^{(n)} \equiv E_{tot}^{(n)}(x) - \left[\left(1 - \frac{x}{3}\right)E_{tot}^{x=0} - \frac{x}{3}E_{tot}^{x=3}\right], \quad (20)$$

where $E_{tot}^{x=0}$ and $E_{tot}^{x=3}$ correspond to the total energy of the pure phases. Formation energies are zero at the composition extremes, when $x = 0$ or $x = 3$. To ensure that this is the case, from each equation (n) in Eq. (16) we subtract the energy corresponding to the linearly weighed average between $E(x = 0)$ and $E(x = 3)$ at $x(n)$. Using Eq. (17), this yields

$$\overline{E}_{form}^{(n)} = \left(\overline{A}_n - \frac{x_{td}^{(n)} + x_{oh}^{(n)}}{3}\overline{B} - \overline{C}\right)\overline{M}, \quad (21)$$

where $\overline{B} = \{1, 0, 2, 0, -1, 0, -2, 4, 0, -4\}$ and $\overline{C} = \{0, 0, 0, 0, 1, 0, 2, 0, 0, 4\}$. This equation is solved similarly to Eq. (19).

The only means to improve the prediction performance of the EMF model is to increase the training set size, compared to the various different hyperparameters associated with training the ANN or MTP. Unlike the ANN and MTP models, the implementation of the EMF model in this work does not depend on the exact atomic positions or cell volume of the relaxed structures. While the classic Miedema model involves additional terms, such as elastic strain and lattice structure information [5,36,37], the EMF model in this work depends only on the oh or td site occupancy of the Co or Mn species.

IV. RESULTS AND DISCUSSION

A. ML prediction of formation energies

1. Development of the training procedure

Training, validation, and hold-out sets. The initial DFT training set from Ref. [1] contains 540 relaxed structures of A and B type and 16 C type (see Sec. II). Another 496 C-type structures were calculated with the same settings as in Ref. [1]: spin-polarized DFT calculations with the Vienna *ab initio* simulation package (VASP) [38,39] and the strongly constrained and appropriately normed (SCAN) meta-generalized gradient approximation (meta-GGA) functional [40]. 115 of the C-type structures contained 12 oh Co and 2 td Co. The remaining 381 C-type structures calculated are

randomly selected with various compositions. The total DFT data set contains 1052 structures. When predicting energies for the full configuration space of set C, the entire DFT data set described above is used to train and validate the ML model. However, the design of the training procedure (outlined in the following subsections below) involved splitting this full training set into various training, validation, and hold-out sets.

Imbalanced data sets. The training data across the full composition range contain many more possible structures at intermediate compositions than at the composition extremes. This can pose a challenge when wanting to accurately predict energies of structures that are less well represented in the training data. As a test, the DFT data for A- and B-type structures from Ref. [1] was randomly split into a training set with 360 structures, and a test set with the remaining ones. The initial training data for set A- and B-type structures contained only the final, relaxed structures. Therefore, there was insufficient force information from the relaxation trajectory to fit interatomic potentials for the systems, as is often done in other ML studies to predict the full potential energy surface (PES) [19,27,41,42]. Only the basins of the PES for each mixed phase were of interest for this study. Therefore, instead ML models were trained to map initial, unrelaxed atomic configurations onto final, relaxed energies based on DFT data of similar structures.

The different hyperparameters associated with each of the ML methods (see Methods section) were trialled to determine those which resulted in predictions for the validation set with the smallest RMSE. The best performance models were then used to predict total energies, and from them formation energies $E_{form}^{(n)}$ via Eq. (20). Plots of formation energy as a function of composition are shown in Figs. 2(a)–2(c). The ANN method yields a smaller RMSE than MTP (3.1 vs 4.6 meV per atom). However, it performs worse at the composition extremes than MTP, where the formation energy even drops below the dashed line marking zero energy at certain compositions.

In order to better balance the sampling for the ANN training, we trialled generating artificial data for the minority data based on k-nearest neighbors using the SMOTE algorithm [43]. However, the structures that this method generated were often unphysical. This is likely to be due to the small sample size and high dimensionality of the data with each atom type, its coordinates as well as the types and coordinates of its neighbors within the cutoff radius all being used to define the representation of the structure. We found that the predictive performance of the ANN at composition edges improves by weighting higher (or oversampling) the minority data during training [Fig. 2(b)]. This however slightly increases the RMSE from 3.1 to 3.6 meV per atom.

Despite the low RMSEs obtained with the ANN, based on the better qualitative agreement of the MTP predictions for this imbalanced data set without the need for any data pretreatment, we do not proceed further with the ANN method for this study. The optimal pretreatment of training data is an open research question [44] and beyond the scope of this work. However, the improvements observed from oversampling imply that further investigations into data pretreatment procedures could result in a very good predictive performance with the ANN. The optimal choice of ML method and data
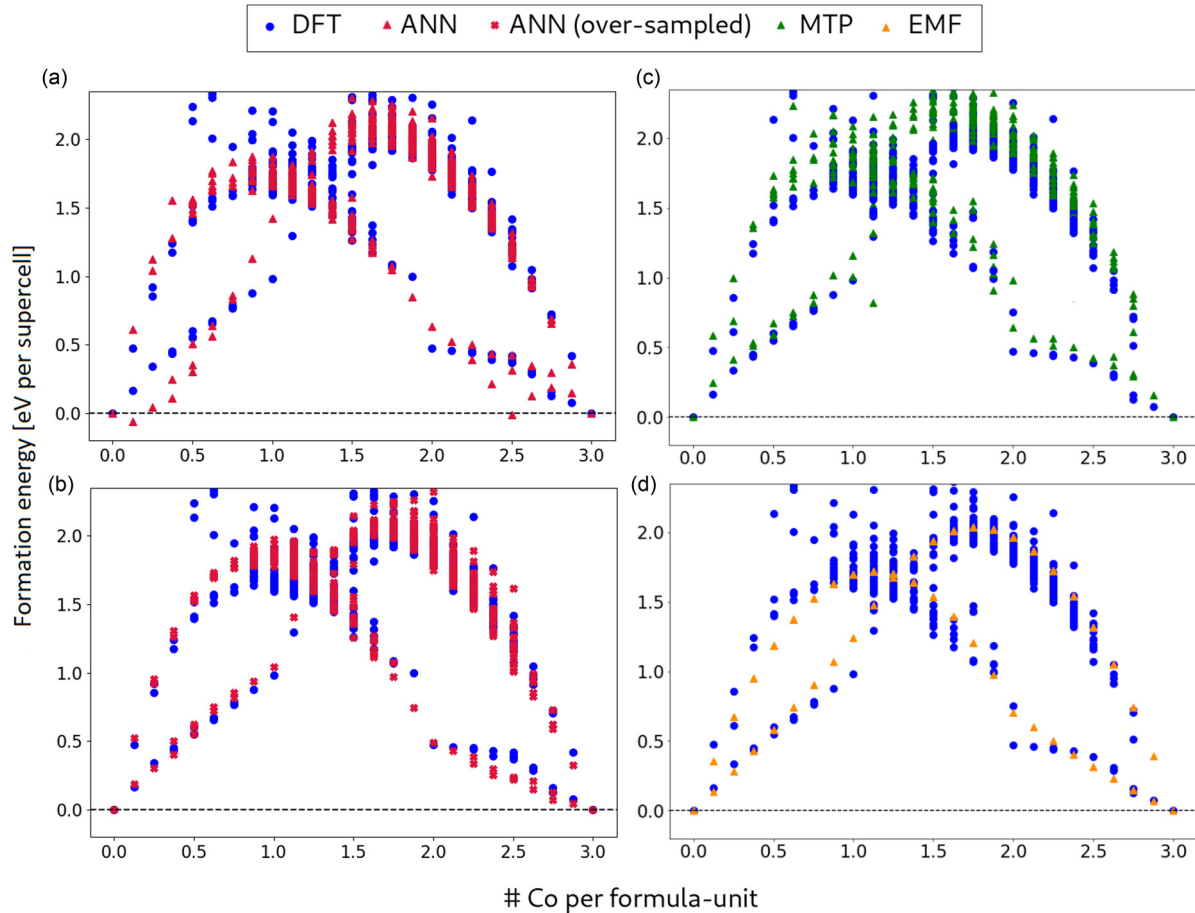
FIG. 2. Predicted formation energies and reference DFT data for set A- and B-type configurations of $(Co_xMn_{1-x})_3O_4$. All models are trained with 360 structures of the total data set shown (540 structures). Structures not included in the training set are used as the hold-out set to calculate the RMSE in units of meV per atom. (a) Predictions from the ANN with a RMSE of 3.1. (b) Predictions from the ANN with an "oversampled" training set and a RMSE of 3.6. (c) Predictions from the MTP method with a RMSE of 4.6. (d) Predictions from the EMF model with a RMSE of 2.7.

