REVIEW ARTICLE   OPEN

Check for updates

Magnetic Moment Tensor Potentials for collinear spin-polarized materials reproduce different magnetic states of bcc Fe

Ivan Novikov [ID][1,2✉], Blazej Grabowski [ID][2], Fritz Körmann[3,4] and Alexander Shapeev[1]

We present the magnetic Moment Tensor Potentials (mMTPs), a class of machine-learning interatomic potentials, accurately reproducing both vibrational and magnetic degrees of freedom as provided, e.g., from first-principles calculations. The accuracy is achieved by a two-step minimization scheme that coarse-grains the atomic and the spin space. The performance of the mMTPs is demonstrated for the prototype magnetic system bcc iron, with applications to phonon calculations for different magnetic states, and molecular-dynamics simulations with fluctuating magnetic moments.

INTRODUCTION

Magnetic contributions are essential for modeling magnetic materials as they critically affect phase stability[1–3], vibrational properties[4–6], interstitial energies[7], local[8,9] and extended defects[10,11], and kinetics[12,13]. Taking the magnetic degrees of freedom properly into account is a prerequisite for computationally-aided design and development of a large number of technologically relevant materials, ranging from various steels for construction and safety applications[1–6,8–13] to hard magnets for applications in electrical transportation and renewable energy technologies[14,15].

One of the most popular computational methods, capable of capturing magnetism, are first-principles calculations realized by density functional theory (DFT). DFT calculations are, however, computationally expensive and limited to small system sizes and to a small number of magnetic configurations. DFT calculations that sample the magnetic degree of freedom explicitly, as needed for, e.g., lattice vibrations or vacancy formation energies in magnetically excited states, are therefore available only for very few selected cases.

Recent progress in machine-learning potentials has significantly accelerated accurate simulations of materials and molecules[16–26]. Such potentials express the interatomic energy as a function of atomic positions alone. Ignoring the electronic degrees of freedom, yet assuming a very flexible functional form for the interatomic energies, machine-learning potentials feature near-quantum mechanical accuracy at a computational efficiency of the order of classical interatomic potentials[27]. However, by ignoring the electronic degrees of freedom such potentials cannot distinguish different magnetic states, simply because different magnetic states feature different energies and the functional form of machine-learning potentials prohibits capturing such a magnetically-induced energy variation. In this paper, we introduce a strategy to overcome this fundamental shortcoming.

RESULTS

Magnetic moment tensor potential

The starting point is a given set of energies, $E^{DFT}(R, S)$, which include the magnetic degree of freedom, e.g., computed via DFT, and where R and S denote a set of atomic coordinates and corresponding atomic spins. There are various ways to compute these energies from DFT, e.g., via fully relaxing the spin degree of freedom or, if one is interested in a broader sampling of $E^{DFT}(R, S)$, via constrained spin calculations[28–32]. We on purpose do not discuss in this work the different approaches available and their corresponding challenges to computing $E^{DFT}(R, S)$ since the main focus here is on an efficient parametrization for a given $E^{DFT}(R, S)$. We utilize standard spin-polarized DFT calculations where the local atomic moments are differently initialized while their longitudinal component is fully relaxed. The different magnetic configurations sampled are discussed further below. We note, however, that the proposed machine-learning potential can be straightforwardly applied with, e.g., constrained spin calculations.

The heart of the proposed approach is to approximate the energy $E^{DFT}(R, S)$ with Moment Tensor Potentials (MTPs)[33,34] the idea of which is to expand the energy locally as a polynomial of its degrees of freedom, corrected in order to allow for a finite cutoff of the potential. We note that there are other functional forms allowing for approximation of E^{DFT} as a function of enriched degrees of freedom[35–37]. A similar functional form as utilized in MTPs has been recently employed within the atomic cluster expansion (ACE)[38]. Both approaches feature a complete basis of invariant polynomials that differ only in the representation of the angular terms; MTP uses tensors while ACE uses spherical harmonics.

