



OPEN

## Constrained DFT-based magnetic machine-learning potentials for magnetic alloys: a case study of Fe–Al

Alexey S. Kotykhov<sup>1,2</sup>, Konstantin Gubaev<sup>3</sup>, Max Hodapp<sup>4</sup>, Christian Tantardini<sup>5,6,7</sup>✉, Alexander V. Shapeev<sup>1</sup> & Ivan S. Novikov<sup>1,2</sup>✉

We propose a machine-learning interatomic potential for multi-component magnetic materials. In this potential we consider magnetic moments as degrees of freedom (features) along with atomic positions, atomic types, and lattice vectors. We create a training set with constrained DFT (cDFT) that allows us to calculate energies of configurations with non-equilibrium (excited) magnetic moments and, thus, it is possible to construct the training set in a wide configuration space with great variety of non-equilibrium atomic positions, magnetic moments, and lattice vectors. Such a training set makes possible to fit reliable potentials that will allow us to predict properties of configurations in the excited states (including the ones with non-equilibrium magnetic moments). We verify the trained potentials on the system of bcc Fe–Al with different concentrations of Al and Fe and different ways Al and Fe atoms occupy the supercell sites. Here, we show that the formation energies, the equilibrium lattice parameters, and the total magnetic moments of the unit cell for different Fe–Al structures calculated with machine-learning potentials are in good correspondence with the ones obtained with DFT. We also demonstrate that the theoretical calculations conducted in this study qualitatively reproduce the experimentally-observed anomalous volume-composition dependence in the Fe–Al system.

Magnetism is important to be explicitly taken into account for the successful computational prediction of many properties in single-component metals<sup>1–6</sup> and multi-component alloys<sup>7–15</sup>. In particular, magnetic properties of the constituting elements of alloys affect the phase stability<sup>7–10</sup>. Furthermore, magnetism can be responsible for the unusual properties like negative thermal expansion<sup>11,12</sup>, anomalous volume-composition dependence<sup>13</sup>, and the so-called “half-metallic behavior” in perovskites<sup>14</sup> or full-Heusler alloys<sup>15</sup>. In such multi-component alloys not only magnetism is a more complex *physical* phenomenon as compared to the single-component materials, the presence of magnetism in multi-component alloys also leads to extra difficulties for its *computational* studies. Steel, a workhorse of heavy industry, is an example of such a material; it has immense variety of applications and corresponding sub-types, possessing magnetic properties due to Fe being presented in their content. Steels, as well as other Fe-based alloys, can be characterized by different magnetic orders for which the electronic ground state can be substantially different. One of the most widely used methods for simulations of ground state properties of condensed matter is density functional theory (DFT). In magnetism, however, we are often interested in excited state properties (e.g., excited magnetic moments), while, strictly speaking, DFT is a theory for the electronic ground state: the fundamental theorem of DFT relies on a minimization of the energy in the functional space of many-body electronic wavefunction, which is the Slater determinant that avoid the electronic localization due to the spreading of all electrons on the entire orbitals (i.e., Kohn–Sham equations).

<sup>1</sup>Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Bolshoy Boulevard 30, Moscow 143026, Russian Federation. <sup>2</sup>Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow Region 141701, Russian Federation. <sup>3</sup>University of Stuttgart, Postfach 10 60 37, 70049 Stuttgart, Germany. <sup>4</sup>Materials Center Leoben Forschung GmbH (MCL), Leoben, Austria. <sup>5</sup>Hylleraas Center, Department of Chemistry, UiT The Arctic University of Norway, Langnes, PO Box 6050, 9037 Tromsø, Norway. <sup>6</sup>Department of Materials Science, Rice University, Houston, TX 77005, USA. <sup>7</sup>Institute of Solid State Chemistry and Mechanochemistry SB RAS, ul. Kutateladze 18, Novosibirsk 630128, Russian Federation. ✉email: christiantantardini@gmail.com; i.novikov@skoltech.ru

To overcome this problem constraints are introduced in the Kohn–Sham equations, either as soft constraints contributing a penalty function to the total energy<sup>16–18</sup> or, as recently proposed, hard constraints solved self-consistently using the Lagrangian multiplier method<sup>19</sup>.

Despite the success of different flavours of DFT calculations in investigating different materials during the last 30 years they are still computationally expensive even when used them on modern supercomputers. Moreover, in computational studies of magnetic materials, including magnetic moments as an additional degree of freedom in the calculation of DFT energy, the computational time increases drastically. To overcome the limitations of DFT in size and time, requires alternative approaches. Effective interaction model (EIM) was proposed in Refs.<sup>20,21</sup> and applied to the Fe–Ni system. One approach that has gained massive interest in the computational materials community over the last 15 years are machine-learning interatomic potentials (MLIPs)<sup>22–32</sup>. The main idea behind MLIPs is their ability to avoid running the full DFT simulation by interpolating between a relatively small training set of carefully selected single-point DFT calculations. A MLIP that has been trained on configurations that are representative for the entire configurational space appearing in a simulation then approximates local DFT energies and forces with, in principle, arbitrary accuracy—contrary to (semi-)empirical interatomic potentials. The cited MLIPs (see e.g.<sup>22–32</sup>) have been developed for non- or ferromagnetic materials, which do not require taking into account magnetic degrees of freedom explicitly in the functional of MLIPs. In order to approximate DFT energies that strongly depend on the magnitude and direction of the magnetic field we must enrich the functional form of MLIPs with magnetic moment features which has been proposed in a number of works<sup>33–38</sup>. The open problem, however, is the generation of suitable training sets. We emphasize that in addition to the configurations with non-equilibrium atomic positions and lattice parameters we also need configurations with non-equilibrium magnetic moments for the proper fitting of magnetic MLIP. We discuss the creation of such a training set in the present work.

In this paper we generalize a single-component magnetic Moment Tensor Potential<sup>35</sup> (one of the recently developed MLIPs with magnetic degrees of freedom, mMTP) to the case of multi-component magnetic materials. Here, we use the recently developed cDFT approach<sup>19</sup> to evaluate the magnetism of Fe–Al system by first-principles calculations being such approach seen to efficiently describe magnetism in considering a potential-based formulation of the self-consistency problem<sup>19</sup>. These cDFT data will be subsequently used to generate a training set for fitting an mMTP, i.e. in addition to the configurations with equilibrium magnetic moments we also include the ones with non-equilibrium (or, perturbed) magnetic moments in the training set. An mMTP fitted to such a training set allows for equilibrating magnetic moments of a configuration starting from the perturbed ones. In other words, we consider magnetic moments on the same grounds as atomic positions: we generate a training set with both relaxed and perturbed atomic positions, fit a potential, and, finally, relax both (magnetic and “positional”) degrees of freedom of the structures with the trained potential.