preparation procedure may also be dependent upon the particular system under investigation and the available data. For example, the prediction accuracy may differ considerably for a large, balanced data set where it is desirable to reduce the impact of outliers in the data set on the model.

The third method trialled is the EMF model shown in Fig. 2(d). This method gave the best RMSE of all three methods (2.7 meV per atom) and did not suffer from the visibly poor predictions at the composition edges as with the ANN. It is likely that applying the constraint of zero formation energy at the composition extremes in the EMF model is responsible for this improvement. However, the EMF predictions do not reproduce well the spread in the DFT data at each composition.

Assessment of prediction capabilities. Figure 1 shows a substantial variation in the lattice parameters of the $(Co_xMn_{1-x})_3O_4$ supercells as a function of the number of Co. It has been shown that cluster expansion predicted energies are typically less accurate when the systems undergo substantial atomic relaxation [45]. In the data set for $(Co_xMn_{1-x})_3O_4$, the total energy of each supercell is most strongly dominated by the total number of the substituting species, i.e., number of Co, in the supercell. However, finer energy differences between

structures depend on the fraction of Co on td or oh sites and, even more subtly, the variations in atomic arrangements between structures with the same number of Co on the same type of crystallographic sites. As our EMF model takes only the number of oh and td Co as its representation of the structure, it is unable to distinguish between different structures having the same number of oh and td Co.

The ANN and MTP methods are able to distinguish the energies of structures with the same number of oh and td Co only if the unrelaxed ionic coordinates in the input structures are scaled by the relaxed lattice parameters. Figure 3 shows the predicted MTP versus DFT energies when using a completely unrelaxed structure as the input (a), and when the unrelaxed ionic coordinates are scaled by the relaxed lattice parameters (b). The Pearson correlation coefficient increases from -0.0398 to 0.778 and the RMSE decreases from 6.14 to 0.969 meV per atom when using the relaxed lattice parameters. Therefore it is important to obtain estimates of relaxed lattice parameters for set C structures before predicting their total energy with the MTP model, as described in the next subsection.

Reference [45] highlighted the crucial influence of atomic relaxation in the accuracy of the predicted formation energies
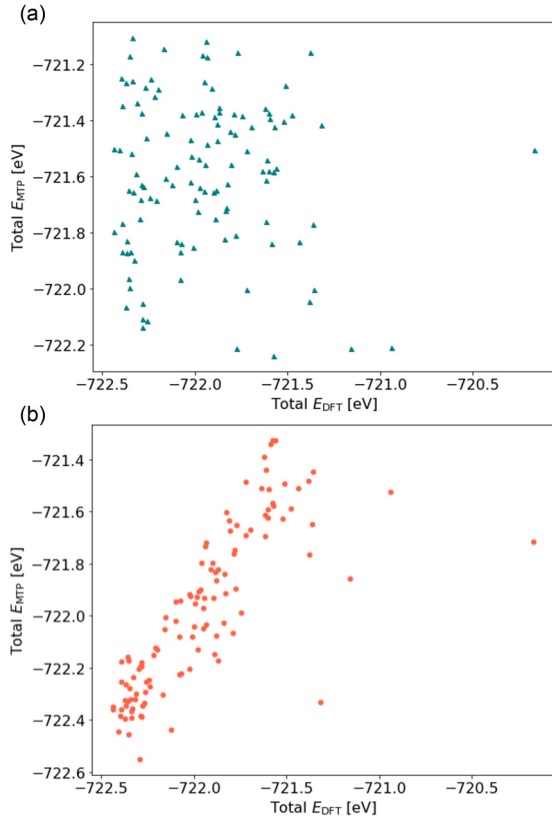
FIG. 3. Correlation between reference DFT energies and MTP predicted energies of the set C 12 oh 2 td hold-out set with (a) completely unrelaxed structures as inputs and (b) structures with unrelaxed ionic coordinates scaled by relaxed lattice parameters.

of mixed phases using cluster expansion. It appears that a similar phenomenon is at play in our particular use of ML, i.e., training only on fully relaxed configurations, whereby

strong relaxations lead to decreased prediction capability. Nonetheless, this does not prevent the accurate prediction of the solubility gap (Sec. IV B), or to calculate specific site Co occupation probabilities that are in principle verifiable by experiment [14]. The fact that the training set is one order of magnitude smaller than one would need to train a full-fledged interatomic potential makes this approach useful and attractive when the availability of computed *ab initio* data is limited.

Prediction of relaxed lattice parameters. Studies on alloys often use Vegard's [24,46] empirical law of a linear relationship at constant temperature between the lattice parameters and the alloy concentration, resulting from the different sizes of the substituting atoms. However, for $(Co_xMn_{1-x})_3O_4$, factors other than concentration also affect the lattice parameters for a given Co:Mn ratio. The occupation of the oh sites strongly influences the lattice parameters (Sec. II). Also, increasing Mn content does not smoothly distort the cubic structure into the tetragonal cell shape of Mn_3O_4. DFT training data in Fig. 4(d), and experimental measurements for this system [13] show a composition-dependent phase transition of the ground state structure from the tetragonal spinel of Mn_3O_4 to the cubic spinel of pure Co_3O_4. The composition at which the ground state phase transition occurs differs between the A-, B-, and C-type structures [Figs. 4(a)–4(c)].