In our approach the total interaction energy is partitioned into contributions of individual local atomic environments:

$$E^{mMTP} = \sum_{i=1}^{N} V(\mathfrak{n}_i), \tag{1}$$

where $\mathfrak{n}_i$ is the neighborhood of the i'th atom and N is the number of atoms in the atomic configuration. In the present paper the

[1]Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Nobel St. 3, Moscow 143026, Russia. [2]Institute for Materials Science, University of Stuttgart, Pfaffenwaldring 55, 70569 Stuttgart, Germany. [3]Computational Materials Design, Max-Planck-Institut für Eisenforschung GmbH, D-40237 Düsseldorf, Germany. [4]Materials Science and Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands. ✉email: i.novikov@skoltech.ru

degrees of freedom are atomic positions $\mathbf{R} = \{\mathbf{r}_i, i = 1, \ldots, N\}$ and spins $S = \{s_i, i = 1, \ldots, N\}$ as opposed to the originally developed MTPs[33,34] in which the potential energy depends only on atomic positions. The atomic neighborhood of the i'th atom, $\mathfrak{n}_i$, is hence described by the relative interatomic positions $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$, the spin of the central atom, s_i, and the spins of the neighboring atoms s_j, formally

$$\mathfrak{n}_i = \{(\mathbf{r}_{ij}, s_i, s_j) : j = 1, \ldots, N_{nb}^i\},$$

where N_{nb}^i is the number of neighbors of the i'th atom.

The expansion of the function V is:

$$V(\mathfrak{n}_i) = \sum_a \xi_a B_a(\mathfrak{n}_i),$$

where $\boldsymbol{\xi} = \{\xi_a\}$ are the "linear" parameters to be optimized. The function V is assumed to be an arbitrary polynomial of the corresponding degrees of freedom, modified so that instead of the polynomial growth the potential V vanishes beyond some cutoff distance. The potential is expanded via basis functions B_a defined through the so-called moment tensor descriptors

$$M_{\mu,\nu}(\mathfrak{n}_i) = \sum_{j=1}^{N_{nb}^i} f_\mu(|\mathbf{r}_{ij}|, s_i, s_j) \underbrace{\mathbf{r}_{ij} \otimes \ldots \otimes \mathbf{r}_{ij}}_{\nu \text{ times}}, \qquad (2)$$

where "$\otimes$" is the outer product of vectors, and, thus, the angular part $\mathbf{r}_{ij} \otimes \ldots \otimes \mathbf{r}_{ij}$ is a tensor of ν'th rank. The function $f_\mu(|\mathbf{r}_{ij}|, s_i, s_j)$ is a polynomial of $|\mathbf{r}_{ij}|$, s_i and s_j, modified for a finite cutoff radius. It has the form:

$$f_\mu(|\mathbf{r}_{ij}|, s_i, s_j) = \sum_{\zeta=1}^{N_\varphi} \sum_{\gamma=1}^{N_\psi} \sum_{\beta=1}^{N_\psi} c_\mu^{\beta,\gamma,\zeta} \psi_\beta(s_i) \psi_\gamma(s_j) \varphi_\zeta(|\mathbf{r}_{ij}|)(r_{cut} - |\mathbf{r}_{ij}|)^2, \qquad (3)$$

where $\mathbf{c} = \{c_\mu^{\beta,\gamma,\zeta}\}$ are the "radial" parameters to be optimized, N_φ is the number of polynomial basis functions $\varphi_\zeta(|\mathbf{r}_{ij}|)$ on the interval $[r_{min}, r_{cut}]$, where r_{min} is the minimal distance between atoms and r_{cut} is the cutoff radius beyond which atoms do not interact. The term $(r_{cut} - |\mathbf{r}_{ij}|)^2$ ensures a smooth vanishing of the potential for $|\mathbf{r}_{ij}| > r_{cut}$. The other functions, $\psi_\beta(s_i)$ and $\psi_\gamma(s_j)$, are the polynomial basis functions of the local spins of the central and neighboring atoms, respectively. The number of these spin basis functions is N_ψ. They are defined on the interval $[s_{min}, s_{max}]$, where the values s_{min} and s_{max} are the minimal and maximal local magnetic moments in the system being investigated.

The mMTP basis functions B_a are defined as all possible contractions of $M_{\mu,\nu}(\mathfrak{n}_i)$ to a scalar, e.g.,

$$M_{1,0}(\mathfrak{n}_i), \; M_{0,1}(\mathfrak{n}_i) \cdot M_{1,1}(\mathfrak{n}_i), \; M_{3,2}(\mathfrak{n}_i) : M_{1,2}(\mathfrak{n}_i), \; \ldots,$$