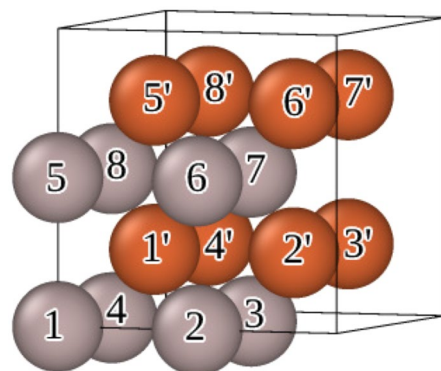
We test our multi-component mMTPs on the system of Fe–Al with different concentrations and positions of Al and Fe. We demonstrate that the formation energies, the equilibrium lattice parameters, and the total magnetic moments of the unit cell predicted with the fitted mMTPs for various compounds of the Fe–Al alloy are in a good correspondence with the DFT ones. We also demonstrate that the theoretical calculations conducted in this paper reproduce the anomalous volume–composition dependence in the Fe–Al system shown in Ref.<sup>13</sup> and experimentally observed<sup>39</sup>.

## Results and discussion

### Training set

The training set for magnetic Moment Tensor Potential (mMTP) fitting consists of different configurations of bcc Fe–Al in which the concentration of Al varies from 0 to 50%. We additionally include pure bcc–Al in the training set. The configurations in the training set consist of 16 atoms ( $2 \times 2 \times 2$  conventional bcc cell). There are altogether 2012 configurations in the training set constructed as detailed below.

For constructing the training set we start from the 23 “parent” configurations taken from Ref.<sup>13</sup>. These configurations differ the way Al occupies the supercell sites as given in Table 1. The notations for the supercell sites (as used in the first column of Table 1 where they denote the locations of the Al atoms) are shown in Figure 1.



**Figure 1.** The  $2 \times 2 \times 2$  16-atomic supercell with the notations of atomic sites in Fe–Al structures.

| Structure                                     | Step 1 | Step 2 | Step 3.1 | Step 3.2 |
|---|--------|--------|----------|----------|
| Bcc-Al  | 3      | 15     | 75       | 14       |
| Bcc-Fe  | 3      | 14     | 60       | 14       |
| Fe <sub>15</sub> Al <sub>1</sub> -1           | 3      | 15     | 75       | 14       |
| Fe <sub>14</sub> Al <sub>2</sub> -11'         | 7      | 15     | 15       | 14       |
| Fe <sub>14</sub> Al <sub>2</sub> -15          | 7      | 14     | 69       | 14       |
| Fe <sub>14</sub> Al <sub>2</sub> -13          | 5      | 14     | 69       | 14       |
| Fe <sub>14</sub> Al <sub>2</sub> -17          | 3      | 15     | 75       | 14       |
| Fe <sub>13</sub> Al <sub>3</sub> -124         | 12     | 14     | 67       | 14       |
| Fe <sub>13</sub> Al <sub>3</sub> -157         | 11     | 14     | 56       | 14       |
| Fe <sub>13</sub> Al <sub>3</sub> -368         | 7      | 15     | 75       | 14       |
| Fe <sub>12</sub> Al <sub>4</sub> -1245        | 5      | 11     | 55       | 14       |
| Fe <sub>12</sub> Al <sub>4</sub> -1247        | 11     | 12     | 58       | 14       |
| Fe <sub>12</sub> Al <sub>4</sub> -1357        | 10     | 15     | 45       | 14       |
| Fe <sub>12</sub> Al <sub>4</sub> -1368        | 3      | 8      | 39       | 14       |
| Fe <sub>11</sub> Al <sub>5</sub> -13681'      | 5      | 6      | 30       | 14       |
| Fe <sub>11</sub> Al <sub>5</sub> -12457       | 5      | 15     | 75       | 14       |
| Fe <sub>10</sub> Al <sub>6</sub> -124568      | 0      | 0      | 0        | 0        |
| Fe <sub>10</sub> Al <sub>6</sub> -123468      | 0      | 0      | 0        | 0        |
| Fe <sub>10</sub> Al <sub>6</sub> -12341'5'    | 18     | 15     | 72       | 14       |
| Fe <sub>10</sub> Al <sub>6</sub> -124678      | 4      | 15     | 74       | 14       |
| Fe <sub>9</sub> Al <sub>7</sub> -1234568      | 6      | 15     | 75       | 14       |
| Fe <sub>8</sub> Al <sub>8</sub> -12345678     | 3      | 14     | 70       | 14       |
| Fe <sub>8</sub> Al <sub>8</sub> -13682'4'5'7' | 3      | 15     | 69       | 14       |
| Total:  | 134    | 286    | 1298     | 294      |

**Table 1.** Construction of the training set. We first conduct full geometrical optimization (relaxation) of structures and get fully equilibrium structures and the ones from the relaxation path (Step 1). Next we displace atomic positions and lattice vectors and run DFT calculations with the optimization of magnetic moments (Step 2). Finally, we perturb magnetic moments in the configurations obtained after Steps 1 and 2 and get the configurations with non-equilibrium magnetic moments (Steps 3.1 and 3.2). The numbers of converged configurations for each step are given in the table. The total number of converged configurations is 2012. Number of configurations in the training set: 2012.

