Collinear-spin machine learned interatomic potential for Fe7Cr2Ni alloy

Lakshmi Shenoy, Christopher D. Woodgate, Julie B. Staunton, Albert P. Bartók, Charlotte Becquart, Christophe Domain, James R. Kermode

Collinear-spin machine learned interatomic potential for Fe$_7$Cr$_2$Ni alloy

Lakshmi Shenoy [1,*] Christopher D. Woodgate [2] Julie B. Staunton [2] Albert P. Bartók [1,2] Charlotte S. Becquart [3]
Christophe Domain [4] and James R. Kermode [1]

[1]*Warwick Centre for Predictive Modelling, School of Engineering, University of Warwick, Coventry CV4 7AL, United Kingdom*
[2]*Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom*
[3]*Univ. Lille, CNRS, INRAE, Centrale Lille, UMR 8207 - UMET - Unité Matériaux et Transformations, F-59000 Lille, France*
[4]*Electricite de France, EDF Recherche et Developpement, Departement Materiaux et Mecanique des Composants,
Les Renardieres, F-77250 Moret sur Loing, France*

We have developed a machine learned interatomic potential for the prototypical austenitic steel Fe$_7$Cr$_2$Ni, using the Gaussian approximation potential (GAP) framework. This GAP can model the alloy's properties with close to density functional theory (DFT) accuracy, while at the same time allowing us to access larger length and time scales than expensive first-principles methods. We also extended the GAP input descriptors to approximate the effects of collinear spins (spin GAP), and demonstrate how this extended model successfully predicts structural distortions due to antiferromagnetic and paramagnetic spin states. We demonstrate the application of the spin GAP model for bulk properties and vacancies and validate against DFT. These results are a step towards modeling the atomistic origins of ageing in austenitic steels with higher accuracy.

I. INTRODUCTION

Austenitic stainless steels are key structural materials with high mechanical strength and corrosion resistance. They have many applications, from everyday tools such as household appliances and cars, to structural components of buildings, bridges, and industrial machines [1]. Their desirable properties come from alloying Fe with a concentration of 16–20% Cr for rust resistance, and about 10% Ni to maintain a crack-resistant face-centred-cubic crystal structure at all temperatures where it remains solid (< 1800K) [2]. Specific properties of the steel can be fine tuned by adding small concentrations ($< 1\%$) of other solutes such as Mo, Mn, P, C etc., depending on the application of interest [3]. However, the core properties come from the Fe-Cr-Ni mix, so in this paper we focus on the prototypical austenitic steel with composition Fe$_7$Cr$_2$Ni.

A key application of austenitic steels is in nuclear power plants. The austenitic steel grades 304 and 316, which have a composition very close to the model alloy Fe$_7$Cr$_2$Ni, are used in the primary containment barriers of light-water nuclear reactor pressure vessels (RPV) [4]. The RPV interiors are exposed to constant neutron bombardment and a wide range of thermal fluctuations from cryogenic temperatures during maintenance to operating temperature of $\sim$600 K. Over

time, these harsh conditions trigger the formation of radiation-induced point defects in the RPV interiors, which eventually cluster to form larger defects such as voids and dislocation loops. Macroscopically, these manifest as creep, swelling, stress corrosion, and other ageing effects, which are safety concerns [5]. To improve RPV designs and facilitate mitigation efforts, there is interest in understanding the mechanisms behind these ageing phenomena.

Experimental data for aging phenomena are limited due to the challenges in conducting controlled experiments in extreme conditions, and also challenges in measuring atomic quantities experimentally [6–8]. We therefore require computational methods to simulate these phenomena across length scales, to fill in the gaps between the available experimental data, and to explain the observed aging effects.

There are various tiers of first-principles methods available for materials modeling, with higher accuracy being accompanied by (prohibitively) higher computational costs. While more complex phenomena such as nuclear cascades or spectroscopic excitations require higher levels of theory, the first principles method of choice for modeling mechanical properties as in our case is density functional theory (DFT), in which the electronic structure of materials is calculated from approximations of quantum mechanics [9]. There have been efforts to study the Fe$_7$Cr$_2$Ni alloy using DFT [5,10–12]. However, as the computational cost of DFT calculations scales cubically with system size, these studies were restricted to supercells containing up to 256 atoms. These supercells are too small to model extended defects such as voids or dislocations. Also, as DFT calculations are computationally expensive, these studies were restricted to limited statistics, which is not ideal for modeling the variations due to chemical complexity of concentrated alloys.

---

*lakshmi.shenoy@warwick.ac.uk

Computationally more affordable alternatives for atomistic modeling of materials that allow us to simulate larger length and time scales are classical interatomic potentials. These models have a fixed mathematical form and are fit to experimental and DFT data for the material properties of interest. For metallic systems, the main classical models used are the embedded atom models (EAM) [13], the modified EAM [14], and the Finnis Sinclair model [15]. While there have been many efforts to study alloys using classical interatomic potentials [16–20], these studies are restricted to just a few properties owing to the fixed mathematical form of empirical models. For instance, Bonny *et al.* developed a FeCrNi EAM for plasticity [18] and another FeCrNi EAM for point defects [19] (referred to as EAM-13 throughout this paper), but neither can simulate both types of defects as they were trained for their respective target properties. We would ideally like to have a model that can simulate multiple defects and their interactions.

A successful emerging approach that can address the shortcoming of both DFT and EAM approaches is machine learned interatomic potentials (MLIPs) [21]. MLIPs are trained on a database of first-principles data, generally from DFT. Their highly flexible mathematical forms allow them to interpolate the training data smoothly and so fit the potential energy surface with near DFT accuracy, while at the same time being much cheaper to use for predictions than first-principles methods directly. A major advantage of the MLIPs over classical potentials like EAM is that they can be trained on multiple defects, by including representative configurations in the training database. Moreover, the accuracy and scope of an MLIP can be improved iteratively by appending new training configurations to its training database, chosen either manually using knowledge about the system of interest or using active learning schemes [22].

A range of MLIP approaches have enjoyed considerable success, including kernel methods [23] linear and nonlinear expansions of the potential in a polynomial basis such as the atomic cluster expansion (ACE) [24] or moment tensor potential (MTP) [25] frameworks, and most recently message-passing neural networks [26–28]. In this study, we use the Gaussian approximation potential (GAP) framework, a kernel based method [23]. The GAP model has physically motivated kernel choices and hyperparameters, making it intuitive to understand. It uses Gaussian process regression to fit the potential energy surface; the Bayesian interpretation of this procedure allows us to derive error estimates for GAP predictions, which is a useful feature of this choice of MLIP. The GAP model has been fitted and validated for many elemental materials such as Si [29], W [30], Fe [31], and C [32], demonstrating the effectiveness of this model. More recently, it has also been used for binary alloys [33] and the ternary alloy Ge-Sb-Te [34]. In this paper, we fit and validate a GAP model for Fe_7Cr_2Ni. We also extend the GAP descriptors to incorporate the effects of the spin-polarized electronic structure of the alloy in terms of collinear spins associated with the Fe atoms, and demonstrate how this spin GAP model improves predictions compared to the standard GAP model.

This paper is organized as follows: Section II outlines how the training database was assembled, and includes details of

TABLE I. Number of training and testing configurations for each subgroup DB*x*, comprising either 256-atom bulk or 255-atom vacancy configurations.

	Config. type	Temperature	Train	Test
DB1	relaxed bulk	0 K	111	12
DB2	sheared bulk	$\sim$0 K	190	75
DB3	relaxed vacancy	0 K	36	8
DB4	vacancy optimisation	$\sim$0 K	72	20
DB5	bulk MD	1000/1500/2000 K	60	13
DB6	vacancy MD	1000/1500/2000 K	71	20
DB7	bulk MD quench	1000 K $\rightarrow$ 0 K	38	4
DB8	vacancy MD quench	1000 K $\rightarrow$ 0 K	44	28
	element substitutions	0K	0	42

the standard GAP and spin GAP fitting procedures. Section III discusses the main results for bulk and point defect properties of Fe_7Cr_2Ni using the spin GAP. The results are validated against DFT and experiments, and compared with EAM and standard GAP predictions.

II. METHODOLOGY

A. Training Database

The first step to training a MLIP is to assemble a training database of high accuracy data relevant to properties of interest. Our training database comprises configurations relevant to the bulk, finite temperature, and point defect properties of the alloy. It can be categorized into eight subdatabases (DB*x*) based on configuration type, as summarized in Table I. Figure 1 shows the kernel principal component analysis (KPCA) of the training database configurations using the SOAP kernel [35], colored by subdatabases. The horizontal axis KPCA component can be interpreted as being roughly correlated with the temperature of the configurations, while the vertical axis KPCA component is correlated with strain.

Perfect-lattice bulk and monovacancy supercells (containing 256 and 255 atoms respectively) with different element
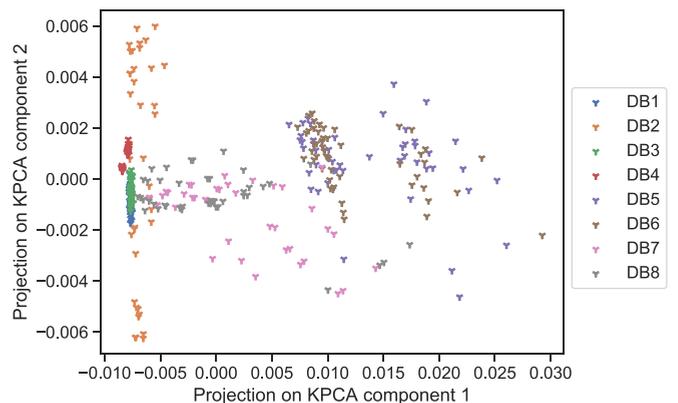
FIG. 1. Kernel principal component analysis (KPCA) of training data in the three-species SOAP descriptor space. This approach allows us to identify gaps in the training configuration space; e.g., DB7-8 were included to sample the gap between the DB1-4 and DB5-6 clusters.
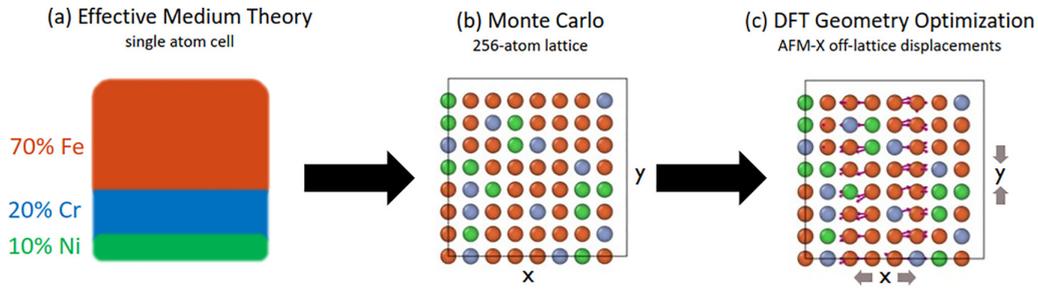
FIG. 2. Protocol for generating bulk and vacancy training structures that sample different elemental configurations, with realistic short-range ordering derived from effective medium (EM) theory. In this case, the DFT relaxed off-lattice displacements shown by the red arrows of step 3 are for AFM layering along the $\hat{x}$ axis (AFM-X).