where "$\cdot$" is the dot product of two vectors, and ":" is the Frobenius product of two matrices. In principle, an infinite number of such mMTP basis functions could be constructed. In order to choose which basis functions to include in practice in the mMTP, we introduce the so-called level of each descriptor, $\text{lev}M_{\mu,\nu} = 2 + 4\mu + \nu$, choose a certain lev_{max}, and include in the mMTP each basis function with $\text{lev}B_a \leq \text{lev}_{max}$ (see Novikov et al.[39] for details). Thus, the number of the "linear" parameters $\boldsymbol{\xi}$ depends on lev_{max}, which also determines the number of radial functions, N_μ. The number of the "radial" parameters $\mathbf{c}$ is equal to $N_\mu N_\varphi N_\psi^2$. We denote all free parameters of an mMTP collectively by $\boldsymbol{\theta} = \{\boldsymbol{\xi}, \mathbf{c}\}$, and the total interaction energy by $E^{mMTP} = E^{mMTP}(\boldsymbol{\theta}; \mathbf{R}, S)$.

We note that the mMTP formalism contains the Heisenberg model as a special, limiting case. In particular, first-degree polynomials have to be utilized for $\psi_\beta(s) = \psi_\gamma(s) = s$ in Eq. (3), and φ_ζ needs to "encompass" (i.e., be nonzero at) the nearest neighbors only. Such a choice of terms in the expansion Eq. (3) also leads to a model similar to the one proposed in Nikolov et al.[37], except that in the latter case the full vectorial spins were

considered. Moreover, the biquadratic terms, $(s_i s_j)^2$[40,41], adopted by Nikolov et al.[37], arise naturally when $M_{\mu,0}$ is constructed with such choices of ψ_β and φ_ζ and gets multiplied by itself. Then the radial parameters $c_\mu^{\beta,\gamma,\zeta}$ correspond to the coupling constants as obtained from DFT data.

The free parameters $\boldsymbol{\theta}$ in our approach are found by fitting E^{mMTP} to DFT data. We consider a training set including K magnetic configurations $(\mathbf{R}^{(k)}, S^{(k)})$ with known DFT energies E^{DFT}, DFT forces $\mathbf{f}_i^{DFT}$ on every atom i, and a 3×3 tensor of DFT stresses σ^{DFT} and minimize the objective function:

$$\sum_{k=1}^K \left[w_e |E^{mMTP}(\boldsymbol{\theta}; \mathbf{R}^{(k)}, S^{(k)}) - E^{DFT}(\mathbf{R}^{(k)}, S^{(k)})|^2 \right.$$
$$+ w_f \sum_i |\mathbf{f}_i^{mMTP}(\boldsymbol{\theta}; \mathbf{R}^{(k)}, S^{(k)}) - \mathbf{f}_i^{DFT}(\mathbf{R}^{(k)}, S^{(k)})|^2$$
$$\left. + w_s |\sigma^{mMTP}(\boldsymbol{\theta}; \mathbf{R}^{(k)}, S^{(k)}) - \sigma^{DFT}(\mathbf{R}^{(k)}, S^{(k)})|^2 \right]$$

where $|\cdot|$ is the length of a vector or the Frobenius norm of a matrix. The optimization of the parameters is carried out using an iterative quasi-Newton optimization method, specifically, the Broyden-Fletcher-Goldfarb-Shanno algorithm (BFGS) starting with a random initial guess. As opposed to mMTP, the energy of the non-magnetic MTP, proposed in our earlier works, does not depend on spins, i.e. $E^{MTP} = E^{MTP}(\boldsymbol{\theta}; \mathbf{R})$, and the functions $f_\mu(|\mathbf{r}_{ij}|)$ do not include spins.

Convergence of magnetic MTP

We first analyze the convergence behavior of the magnetic and non-magnetic MTP toward DFT energies as the number of parameters is increased. The convergence was measured on a hold-out set of about 1000 configurations not participating in the fitting of the potentials. Figure 1 shows that the mMTP exhibits a steady convergence, while the non-magnetic MTP does not. This reiterates our original motivation: the space of atomic positions ($\mathbf{R}$) is not the right one for approximating the quantum-mechanical energy, but enriched with spins, ($\mathbf{R}, S$), this becomes a suitable space for that purpose.

Based on the convergence tests, we have chosen a well converged $\text{lev}_{max} = 24$ for the subsequent tests. For both MTP and mMTP we took $N_\varphi = 12$ polynomial functions of the atomic positions with $r_{min} = 2$ Å, $r_{cut} = 5.5$ Å. For the mMTP we took $N_\psi = 2$ polynomial functions of the local magnetic moments with $s_{min} = -3.5\,\mu_B$ and $s_{max} = 3.5\,\mu_B$. The total number of MTP parameters was 937 while that of mMTP was 1153. The weights in the objective function were $w_e = 1$, $w_f = 0.01$, and $w_s = 0.001$.