For generating the configurations for the training set we first conduct full geometrical optimization (relaxation) of the structures (i.e., we minimize energies with respect to atomic positions, lattice vectors, and magnetic moments) from the left column in Table 1. We took the Fe magnetic moments of  $3 \mu_B$  and the Al magnetic moments of  $0 \mu_B$  as an initial guess. After the minimization we added the configurations from the relaxation path to the training set (see Step 1 in Table 1). After that we compress and extend lattice vectors of equilibrium configurations by 1%. Next, we take each of these configurations, and we apply random displacements to each coordinate of each atomic position in the range from  $-0.1$  to  $0.1 \text{ \AA}$ ; we do it five times and obtain five configurations from each equilibrium, compressed and extended configuration. For the resulting displaced configurations we optimize magnetic moments (see Step 2 in Table 1). In Step 3, as opposed to Steps 1 and 2 in which we conduct DFT calculations without any constraints, we impose hard constraints on magnetic moments to obtain the configurations with non-equilibrium magnetic moments. In Step 3.1 we take each configuration from the previous step and five times randomly perturb each equilibrium atomic magnetic moment by increasing or decreasing this value by at most 15%, and calculate these configurations with cDFT (see Step 3.1 in Table 1). As a result, the maximum deviation from the equilibrium magnetic moment for the Fe atom could reach  $0.4 \mu_B$ , for the Al atom the maximum deviation could be  $0.01 \mu_B$  (in the configurations with both Fe and Al atom). In the case of pure Al we randomly perturb magnetic moments by the values from the range  $(-0.03, 0.03) \mu_B$ . Finally, we randomly perturb equilibrium magnetic moments in the configurations obtained in Step 1 by increasing or decreasing their value by at most 50% and also conduct cDFT calculations for the configurations with non-equilibrium magnetic moments, but with the equilibrium positions and lattice vectors (see Step 3.2 in Table 1). As a result, the maximum deviation from the equilibrium magnetic moment for the Fe atom could reach  $1.3 \mu_B$ , for the Al atom the maximum deviation could be  $0.04 \mu_B$  (in the configurations with both Fe and Al atom). In the configurations with pure Al we randomly perturb magnetic moments by the values from the range  $(-0.1, 0.1) \mu_B$ . Thus, we obtain the training set including fully equilibrium configurations and the ones with perturbed atomic positions, lattice vectors, and magnetic moments. Number of configurations converged for each step and each structure is given in Table 1. The training set contains the total of 2012 configurations. We note that around 80% of Fe-Al configurations converged during the DFT (cDFT) calculations with the ABINIT code. We also emphasize that all the cDFT calculations with perturbed magnetic moments and equilibrium atomic positions and lattice parameters (at Step 3.2) were converged.

For the verification of the fitted mMTPs we also generate a “verification” set. We start with the configurations obtained in Step 1 of constructing the training set, and perform Steps 2 and 3.2 (i.e., we perturb atomic positions and lattice parameters and optimize magnetic moments (Step 2) and we perturb magnetic moments at equilibrium lattice parameters and atomic positions (Step 3.2)). We thus obtain additional 336 configurations for the verification of the fitted mMTPs. Unlike for the training set, we omitted Step 3.1 for the purpose of generating the verification set (and we therefore do not call it “validation set”), because we are mostly concerned about the accuracy on equilibrium configurations.

### Fitting and verification of the ensemble of mMTPs

The results of fitting and verification of the ensemble including five mMTPs are given in Table 2. It could be seen that the uncertainty in energy, force, and stress errors is small compared to the magnitude of the errors, and the fitting errors and the errors on the “verification” set are reasonable and typical for the periodic crystal systems. We use the fitted ensemble of five mMTPs for further computations of the values of interest because from Table 2 we see no overfitting of the ensemble of five mMTPs. However, it should be noted that the force error on the “verification” set is smaller than the fitting force error and the stress error on the “verification” set is two times larger than the fitting stress error. Both the difference between force and stress errors are beyond the 95 % confidence interval. The reason may be that we constructed the “verification” set without Step 3.1 as opposed to the training set. In order to check it out we carry out an additional fivefold cross-validation.

### Fivefold cross-validation

For the fivefold cross validation we combine the training and “verification” sets and create the total set of 2328 configurations. Next we split this set into five non-overlapping parts and carry out cross-validation, i.e. we fit mMTPs on four parts and validate it on the fifth part. The results are given in Table 3. From the table we conclude that the fitting and validation errors are close to each other and are within 99% confidence interval, i.e., the training set was constructed correctly.

### Formation energy, lattice parameter, and total magnetic moment for different Fe–Al compounds

For testing the predictive power of the ensemble of fitted mMTPs we compare the formation energies, the equilibrium lattice parameters, and the total magnetic moments of the unit cell predicted with the mMTPs and DFT (implemented in ABINIT) for different compounds of Fe–Al. Uncertainty of the errors estimation in all the mentioned quantities in figures and tables are provided with the 95 % confidence interval (i.e., 2- $\sigma$  interval).

Formation energy is given in Fig. 2. The ensemble of mMTPs correctly reproduces the formation energy trend calculated with DFT: it decreases when the concentration of Al increases. The maximum error of the formation energy prediction is around 20 meV/atom for the  $\text{Fe}_8\text{Al}_8$ -13682'4'5'7' structure. From Fig. 2 we find the configurations with minimum energy for a given concentration of Al, i.e., closest to the convex hull: pure bcc-Fe,  $\text{Fe}_{15}\text{Al}_1$ -1,  $\text{Fe}_{14}\text{Al}_2$ -13,  $\text{Fe}_{13}\text{Al}_3$ -368,  $\text{Fe}_{12}\text{Al}_4$ -1368,  $\text{Fe}_{11}\text{Al}_5$ -12457,  $\text{Fe}_{10}\text{Al}_6$ -124678,  $\text{Fe}_9\text{Al}_7$ -1234568, and  $\text{Fe}_8\text{Al}_8$ -12345678. We provide further results for these configurations below.

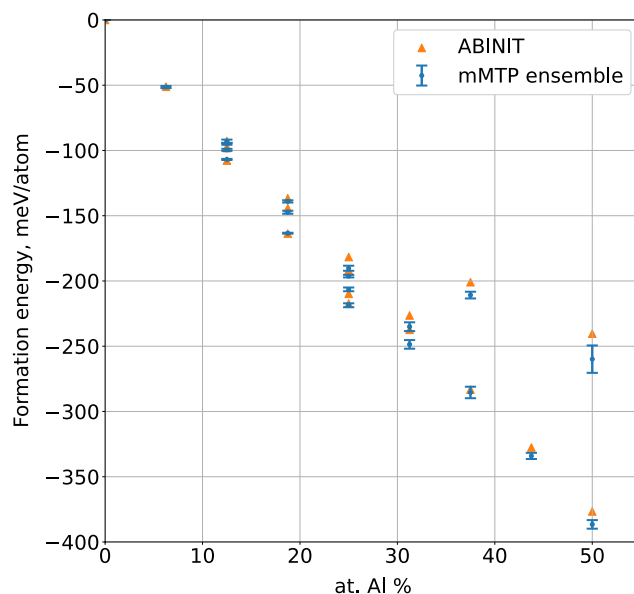
In Fig. 3 we compare the equilibrium lattice parameters calculated with the mMTPs and DFT. As for the formation energies, the mMTP- and DFT-provided equilibrium lattice parameters are close to each other, the maximum error in their prediction is about 0.014 Å (for the  $\text{Fe}_{11}\text{Al}_5$ -12457 structure). We also observe the anomalous lattice parameter/composition dependence in the Fe–Al structures. We, first, see the nearly linear increase of the lattice parameter up to the Al concentration of 18.75 %. When the Al concentration is between

|              | Energy error (meV/atom) | Force error (meV/Å) | Stress error (GPa) |
|--------------|-------------------------|---------------------|--------------------|
| Fitting      | $3.65 \pm 0.13$         | $60 \pm 3$          | $0.42 \pm 0.06$    |
| Verification | $4.03 \pm 0.05$         | $48 \pm 4$          | $0.79 \pm 0.10$    |