distributions that capture relevant short-range ordering were generated using an effective medium theory (EM) as described in Sec. II B. These were fully relaxed as shown in the schematic of Fig. 2, using DFT calculations as described in Sec. II C, to generate the training subdatabases DB1 and DB3 respectively. Random strains were applied to a subset of DB1 samples, to generate DB2, aimed at providing reference material, which can inform an accurate description of the alloy's elastic behavior. Snapshots from the DFT geometry optimisation trajectory of the EM monovacancy structures were compiled to form DB4, to provide details of the potential energy surface close to 0 K.

The EM structures were generated to capture short-range ordering at temperatures of 600 K, 1000 K, 1500 K, and 2000 K, respectively. We ran *ab initio* molecular dynamics (MD) for the 1000–2000 K configurations from DB1 and DB3 at the temperatures they were generated for, and uncorrelated snapshots from these MD trajectories were compiled to form DB5 and DB6, respectively. In Fig. 1, we can see that the high-temperature DB5-6 samples sit quite far apart from the low-temperature DB1-4 samples along the first KPCA component. To aid the GAP model with better interpolation of training data, we need to populate such gaps between training data clusters along the principal axes of the kernel space. To this end, we extend the database by including snapshots from the geometry optimisation of a few bulk and monovacancy MD structures, comprising DB7 and DB8, respectively. These configurations are seen to sample the space between low- and high-temperature clusters in the KPCA plot Fig. 1, and their inclusion decreased the errors of our models. The GAP potentials (both with and without the Fe spin extensions) were trained on this combined database DB1-8 as described in Sec. II D.

B. Effective Medium Theory

Past DFT studies on FeCrNi have modelled the alloy using special quasirandom structures (SQS) [5,10,11]. A limitation of using SQS for magnetic materials is that in aiming to produce local environments reflective of the random alloy, they neglect atomic short-range order. It would be more effective to train a model using configurations in which temperature-dependent short-range order is present, especially in the case of point defects where the energetics are correlated to the concentration of species decorating the defect [5].

To obtain atomic configurations with physically motivated atomic short-range order, we draw samples from equilibrated, lattice-based Monte Carlo (MC) simulations, as depicted by the first step in Fig. 2. A configuration is described by a set of lattice site occupancies $\{\xi_{i\alpha}\}$, where $\xi_{i\alpha} = 1$ if site i is occupied by an atom of chemical species α, and 0 otherwise. The internal energy is described by a simple pairwise Bragg-Williams Hamiltonian [36,37],

$$H = \sum_{i\alpha;j\alpha'} V_{i\alpha;j\alpha'} \xi_{i\alpha} \xi_{j\alpha'}, \tag{1}$$

where the atom-atom interchange parameter $V_{i\alpha;j\alpha'}$ describes the energy associated with an atom of species i on site α interacting with an atom of species j on site α'. We assume these interchange parameters are isotropic, homogeneous, and have finite range, which simplifies Eq. (1).

The $V_{i\alpha;j\alpha'}$ are obtained using the $S^{(2)}$ theory for multicomponent alloys [38], the details and implementation of which have been discussed extensively in earlier papers [39–42]. The theory uses the Korringa-Kohn-Rostoker (KKR) formulation of DFT, with disorder described via the coherent potential approximation (CPA) [43], producing an effective medium representing the electronic structure of the disordered alloy. Our method captures the effects of an ordered magnetic state, in our case the single-layer antiferromagnetic (AFM) state, on the short-range-order state [41]. It is also possible to treat vacancies in this formalism at low concentrations by including a chemical species with no associated electrons, and no nuclear charge.

In this paper, we used the HUTSEPOT code [44] to construct the self-consistent potentials of DFT within the KKR-CPA formalism. The interchange parameters are fitted to the first four coordination shells, and they are tabulated in Appendix A for reference. The MC simulations used a cells of 256 atoms with periodic boundary conditions applied in all three directions. Simulations were equilibrated at the desired temperature and then decorrelated samples were drawn 25 600 MC steps apart, i.e., 100 MC steps per lattice site. The samples drawn from high-temperature simulations will have little to no short-range order, while those drawn at lower temperatures will have a degree of short-range order present. These lattice-based EM configurations are then relaxed with DFT using setting described in the next section, to get the cell distortion and off-lattice atomic displacements as depicted in the last step of Fig. 2.

C. DFT Settings

The Vienna *Ab initio* Software Package (VASP) [45–47] was used to compute ground-state DFT structure, energy, atomic forces, and cell stresses for the training database. MLIPs can only be as accurate as the data they are trained on, so the choice of DFT settings cap the achievable accuracy. A key choice when setting up DFT calculations is which exchange-correlation (XC) functional to use. For structural properties of transition metals, the Perdew–Burke–Ernzerhof [48] flavor of generalized gradient approximation (GGA) functionals is well established in the literature as a good choice, and so this is the XC functional we use. However, for more complex phenomena involving electronic excitations, we would require beyond-GGA approaches such as hybrid exchange, or even beyond-DFT methods such as the GW approach, and so we cannot hope to model such phenomena using a MLIP trained on GGA-based DFT.

Depending on the range of atomic interactions that we hope to model, the choice of pseudopotential would also affect the DFT accuracy. For example, modeling nuclear cascades would require very short-range interactions that benefit from all-electron DFT codes or semicore pseudopotentials [49]. However, since we are targeting mechanical properties, the key players are the outer shell electrons of the alloying elements, so the accuracy we require is not compromised by using softer pseudopotentials for the core electrons. We use the standard VASP pseudopotentials based on the projector augmented wave (PAW) method [50], with eight, six, and ten valence electrons for Fe, Cr, and Ni respectively.

As Fe and Ni are magnetic materials, and Cr has shown a spin-density wave at low temperatures [51,52], it is essential to turn on the spin setting for more accurate DFT calculations of this alloy. Piochaud *et al.* reported that their tests with noncollinear spins generally relaxed to collinear arrangements [5], so we constrain our systems to collinear spin. Also, while austenitic steel is paramagnetic (PM) at operation temperatures of nuclear reactors, single-layer antiferromagnetic (AFM) ordering was found to be the most stable ground state [5]. To ensure we get ground-state properties right, we impose an AFM ordering for DFT relaxations of our training samples. Initial magnetic moment magnitudes are overestimated to allow for relaxation, and are set at 3.0, 2.0, and 1.0 μ_B for Fe, Cr, and Ni respectively. These atomic magnetic moments evolve during DFT electronic minimisation, and are computed in VASP by integrating the spin density within a sphere centered around the atom of interest.

A Monkhorst-Pack **k**-point mesh [53] of size $3 \times 3 \times 3$ was used to sample the Brillouin zone for all the training samples. Training data was computed with a plane-wave energy cutoff of 600 eV and an electronic self-consistency criterion of 10^{-7} eV. Preliminary steps, i.e., initial geometry optimisations and *ab initio* MD trajectories, were done using a lower self-consistency criterion of 10^{-4} eV. Electronic minimisation were performed using the preconditioned RMM-DIIS algorithm, while ionic position optimisations were done using the conjugate gradient method. The *ab initio* MD trajectories were run for the NVT ensemble, using the Nose-Hoover thermostat [54,55] and timesteps of either 5 fs (for T = 1000 K, 1500 K) or 2 fs (for T = 2000 K). The initial $\sim$200 steps (exact number

depending on when the total energy stabilizes) of the trajectories were discarded, and uncorrelated samples were taken from the remaining 1000–1500 steps of each trajectory. These were then evaluated with the higher precision DFT settings before being included in the training database.

D. Potential Fitting

This section outlines how the GAP fits the Born-Oppenheimer potential energy surface (PES) by mapping the DFT training structures to their DFT energies, forces, and stresses. As with other interatomic potentials, the GAP assumes that the energy E_{N_c} of a configuration with N_c atoms can be written as

$$E_{N_c} = \sum_{i=1}^{N_c} \epsilon(\hat{\mathbf{q}}_i) \tag{2}$$

where $\epsilon(\hat{\mathbf{q}}_i)$ is the energy contribution from atom i in the configuration, and $\hat{\mathbf{q}}_i$ is a descriptor vector that captures the unique features of the atom i environment. The atomic energies are modelled as a linear combination of kernels [56] given by

$$\epsilon(\hat{\mathbf{q}}_i) = \sum_{j}^{M} \alpha_j K(\hat{\mathbf{q}}_j, \hat{\mathbf{q}}_i) = \mathbf{k}_i^T \boldsymbol{\alpha} \tag{3}$$

where the kernel basis functions $K(\hat{\mathbf{q}}_j, \hat{\mathbf{q}}_i)$ capture the similarity between environment i and the M environments of the basis set. As GAP training databases are generally large (e.g., $N \sim 160\,000$ atomic environments in this paper), the usual strategy is to use only a small portion $M \ll N$ of the training set for the basis set. This makes the fit computationally tractable, while still retaining accuracy as the training database contains many groups of similar configurations that can be sparsified without much loss of information. These M sparse points are selected by CUR decomposition, which searches for maximally dissimilar environments [29], ensuring that the chosen basis is representative of the variety across the full training database.