For each model, we fitted five potentials and selected the best (with the least training error). The validation root-mean-square errors are shown in Table 1. We can see that adding local
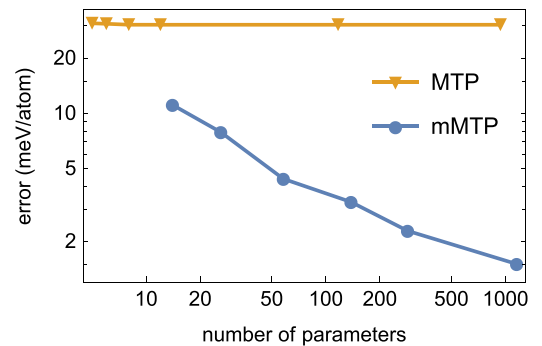

Fig. 1 Convergence of the magnetic potential, mMTP, with respect to DFT energies and lack of convergence for the non-magnetic MTP. The graph indicates that most variation of the energy on the training set is in the magnetic degrees of freedom that can be captured only by the magnetic potential.

Table 1. The best non-magnetic and magnetic MTPs.

Model	Level	Number of parameters	Energy error meV/atom	Force error meV/Å (%)	Stress error GPa (%)
MTP	24	937	30.4	195 (27.3)	0.542 (13.3)
mMTP	24	1153	1.5	64 (9.0)	0.087 (2.2)

Despite the fact that the magnetic MTP has only a small increase in the number of parameters, its accuracy is much higher than that of the non-magnetic MTP.
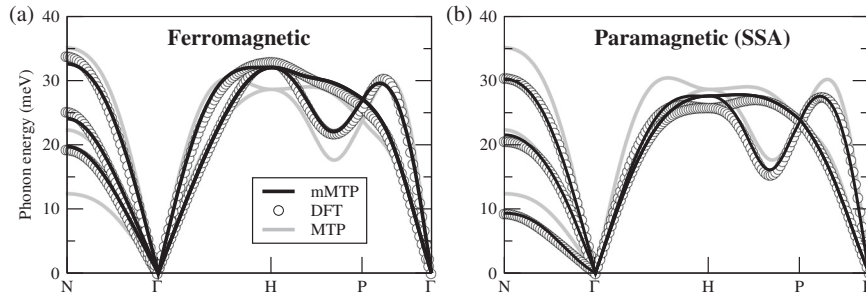

Fig. 2 Phonon spectra computed for different magnetic states. Magnetic MTP reproduces spectra both for **a** the ferromagnetic and **b** paramagnetic state (modeled using the SSA approach) based on the DFT calculations, whereas non-magnetic MTP cannot distinguish the two states and produces results between the ferromagnetic and paramagnetic ones.

magnetic moments to the potential as additional degrees of freedom does not significantly increase the number of parameters, but greatly improves the accuracy of training.

Phonon spectra prediction

We next evaluate the performance of the best optimized MTP and mMTP potentials to predict phonon spectra of different magnetic states. We consider two extreme scenarios representing the limits of magnetic configurations, namely the ferromagnetic state, in which all spins are aligned parallel and a paramagnetic state, treated in the adiabatic limit of fast fluctuating spins. Since the phonon energies were derived from small perturbations (utilizing the small displacement method), this test is a very sensitive measure to detect how well even very small variations in interatomic forces can be captured. The results for the ferromagnetic case for both potentials are shown in Fig. 2a in comparison with the data directly obtained from DFT. The agreement between the mMTP and the DFT data is excellent whereas the non-magnetic MTP shows significant deviations, in particular around the N-point. The deviations for the non-magnetic MTP are a direct consequence of the training database which also includes magnetically disordered configurations responsible for pronounced phonon softening as discussed in the following.

To compute the phonon spectra in the paramagnetic regime we utilize the spin-space averaging (SSA) method[6]. In this approach, effective interatomic forces can be defined by averaging over various disordered magnetic configurations weighted by a Boltzman distribution. For the actual averaging we utilized the crystal symmetries as proposed in refs. [5,6] and performed the SSA using a single random magnetic configuration for which each atom is displaced in each cartesian direction. This provides a large number of locally inequivalent magnetic configurations (i.e., $54 \cdot 3 = 162$ configurations for the employed supercell). This procedure was shown to be robust with respect to the actually chosen random magnetic configuration as discussed in Körmann et al.[6].