The lack of force information in the initial training data set prevents us from relaxing cell shapes with an MTP potential. Instead, we use all of the DFT training data (Sec. IV A 1) to estimate lattice parameters based on the number of oh and td Co in the supercell. The training data is classified as cubic when all three lattice parameters are the same within a tolerance of 0.05 Å, and tetragonal otherwise. This gave 824 cubic and 228 tetragonal structures. Their formation energies as a function of composition are displayed in Fig. 4(e). Best-fit quadratic surfaces are then fit to cubic and tetragonal data separately for each lattice parameter as a function of the number of Co on td
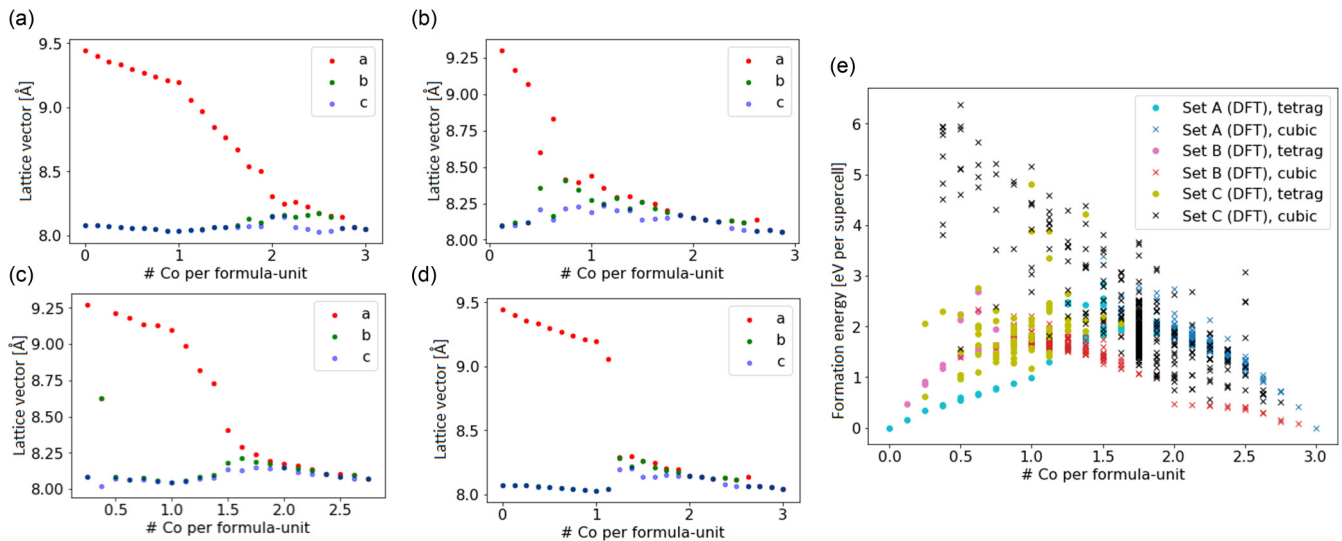

FIG. 4. Lattice parameters (denoted *a*–*c* from largest to smallest) as a function of the number of Co atoms in the supercell from the training DFT data for the lowest-energy structures that are (a) A type, (b) B type, (c) C type, and (d) all three sets combined. Structures are identified as cubic when all lattice parameters are equal, within a tolerance of 0.05 Å. (e) Formation energy as a function of composition for the same DFT data with structures identified as cubic (tetragonal) denoted by crosses (filled circles).

and oh sites in the supercells. For cases with several structures with the same number of oh and td Co, the minimum energy structure is used in the fit. These surfaces are then used to predict the lattice parameters for cubic and tetragonal structures, both when sampling the full configuration space of set C, and also to replace the true relaxed lattice parameters of the training set. All best-fit surfaces are included in the SM [14] (Secs. 2 and 3).

For both tetragonal and cubic structures, the largest weighted mean error in the fits was for the largest lattice parameter a, which is the one varying the most as a function of composition. With this method any structure with the same number of oh and td Co will be assigned the same volume. However, the distribution of cell volumes in the DFT data for C-type structures shows substantial variations, even between structures with the same number of oh and td Co (see Sec. 4 of the SM [14]). Cell volumes of the 12 oh 2 td Co set have a range of 9.78 Å^3 compared to 15.44 Å^3 for the set of all C-type structures with 14 Co, with the largest variance in the lattice parameters being that of the largest lattice parameter a. On this basis, the ability to distinguish structures with the same number of oh and td Co may be beyond the capabilities of the current model. However, the ability to distinguish structures with the same number of Co but a different td:oh ratio is still attainable, and important when later accounting for different sources of entropy [10].

Additional tetragonal training data. The Co-poor side of Fig. 4(e) shows coexistence between tetragonal and cubic C-type structures, with the latter corresponding to higher energies. This behavior is consistent with the tetragonal to cubic phase transition in the pure Mn_3O_4 compound [12]. However, there is no such phase transition in the pure Co_3O_4 compound. For tetragonal structures, beyond the Co-poor range, our method of creating structures based on estimated lattice vectors will create some tetragonal phases that are not represented in the DFT training data, which could lead to extreme extrapolation by the ML model.

Choosing a hard cut-off composition for the existence of tetragonal structures based on the available DFT training data would be a large approximation. Instead, additional tetragonal structures are calculated with DFT and added to the training set so that some of these types of structure are present in the training data. The structures were selected by training a MTP for tetragonal structures with just the available DFT training data and predicting energies for thousands of randomly generated tetragonal C-type structures. The structures at each composition 5–15 Co in the supercell that were predicted to have the lowest energy by the MTP model were then selected for the additional DFT calculations. The tetragonal MTP was then retrained with this additional data for all subsequent parts of this work.

Energy predictions when training with the estimated lattice parameters. To assess the level of discrimination afforded by the method, first MTP was trained with all cubic data except those corresponding to $(n_{oh}, n_{td}) = (12, 2)$, which formed the hold-out set. $n_{td(oh)}$ stands for the number of Co per supercell in td(oh) sites. This test yields an RMSE of 3.61 meV per atom, but it does not achieve good correlation between reference DFT energies and model predictions. This suggests that one would need a fully trained interatomic potential to

TABLE I. RMSE in meV per atom and Pearson correlation coefficient between reference DFT energies and MTP predictions for cubic hold-out set C-type structures with different total number of Co in the supercell.

# Co	13	14	15	16	17	18
RMSE	6.06	2.46	3.57	1.40	5.40	1.05
Pearson coeff.	0.278	0.0758	0.787	0.878	0.313	0.934

distinguish between structures sharing the same (n_{td}, n_{oh}), which is not possible with the available amount of training data.