Moving to the left-hand side of Eq. (3), DFT does not give direct access to the atomic energies ϵ, but rather, it gives us total energies E_{N_c} of the training configurations, along with atomic forces and virial stresses. As forces and stress are properties that can be derived from atomic energies (e.g., $\partial/\partial x$ for an atomic force along the Cartesian x direction), we can back-calculate the unknown N component vector $\boldsymbol{\epsilon}$ of atomic energies, from the known D component vector $\mathbf{y}$ of DFT data, by the relation $\mathbf{y} = \mathbf{L}^T \boldsymbol{\epsilon}$, where $\mathbf{L}$ is a linear differential operator $\mathbf{L}$ [30]. This allows us to derive the covariance kernel of the unknown target atomic energies ($\mathbf{K}_{NN}$) from the covariance kernel of the known DFT training data ($\mathbf{K}_{DD}$) using the relation $\mathbf{K}_{DD} = \mathbf{L}^T \mathbf{K}_{NN} \mathbf{L}$. The sparse ridge-regression solution to Eq. (3) regularized by a diagonal tolerance matrix $\Lambda = \sigma_v^2 \mathbf{I}$ is

$$\boldsymbol{\alpha} = [\mathbf{K}_{MM} + \mathbf{K}_{MN} \mathbf{L} \Lambda^{-1} \mathbf{L} \mathbf{K}_{NM}]^{-1} \mathbf{K}_{MN} \mathbf{L} \Lambda^{-1} \mathbf{y}. \tag{4}$$

With these optimized coefficients $\boldsymbol{\alpha}$, we can make predictions for new atomic environments $\hat{\mathbf{q}}_*$ using Eq. (3). In this paper, we use the smooth overlap of atomic positions (SOAP)

TABLE II. Parameters for the collinear-spin GAP.

Parameter	Two-body	Many-body
Kernel choice	Gaussian	SOAP
Radial cutoff	6.0 Å	6.0 Å
Cutoff transition width	1.0 Å	1.0 Å
δ	1.0 eV	0.1 eV
Sparse points	100	800
θ	1.0	
n_{max}, l_{max}		10,10
ζ		4
σ_{atom}		1.0 Å

descriptor

$$\mathbf{q}_i = \sum_{m=-l}^{l} (c^i_{nlm})^* c^i_{nlm} \quad ; \quad \hat{\mathbf{q}}_i = \frac{\mathbf{q}_i}{|\mathbf{q}_i|} \tag{5}$$

where c_{nlm} are expansion coefficients for the atomic neighbor density power spectrum [35]. For computational tractability, this sum is truncated at radial-basis functions up to order $n < n_{max}$ and spherical harmonics up to degree $l < l_{max}$. The SOAP descriptor is suitable for describing atomic environments since it is translationally, rotationally, and permutationally invariant [35]. The corresponding covariance kernel is

$$K(\hat{\mathbf{q}}_i, \hat{\mathbf{q}}_j) = \delta^2 |\hat{\mathbf{q}}_i \cdot \hat{\mathbf{q}}_j|^\zeta \tag{6}$$

where δ is a factor that corresponds to the spread in energy of the training dataset, and ζ is tuned to amplify variations in the dot product, to pick out fine features.

While the SOAP kernel is good at capturing smoothly varying details of the PES, it is not as effective in capturing sharp features as well, such as exchange repulsion when atoms get close to each other, or phase transitions [30]. It has been found to be useful to use multiple kernels targeting different scales, so that each kernel does not have to simultaneously fit vastly different features [29]. We hence use simple two-body kernels as well, to roughly capture the main trends of the PES, so that the SOAP kernels can focus on finer details. The two-body descriptor is just the pairwise separation between atoms, and we use the squared-exponential form

$$K(\hat{q}_i, \hat{q}_j) = \delta^2 \exp\left(-\frac{1}{2}\left(\frac{\hat{q}_i - \hat{q}_j}{\theta_j}\right)^2\right) \tag{7}$$

for the two-body kernel, where θ_j is a lengthscale hyperparameter. For a three-species alloy, we require six two-body kernels (one for each pair of species), and three SOAP kernels (one for each species). The kernel hyperparameters of Eqs. (5), (6), and (7) that we use are summarized in Table II. In addition these, we use $\sigma_\nu^{energy} = 0.001$ eV/atom, $\sigma_\nu^{force} = 0.1$ eV/Å, and $\sigma_\nu^{virial} = 0.05$ eV/atom for the regularization described in Eq. (4), values chosen based on the accuracy of our DFT training data. The MPI parallel `gap_fit` code from the QUIP software package [57,58] was used to fit the GAP on the full dataset of Table I.

The cumulative distribution functions of the training and test energy errors of the resulting nonmagnetic GAP are
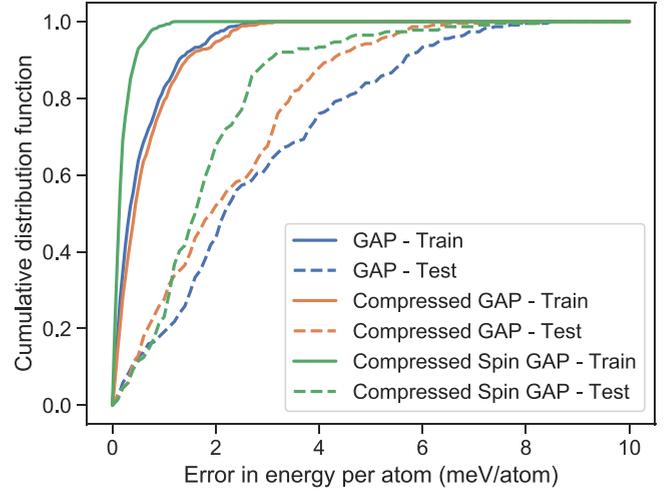
FIG. 3. Cumulative distribution functions of energy training (solid lines) and test (dashed lines) errors for the standard SOAP + GAP model (blue), a GAP model using compressed SOAP descriptors (orange) and a GAP model with an additional species included to differentiate between spin-up and spin-down Fe sites (spin GAP; green). Compression does not significantly reduce accuracy (orange vs blue), and the spin GAP model (green) performs significantly better.

shown in Fig. 3. Similar plots for the force and stress errors have been included in Appendix B. The GAP has energy training and test error within 2 meV/atom and 6 meV/atom respectively, which are comparable to the differences in DFT energies due to choice of AFM layering direction. The train and test force errors are below 200 meV/Å for low-temperature configurations and 400 meV/Å for high-temperature configurations. The relatively large force errors for low-temperature configurations are because the GAP cannot fit the directional spin-induced face-centered tetragonal (FCT) ground-state structure, as the GAP does not have spin degrees of freedom. The stress training and test errors are within 2 GPa and 4 GPa respectively, which are small as they are less than 2% of the typical elastic constant values that are about ∼200GPa.

E. Spin Model

As the SOAP and two-body descriptors do not have any information on spin, the standard GAP is spin-agnostic and cannot capture the tetrahedral deformation of AFM layering, and this limits the accuracy of its predictions. To address this, we introduce a fictitious fourth species to the GAP model, such that we have different descriptors for up spin (Fe ↑) and down spin (Fe ↓). This expanded descriptor set gives the model the flexibility to learn the tetrahedral deformation of AFM layering, and could be extended to the PM phase too. We henceforth refer to this 4-species potential as the spin GAP.

As the SOAP descriptor length scales quadratically with number of species, slowing down training and evaluation times prohibitively, we needed to use descriptor compression methods to make the training computationally tractable, such as those introduced in [59]. We use combined radial and species compression with tensor product coupling across 300

TABLE III. Lattice parameters and elastic constants of the FeCrNi alloy computed using DFT, GAP, spin GAP, and EAM-13 at 0 K. For comparison, DFT results for AFM Fe at 0 K and experimental values for finite temperatures have also been included. Dashes indicate that the value is equal to the previous value by symmetry.

	DFT-Fe	DFT	GAP	Spin GAP	EAM13	Expt (> 0 K)
Phase	FCT	FCT	FCO	FCT	FCC	∼FCC [a]
a (Å)	3.42	3.50	3.49	3.50	3.55	3.52–3.59 [a]
b (Å)	–	–	3.53	–	–	–
c (Å)	3.68	3.58	3.56	3.58	–	–
C_{11}, C_{22}, C_{33} (GPa)	321, –, 254	259, –, 249	299, 243, 232	272, –, 268	333, –,	204–226 [b]
C_{12}, C_{13}, C_{23} (GPa)	246, 96, –	164, 124, –	175, 163, 151	181, 126, –	196, –, –	132–134 [b]
C_{44}, C_{55}, C_{66} (GPa)	166, –, 262	147, –, 191	155, 171, 186	148, –, 188	161, –, –	111–122 [b]

[a]Temperature range 750–900 K. Ref. [60].
[b]Temperature range 4–295 K. Ref. [61,62].

mixed radial and species channels. To check the effects of compression, a compressed version of the three-species GAP was also fit. Figure 3 shows that the compressed GAP has comparable errors to the standard GAP, verifying that there is no significant loss of accuracy when compression is used for this system. The spin GAP, which also uses compressed descriptors, has significantly improved training and test errors over those of the standard GAP. Typical energy errors decrease by 1 meV/atom, force errors by about 50 meV/Å, and stress training errors decrease by about 0.2 GPa.

III. RESULTS

A. Bulk properties

The ground-state bulk properties of the FeCrNi alloy predicted by the GAP and the spin GAP are summarized in Table III, along with the corresponding DFT, EAM, and experimental values for comparison, as well as DFT for pure Fe in an AFM spin state. The DFT results for the ground-state structure of both AFM-Fe and AFM-FeCrNi are FCT, with the longer lattice constant corresponding to the AFM layering direction. This tetragonal distortion is more significant in pure AFM Fe (0.26 Å) compared to the alloy (0.08 Å), as the high concentration of Cr and Ni in the alloy lattice reduce the magnetism-induced geometry effects. Relaxation of EM structures with EAM-13 gives a lattice constant within the DFT and experimental range; however, EAM-13 fails to predict the lattice distortion, leading to a symmetric face centered cubic (FCC) lattice instead, as seen from the first set of bars in Fig. 4(a).