The resulting DFT-based phonon spectrum shown in Fig. 2b features a pronounced softening at the N-point[6]. This softening is related to the decrease of the elastic constants and constitutes an important precursor of the structural transformation in iron. The non-magnetic MTP cannot distinguish the underlying atomic

forces in these different magnetic states from the ferromagnetic forces. This is the reason why the MTP phonon spectrum for the paramagnetic state shown in Fig. 2b is exactly the same as the one in Fig. 2a for the ferromagnetic state. The non-magnetic MTP spectra fall in-between the ferromagnetic and paramagnetic solutions and hence do not quantitatively reproduce the DFT data in either regime. In contrast, applying the SSA approach with the mMTP reveals an excellent agreement with the DFT data, reproducing quantitatively important characteristics such as, e.g., the decrease of the phonon energies near the N-point and along the H-P path.

Disordered-local-moment molecular-dynamics simulations

To evaluate the performance of the mMTP at finite temperatures and larger atomic displacements, we have performed molecular dynamics (MD) simulations. The temperature was set to 800 K and the lattice constant to 2.9 Å. To sample not only the vibrational degrees of freedom but the spin space and in particular the coupling between vibrations and spins, we have performed disordered-local-moment MD (DLM-MD) simulations[42]. Further, in order to explicitly validate the mMTP against DFT, we have utilized the concept of thermodynamic integration, similarly as used in the TU-TILD+MTP method previously[43]. Specifically, we have introduced a linear coupling between DFT and mMTP forces,

$$F_\lambda = \lambda F^{DFT} + (1 - \lambda)F^{mMTP}, \tag{4}$$

with the coupling constant λ and DFT and mMTP forces F^{DFT} and F^{mMTP}. The coupled forces F_λ were used for evolving the DLM-MD trajectories. The mMTP in this test was fitted to 'pure' DFT calculations (i.e., nominally corresponding to $\lambda = 1$) and tested independently for a new set of calculations at $\lambda = 0, 0.5, 1$. To render the DFT calculations feasible we employed a 16-atom supercell for the DLM-TI calculations; cross-checks for a 54-atom supercell showed similar results. Further details are given in the Methods section.

Figure 3 highlights the excellent performance of the mMTP. In the left panel, we observe that the mMTP energies fall almost on top of the DFT energies; the root-mean-square error (RMSE) is only 2.0 meV/atom—of the same order as obtained previously for non-magnetic systems[43]. The middle panel clarifies that the best possible non-magnetic MTP is almost an order of magnitude away
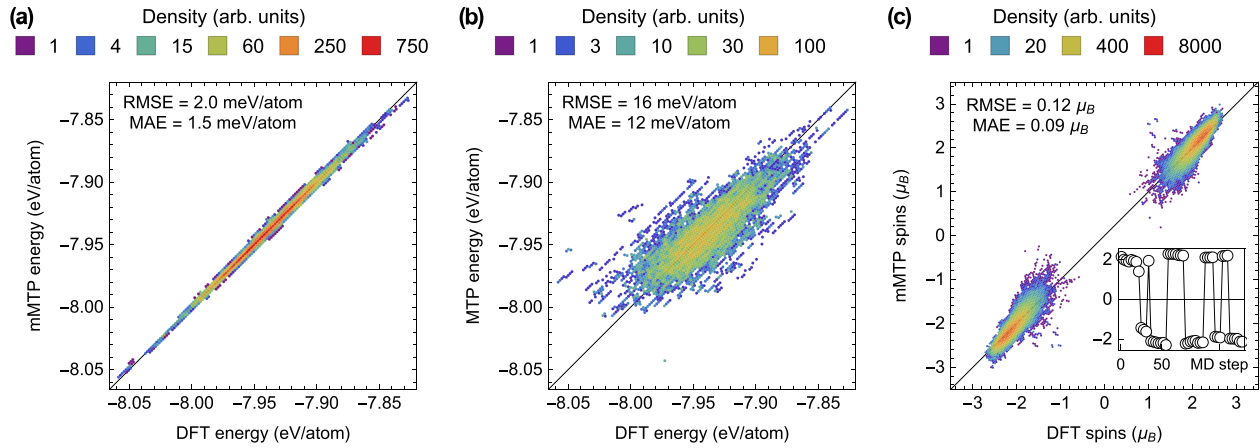
Fig. 3 Energy and spin correlations between different models obtained from DLM-TI calculations at 800 K. It demonstrates that the magnetic MTP predicts both **a** DFT energies, and **c** spins with high accuracy, whereas **b** non-magnetic MTP is almost an order of magnitude worse in accuracy.

in terms of energy accuracy, with an RMSE of 16 meV/atom. The right panel of Fig. 3 shows the spin correlation between mMTP and DFT, which of course only the magnetic version of the MTP is capable to reproduce. We observe an RMSE of 0.12 μ_B, which is ~5% of the magnitude of the absolute spin.