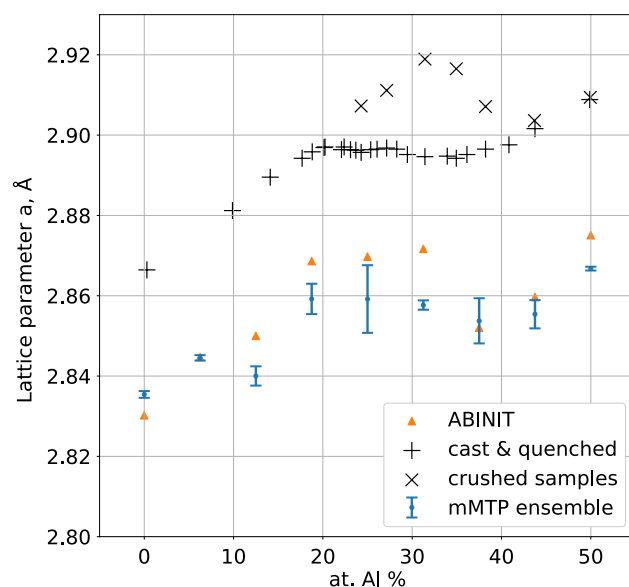
**Table 2.** Fitting root-mean-square errors (RMSEs) and RMSEs obtained on the “verification” set for the ensemble of five mMTPs and uncertainty of the errors estimation (we provide it with the 95 % confidence interval, i.e., 2- $\sigma$  interval). The obtained errors are reasonable and typical for the periodic crystal systems. The ensemble of fitted mMTPs was not overfitted.

|            | Energy error (meV/atom) | Force error (meV/Å) | Stress error (GPa) |
|------------|-------------------------|---------------------|--------------------|
| Fitting    | $2.8 \pm 0.7$           | $57 \pm 4$          | $0.36 \pm 0.06$    |
| Validation | $3.1 \pm 1.7$           | $51 \pm 4$          | $0.33 \pm 0.06$    |

**Table 3.** Fivefold cross-validation of mMTPs. We provide uncertainty of the errors estimation within the 99% confidence interval, i.e., 3- $\sigma$  interval. Fitting and validation errors are close to each other and are within 99% confidence interval.



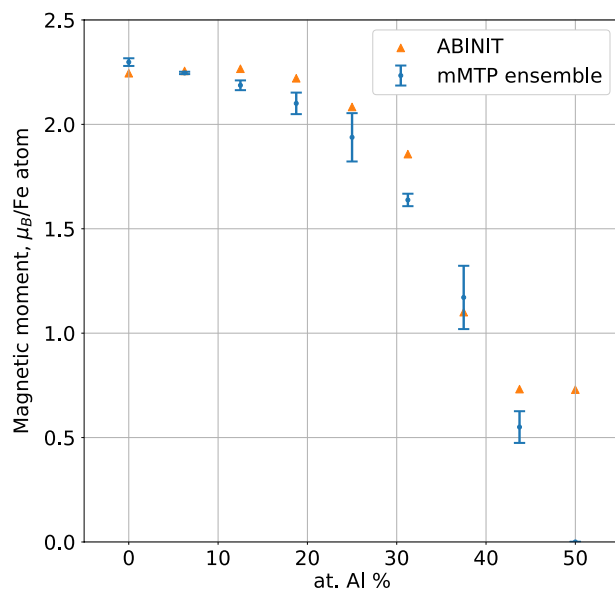
**Figure 2.** Formation energies for different compounds of Fe–Al predicted with the ensemble of mMTPs and DFT. The trend of formation energy was correctly reproduced with the ensemble of mMTPs, maximum error in formation energy prediction is 20 meV/atom for the  $\text{Fe}_8\text{Al}_8-13682'4'5'7'$  structure.



**Figure 3.** Lattice parameter dependencies on concentration of Al in the Fe–Al compounds computed with mMTP and DFT, and obtained experimentally (the experimental data are taken from Ref.<sup>39</sup>). The theoretical (mMTP and DFT) lattice parameters are close to each other: the maximum error in their prediction is 0.014 Å. Both the theoretical and the experimental dependencies of lattice parameters on Al concentration are anomalous: they are nonlinear at more than 18.75 % of Al concentration (at more than 31.25 % for the crushed samples).

18.75 % and 31.25 % the values of lattice parameters are close to constant. Next (between 31.25 % and 37.5 %), we see the decrease in lattice parameter and, finally, another increase up to 50 %.

In Fig. 3 we also provide the experimental lattice parameters from the paper<sup>39</sup>. The experimental work<sup>39</sup> has somewhat different lattice parameter dependencies for the differently processed samples (cast & quenched vs crushed ones), yet they both are anomalous, i.e., there is no linear dependency at more than 18.75 % Al concentration for the crushed samples (and at more than 31.25 % Al concentration for the cast & quenched ones). The origin of this anomaly itself can be successfully attributed to the change in magnetic moments of Fe atoms (see Fig. 4), as was assumed in Ref.<sup>39</sup> and theoretically verified in Ref.<sup>13</sup>, and in this paper.



**Figure 4.** Total magnetic moment of unit cell divided by the number of Fe atoms for different compounds of Fe–Al computed with the ensemble of mMTPs and DFT. Both the mMTPs and DFT predict the decrease in total magnetic moment when the concentration of Al increases.

Compositional dependencies of the total magnetic moment of the 16-atomic unit cell divided by the number of Fe atoms obtained with the ensemble of mMTPs and DFT are shown in Figure 4. The overall agreement between the mMTPs and DFT is very good: we see that the total magnetic moment of the unit cell decreases when the concentration of Al increases. The values of the total magnetic moments obtained with the mMTPs and DFT are also close to each other except for the  $\text{Fe}_8\text{Al}_8$ -12345678 structure: all the fitted mMTPs in the ensemble give zero magnetic moment whereas DFT gives  $0.7 \mu_B$ . Except for this discrepancy for the  $\text{Fe}_8\text{Al}_8$ -12345678 structure, from the results of this subsection we conclude that the ensemble of mMTPs fitted to the DFT data essentially reproduces the variations in formation energies, lattice parameters, and total magnetic moments calculated with DFT.