In comparison, the standard GAP predicts that a distorted structure is more energetically favorable than a symmetric FCC structure, as it is trained on relaxed DFT structures that have FCT lattices with zero atomic forces. However, since the standard GAP has no spin degrees of freedom, it does not have the framework to process the spin inputs required to elongate one direction over the other two. Instead, the GAP training correlates the distortion to minor differences in the alloy composition along the different axes in the relaxed training set geometries, and so predicts a face-centred orthorhombic (FCO) relaxed lattice based on composition along the different axes of test structures. As seen from the second set of bars in Fig. 4(a), the three lattice constants of the GAP relaxed FCO structure are equally spaced in a similar range to that

of the two DFT relaxed FCT lattice constants. The spin GAP, on the other hand, has a more flexible model form that can capture the FCT distortion associated with alternating Fe ↑ and Fe ↓ layers. The third set of bars in Fig. 4(a) show that the spin GAP successfully predicts lattice constants in excellent agreement with DFT, and the experimental values in Table III.

The cumulative distribution functions of the nonzero elastic matrix components, computed using the `matscipy` PYTHON package [63], of EAM-13, spin GAP, and DFT are shown in Fig. 4(b). As the EAM predicts an FCC structure, it has only three distinct nonzero elastic constants $\{C_{11}, C_{12}, C_{44}\}$. The DFT and spin GAP predict an FCT structure, and so have six distinct nonzero elastic constants $\{C_{11}, C_{33}, C_{12}, C_{13}, C_{44}, C_{66}\}$. The standard GAP on the other hand has nine distinct nonzero elastic constants as it predicts an FCO structure, the mean values of which are in Table III.

As seen in the first two sets of bars in Fig. 4(b), the spin GAP slightly overestimates C_{11} and C_{33} by ∼8% each, but gets their relative magnitude right. This is a significant improvement from EAM-13, which overestimates the DFT C_{11} by 30%. The EAM-13 C_{11} is in fact closer to that of pure AFM Fe reported in Table III, indicating that EAM-13 does not capture the softening of this elastic mode due to alloying elements Cr and Ni, while the spin GAP does capture this softening. Moving to the off-diagonal elastic components, the spin GAP C_{13} agrees very well with DFT, and is an improvement over the EAM-13 that overestimates it by 20%. The spin GAP performs less well with C_{12}, overestimating it by 10%, but again this is a significant improvement over EAM-13, which overestimates this component by 58%. Lastly, the spin GAP C_{44} and C_{66} match very well with DFT as seen in the last two sets of bars in Fig. 4(b), while EAM-13 predicts an averaged value between the DFT $[C_{44}, C_{66}]$ range. Overall, the GAP elastic constants agree well with DFT. The 8–10% overestimation of modes $\{C_{11}, C_{33}, C_{13}\}$ indicate that there are still errors in getting the behavior of the long vs short axis of the spin-induced tetrahedral deformation. This is likely because our spin model is still quite a simple approximation to the full physics of the alloy—it excludes contributions from Cr and Ni spins and ignores spin magnitudes.

Given that the spin GAP predicts the different elastic constants well, it unsurprisingly also predicts the equation of state reasonably well. Figure 5(a) compares the three spin GAP E-V curves (AFM spin layering along $\hat{\mathbf{X}}$, $\hat{\mathbf{Y}}$, and $\hat{\mathbf{Z}}$ directions) of a 256-atom test configuration with respect to its DFT EV
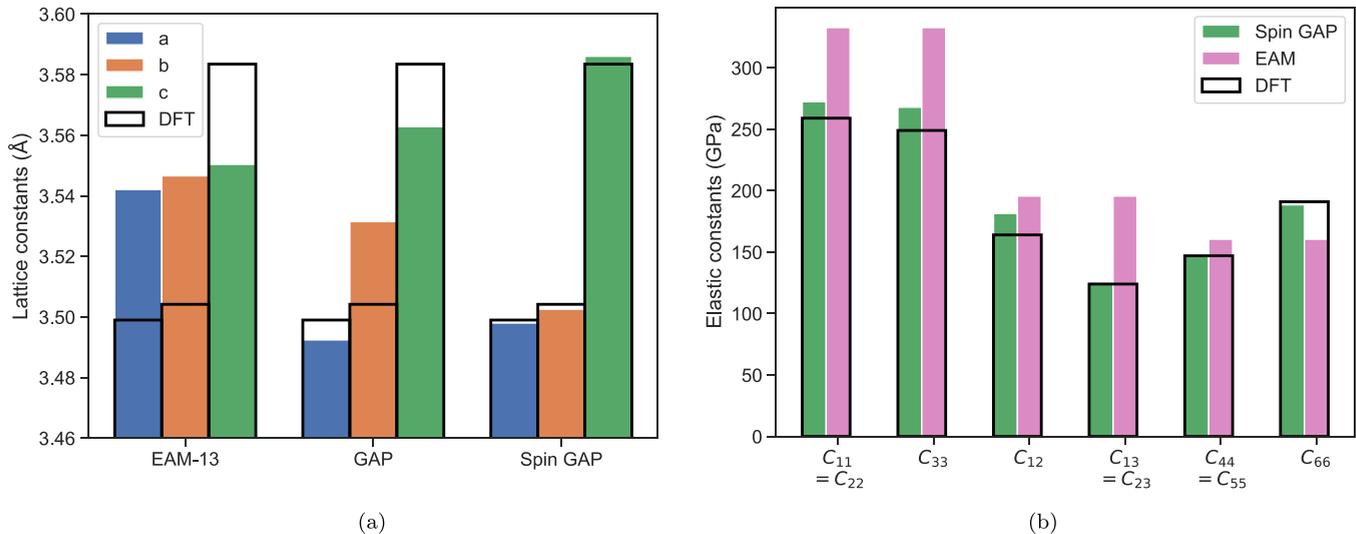
FIG. 4. (a) Lattice constants (a), (b), (c) averaged over 400 configurations predicted by EAM, GAP, and spin GAP for the AFM state, with respect to the DFT AFM lattice constants averaged over 111 configurations. (b) The six independent elastic constants of FCT structures predicted by spin GAP (green) and EAM (pink) from the average over 100 configurations each, compared to the DFT reference values computed for one configuration.

curves. We can see than the spin GAP EV curves fall within the range of the DFT EV curves, even if they do not match exactly. The spin GAP also gets the relaxed volume (E-V curve minimum) right, whereas the EAM-13 overestimates the relaxed volume as it is restricted to the FCC lattice. The bulk modulii computed from fitting an equation of state to these spin GAP E-V curves are in very good agreement with DFT, overestimating it by just 4 GPa ($\sim$2%), as seen in Fig. 5(b) (green distribution vs black dashed DFT reference), whereas the EAM-13 (pink distribution) overestimates the bulk modulus by 21%.

As the spin GAP is orders of magnitude faster than DFT, it allows for much more sampling of different random configurations, or larger supercells than the 256-atom ones used for DFT. Figure 6 shows the mean and spread (one standard deviation) of the spin GAP relaxed energy distributions at different supercell sizes, computed over 40 different configurations for each size. We see that the spin GAP relaxed energy converges for supercells with more than 200 atoms, and its converged energy is in good agreement with the DFT reference values (black crosses) within the converged size range. For smaller sizes, as is the case when applying DFT to random alloys, finite-size effects are significant leading to unconverged energy predictions. Also, the distributions at smaller sizes are broader, because each configuration's energy is computed over a smaller set of random local environments, leading to a larger variance between different configurations.

For the cell size used for training (256-atoms), the full distribution of spin GAP relaxed energies is in Fig. 10(a), and we can see that its peak agrees very well with the DFT mean. The spin-agnostic GAP on the other hand underestimates the ground-state energy by about 2 meV/atom. We attribute this underestimation to the fact that, while the GAP correlates the tendency to distort to minor changes in composition, it associates FCT directionality to noise as it does not have spin degrees of freedom. The GAP train-

ing regularizes the forces of the low energy DFT training structures, and extrapolates to predicting FCO as a lower-energy structure than FCT. The spin GAP overcomes this model form error by associating the tetragonal distortion with the AFM layering.

B. Chemical potentials

An essential ingredient for accurately computing energetics of point defects and their clusters is the chemical potential μ of each element in the alloy. This is the energy required to move an atom of the element from vacuum (infinitely far away from any interactions) to its bulk lattice site in the alloy supercell. The energy $E^{A \to B}$ required to substitute an element A with another element B at a lattice site in an alloy is equal to

$$E^{A \to B} = \mu^{B} - \mu^{A}. \qquad (8)$$

This gives three equations for A,B $\in$ {Fe,Cr,Ni}, of which only two are independent. To solve for the three unknown chemical potentials, we require three independent equations, and we can get this third equation from the sum of chemical potentials of the N atoms in a configuration, which gives the bulk energy E_N of that configuration. For Fe_7Cr_2Ni, this is expressed by the equation

$$E_N = (0.7\mu^{Fe} + 0.2\mu^{Cr} + 0.1\mu^{Ni})N. \qquad (9)$$

The chemical potentials μ^{Fe}, μ^{Cr}, and μ^{Ni} can be determined for any lattice site in a cell by computing the two relevant substitution energies, and then solving the system of equations defined by Eqs. (8) and (9). For the spin GAP, we solve this three-species system of equations twice: first for {Fe $\uparrow$, Cr, Ni}, and then for {Fe $\downarrow$, Cr, Ni}, and we then take the average of the two resulting chemical potentials per species to get μ^{Fe}, μ^{Cr}, and μ^{Ni}.
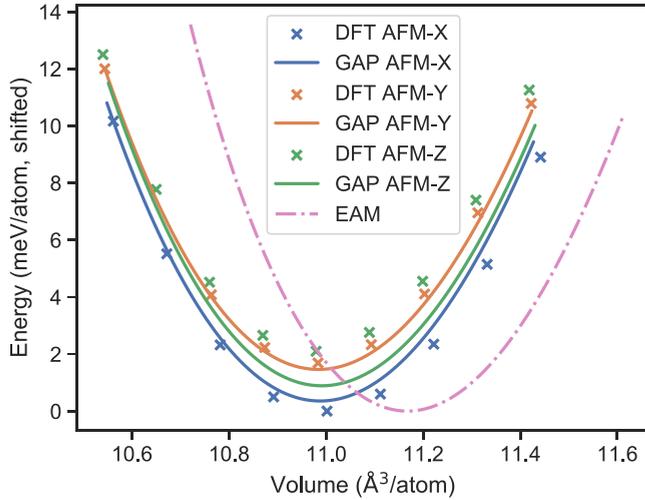
(a)
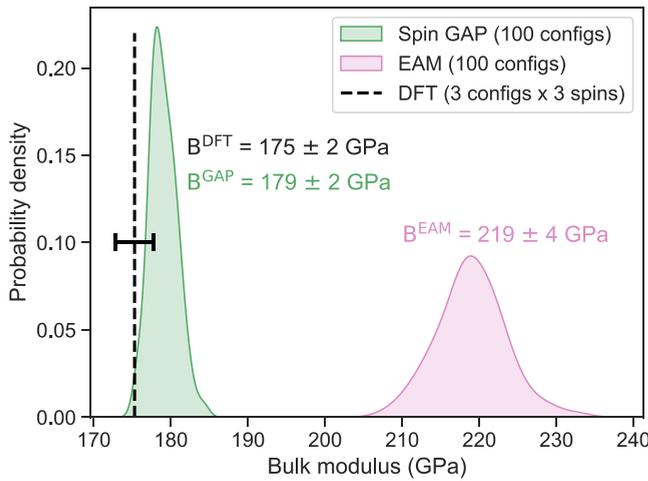
(b)

FIG. 5. (a) Energy volume (E-V) curves predicted by DFT (crosses) and spin GAP (solid lines) with AFM layering in $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ directions, and the corresponding spin-agnostic EAM-13 (dotted-pink line). (b) Bulk modulus computed from a Birch-Murnaghan equation of state fit to the E-V curves.