We stress that Fig. 3 includes values for all the investigated coupling constants $\lambda = 0.0, 0.5, 1.0$. Looking at each λ value separately, the correlations are in fact very similar. This means that there is no difference in the correlation, if we use pure DFT forces (cf. Eq. (4)), pure mMTP forces, or DFT-mMTP coupled forces to evolve the MD. This hence allows one to perform a full thermodynamic integration from the mMTP to DFT and compute the respective free energy difference, which is however beyond the scope of the present work.

DISCUSSION

We have developed the mMTPs, a class of magnetic machine-learning interatomic potentials capable of simultaneously and accurately approximating spin and atomic degrees of freedom. This has been achieved by utilizing a two-step minimization scheme for the spin and atomic configurational space. Applying the mMTP to DFT-derived data for the prototypical bcc iron system reveals that the mMTPs are capable to quantitatively approximate local magnetic moments, energies, and forces for various magnetic states (see the Supplementary Material for further tests). A number of applications such as the computation of phonon spectra in ferro- and paramagnetic states, as well as MD simulations including spin-flips, demonstrate that mMTPs provide near DFT accuracy without significantly losing the computational efficiency of classical interatomic potentials.

METHODS

Derivation of mMTP

Here we derive the form of the MTP as a function of relative atomic positions $\mathbf{r}_{ij}$ and vectorial magnetic moments, $\mathbf{s}_i$ and $\mathbf{s}_j$. Following the logic of the original paper introducing the MTP[33], an arbitrary polynomial of the positions and magnetic moments can be represented as all possible contractions of the following Moment Tensors,

$$M_{\zeta,\nu,\beta,\gamma,\xi,\eta} = \sum_{j=1}^{N_{nb}^i} Q_\zeta(|\mathbf{r}_{ij}|)|\mathbf{s}_i|^\beta|\mathbf{s}_j|^\gamma \ (\mathbf{r}_{ij}^{\otimes \nu}) \otimes (\mathbf{s}_i^{\otimes \xi}) \otimes (\mathbf{s}_j^{\otimes \eta}), \quad (5)$$

where Q_ζ is the ζ-th radial basis function and by definition

$$\mathbf{v}^{\otimes n} = \underbrace{\mathbf{v} \otimes ... \otimes \mathbf{v}}_{n \text{ times}}$$

for an arbitrary vector $\mathbf{v}$. We note that Eq. (5) directly corresponds to ref. [36], Eq. 26 with spherical harmonics $\mathbf{Y}_l^m(\mathbf{v}/|\mathbf{v}|)$ instead of tensors $\mathbf{v}^{\otimes n}$ ($|l| \leq m \leq n$).

We next assume scalar-valued spins, i.e., that it is sufficient to consider $\xi = \eta = 0$ (and adsorb the sign of the spin into the radial part):

$$M_{\zeta,\nu,\beta,\gamma} = \sum_{j=1}^{N_{nb}^i} Q_\zeta(|\mathbf{r}_{ij}|)s_i^\beta s_j^\gamma \ (\mathbf{r}_{ij}^{\otimes \nu}). \quad (6)$$

We could choose to directly expand the energy over different contractions of tensors $M_{\zeta,\nu,\beta,\gamma}$, but instead we combine different products of radial and spin basis functions, $Q_\zeta(|\mathbf{r}_{ij}|)s_i^\beta s_j^\gamma$, into the functions $f_\mu(\mathbf{r}_{ij}, s_i, s_j)$ with coefficients $c_\mu^{\beta,\gamma,\zeta}$ that are found from data, as explained in the main text of the manuscript. Note that in this work we use Chebyshev polynomials $\psi_\beta(s)$ instead of monomials s^β.