On the other hand, when the hold-out sets contain different n_{td}/n_{oh} ratios, the cubic-trained MTP displays increasingly larger Pearson correlation and smaller RMSE's as the total number of Co is increased (Table I). An exception is the 17 Co data set, whose reduction in correlation is caused by a single outlier [Fig. 5(a)]. From approximately 16 Co in the supercell (2 per formula unit), the minimum energy structures for all three substitution schemes are all cubic (Fig. 4). Therefore the
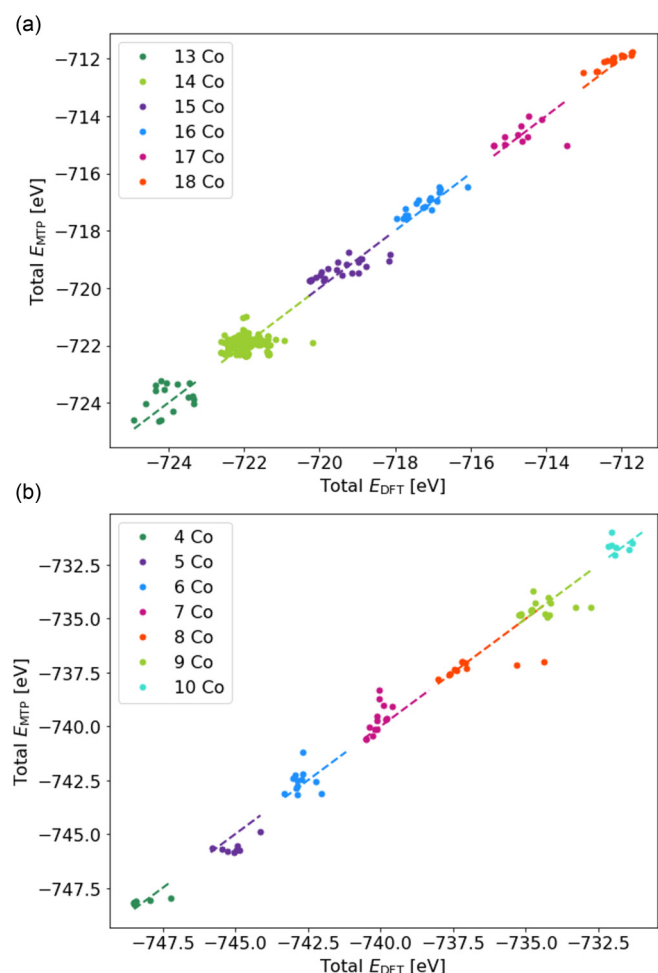

FIG. 5. Scatter plots for calculated DFT total energies against MTP predictions for hold-out sets of C-type structures with different total number of Co in the supercell for (a) cubic structures and (b) tetragonal structures.

TABLE II. RMSE in meV per atom and Pearson correlation coefficient between reference DFT energies and MTP predictions for tetragonal hold-out set C-type structures with different total number of Co in the supercell.

# Co	4	5	6	7	8	9	10
RMSE	2.38	7.51	7.52	8.29	16.9	8.58	4.63
Pearson coeff.	0.889	0.726	0.0160	0.661	0.597	0.249	−0.133

volume-composition relationship is much simpler and likely to be better represented by our surfaces of best fit.

For cubic structures with less than 15 Co in the supercell, correlations were typically poorer. For the 14 Co hold-out set, this is due to a large portion of the set being structures with the same oh:td ratio for which, as already discussed, the model with estimated lattice parameters does not yield well-correlated energy predictions. Furthermore, as can be seen from Fig. 4, between the different substitution schemes there is much more variation in cell parameters at intermediate compositions where the phase transition occurs at different compositions for the different substitution schemes. Energy predictions for cubic structures at intermediate compositions are therefore likely to be less reliable than those in the "clearly cubic" regime for >15 Co in the supercell.

Similar hold-out tests were also performed for tetragonal C-type structures and a MTP trained only on tetragonal data [Fig. 5(b) and Table II]. In this case, better correlation was typically achieved for Co-poor structures, i.e., in the "more

tetragonal" regime. However tetragonal structures have considerably more variation in lattice parameters than cubic ones, even for as few as 4 Co in the supercell in set B [Fig. 4(b)]. It is therefore likely that the lattice parameter estimates are less accurate for tetragonal structures, especially those compositionally closer to being B type. This is reflected by the relatively large RMSEs presented in Table II. The correlation substantially decreases upon increasing number of Co in the supercell.

Intermediate compositions have a complex composition-dependent phase coexistence and substantial cell volume variation. There, accurate energy predictions likely can only be obtained using a trained full potential capable of performing structural relaxation. However, for very Co-poor structures and those with >15 Co in the supercell, energy predictions with estimated lattice parameters are expected to be more reliable. In all hold-out set tests performed in this section, the settings for MTP training that achieved the best RMSE were the level eight (i.e., with lev$_{max}$ = 8, see Sec. III B) with weights of 1.0 for energy and 0.0 for stress and forces (due to lack of force information in initial training data). The cutoff radius was R_{cut} = 5 Å. These settings were therefore used in all subsequent parts of this work for sampling the configuration space of set C-type structures.

2. Filling in the configuration space

Random sampling of set C. To sample the millions of possible configurations of the 56-atom supercell for C-type structures, we generate random structures and only retain
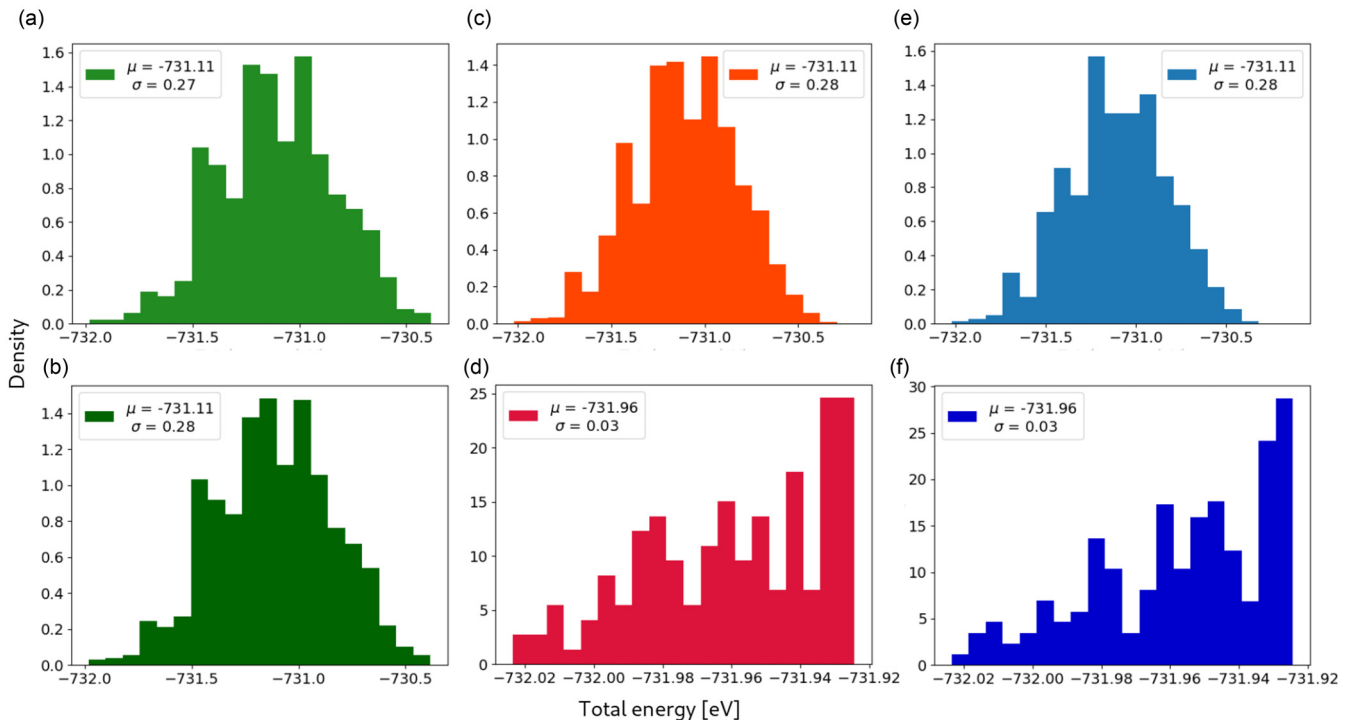