This procedure was carried out for randomly chosen lattice sites of 800 different configurations using the spin GAP. The distributions of substitution energies are shown in Fig. 7(a), with reference DFT minimum values $E_{\mathrm{subs}}^{\mathrm{A} \to \mathrm{B}}$ from [5]. We see that the minimum substitution energies from literature (using the same DFT settings as this paper) agree well with the minima of the corresponding spin GAP substitution energy distributions. The resulting chemical potential distributions are shown in the violin plots of Fig. 7(b). DFT chemical potentials averaged over five configurations are also labeled using black crosses. We can see that the spin GAP $\langle \mu^{\mathrm{Fe}} \rangle$ agrees very well with DFT. The spin GAP $\langle \mu^{\mathrm{Ni}} \rangle$ and $\langle \mu^{\mathrm{Cr}} \rangle$ are slightly offset from the corresponding DFT averages by $+2\%$ and -2% respectively; this is probably a result of not accounting fully for the magnetic nature of Ni and Cr in this alloy.
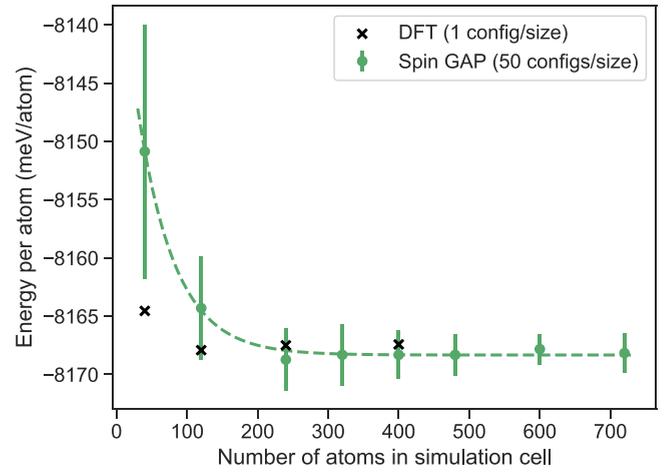
FIG. 6. Convergence of spin GAP ground-state energy with respect to supercell size. The green dots and bars show the mean and standard deviation respectively, of the spin GAP energy distribution of 40 configurations at each cell size. The spin GAP energy converges to 1 meV/atom for cells with > 200 atoms, and the converged energy agrees well with DFT (black crosses).

C. Monovacancies

To compute vacancy formation energies, we start with a relaxed bulk configuration of N atoms and energy E_N, remove one atom of a given species "el", and then carry out a fixed-cell relaxation of the atomic positions to get the energy E_{N-1}^{el} of the vacancy configuration. As the vacancy configuration has one atom less than the corresponding bulk system, the comparison of the two configurations requires the chemical potential μ^{el} of the atom removed to form the vacancy, giving the formula

$$E_{\mathrm{vac}}^{\mathrm{el}} = \mu^{\mathrm{el}} + E_{N-1}^{\mathrm{el}} - E_N \qquad (10)$$

where quantities $(\mu^{\mathrm{el}} + E_{N-1}^{\mathrm{el}})$ and E_N both have N atoms each. We first compute the intermediate step $(E_{N-1}^{\mathrm{el}} - E_N)$ for 800 different configurations using the spin GAP and the results are shown in Fig. 7(c). In this case, the spin GAP result for Cr agrees well with DFT, whereas those for Fe and Ni are underestimated. These offsets are consistent with the spin GAP underestimating $E^{\mathrm{Fe} \to \mathrm{Ni}}$ slightly, as seen previously in Fig. 7(a).

In fact, the small species-dependent errors in μ^{el} and $E_{N-1}^{\mathrm{el}} - E_N$ cancel each other out to give a consistent behavior in vacancy energy E_{vac}, irrespective of which element was removed to form the vacancy. This is evident from the consistent mean and spread of the E_{vac} energy distributions denoted by the bars in Fig. 7(d). This cancellation of errors makes sense as μ^{el} involves introducing an atom into the bulk whereas $E_{N-1}^{\mathrm{el}} - E_N$ involves the opposite procedure of removing an atom from the bulk structure. The resulting prediction of $E_{\mathrm{vac}}^{\mathrm{SpinGAP}} = 1.80$ eV is in good agreement with the DFT value of $E_{\mathrm{vac}}^{\mathrm{DFT}} = 1.96$ eV from the literature [5]. This prediction for E_{vac} is significantly more accurate than the standard spin-agnostic GAP trained on the same database, which predicts $E_{\mathrm{vac}}^{\mathrm{GAP}} = 1.6$ eV. The spin GAP also performs marginally better than EAM-13, which gives a mean prediction of $E_{\mathrm{vac}} = 1.77$ eV as seen in Fig. 7(d). We attribute the
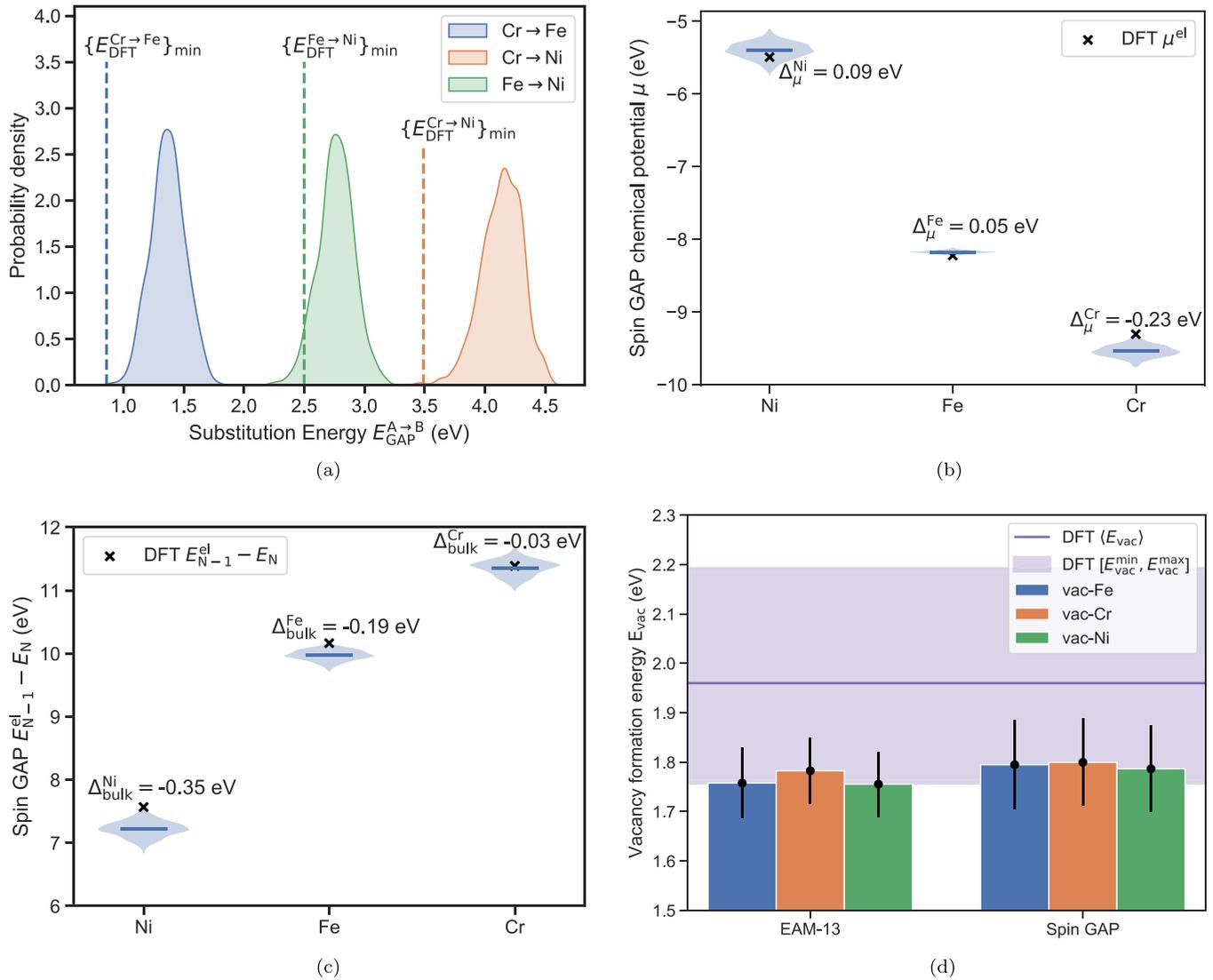
FIG. 7. (a) Distribution of substitution energies for elemental swaps predicted by spin GAP (100 configurations each) and corresponding DFT minima (dashed lines) computed from 40 DFT calculations performed by Piochaud *et al.* [5]. Distribution of (b) chemical potentials μ and (c) energy differences $E_{N-1} - E_N$ predicted by the spin GAP from 800 configurations (256-atom) with respect to DFT means from five configurations (256-atom). (d) Mean and twice the standard deviations of spin GAP and EAM-13 vacancy formation energies E_{vac} of 200 different EM configurations respectively, split by vacancy element type. The DFT mean and range of E_{vac} are taken from [5].

remaining error of –0.2 eV in the spin GAP prediction to the incomplete magnetic model, as we do not account for spin polarization of the electron density of Cr and Ni or for variations in Fe magnetic-moment magnitudes near to defect sites. A DFT study on this alloy finds that the DFT vacancy formation energy drops from 1.98 eV down to 1.80 eV when they switch from imposing paramagnetic spins to ferrimagnetic spins on the same chemical composition, showing that changes in the magnetic state alone can indeed lead to an order of 0.2 eV changes in E_{vac} [12], further suggesting that the spin GAP's underestimation of E_{vac} is likely due to limitations of our spin model.