DFT details

All DFT calculations were performed with VASP[44–47] utilizing the projector augmented wave (PAW) method[48] and the generalized gradient approximation[49]. For the training set of the 54-atom supercell, we considered 70 atomic configurations generated from an initial ferromagnetic MD at 1000 K. For each of these atomic configurations 200 different arrangements of magnetic spins have been initialized of which 67% converged under the high cutoff energy of 500 eV and k-point density of 11,664 k-points × atoms ($6 \times 6 \times 6$ grid) chosen in combination with a convergence criterion of 10^{-7} eV per supercell to ensure high-accurate DFT data. This resulted into in total 9351 calculations. To impose spin-inversion symmetry we added the same number of configurations to the training with reversed spin directions. The DFT calculations were performed at a lattice parameter of 2.9 Å corresponding to the experimental value near the Curie temperature. The phonon calculations have been performed utilizing the finite-displacement method with a displacement of 0.02 Å and utilizing the same set of technical parameters.

Disordered-local moment thermodynamic integration from mMTP to DFT

To sample the paramagnetic state at finite temperatures within the framework of thermodynamic integration (TI), we have employed the disordered-local-moment (DLM) MD[42]. The local magnetic moments were flipped randomly every 10 fs (=10 MD steps) such that half of the moments were pointing up and the other half down. The timestep for the MD was set to 1 fs; small enough to sample well the time development of the magnetic moments also within the 10 fs time intervals. The temperature was controlled by the Nose thermostat[50]. Usage of the Nose thermostat was critical; tests with the Langevin thermostat showed that it cannot stabilize the temperature well due to the additional impact of the spin flips on the energy of the system.

Spin-polarized DFT calculations in general and DLM calculations for Fe in particular are very prone to convergence problems, due to a flat energy landscape with many local minima as a function of spin state. Therefore, for the calculation of the DFT energy and forces during the MD, a very tight convergence criterion of 10^{-7} eV per supercell was set, in order to enforce sampling of the original DFT energy landscape that served as the input to the magnetic MTP fitting. To nevertheless allow for an efficient DFT MD, we have restricted the number of electronic iteration steps (typically to 40). Not fully converged DFT calculations were omitted from the comparison to the magnetic mMTP. Likewise, DFT calculations featuring local moment flips with respect to the mMTP data were not considered in the comparison.

To increase the efficiency of the DFT DLM-MD simulations we found it necessary to turn off the wave function extrapolation (both linear and quadratic); the reason for this lying in the randomization of the spins along the MD trajectory. A further efficiency increase was achieved by equilibrating the MD at the temperature of interest by utilizing the efficient mMTP. In this way, the part of the MD involving the expensive DFT calculations started directly on a well-equilibrated trajectory.

The DLM-TI was performed at a temperature of 800 K and at a lattice constant of 2.9 Å. A supercell of $2 \times 2 \times 2$ (in terms of the conventional bcc unit cell) with 16 atoms was utilized. A dense k-point sampling of $6 \times 6 \times 6$ corresponding to 3456 k-points $\times$ atom, a plane wave cutoff of 500 eV, and Fermi-Dirac smearing were used for the DFT calculations. For the mMTP calculations, initial magnetic moments were set according to the DFT moments. Then, for every mMTP energy and force calculation, the magnetic moments were fully relaxed based on the mMTP energetics. Coupling constants of $\lambda = 0.0, 0.5, 1.0$ were utilized. At each coupling constant two different random seeds were used to generate distinct trajectories. In total, more than 22,000 MD steps (22 ps) were conducted to generate statistically highly reliable correlation plots as shown in Fig. 3 of the main text. Test calculations for a larger $3 \times 3 \times 3$ supercell with 54 atoms turned out to be computationally highly demanding due to the strict DFT convergence parameters. Corresponding results indicate however a similar performance of the mMTP also for the larger supercell.

DATA AVAILABILITY

The authors declare that all data supporting the findings of this study are available from the corresponding author on reasonable request.