FIG. 6. MTP total energy predictions for random samples of set C-type structures of (Co$_x$Mn$_{1-x}$)$_3$O$_4$ with 10 Co in the supercell. (a) 1000 attempts batch. (b) Batch from (a) augmented with a second 1000 batch. (c) 100 000 attempts batch. (d) Lowest 0.1 eV of distribution in (c). (e) 2 million attempts batch. (f) Lowest 0.1 eV of distribution in (e). The mean (μ) and standard deviation (σ) of each distribution is given in the legend.

unique configurations. Figure 6 compares the distribution of total energies for configurations with 10 Co in the supercell predicted by MTP for random batches of 1000, 100 000, and 1 949 176 unique structures (the limit of possible configurations for this particular composition is 1 961 256). The lowest energy side of the distributions deserves particular attention, as these configurations contribute the most to the total free energy [Figs. 6(d) and 6(f)]. The mean and standard deviation of the distribution for the sample of 100 000 structures [Figs. 6(c) and 6(d)] has converged to that of the (almost) complete sample [Figs. 6(e) and 6(f)]. Therefore, for all compositions of set C, the generation of 100 000 unique structures is attempted (noting that at some compositions there are not as many as 100 000 unique structures in the total configuration space). Total energies are then predicted for these structures with MTP and EMF models to sample the full composition space of set C.

Recovering the symmetry degeneracy of the training data. During the generation of the training data for set A- and B-type structures in Ref. [1], the software package CASM [47] was used to select only symmetrically unique structures in order to reduce the total number of necessary DFT calculations. However, this introduces a bias into the data. A similar approach is adopted by the SOD software package [48] to reduce the number of calculations to perform when modeling disordered solids, but in this case the degeneracy is retained to allow the computation of "entropy-reduced" energy. For the set A- and B-type training data, it was necessary to first recover the symmetry degeneracy to remove the bias when sampling the configuration space and we outline our procedure for this in the SM [14] (Sec. 5).

Scaling data by combinatorial space. To scale by the total number of possible combinatorial substitutions, the data are grouped by the total number of Co atoms in the supercell and then by the total number of Co on oh sites. A scaling factor is then determined by finding the factor necessary to scale up the total group size to be equal to the total combination space for that particular number of Co and Co-on-oh-sites,

$$c_{\text{tot}} = \frac{n_{\text{tot}}^{\text{td}}!}{n_{\text{Co}}^{\text{td}}!\left(n_{\text{tot}}^{\text{td}} - n_{\text{Co}}^{\text{td}}\right)!} \frac{n_{\text{tot}}^{\text{oh}}!}{n_{\text{Co}}^{\text{oh}}!\left(n_{\text{tot}}^{\text{oh}} - n_{\text{Co}}^{\text{oh}}\right)!}, \qquad (22)$$

where $n_{\text{tot}}^{\text{td}}$ is the total number of td sites in the 56 atom supercell, and $n_{\text{Co}}^{\text{td}}$ is the number of the td sites occupied by Co. Similarly, $n_{\text{tot}}^{\text{oh}}$ is the total number of oh sites and $n_{\text{Co}}^{\text{oh}}$ is the number of these sites occupied by Co.

For the randomly generated set C structures, this simply results in multiplying each structure by the scaling factor. However, for set A and B, which were set up to be symmetrically distinct structures, instead the symmetry degeneracies of all structures with the same Co count and Co-on-oh-sites count are summed to determine the total group size for each Co count and Co-on-oh-sites count. The symmetry degeneracies are then scaled so that the total symmetry degeneracy for each group equals the total combination space for the particular number of Co and Co-on-oh-sites. This preserves the weight of each structure in set A and B, where some structures would have had more equivalent structures if they had been generated randomly. Note that not all configurations generated by CASM were successfully relaxed, but by scaling

the structures that did relax by configuration space, essentially an average is being taken over the structures that did relax as an approximation for missing structures from the training set.

B. Prediction of free energies

The thermodynamic stability of alloys, or other mixed phases such as $(\text{Co}_x\text{Mn}_{1-x})_3\text{O}_4$, depends on minimizing the Gibbs free energy,

$$G(N, P, T; \vec{x}) = H - TS, \qquad (23)$$

where N is the number of atoms, P is the pressure, T is the temperature, H is the enthalpy, S is the entropy, and $\vec{x}$ is a vector representing the full set of molar fractions of the alloying species [5]. For solids at atmospheric pressure, G can be approximated to the Helmholtz free energy F [49,50]. As shown in Eq. (23), the free energy is reduced by S. At low temperatures, the product of TS is small, so the magnitude of the free energy is dominated by H. However, at higher temperatures, there is a stronger reduction of the free energy by the entropic term.

There are many different possible contributions to the entropic term, such as the different chemical substitutions within the alloy (S_{chem}), vibrational entropy (S_{vib}), and degrees of freedom due to electronic and magnetic excitations (S_{elec} and S_{mag}, respectively) [5], giving

$$S = S_{\text{chem}} + S_{\text{vib}} + S_{\text{elec}} + S_{\text{mag}} + \cdots . \qquad (24)$$

From a statistical-mechanics perspective, the reduction in the free energy by an increased number of states of the system (from various entropic contributions) can be understood as the phase enclosing more states in its phase space being more likely to be visited as the system undergoes microscopic transitions and hence has an increased stability relative to other phases [49]. The stabilization of multicomponent alloys due to the entropy of mixing from chemical substitution is a fundamental concept of high entropy alloys (HEAs) and there are a number of reviews on this particular subject such as Refs. [5,51,52]. $(\text{Co}_x\text{Mn}_{1-x})_3\text{O}_4$ does not technically meet the specifications to be considered a HEA, such as containing five or more elements in nearly equal atomic ratios [52]. However, this does not eliminate the possibility of stabilization of this system from configurational or other entropy sources.