On the other hand, a recently reported fully noncollinear magnetic atomic cluster expansion (ACE) model for iron also makes similar magnitudes errors in vacancy formation energies (errors of 0.4 eV and 0.2 eV when trained without and with extended defect sub-databases respectively), despite being a much more rigorous magnetic model [64]. In their case, the errors probably come from challenges in sampling their multiplicatively larger spin space (as spin magnitudes and angles are included). Whereas in our case, our simpler spin model comes with a relatively constrained spin space that is much easier to sample well. It is a promising result that we get comparable performance, suggesting that our spin model captures the main magnetic contributions despite being relatively simple. It would be interesting to try relaxing a few of the assumptions of our simple spin model while still keeping the spin space manageable, but that is beyond the scope of this paper and will be investigated in future works.

TABLE IV. Linear correlation of vacancy formation energy with respect to number of first nearest neighbours (#1nn) of type Cr and Ni respectively, from the literature [5,10] (columns 2 and 3) and our paper (column 5), and for $(E_{N-1}^{Fe} - E_N)$ from our paper (column 4).

#1nn	E_{vac} ([5])	E_{vac} ([10])	$(E_{N-1} - E_N)$	E_{vac}
Cr	0.05	0.02	0.02	0.01
Ni	–0.04	–0.04	–0.01	–0.004

D. Atomic Short Range Order

DFT studies [5,10] conclude that vacancies in the Fe_7Cr_2Ni are more stable in Ni-rich environments and less stable in Cr-rich environments. The gradients reported in the literature for the linear correlation between E_{vac} and number of Cr and Ni neighbours respectively have been included in Table IV for reference. The two studies agree well on the correlation with Ni neighbors, whereas Manzoor *et al.* find a weaker correlation with respect to Cr neighbors compared to Piochaud *et.al.*, which they attribute to the fact that they have four times the amount of sampling [10]. In both cases, however, they use constant precomputed values for the chemical potentials, so trends in their vacancy formation energy Eq. (10) are effectively proportional to trends in just $E_{N-1}^{el} - E_N$.

The computational expense of calculating chemical potentials, i.e., two fixed-cell relaxations of element substitutions Eq. (8) for each vacancy, makes it impractical to compute it for every vacancy using DFT, but this is now possible to run cheaply with the spin GAP. Figure 8 from left to right shows the trends in energy difference $E_{N-1}^{el} - E_N$, the corresponding chemical potentials, and the vacancy formation energies with respect to number of Cr neighbors as predicted by the spin GAP. These results show that the chemical potential is also

correlated to the number of Cr neighbors, with the opposite slope to the energy difference $E_{N-1}^{el} - E_N$. Computing vacancy formation energy involves adding up these two quantities [see Eq. (10)] and so their opposing slopes offset each other, reducing the overall magnitude of positive correlation between E_{vac} and number of Cr neighbors. Therefore, while we get consistent gradient of 0.02 eV/Cr for our $E_{N-1}^{el} - E_N$ and the E_{vac} from the literature (see values in Table IV), the inclusion of variations in chemical potentials decreases the overall correlation. This makes sense as the chemical potential accounts for majority of the heterogeneity in the vacancies neighborhood. The remnant positive correlation is likely because of the volume of the vacancy, which has previously been seen to affect stability [10]. As Cr atoms are the smallest species in the alloy, they relax less into the vacancy, leading to larger vacancies that are less stable and so have a higher formation energy.

The inverse trends (but to a weaker extent) of the three subplots of Fig. 8 are seen with respect to number of Ni neighbors, the gradients of which are reported in Table IV. The spin GAP predicts a smaller negative correlation of -0.01 eV/Ni for $E_{N-1}^{el} - E_N$ compared to the –0.04 eV/Ni values in the literature. Similar to the discrepancy between the two DFT studies for Cr trends, our lower correlation could be due to more variety in sampling. Piochaud *et al.* generate all their vacancy structure by removing different atoms from one 256-atom SQS structure, and Manzoor *et al.* do the same over two 256-atom SQS structures, leading to a large overlap in the neighborhood of the vacancies they consider. On the other hand, our spin GAP results were computed for vacancies introduced in 200 different starting structures from effective medium theory, and so sample a larger variety in long range vacancy environments. Despite the weaker correlations, we do still get a small negative relation of –0.004 eV/Ni for E_{vac} and number of Ni neighbors. Similar to the previous reasoning on
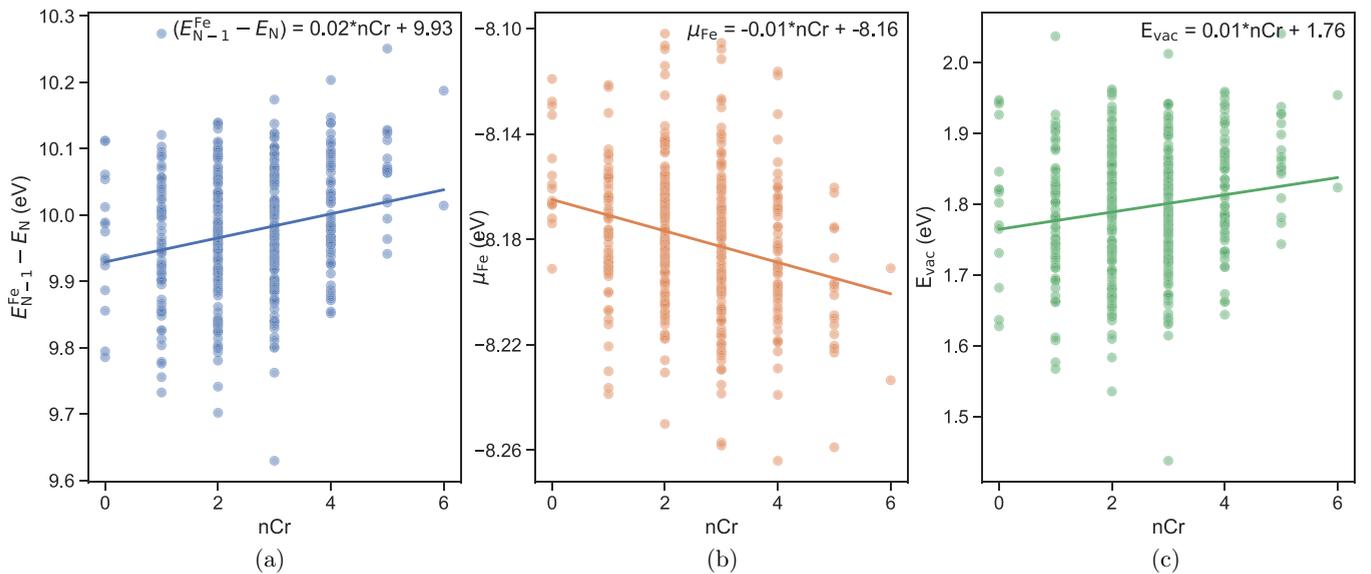
FIG. 8. Linear correlation between (a) energy difference of vacancy vs bulk configurations $E_{N-1}^{Fe} - E_N$ for vacancies generated by removing an iron atom each from the 200 different bulk configurations, (b) corresponding iron chemical potentials μ_{Fe} for those vacancy sites, and (c) corresponding vacancy formation energy E_{vac} computed by adding the quantities of (a) and (b), all with respect to number of Cr atoms amongst the 12 nearest neighbors of the vacancy.
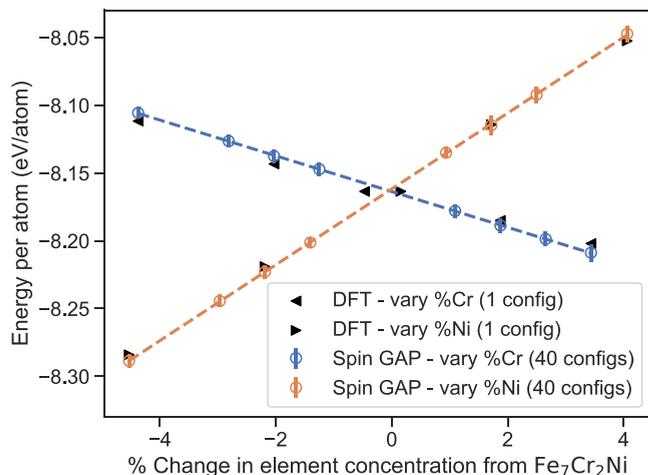
FIG. 9. Variation in ground-state energy per atom of bulk alloy with variation of Cr → Fe concentrations (blue) and Ni → Fe concentrations (orange), where 0% refers to the target 70:20:10 ratio of Fe:Cr:Ni. We see good agreement with DFT values (black triangles) for this range of ±4% variation in composition.

stability of vacancy volumes, Ni atoms are the largest species in the alloy, so having more Ni neighbors leads to smaller vacancy volumes, which are marginally more stable.

E. Alloy Composition

In the 304 and 316 austenitic stainless steel grades, the typical Cr and Ni concentrations can vary up to ±4% from the model 70:20:10 ratio of Fe, Cr, and Ni in our training set. It would hence be useful to check if our spin GAP can extrapolate to small variations in the alloying ratios. Figure 9 shows the results for bulk relaxation for configurations within ±4% variations in Cr (blue) and Ni (orange) concentrations. We see very good agreement with the DFT reference values in this range, validating that the spin GAP can indeed be used to extrapolate to compositions about the training ratio. This is because the GAP framework breaks down each configuration into local atomic environments (LAEs) as in Eq. (2), and the training set is likely to have a few Cr or Ni rich/poor LAEs since we sampled a large number of EM configurations, and these LAEs inform the potential for variations in alloy composition about the target ratio.