## Conclusion

In this paper we proposed the machine-learning interatomic potential with magnetic degrees of freedom (magnetic Moment Tensor Potential, mMTP) for prediction the properties of magnetic alloys. This potential was trained on data obtained with the recently developed method of cDFT calculations<sup>19</sup> that allows us to compute energies of configurations with non-equilibrium (excited) magnetic moments and, thus, to consider magnetic moments as degrees of freedom along with atomic positions, atomic types, and lattice vectors. We verify the developed magnetic multi-component machine-learning potentials on the Fe–Al system. We, first, created a training set including fully equilibrium atomic positions, lattice vectors, magnetic moments and the perturbed ones for different concentrations of Fe and Al in the Fe–Al system. Next, we fitted the ensemble of five mMTPs to the cDFT data and compared the dependencies of formation energies, equilibrium lattice parameters, and total magnetic moments of unit cell on the concentration of Al atoms predicted with the ensemble of mMTPs and DFT. We concluded that the mentioned mMTP and DFT differences are minor. Both mMTPs and DFT reproduced anomalous volume-composition dependence in the Fe–Al system obtained theoretically in the previous studies and has been experimentally observed. The main difference between the mMTP and DFT results was found for the  $\text{Fe}_8\text{Al}_8$ -12345678 structure: the ensemble of mMTPs gave the local minimum with zero magnetic moments for the Fe atoms whereas DFT predicted the minimum with magnetic moments of  $0.7 \mu_B$  for the Fe atoms. Nevertheless, the rest of the results obtained with DFT and mMTP are in good correspondence.

In future, we are planning to develop an active learning algorithm for mMTP. Our confidence of developing an efficient active learning algorithm stems from the fact that with cDFT we are able to treat magnetic moments on the same footing as atomic positions; and for learning on atomic positions/geometries various successful active learning algorithms already exist. With active learning, we see in principle no obstacle in applying our methodology to predict material defect properties of other multi-component systems, e.g., of high-entropy alloys. In Ref.<sup>40</sup> we have developed an MTP-based algorithm for computing stacking fault energies, surface energies, and elastic constants, for (non-magnetic) bcc random alloys and used it to predict ductility of Mo–Nb–Ta over the entire composition space. Extending it to the case of magnetic alloys should be straightforward once we have developed an active learning algorithm, and this will allow us to screen for new materials with exotic mechanical properties over a much larger space of (magnetic and non-magnetic) metallic alloys than the space that can currently be approached with today's state-of-the-art methods. Finally, active learning will allow us to train mMTP during molecular dynamics simulations and apply it to investigating the processes and predicting the properties of magnetic materials at finite temperature.



## Methodology

### Magnetic multi-component moment tensor potential (mMTP)

The concept of magnetic multi-component Moment Tensor Potential (mMTP) presented in the current research is based on the previously developed non-magnetic MTP for multi-component systems<sup>41,42</sup> and magnetic MTP for single-component systems<sup>35</sup>.

The mMTP potential is local, i.e., the energy of the atomistic system is a sum of energies of individual atoms:

$$E = \sum_{i=1}^{N_a} E_i, \quad (1)$$

where  $i$  stands for the individual atoms in an  $N_a$ -atom system. We note that any configuration includes lattice vectors  $\mathbf{L} = \{l_1, l_2, l_3\}$ , atomic positions  $\mathbf{R} = \{\mathbf{r}_1, \dots, \mathbf{r}_{N_a}\}$ , types  $Z = \{z_1, \dots, z_{N_a}\}$  (we also denote  $N_{\text{types}}$  by the total number of atomic types in the system), and magnetic moments  $M = \{m_1, \dots, m_{N_a}\}$ . The energy of the atom  $E_i$ , in turn, has the form:

$$E_i = \sum_{\alpha=1}^{\alpha_{\max}} \xi_{\alpha} B_{\alpha}(\mathbf{n}_i), \quad (2)$$

where  $\xi = \{\xi_{\alpha}\}$  are the “linear” parameters to be optimized and  $B_{\alpha}$  are the so-called *basis functions*, which are contractions of the *descriptors*<sup>25</sup> of atomistic environment  $\mathbf{n}_i$ , yielding a scalar. The  $\alpha_{\max}$  parameter can be changed to provide potentials with different amount of parameters<sup>35</sup>.

The descriptors are composed of the radial part, i.e., the scalar function depending on the interatomic distances and atomic magnetic moments, and the angular part, which is a tensor of rank  $\nu$ :

$$M_{\mu,\nu}(\mathbf{n}_i) = \sum_j f_{\mu}(|\mathbf{r}_{ij}|, z_i, z_j, m_i, m_j) \underbrace{\mathbf{r}_{ij} \otimes \dots \otimes \mathbf{r}_{ij}}_{\nu} \text{ times}, \quad (3)$$

where  $\mathbf{n}_i$  stands for the atomic environment, including all the atoms within the  $R_{\text{cut}}$  distance (or less) from the central atom  $i$ ,  $\mu$  is the number of the radial function,  $\nu$  is the rank of the angular part tensor,  $|\mathbf{r}_{ij}|$  is the distance between the atoms  $i$  and  $j$ ,  $z_i$  and  $z_j$  are the atomic types,  $m_i$  and  $m_j$  are the magnetic moments of the atoms.

The radial functions are expanded in a basis of Chebyshev polynomials:

$$f_{\mu}(|\mathbf{r}_{ij}|, z_i, z_j, m_i, m_j) = \sum_{\zeta=1}^{N_{\phi}} \sum_{\beta=1}^{N_{\psi}} \sum_{\gamma=1}^{N_{\psi}} c_{\mu,z_i,z_j}^{\zeta,\beta,\gamma} \phi_{\zeta}(|\mathbf{r}_{ij}|) \psi_{\beta}(m_i) \psi_{\gamma}(m_j) (R_{\text{cut}} - |\mathbf{r}_{ij}|)^2. \quad (4)$$