All available training data was used to train a MTP and EMF model, where for the MTP model the training data was split into tetragonal and cubic structures. While our EMF model takes only number of oh and td Co in the supercell as inputs for describing the system, it was necessary to estimate lattice parameters for all of the randomly generated structures used to sample the configuration space of set C with the MTP model, as described earlier. The EMF model involved ten fitting parameters compared to 156 for the MTP model when training with the 8g basis set and three different atomic species. These trained models were used to predict the energies of the full composition space of set C using the randomly generated structures and scaling by combination space. As the EMF model is unable to distinguish structures with the same number of oh and td Co in the supercell, but with different cell shapes and energies, only the minimum energy structure in the training set for each number of oh and td Co is used to train
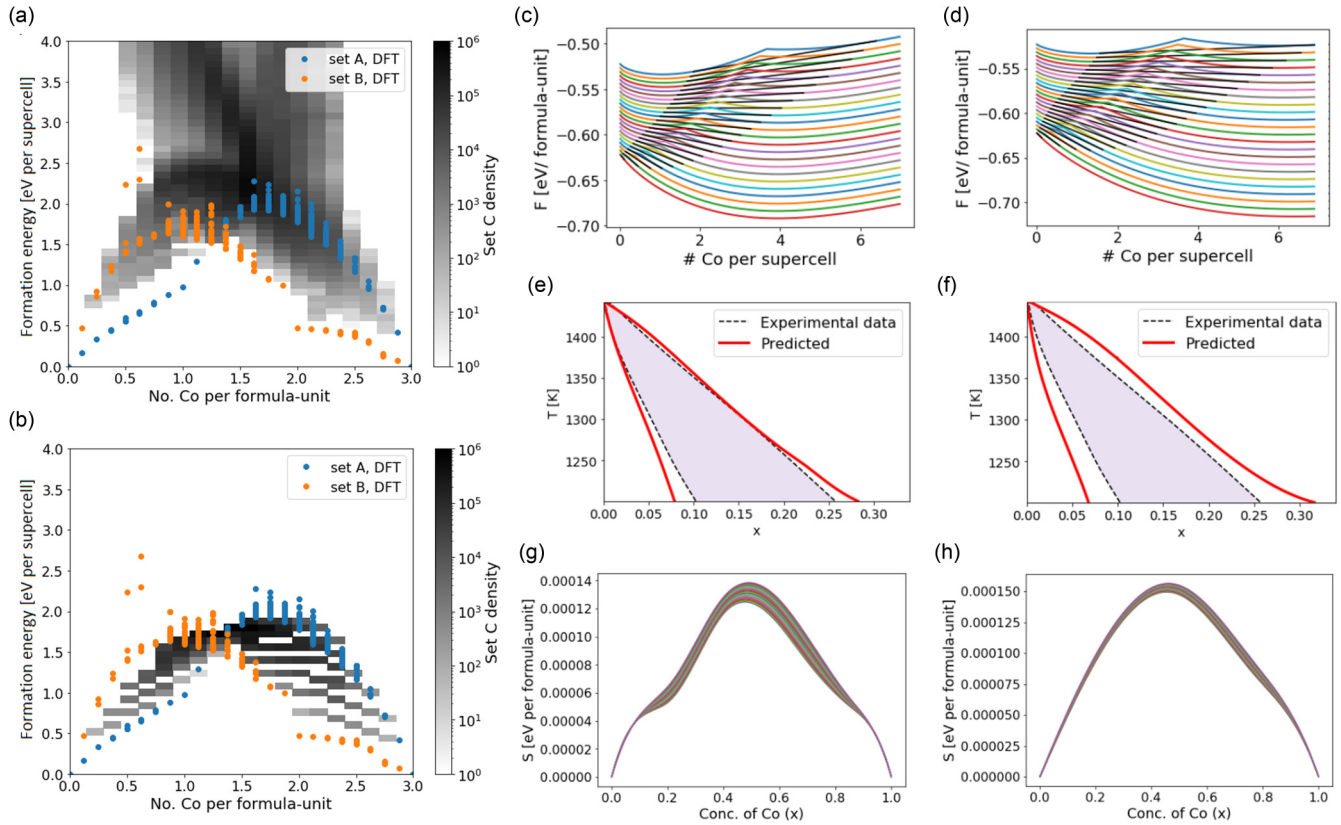
FIG. 7. Formation energy as a function of Co composition in supercells of (Co$_x$Mn$_{1-x}$)$_3$O$_4$ with set C predicted energies shown in black from (a) the MTP model and (b) the EMF model. Free energy curves calculated across a temperature range of 1200–1440 K including: magnetic, vibrational, and configurational entropy as obtained from (c) the MTP model sampling of set C and (d) that of the EMF model. Solubility gaps obtained from the tangents of the free energy curves [black lines in (c) and (d)] are shown in (e) and (f), respectively, where experimental data is taken from Ref. [13]. Configurational entropy extracted from each model is shown in (g) for MTP and (h) for EMF.

the model. When predicting the free energies with the EMF model, the scaling factor for set C structures is divided by two as with the EMF model there is not a cubic and tetragonal version of the same structure. With this model, there is only the minimum energy phase at each composition.

Total energies for C-type configurations of (Co$_x$Mn$_{1-x}$)$_3$O$_4$ predicted by the MTP and EMF models were used to produce plots of formation energy vs composition via Eq. (20), which are shown in Figs. 7(a) and 7(b) for MTP and EMF models, respectively. It can be seen that the EMF model produces much more discretised energy predictions due to its inability to distinguish structures with the same number of oh and td Co. These formation energies were used to calculate F as a function of temperature and Co concentration x, including also magnetic entropy, vibrational entropy, and a correction to the configurational entropy at the composition edges (due to finite-size limitations of the supercells). These methods are outlined in Ref. [10], where the importance of each of these contributions was demonstrated. Calculated F curves from MTP and EMF are shown in Figs. 7(c) and 7(d), respectively.

The calculated F from each model is then used to compute the SG (or phase coexistence region) between tetragonal (H) and cubic (S) phases of (Co$_x$Mn$_{1-x}$)$_3$O$_4$. This method is again described in Ref. [10]. Calculated SGs from MTP and EMF are shown in Figs. 7(e) and 7(f), respectively, and compared to

experimental data from Ref. [13]. Both methods gave similar SGs in this low-Co concentration range, but the EMF model predictions diverge more from the experimental SG, especially at very Co-poor compositions and higher temperatures.

Figures 7(g) and 7(h) show the extracted configurational entropy from MTP and EMF models, respectively. This is extracted from the predicted configurational free energy F_{config}, via

$$S_{\mathrm{config}} = -\left(\frac{\partial F_{\mathrm{config}}}{\partial T}\right)_V. \qquad (25)$$

The most noticeable difference between the configurational entropy predictions by the two models is the Co-poor side where the EMF model gives larger configurational entropies than with the MTP model at the same Co concentrations.

Based on the better agreement between the MTP prediction and experimental data for the SG in the Co-poor region of the phase diagram, and from comparing the set C energy predictions to the DFT data for C-type structures [shown in Fig. 4(e)], the MTP predictions appear to reproduce the known features of the system more accurately. One may also ask how classical force fields would perform versus ML predictions. However, it is hard to make a comparison between classical force fields and machine learned potentials on an equal footing because of the different methodologies and philosophies

of their development: the classical potentials are typically developed through many loops of trial-and-error, while the machine learned potentials are trained automatically on an *ab initio* database generated ad hoc for the present problem. Some explicit comparisons can be found in Refs. [53,54].