The gradients of the linear trends in Fig. 9 are 0.0132 eV/%Cr and 0.0280 eV/%Ni. These convert to 1.32 eV and 2.80 eV for single atom Cr → Fe and Ni → Fe swaps respectively, which are consistent with the respective mean substitution energies of Fig. 7(a). This linear trend indicates that for small variations in alloying concentrations, the substitution energies do not change much. The chemical potential, however, would vary, since their computation involves the bulk energy as well via Eq. (9), and these bulk energies vary linearly as seen in Fig. 9. Since we have accurate bulk and substitution energies for the ±4% concentration range, we can expect all derivative properties computed in the previous results sections to follow through reliably.
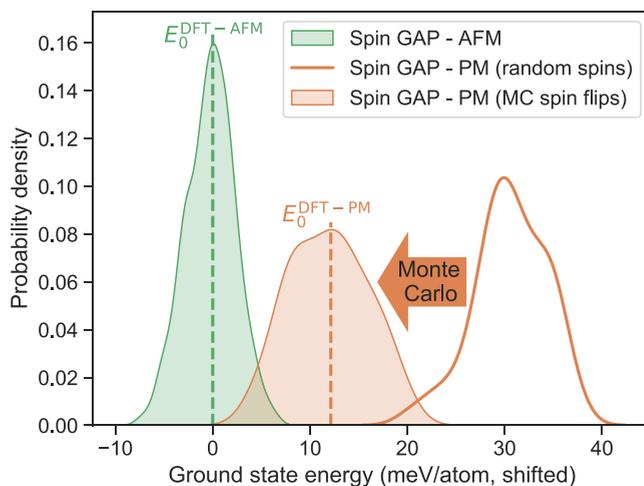
(a)
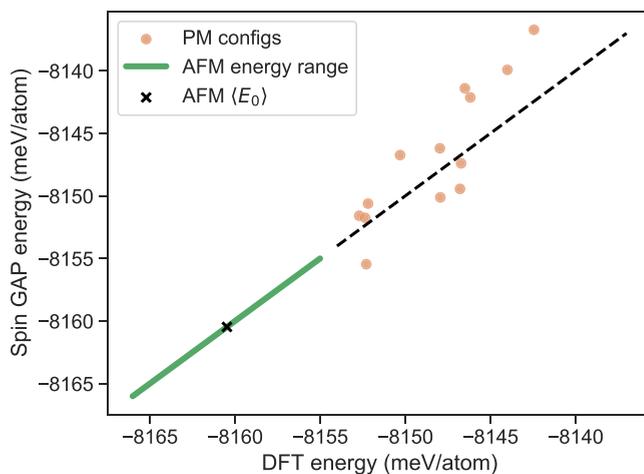
(b)

FIG. 10. (a) DFT mean values (dashed lines) and spin GAP distributions for ground-state energies for AFM (green) vs PM states using uniform random 50% up-down iron spins (unshaded orange) and after Monte Carlo optimisation with spin flip moves (shaded orange), comprising 400, 50, and 50 configurations respectively. (b) Correlation plot of spin GAP relaxed energies for 13 PM reference DFT configurations with respect to their DFT relaxed energies.

F. Paramagnetism

Although the spin GAP was trained only on AFM DFT data, in this section we assess how well it extrapolates to the paramagnetic (PM) state. Comparing the magnetic ground-state energies, Fig. 10(a) shows that the spin GAP correctly predicts the PM ground state of this alloy to be of higher energy, and hence less stable at 0 K, than the AFM ground state. The unshaded orange distribution comprises energies of the starting configurations initialized with perfectly random 50% up and down spin iron atoms, whereas the shaded orange distribution comprises their equilibrated energies after using Metropolis Monte Carlo with spin flip moves to optimize their paramagnetic spin states. We see that the equilibrated PM energy distribution agrees very well with the mean PM energy from DFT marked by the orange-dashed line in Fig. 10(a),
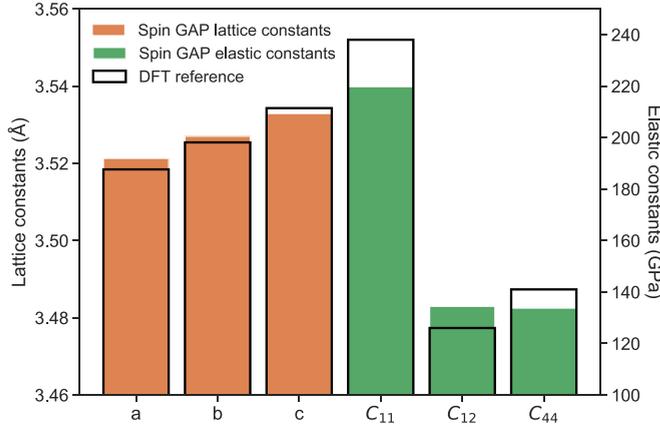
FIG. 11. Paramagnetic lattice constants (orange bars) and elastic constants (green bars) of the spin GAP averaged over 50 configurations, compared to their respective DFT reference values (black bars). The DFT reference for lattice constants are an average over our 13 DFT PM test configurations, while those for the elastic constants are from the study by Antillon *et al.* [12].

showing that the spin GAP can be used to generate reliable PM spin state configurations despite not being specifically trained for it. Figure 10(b) shows the direct correlation plot for energies of the 13 paramagnetic DFT test structures, where spin GAP was used to evaluate these test structure with the same spin state as that from DFT. We see that the spin GAP reproduces the DFT energies of these paramagnetic configurations very well, as all the points lie along the diagonal with the maximum deviation being 6 meV/atom.

For the PM ground-state structure, DFT predicts a FCC structure with an average lattice constant of 3.53 Å. The spin GAP accurately predicts the paramagnetic FCC structure, with lattice constants that agree remarkably well with DFT, as seen from the orange bars of Fig. 11. The lattice constants of all samples were sorted in ascending order and assigned to *a*, *b*, and *c*, respectively, and Fig. 11 shows their mean values. The small ∼0.005 Å difference between the three PM lattice constants is probably due to minor variations in ratio of the three species along the three axes, and this too is consistent between the spin GAP and DFT. As seen from the green bars in Fig. 11, the spin GAP agrees reasonably well with the paramagnetic elastic constants as well, agreeing within 5–8% to their respective DFT reference values from the study by Antillon *et al.* [12]. The fact that the spin GAP reproduces these properties is a promising result, as the paramagnetic state is relevant to modeling the temperature range of most applications of Fe_7Cr_2Ni based austenitic steels.

IV. CONCLUSIONS

In summary, we have developed two machine learning interatomic potentials—a standard GAP and a collinear spin GAP—for a model Fe_7Cr_2Ni alloy representative of austenitic steels. They are each trained on the same DFT database containing 159 k atomic environments, and validated against as independent test set. We demonstrate the shortcomings

of the standard GAP in not being able to capture the ground-state tetragonal distortion due to AFM layering observed in this alloy, as the standard GAP formalism cannot account for magnetic spins. We then proposed an extended model incorporating spin to correct for this issue, and verified that the spin GAP predicts the ground-state structure, energy, elastic properties, and vacancies in very good agreement with DFT.

Atomistic modeling using this spin GAP could be used to compute high accuracy inputs to larger scale microstructural models. Our spin GAP is seen to extrapolate well to small variations in alloy composition making is useful for atomistic studies of more variety in austenitic steel grades. The spin model is also seen to describe the paramagnetic state well, which is relevant to modeling the alloy at operating temperatures of the steel components in nuclear and industrial machinery. Beyond the properties validated in this paper, it could also be extended via iterative re-training to model more phenomena of interest such as diffusion, phase transformations or responses to radiation. For instance, by adding and testing the description of grain boundary structures, the spin GAP could be used to study segregation of Cr at grain boundaries under radiation. Another avenue for future work would be to extend this spin GAP to study hydrogen embrittlement in austenitic steel, relevant for applications in future fuel pipelines. This spin GAP is a good starting point for efforts at modeling austenitic steel with close to DFT accuracy.

The DFT training data, interatomic potential and property predictions supporting this work are freely available from Zenodo [65].

TABLE V. Fitted pairwise atom-atom interchange parameters obtained from the effective medium theory for the composition Fe_7Cr_2Ni, which are used in lattice-based Monte Carlo simulations to sample configuration space. All energies are in units of meV.

$V_{\alpha\alpha'}^{(1)}$	Fe	Cr	Ni
Fe	0.85	−5.28	4.58
Cr	−5.28	23.44	−9.95
Ni	4.58	−9.95	−12.18
$V_{\alpha\alpha'}^{(2)}$	Fe	Cr	Ni
Fe	0.12	−0.28	−0.29
Cr	−0.28	2.80	−3.64
Ni	−0.29	−3.64	9.33
$V_{\alpha\alpha'}^{(3)}$	Fe	Cr	Ni
Fe	−0.03	−0.16	0.53
Cr	−0.16	1.41	−1.70
Ni	0.53	−1.70	−0.30
$V_{\alpha\alpha'}^{(4)}$	Fe	Cr	Ni
Fe	−0.06	0.51	−0.59
Cr	0.51	−1.03	−1.50
Ni	−0.59	−1.50	7.15
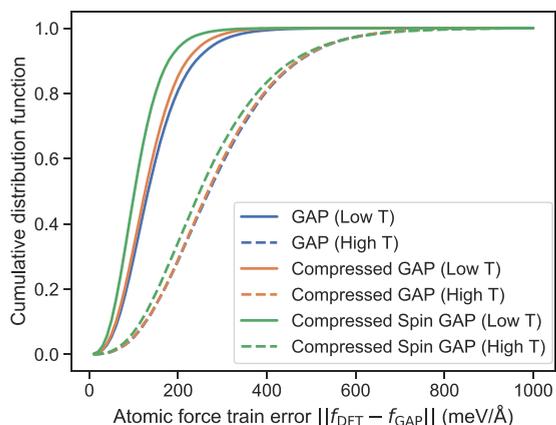
(a)
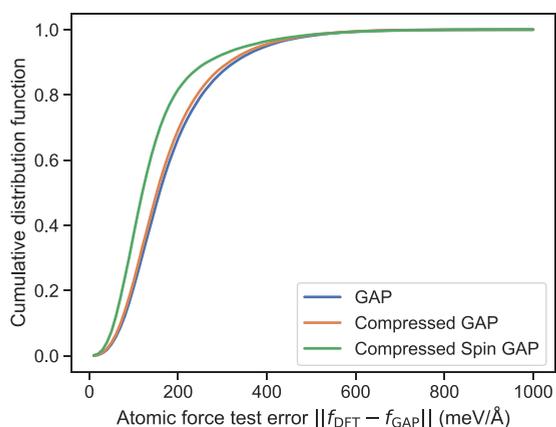
(b)
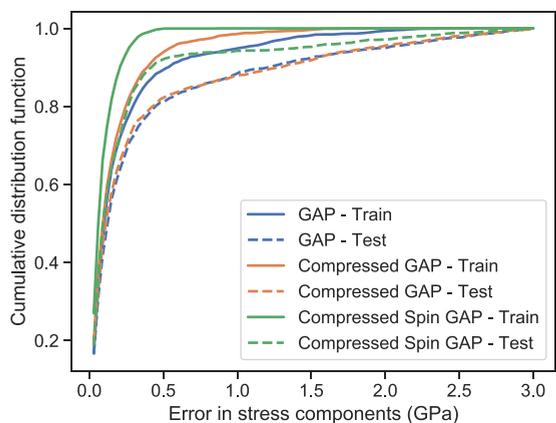
(c)