Here  $\mathbf{c} = \{c_{\mu,z_i,z_j}^{\zeta,\beta,\gamma}\}$  are the “radial” parameters to be optimized, each of the functions  $\phi_{\zeta}(|\mathbf{r}_{ij}|)$ ,  $\psi_{\beta}(m_i)$ ,  $\psi_{\gamma}(m_j)$  is a Chebyshev polynomial of order  $\zeta$ ,  $\beta$  and  $\gamma$  correspondingly, taking values from  $-1$  to  $1$ . The function  $\phi_{\zeta}(|\mathbf{r}_{ij}|)$  yields the dependency on the distance between the atoms  $i$  and  $j$ , while the functions  $\psi_{\beta}(m_i)$  and  $\psi_{\gamma}(m_j)$  yield the dependency on the magnetic moments of the atoms  $i$  and  $j$ , correspondingly. The arguments of the functions  $\phi_{\zeta}(|\mathbf{r}_{ij}|)$  are on the interval  $(R_{\min}, R_{\text{cut}})$ , where  $R_{\min}$  and  $R_{\text{cut}}$  are the minimum and maximum distance, correspondingly, between the interacting atoms. The functions  $\psi_{\beta}(m_i)$  and  $\psi_{\gamma}(m_j)$  are of the same structure, which we explain for the case of the former one. The argument of the function  $\psi_{\beta}(m_i)$  is the magnetic moment of the atom  $i$ , taking the values on the  $(-M_{\max}^{z_i}, M_{\max}^{z_i})$  interval. The value  $M_{\max}^{z_i}$  itself depends on the type of atom  $z_i$ , and is determined as the maximal absolute value of the magnetic moment for atom type  $z_i$  in the training set. Similar to the conventional MTP, the term  $(R_{\text{cut}} - |\mathbf{r}_{ij}|)^2$  provides smooth fading to 0 when approaching the  $R_{\text{cut}}$  distance, in accordance with the locality principle (1).

We note that magnetic degrees of freedom  $m_i$  from (4) are collinear, i.e., they can take negative or positive values as projection onto the  $Z$  axis (though the choice of the axis is arbitrary). This way, in comparison to non-magnetic atomistic systems with  $N$  atoms, in which the amount of degrees of freedom equals  $4N$  (namely  $3N$  for coordinates and  $N$  for types), for the description of magnetic systems additional  $N$  degrees of freedom are introduced, standing for the magnetic moment  $m_i$  of each atom. The amount of parameters entering the radial functions (Eq. 4) also increases in mMTP compared to the conventional MTP<sup>41,42</sup>. Namely, in MTP this number equals  $N_{\mu} \cdot N_{\phi} \cdot N_{\text{types}}^2$ , while in mMTP it is  $N_{\mu} \cdot N_{\phi} \cdot N_{\text{types}} \cdot N_{\psi}^2$ . Thus, if we take  $N_{\psi} = 2$  (which is used in the current research), the amount of the parameters entering the radial functions would be four times more in mMTP than in MTP.

We denote all the mMTP parameters by  $\theta = \{\xi, \mathbf{c}\}$  and the total energy (1) of the atomic system by  $E = E(\theta) = E(\theta; M) = E(\theta; \mathbf{L}, \mathbf{R}, Z, M)$ .

### Magnetic symmetrization of mMTP

The tensor (Eq. (4)) includes collinear magnetic moments in its functional form. However, it is not invariant with respect to the inversion of magnetic moments, i.e.,  $E(\theta; M) \neq E(\theta; -M)$ , while both original and spin-inverted configurations should yield the same energy due to the arbitrary orientation of the projection axis, which we further call the *magnetic symmetry*.

We use data augmentation followed by explicit symmetrization with respect to magnetic moments to train a symmetric mMTP as we discuss below. Assume we have  $K$  configurations in the training set with DFT energies  $E_k^{\text{DFT}}$ , forces  $\mathbf{f}_{i,k}^{\text{DFT}}$ , and stresses  $\sigma_{ab,k}^{\text{DFT}}$  ( $a, b = 1, 2, 3$ ) calculated. We find the optimal parameters  $\bar{\theta}$  (fit mMTP) by minimizing the objective function:

$$\begin{aligned} & \sum_{k=1}^K \left[ w_e \left| \frac{E_k(\boldsymbol{\theta}; M) + E_k(\boldsymbol{\theta}; -M)}{2} - E_k^{\text{DFT}} \right|^2 \right. \\ & + w_f \sum_{i=1}^{N_a} \left| \frac{f_{i,k}(\boldsymbol{\theta}; M) + f_{i,k}(\boldsymbol{\theta}; -M)}{2} - f_{i,k}^{\text{DFT}} \right|^2 \\ & \left. + w_s \sum_{a,b=1}^3 \left| \frac{\sigma_{ab,k}(\boldsymbol{\theta}; M) + \sigma_{ab,k}(\boldsymbol{\theta}; -M)}{2} - \sigma_{ab,k}^{\text{DFT}} \right|^2 \right], \end{aligned} \quad (5)$$

where  $w_e$ ,  $w_f$ , and  $w_s$  are non-negative weights. By minimizing (5) we find such optimal parameters  $\bar{\boldsymbol{\theta}}$  that yield  $E_k(\bar{\boldsymbol{\theta}}; M) \approx E_k(\bar{\boldsymbol{\theta}}; -M)$ ,  $k = 1, \dots, K$  (the same fact takes place for the mMTP forces and stresses), i.e., we symmetrize the training set to make mMTP learn the required symmetry from the data itself—this is called *data augmentation*.

Next, we modify mMTP to make the energy used for the simulations (e.g., relaxation of configurations) to satisfy the exact symmetry:

$$E^{\text{symm}}(\bar{\boldsymbol{\theta}}; M) = \frac{E(\bar{\boldsymbol{\theta}}; M) + E(\bar{\boldsymbol{\theta}}; -M)}{2}. \quad (6)$$

That is, we substitute the mMTP energy (1) into (6) and get a functional form which satisfies the exact identity  $E^{\text{symm}}(\bar{\boldsymbol{\theta}}; M) = E^{\text{symm}}(\bar{\boldsymbol{\theta}}; -M)$  for any configuration. We also note that  $E(\bar{\boldsymbol{\theta}}) \approx E^{\text{symm}}(\bar{\boldsymbol{\theta}})$ .

### Constrained density functional theory calculations

We use the cDFT approach with hard constraints (i.e., Lagrange multiplier) as proposed by Gonze et al. in Ref.<sup>19</sup>. One way to formulate it is to first note that in a single-point DFT calculation we minimize the Kohn-Sham total energy functional  $E[\rho; \mathbf{R}]$  with respect to the electronic density  $\rho = \rho(r)$  (here  $\rho$  combines the spin-up and spin-down electron densities), keeping the nuclei position  $\mathbf{R}$  fixed. In other words, we solve the following minimization problem:

$$E_{\text{DFT}}(\mathbf{R}) = \min_{\rho} E[\rho; \mathbf{R}],$$

and from the optimal  $\rho^* = \arg \min E[\rho; \mathbf{R}]$  we can, e.g., find magnetization  $m(r) = \rho_+^* - \rho_-^*$ , where the subscripts denote the spin-up (+) and spin-down (−) densities. The magnetic moment of the  $i$ th atom can be found by integrating  $m(r)$  over some (depending on the partitioning scheme) region around the atom:

$$m_i = \int_{\Omega_i} m(r) dr. \quad (7)$$

Since the minimizer  $\rho^*$  depends on  $\mathbf{R}$ ,  $m_i$  are also the functions of  $\mathbf{R}$ .