With the experimental data available for the H+S SG, we are only able to investigate the accuracy in the calculation of the phase diagram, and the different ML methods used to sample set C, for the Co-poor side of the phase diagram, up to approximately $x = 0.3$ in $(Co_xMn_{1-x})_3O_4$. Further experimental data for Co-rich phases would provide valuable information for assessing the accuracy of the ML predictions, similar to the cation site occupancy measurements performed in Ref. [55] for a Mn-rich composition. It was demonstrated in Ref. [10] that composition-dependent magnetic entropy and, in particular, the vibrational entropy from a higher energy cubic phase, are vital to accurately reproduce the SG in the Co-poor region. However, on the Co-rich side, where there is no such higher energy phase, vibrational entropy may be less dominant in determining F and hence other factors such as the accuracy in the sampling of C-type configurations which appear to be closer in energy to the ground state on the Co-rich side in Fig. 7(a) may play a stronger role in the calculation of F. In the SM [14] (Sec. 6) we compare the dominant set structure type as a function of temperature and Co composition as predicted by the EMF and MTP models.

V. SUMMARY AND FURTHER WORK

We have explored the use of three different ML methods to sample the full configuration space of $(Co_xMn_{1-x})_3O_4$: ANNs and MTPs, originally developed to implement ML potentials; and a much simpler EMF method, requiring as inputs only the number of substituting species occupying particular types of crystallographic site [i.e., octahedral or tetrahedral in the case of $(Co_xMn_{1-x})_3O_4$].

Our ANN model was capable of achieving some of the lowest validation RMSEs in many test cases. However, in an imbalanced data set, as in the case of $(Co_xMn_{1-x})_3O_4$ with fewer possible configurations at the composition extremes without any data pretreatment, the ANN performed more poorly than the other methods on minority type data. For the MTP model, due to the lack of force information in our initial data set, it was necessary to obtain estimates for the relaxed lattice parameters of the set C structures based on the DFT training data. In the Co-poor composition range of the phase diagram we have compared available experimental

data [13] to model predictions. In this composition range both EMF and MTP models provide reasonably accurate predictions, with the MTP predictions being superior to those of the EMF model at the lowest Co concentrations and highest temperatures.

Studying the Co-rich side of the phase diagram of $(Co_xMn_{1-x})_3O_4$ and the spinel–rock-salt phase coexistence region [13] could provide valuable insights for energy storage via redox reactions [1–3]. To study this part of the phase diagram, further experimental data to check the accuracy of ML energy predictions would be very valuable. For example, the predicted relative occupation of td and oh sites by Co atoms at different temperatures and concentrations becomes nontrivial beyond eight atoms per supercell: as we show in the Supplemental Material [14], MTP and EMF models predict slightly different occupations, which could be experimentally verified to assess the validity of the models [14]. Furthermore, it would be necessary to have additional *ab initio* data for rock-salt structures. It may also be important to consider the effect of a possible partial low-to-high spin-state transition of oh Co in Co_3O_4 [56–59], which we have neglected in this work as it focused on the Co-poor side of the phase diagram, but could impact the magnetic entropy for Co-rich structures.

Software developed in this work for calculating the solubility gap of $(Co_xMn_{1-x})_3O_4$ with the EMF model and with data from the MTP model in this work is available from [60] under a BSD 3-Clause license. Data from *ab initio* calculations used to train machine-learning models in this study is available from [61]. The trained MTP potential files are also provided in the SM [14].

ACKNOWLEDGMENTS

[1] A. Zaki, J. Carrasco, D. Bielsa, and A. Faik, Tunable redox temperature of a $Co_{3-x}Mn_xO_4$ $(0 \leqslant x \leqslant 3)$ continuous solid solution for thermochemical energy storage, ACS Appl. Mater. Interfaces **12**, 7010 (2020).

[2] L. André, S. Abanades, and L. Cassayre, Mixed metal oxide systems applied to thermochemical storage of solar energy: Benefits of secondary metal addition in Co and Mn oxides and contribution of thermodynamics, Appl. Sci. **8**, 2618 (2018).

[3] L. André, S. Abanades, and L. Cassayre, Experimental investigation of Co–Cu, Mn–Co, and Mn–Cu redox materials applied to solar thermochemical energy storage, ACS Appl. Energy Mater. **1**, 3385 (2018).

[4] B. Ehrhart, E. Coker, N. Siegel, and A. Weimer, Thermochemical cycle of a mixed metal oxide for augmentation of thermal energy storage in solid particles, Energy Proc. **49**, 762 (2014).

[5] M. Widom, Modeling the structure and thermodynamics of high-entropy alloys, J. Mater. Res. **33**, 2881 (2018).

[6] C. M. Rost, E. Sachet, T. Borman, A. Moballegh, E. C. Dickey, D. Hou, J. L. Jones, S. Curtarolo, and J.-P. Maria, Entropy-stabilized oxides, Nat. Commun. **6**, 8485 (2015).

[7] Y. Liu, T. Zhao, W. Ju, and S. Shi, Materials discovery and design using machine learning, J. Materiomics **3**, 159 (2017).

[8] G. H. Gu, J. Noh, I. Kim, and Y. Jung, Machine learning for renewable energy materials, J. Mater. Chem. A **7**, 17096 (2019).

[9] H. Wang, Y. Ji, and Y. Li, Simulation and design of energy materials accelerated by machine learning, WIREs Comput. Mol. Sci. **10**, e1421 (2020).

[10] S. K. Wallace, A. van Roekeghem, A. S. Bochkarev, J. Carrasco, A. Shapeev, and N. Mingo, Modeling the high-temperature phase coexistence region of mixed transition metal oxides from *ab initio* calculations, Phys. Rev. Research **3**, 013139 (2021).

[11] K. Momma and F. Izumi, VESTA: A three-dimensional visualization system for electronic and structural analysis, J. Appl. Crystallogr. **41**, 653 (2008).

[12] M. C. Kemei, J. K. Harada, R. Seshadri, and M. R. Suchomel, Structural change and phase coexistence upon magnetic ordering in the magnetodielectric spinel Mn_3O_4, Phys. Rev. B **90**, 064418 (2014).

[13] Y. Golikov, S. Tubin, V. Barkhatov, and V. Balakirev, Phase diagrams of the CoMnO system in air, J. Phys. Chem. Solids **46**, 539 (1985).

[14] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevMaterials.5.035402 for details on the calculation and additional graphs.

[15] D. P. Tabor, L. M. Roch, S. K. Saikin, C. Kreisbeck, D. Sheberla, J. H. Montoya, S. Dwaraknath, M. Aykol, C. Ortiz, H. Tribukait, C. Amador-Bedolla, C. J. Brabec, B. Maruyama, K. A. Persson, and A. Aspuru-Guzik, Accelerating the discovery of materials for clean energy in the era of smart automation, Nat. Rev. Mater. **3**, 5 (2018).

[16] V. L. Deringer, M. A. Caro, and G. Csányi, Machine learning interatomic potentials as emerging tools for materials science, Adv. Mater. **31**, 1902765 (2019).

[17] T. Morawietz and N. Artrith, Machine learning-accelerated quantum mechanics-based atomistic simulations for industrial applications, J. Comput.-Aided Mol. Des. (2020), doi: 10.1007/s10822-020-00346-6.

[18] S. Raschka and V. Mirjalili, *Python Machine Learning* (Packt, Birmingham, UK, 2017).