FIG. 12. Cumulative distributions of (a) training and (b) test errors in forces, and in (c) stresses, for the standard SOAP GAP (blue), GAP with compressed SOAP descriptors (orange), and spin GAP (green).

APPENDIX A: EFFECTIVE MEDIUM THEORY PARAMETERS

Tabulated in Table V are the pairwise atom-atom interchange parameters of Eq. (1) obtained for the Fe_7Cr_2Ni alloy considered in this paper. The lattice-based model assumes interactions are isotropic and homogeneous, and we write $V_{\alpha\alpha'}^{(n)}$ to denote the energy associated with an atom of species α interacting with an atom of species α' at nth nearest-neighbor distance. We find that it is sufficient to fit interactions to the first four-neighbor distances to accurately capture the data produced using the $S^{(2)}$ theory.

APPENDIX B: POTENTIAL TRAINING & TESTING

Figure 12 shows the training and testing errors for forces and stresses for the GAP, compressed GAP, and spin GAP. In all cases, the spin GAP is seen to perform the best (lower errors). The errors at which these trends plateau were discussed previously in Sec. II E. The composition of the database is reported in Table I.

[1] J. G. Parr, A. Hanson, and R. Lula, *Stainless Steel* (American Society for Metals, Metals Park, OH, 1993).

[2] P. Marshall, *Austenitic Stainless Steels: Microstructure and Mechanical Properties* (Elsevier Applied Science Publishing, London, 1984).

[3] H. Ledbetter and M. Austin, Mater. Sci. Eng. **70**, 143 (1985).

[4] G. S. Was, *Fundamentals of Radiation Materials Science* (Springer, Berlin, 2007).

[5] J. B. Piochaud, T. P. C. Klaver, G. Adjanor, P. Olsson, C. Domain, and C. S. Becquart, Phys. Rev. B **89**, 024101 (2014).

[6] O. Dimitrov and C. Dimitrov, J. Nucl. Mater. **105**, 39 (1982).

[7] C. Dimitrov and O. Dimitrov, J. Phys. F **14**, 793 (1984).

[8] M. Benkaddour, D. Bahia, Y. Belkadi, J. Barkani, and A. Benkaddour, Mater. Today: Proceedings **45**, 7359 (2021).

[9] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[10] A. Manzoor, Y. Zhang, and D. S. Aidhy, Comput. Mater. Sci. **198**, 110669 (2021).

[11] Y. Zhang, A. Manzoor, C. Jiang, D. Aidhy, and D. Schwen, Comput. Mater. Sci. **190**, 110308 (2021).

[12] E. Antillon, N. Bernstein, and M. D. Johannes, Phys. Rev. Mater. **6**, 053801 (2022).

[13] R. A. Johnson, Phys. Rev. B **39**, 12554 (1989).

[14] B. Jelinek, S. Groh, M. F. Horstemeyer, J. Houze, S. G. Kim, G. J. Wagner, A. Moitra, and M. I. Baskes, Phys. Rev. B **85**, 245102 (2012).

[15] H. Rafii-Tabar and A. P. Sulton, Philos. Mag. Lett. **63**, 217 (1991).

[16] R. W. Smith and G. S. Was, Phys. Rev. B **40**, 10322 (1989).

[17] M. Grujicic and X. Zhou, Calphad **17**, 383 (1993).

[18] G. Bonny, D. Terentyev, R. C. Pasianot, S. Poncé, and A. Bakaev, Modell. Simul. Mater. Sci. Eng. **19**, 085008 (2011).

[19] G. Bonny, N. Castin, and D. Terentyev, Modell. Simul. Mater. Sci. Eng. **21**, 085004 (2013).

[20] A. Daramola, G. Bonny, G. Adjanor, C. Domain, G. Monnet, and A. Fraczkiewicz, Comput. Mater. Sci. **203**, 111165 (2022).

[21] V. L. Deringer, M. A. Caro, and G. Csányi, Adv. Mater. **31**, 1902765 (2019).

[22] C. van der Oord, M. Sachs, D. P. Kovács, C. Ortner, and G. Csányi, npj Comput. Mater. **9**, 168 (2023).

[23] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010).

[24] R. Drautz, Phys. Rev. B **99**, 014104 (2019).

[25] A. V. Shapeev, Multiscale Model. Simul. **14**, 1153 (2016).

[26] I. Batatia, D. P. Kovács, G. N. C. Simm, C. Ortner, and G. Csányi, Adv. Neural Inf. Process. Syst. **35**, 11423 (2022).

[27] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, Nat. Commun. **13**, 2453 (2022).

[28] R. Lot, F. Pellegrini, Y. Shaidu, and E. Küçükbenli, Comput. Phys. Commun. **256**, 107402 (2020).

[29] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, Phys. Rev. X **8**, 041048 (2018).

[30] W. J. Szlachta, A. P. Bartók, and G. Csányi, Phys. Rev. B **90**, 104108 (2014).

[31] D. Dragoni, T. D. Daff, G. Csányi, and N. Marzari, Phys. Rev. Mater. **2**, 013808 (2018).

[32] V. L. Deringer and G. Csányi, Phys. Rev. B **95**, 094203 (2017).

[33] C. W. Rosenbrock, K. Gubaev, A. V. Shapeev, L. B. Pártay, N. Bernstein, G. Csányi, and G. L. W. Hart, npj Comput. Mater. **7**, 24 (2021).

[34] F. C. Mocanu, K. Konstantinou, T. H. Lee, N. Bernstein, V. L. Deringer, G. Csányi, and S. R. Elliott, J. Phys. Chem. B **122**, 8998 (2018).

[35] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013).

[36] W. L. Bragg and E. J. Williams, Proc. R. Soc. London, Ser. A **145**, 699 (1934).

[37] W. L. Bragg and E. J. Williams, Proc. R. Soc. London, Ser. A **151**, 540 (1935).

[38] S. N. Khan, J. B. Staunton, and G. M. Stocks, Phys. Rev. B **93**, 054206 (2016).

[39] C. D. Woodgate and J. B. Staunton, Phys. Rev. B **105**, 115124 (2022).

[40] C. D. Woodgate and J. B. Staunton, Phys. Rev. Mater. **7**, 013801 (2023).

[41] C. D. Woodgate, D. Hedlund, L. H. Lewis, and J. B. Staunton, Phys. Rev. Mater. **7**, 053801 (2023).

[42] C. D. Woodgate, L. H. Lewis, and J. B. Staunton, arXiv:2401.02809.

[43] J. S. Faulkner, G. M. Stocks, and Y. Wang, *Multiple Scattering Theory Electronic Structure of Solids*, 1st ed. (IOP Publishing, Bristol, 2018).

[44] M. Hoffmann, A. Ernst, W. Hergert, V. N. Antonov, W. A. Adeagbo, R. M. Geilhufe, and H. Ben Hamed, Physica Status Solidi (b) **257**, 1900671 (2020).

[45] G. Kresse and J. Hafner, Phys. Rev. B **49**, 14251 (1994).

[46] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).

[47] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[48] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[49] P. Olsson, C. Becquart, and C. Domain, Mater. Res. Lett. **4**, 219 (2016).

[50] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).

[51] S. Cottenier, B. D. Vries, J. Meersschaut, and M. Rots, J. Phys.: Condens. Matter **14**, 3275 (2002).

[52] R. Hafner, D. Spišák, R. Lorenz, and J. Hafner, Phys. Rev. B **65**, 184432 (2002).

[53] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[54] S. Nosé, J. Chem. Phys. **81**, 511 (1984).

[55] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).

[56] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, MA, 2006).

[57] S. Klawohn, J. P. Darby, J. R. Kermode, G. Csányi, M. A. Caro, and A. P. Bartók, J. Chem. Phys. **159**, 174108 (2023).

[58] S. Klawohn, J. R. Kermode, and A. P. Bartók, Mach. Learn.: Sci. Technol. **4**, 015020 (2023).

[59] J. P. Darby, D. P. Kovács, I. Batatia, M. A. Caro, G. L. W. Hart, C. Ortner, and G. Csányi, Phys. Rev. Lett. **131**, 028001 (2023).

[60] A. Marucco, Mater. Sci. Eng., A **189**, 267 (1994).

[61] H. M. Ledbetter, J. Appl. Phys. **52**, 1587 (1981).

[62] V. Rajevac, M. Hoelzel, S. A. Danilkin, A. Hoser, and H. Fuess, J. Phys.: Condens. Matter **16**, 2609 (2004).

[63] P. Grigorev *et al.*, J. Open Source Software **9**, 93 (2024).

[64] M. Rinaldi, M. Mrovec, A. Bochkarev, Y. Lysogorskiy, and R. Drautz, npj Comput. Mater. **10**, 12 (2024).

[65] https://doi.org/10.5281/zenodo.10577535.