According to the cDFT approach<sup>19</sup>, we now formulate the problem of minimizing  $E[\rho; \mathbf{R}]$  in which not only  $\mathbf{R}$ , but also  $\rho$  is allowed to change only subject to constraints (7):

$$\begin{aligned} E_{\text{cDFT}}(\rho, \mathbf{R}, M) &= \min_{\rho} E[\rho; \mathbf{R}] \\ &\text{subject to } m_i = \int_{\Omega_i} (\rho_+(r) - \rho_-(r)) dr. \end{aligned}$$

The algorithmic details of how this minimization problem is solved, and how the energy derivatives (forces, stresses, torques) are computed, are described in detail in Ref.<sup>19</sup>.

### Computational details

We used the ABINIT code<sup>43,44</sup> for DFT (and cDFT recently developed and described in Ref.<sup>19</sup>) calculations with  $6 \times 6 \times 6$  k-point mesh and cutoff energy of 25 Hartree (about 680 eV). We utilized the PAW PBE method with the generalized gradient approximation. We applied constraints on magnetic moments of all atoms during cDFT calculations.

We fitted an ensemble of five mMTPs with 415 parameters in order to quantify the uncertainty of mMTPs predictions. For each mMTP we took  $R_{\text{min}} = 2.1 \text{ \AA}$ ,  $R_{\text{cut}} = 4.5 \text{ \AA}$ ,  $M_{\text{max}}^{\text{Al}} = 0.1 \mu_B$ , and  $M_{\text{max}}^{\text{Fe}} = 3.0 \mu_B$ . The weights in the objective function (5) were  $w_e = 1$ ,  $w_f = 0.01 \text{ \AA}^2$ , and  $w_s = 0.001$ .

### Data availability

We published the training set at ([https://gitlab.com/ivannovikov/datasets\\_for\\_magnetic\\_MTP/-/blob/main/training\\_set.cfg](https://gitlab.com/ivannovikov/datasets_for_magnetic_MTP/-/blob/main/training_set.cfg)) and the “verification” set at ([https://gitlab.com/ivannovikov/datasets\\_for\\_magnetic\\_MTP/-/blob/main/verification\\_set.cfg](https://gitlab.com/ivannovikov/datasets_for_magnetic_MTP/-/blob/main/verification_set.cfg)).

Received: 13 July 2023; Accepted: 7 November 2023

Published online: 13 November 2023

### References

1. Ruban, A. V. & Razumovskiy, V. I. Spin-wave method for the total energy of paramagnetic state. *Phys. Rev. B* **85**, 174407 (2012).