[19] J. Behler, Atom-centered symmetry functions for constructing high-dimensional neural network potentials, J. Chem. Phys. **134**, 074106 (2011).

[20] A. V. Shapeev, Moment tensor potentials: A class of systematically improvable interatomic potentials, Multiscale Model. Simul. **14**, 1153 (2016).

[21] A. P. Bartók and G. Csányi, Gaussian approximation potentials: A brief tutorial introduction, Int. J. Quantum Chem. **115**, 1051 (2015).

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

[23] F. Legrain, J. Carrete, A. van Roekeghem, S. Curtarolo, and N. Mingo, How chemical composition alone can predict vibrational free energies and entropies of solids, Chem. Mater. **29**, 6220 (2017).

[24] C. Nyshadham, M. Rupp, B. Bekker, A. V. Shapeev, T. Mueller, C. W. Rosenbrock, G. Csányi, D. W. Wingate, and G. L. W. Hart, Machine-learned multi-system surrogate models for materials prediction, npj Comput. Mater. **5**, 51 (2019).

[25] C. W. Rosenbrock, K. Gubaev, A. V. Shapeev, L. B. Pártay, N. Bernstein, G. Csányi, and G. L. W. Hart, Machine-learned interatomic potentials for alloys and alloy phase diagrams, npj Computational Materials **7**, 24 (2021).

[26] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, Machine learning for molecular and materials science, Nature (London) **559**, 547 (2018).

[27] A. S. Bochkarev, A. van Roekeghem, S. Mossa, and N. Mingo, Anharmonic thermodynamics of vacancies using a neural network potential, Phys. Rev. Mater. **3**, 093803 (2019).

[28] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, Phys. Rev. Lett. **98**, 146401 (2007).

[29] J. Behler, Constructing high-dimensional neural network potentials: A tutorial review, Int. J. Quantum Chem. **115**, 1032 (2015).

[30] J. Carrete, M. López-Suárez, M. Raya-Moreno, A. S. Bochkarev, M. Royo, G. K. H. Madsen, X. Cartoixà, N. Mingo, and R. Rurali, Phonon transport across crystal-phase interfaces and twin boundaries in semiconducting nanowires, Nanoscale **11**, 16007 (2019).

[31] A. Shapeev, Accurate representation of formation energies of crystalline alloys with many components, Comput. Mater. Sci. **139**, 26 (2017).

[32] I. S. Novikov, K. Gubaev, E. V. Podryabinkin, and A. V. Shapeev, The MLIP package: Moment tensor potentials with MPI and active learning, arXiv:2007.08555.

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

[34] A. Miedema, P. de Châtel, and F. de Boer, Cohesion in alloys—fundamentals of a semi-empirical model, Physica B+C **100**, 1 (1980).

[35] N. Provatas and K. Elder, *Phase-Field Methods in Materials Science and Engineering* (John Wiley & Sons, New York, 2010).

[36] A. Takeuchi, K. Amiya, T. Wada, K. Yubuta, W. Zhang, and A. Makino, Entropies in alloy design for high-entropy and bulk glassy alloys, Entropy **15**, 3810 (2013).

[37] M. C. Gao, C. S. Carney, O. N. Doğan, P. D. Jablonksi, J. A. Hawk, and D. E. Alman, Design of refractory high-entropy alloys, JOM **67**, 2653 (2015).

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

[39] 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 (1996).

[40] J. Sun, A. Ruzsinszky, and J. P. Perdew, Strongly Constrained and Appropriately Normed Semilocal Density Functional, Phys. Rev. Lett. **115**, 036402 (2015).

[41] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, Machine Learning a General-Purpose Interatomic Potential for Silicon, Phys. Rev. X **8**, 041048 (2018).

[42] J. Behler, Perspective: Machine learning potentials for atomistic simulations, J. Chem. Phys. **145**, 170901 (2016).

[43] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, SMOTE: Synthetic minority over-sampling technique, J. Artif. Intell. Res. **16**, 321 (2002).

[44] J. M. Johnson and T. M. Khoshgoftaar, Survey on deep learning with class imbalance, J. Big Data **6**, 27 (2019).

[45] A. H. Nguyen, C. W. Rosenbrock, C. S. Reese, and G. L. W. Hart, Robustness of the cluster expansion: Assessing the roles of relaxation and numerical error, Phys. Rev. B **96**, 014107 (2017).

[46] A. R. Denton and N. W. Ashcroft, Vegard's law, Phys. Rev. A **43**, 3161 (1991).

[47] CASM Developers, Casmcode: V0.2.1 (2017).

[48] R. Grau-Crespo, S. Hamad, C. R. A. Catlow, and N. H. de Leeuw, Symmetry-adapted configurational modelling of fractional site occupancy in solids, J. Phys.: Condens. Matter **19**, 256201 (2007).

[49] A. van de Walle and G. Ceder, The effect of lattice vibrations on substitutional alloy thermodynamics, Rev. Mod. Phys. **74**, 11 (2002).

[50] R. Feng, P. K. Liaw, M. C. Gao, and M. Widom, First-principles prediction of high-entropy-alloy stability, npj Comput. Mater. **3**, 50 (2017).

[51] E. P. George, D. Raabe, and R. O. Ritchie, High-entropy alloys, Nat. Rev. Mater. **4**, 515 (2019).

[52] D. Miracle and O. Senkov, A critical review of high entropy alloys and related concepts, Acta Mater. **122**, 448 (2017).

[53] J. R. Boes, M. C. Groenenboom, J. A. Keith, and J. R. Kitchin, Neural network and ReaxFF comparison for Au properties, Int. J. Quantum Chem. **116**, 979 (2016).

[54] B. Grabowski, Y. Ikeda, P. Srinivasan, F. Körmann, C. Freysoldt, A. I. Duff, A. Shapeev, and J. Neugebauer, *Ab initio* vibrational free energies including anharmonicity for multicomponent alloys, npj Comput. Mater. **5**, 80 (2019).

[55] J. Popović, M. Jurić, D. Pajić, M. Vrankić, J. Zavašnik, and J. Habjanič, Effect of the cation distribution and microstructure on the magnetic behavior of the $CoMn_2O_4$ oxide, Inorg. Chem. **56**, 3983 (2017).

[56] V. Brabers and A. Broemme, Low-spin-high-spin transition in the Co_3O_4 spinel, J. Magn. Magn. Mater. **104-107**, 405 (1992).

[57] H. C. O'Neill, Thermodynamics of Co_3O_4: A possible electron spin unpairing transition in Co^{3+}, Phys. Chem. Miner. **12**, 149 (1985).

[58] K. Mocala, A. Navrotsky, and D. Sherman, High-temperature heat capacity of Co_3O_4 spinel: Thermally induced spin unpairing transition, Phys. Chem. Miner. **19**, 88 (1992).

[59] M. Chen, B. Hallstedt, and L. J. Gauckler, Thermodynamic assessment of the Co-O system, J. Phase Equilib. **24**, 212 (2003).

[60] https://doi.org/10.5281/zenodo.4133866.

[61] https://dx.doi.org/10.17172/NOMAD/2020.10.14-1.