2. Körmann, F., Dick, A., Grabowski, B., Hickel, T. & Neugebauer, J. Atomic forces at finite magnetic temperatures: Phonons in paramagnetic iron. *Phys. Rev. B* **85**, 125104 (2012).
3. Ikeda, Y., Seko, A., Togo, A. & Tanaka, I. Phonon softening in paramagnetic bcc Fe and its relationship to the pressure-induced phase transition. *Phys. Rev. B* **90**, 134106 (2014).
4. Gorbатов, O., Korzhavyi, P. A., Ruban, A. V., Johansson, B. & Gornostyrev, Y. N. Vacancy–solute interactions in ferromagnetic and paramagnetic bcc iron: Ab initio calculations. *J. Nucl. Mater.* **419**, 248–255 (2011).
5. Bienvenu, B., Fu, C. C. & Clouet, E. Impact of magnetism on screw dislocations in body-centered cubic chromium. *Acta Mater.* **200**, 570–580 (2020).
6. Schneider, A., Fu, C.-C., Soisson, F. & Barreteau, C. Atomic diffusion in  $\alpha$ -iron across the curie point: An efficient and transferable ab initio-based modeling approach. *Phys. Rev. Lett.* **124**, 215901 (2020).
7. Yang, Y. *et al.* Bifunctional nanoprecipitates strengthen and ductilize a medium-entropy alloy. *Nature* **595**, 245–249 (2021).
8. Körmann, F., Hickel, T. & Neugebauer, J. Influence of magnetic excitations on the phase stability of metals and steels. *Curr. Opin. Solid State Mater. Sci.* **20**, 77–84 (2016).
9. Herper, H., Hoffmann, E. & Entel, P. Ab initio full-potential study of the structural and magnetic phase stability of iron. *Phys. Rev. B* **60**, 3839 (1999).
10. Hasegawa, H. & Pettifor, D. Microscopic theory of the temperature-pressure phase diagram of iron. *Phys. Rev. Lett.* **50**, 130 (1983).
11. Song, Y., Shi, N., Deng, S., Xing, X. & Chen, J. Negative thermal expansion in magnetic materials. *Prog. Mater. Sci.* **121**, 100835 (2021).
12. Lu, H. *et al.* Effects of Fe doping on structure, negative thermal expansion, and magnetic properties of antiperovskite  $\text{mn}_3\text{gn}$  compounds. *J. Am. Ceram. Soc.* (2023).
13. Friák, M. & Neugebauer, J. Ab initio study of the anomalous volume-composition dependence in Fe–Al alloys. *Intermetallics* **18**, 1316–1321 (2010).
14. Butt, M. K. *et al.* Structural, electronic, half-metallic ferromagnetic and optical properties of cubic  $\text{MALO}_3$  (M= Ce, Pr) perovskites: A DFT study. *J. Phys. Chem. Solids* **154**, 110084 (2021).
15. Moutassime, M. *et al.* Magnetic properties and half metallic behavior of the full-Heusler  $\text{Co}_2\text{FeGe}$  alloy: DFT and Monte Carlo studies. *J. Solid State Chem.* **304**, 122534 (2021).
16. Wu, Q. & Van Voorhis, T. Constrained density functional theory and its application in long-range electron transfer. *J. Chem. Theory Comput.* **2**, 765–774 (2006).
17. Ghosh, P. & Gebauer, R. Computational approaches to charge transfer excitations in a zinc tetraphenylporphyrin and c 70 complex. *J. Chem. Phys.* **132**, 104102 (2010).
18. Ma, P.-W. & Dudarev, S. L. Constrained density functional for noncollinear magnetism. *Phys. Rev. B* **91**, 054420 (2015).
19. Gonze, X., Seddon, B., Elliott, J. A., Tantardini, C. & Shapeev, A. V. Constrained density functional theory: A potential-based self-consistency approach. *J. Chem. Theory Comput.* **18**, 6099–6110 (2022).
20. Li, K., Fu, C.-C., Nastar, M., Soisson, F. & Lavrentiev, M. Y. Magnetochemical effects on phase stability and vacancy formation in fcc Fe–Ni alloys. *Phys. Rev. B* **106**, 024106 (2022).
21. Li, K., Fu, C.-C., Nastar, M. & Soisson, F. Predicting atomic diffusion in concentrated magnetic alloys: The case of paramagnetic Fe–Ni. *Phys. Rev. B* **107**, 094103 (2023).
22. Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401. <https://doi.org/10.1103/PhysRevLett.98.146401> (2007).
23. Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403. <https://doi.org/10.1103/PhysRevLett.104.136403> (2010).
24. Thompson, A., Swiler, L., Trott, C., Foiles, S. & Tucker, G. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *J. Comput. Phys.* **285**, 316–330. <https://doi.org/10.1016/j.jcp.2014.12.018> (2015).
25. Shapeev, A. V. Moment tensor potentials: A class of systematically improvable interatomic potentials. *Multiscale Model. Simul.* **14**, 1153–1173 (2016).
26. Schütt, K. T. *et al.* SchNet: A continuous-filter convolutional neural network for modeling quantum interactions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17*, 992–1002 (Curran Associates Inc., 2017) (event-place: Long Beach, California, USA).
27. Smith, J. S., Isayev, O. & Roitberg, A. E. ANI-1: An extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **8**, 3192–3203. <https://doi.org/10.1039/C6SC05720A> (2017).
28. Wang, H., Zhang, L., Han, J. & Weinan, E. Deepmd-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* **228**, 178–184 (2018).
29. Pun, G. P. P., Batra, R., Ramprasad, R. & Mishin, Y. Physically informed artificial neural networks for atomistic modeling of materials. *Nat. Commun.* **10**, 2339. <https://doi.org/10.1038/s41467-019-10343-5> (2019).
30. Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **99**, 014104. <https://doi.org/10.1103/PhysRevB.99.014104> (2019).
31. Takamoto, S., Izumi, S. & Li, J. TeaNet: Universal neural network interatomic potential inspired by iterative electronic relaxations. *Comput. Mater. Sci.* **207**, 111280. <https://doi.org/10.1016/j.commatsci.2022.111280> (2022).
32. Batzner, S. *et al.* E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453. <https://doi.org/10.1038/s41467-022-29939-5> (2022).
33. Drautz, R. Atomic cluster expansion of scalar, vectorial, and tensorial properties including magnetism and charge transfer. *Phys. Rev. B* **102**, 024104. <https://doi.org/10.1103/PhysRevB.102.024104> (2020).
34. Nikolov, S. *et al.* Data-driven magneto-elastic predictions with scalable classical spin-lattice dynamics. *npj Comput. Mater.* **7**, 153. <https://doi.org/10.1038/s41524-021-00617-2> (2021).
35. Novikov, I., Grabowski, B., Körmann, F. & Shapeev, A. Magnetic moment tensor potentials for collinear spin-polarized materials reproduce different magnetic states of bcc Fe. *npj Comput. Mater.* **8**, 13 (2022).
36. Domina, M., Cobelli, M. & Sanvito, S. Spectral neighbor representation for vector fields: Machine learning potentials including spin. *Phys. Rev. B* **105**, 214439. <https://doi.org/10.1103/PhysRevB.105.214439> (2022).
37. Yu, H., Zhong, Y., Ji, J., Gong, X. & Xiang, H. Time-reversal equivariant neural network potential and Hamiltonian for magnetic materials. arXiv preprint [arXiv:2211.11403](https://arxiv.org/abs/2211.11403) (2022).
38. Rinaldi, M., Mrovec, M., Bochkarev, A., Lysogorskiy, Y. & Drautz, R. Non-collinear magnetic atomic cluster expansion for iron. arXiv preprint [arXiv:2305.15137](https://arxiv.org/abs/2305.15137) (2023).
39. Taylor, A. & Jones, R. M. Constitution and magnetic properties of iron-rich iron–aluminum alloys. *J. Phys. Chem. Solids* **6**, 16–37 (1958).
40. Novikov, I., Kovalyova, O., Shapeev, A. & Hodapp, M. AI-accelerated materials informatics method for the discovery of ductile alloys. *J. Mater. Res.* **37**, 3491–3504. <https://doi.org/10.1557/s43578-022-00783-z> (2022).
41. Gubaev, K., Podryabinkin, E. V. & Shapeev, A. V. Machine learning of molecular properties: Locality and active learning. *J. Chem. Phys.* **148**, 241727 (2018).
42. Gubaev, K., Podryabinkin, E. V., Hart, G. L. & Shapeev, A. V. Accelerating high-throughput searches for new alloys with active learning of interatomic potentials. *Comput. Mater. Sci.* **156**, 148–156 (2019).
43. Gonze, X. *et al.* The abinit project: Impact, environment and recent developments. *Comput. Phys. Commun.* **248**, 107042 (2020).

44. Romero, A. H. *et al.* Abinit: Overview and focus on selected capabilities. *J. Chem. Phys.* **152**, 124102 (2020).

## Acknowledgements

The authors would like to thank Professor Dr. Xavier Gonze of UCLouvain for useful discussion. Ch.T performed this work under a state task to Institute of Solid State Chemistry and Mechanochemistry of the Siberian Branch of the Russian Academy of Sciences project number 121032500059-4. Ch. T. also thanks the Siberian Super-Computer Center of the Siberian Branch of the Russian Academy of Sciences for computational resources. M.H. gratefully acknowledges the financial support under the scope of the COMET program within the K2 Center “Integrated Computational Material, Process and Product Engineering (IC-MPPE)” (Project No 886385). This program is supported by the Austrian Federal Ministries for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK) and for Labour and Economy (BMAW), represented by the Austrian Research Promotion Agency (FFG), and the federal states of Styria, Upper Austria and Tyrol. This work was supported by Russian Science Foundation (grant number 22-73-10206, <https://rscf.ru/project/22-73-10206/>).

## Author contributions

A.S.K. wrote the first draft and performed the work. M.H., K.G. and A.V.S. participated to the development of machine learning model. Ch.T. participated to the work of cDFT with ABINIT. I.S.N. supervised the entire work. All the authors reviewed the manuscript.

## Funding

Open access funding provided by UiT The Arctic University of Norway (incl University Hospital of North Norway).

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to C.T. or I.S.N.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023