REFERENCES

1. Hasegawa, H. & Pettifor, D. Microscopic theory of the temperature-pressure phase diagram of iron. *Phys. Rev. Lett.* **50**, 130 (1983).
2. 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).
3. 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 (2016).
4. Ruban, A. V. & Razumovskiy, V. I. Spin-wave method for the total energy of paramagnetic state. *Phys. Rev. B* **85**, 174407 (2012).
5. 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).
6. 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).
7. Dudarev, S. & Derlet, P. A 'magnetic' interatomic potential for molecular dynamics simulations. *J. Phys. Condens. Matter* **17**, 7097 (2005).
8. Boukhvalov, D., Gornostyrev, Y. N., Katsnelson, M. & Lichtenstein, A. Magnetism and local distortions near carbon impurity in γ-iron. *Phys. Rev. Lett.* **99**, 247205 (2007).
9. Gorbatov, 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 (2011).
10. L., S. et al. Stacking fault energy of c-alloyed steels: the effect of magnetism. *Acta Mater.* **122**, 72 (2017).
11. Bienvenu, B., Fu, C. C. & Clouet, E. Impact of magnetism on screw dislocations in body-centered cubic chromium. *Acta Mater.* **200**, 570 (2020).
12. H., O. et al. Atomic relaxation around defects in magnetically disordered materials computed by atomic spin constraints within an efficient lagrange formalism. *Phys. Rev. B* **102**, 144101 (2020).
13. Schneider, A., Fu, C.-C., Soisson, F. & Barreteau, C. Atomic diffusion in α-iron across the curie point: An efficient and transferable ab initio–based modeling approach. *Phys. Rev. Lett.* **124**, 215901 (2020).
14. S. H., İ. et al. Ab initio phase stabilities of ce-based hard magnetic materials and comparison with experimental phase diagrams. *Phys. Rev. Mater.* **3**, 084407 (2019).
15. Matsumoto, M. & Akai, H. Calculating curie temperatures for rare-earth perma-nent magnets: Ab initio inspection of localized magnetic moments in d-electron ferromagnetism. *Phys. Rev. B* **101**, 144402 (2020).
16. Behler, J. First principles neural network potentials for reactive simulations of large molecular and condensed systems. *Angew. Chem. Int. Ed.* **56**, 12828 (2017).
17. Deringer, V. L., Caro, M. A. & Csányi, G. Machine learning interatomic potentials as emerging tools for materials science. *Adv. Mater.* **31**, 1902765 (2019).
18. Schütt, K. et al. Schnet: a continuous-filter convolutional neural network for modeling quantum interactions. *Adv. Neural Inf. Process. Syst.* **30**, 991 (2017).
19. Lubbers, N., Smith, J. S. & Barros, K. Hierarchical modeling of molecular energies using a deep neural network. *J. Chem. Phys.* **148**, 241715 (2018).
20. V., J. et al. On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events. *npj Comput. Mater.* **6**, 20 (2020).
21. Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **99**, 014104 (2019).
22. 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 (2017).
23. Jinnouchi, R., Karsai, F. & Kresse, G. On-the-fly machine learning force field gen-eration: application to melting points. *Phys. Rev. B* **100**, 014105 (2019).
24. Cusentino, M. A., Wood, M. A. & Thompson, A. P. Explicit multi-element extension of the spectral neighbor analysis potential for chemically complex systems. *J. Phys. Chem. A* **124**, 5456 (2020).
25. Chmiela, S. et al. Machine learning of accurate energy-conserving molecular force fields. *Sci. Adv.* **3**, e1603015 (2017).
26. Pun, G. P., Batra, R., Ramprasad, R. & Mishin, Y. Physically informed artificial neural networks for atomistic modeling of materials. *Nat. Commun.* **10**, 1 (2019).
27. Zuo, Y. et al. Performance and cost assessment of machine learning interatomic potentials. *J. Phys. Chem. A* **124**, 731 (2020).
28. Dederichs, P. H., Blügel, S., Zeller, R. & Akai, H. Ground states of constrained systems: Application to cerium impurities. *Phys. Rev. Lett.* **53**, 2512 (1984).
29. Stocks, G. et al. Towards a constrained local moment model for first principles spin dynamics. *Philos. Mag. B* **78**, 665 (1998).
30. Singer, R., Fähnle, M. & Bihlmayer, G. Constrained spin-density functional theory for excited magnetic configurations in an adiabatic approximation. *Phys. Rev. B* **71**, 214435 (2005).
31. Kaduk, B., Kowalczyk, T. & Van Voorhis, T. Constrained density functional theory. *Chem. Rev.* **112**, 321 (2012).
32. Ma, P.-W. & Dudarev, S. Constrained density functional for noncollinear mag-netism. *Phys. Rev. B* **91**, 054420 (2015).
33. Shapeev, A. V. Moment tensor potentials: a class of systematically improvable interatomic potentials. *Multiscale Model. Simul.* **14**, 1153 (2016).

34. Gubaev, K., Podryabinkin, E. V., Hart, G. L. W. & Shapeev, A. V. Accelerating high-throughput searches for new alloys with active learning of interatomic potentials. *Comp. Mater. Sci.* **156**, 148 (2019).

35. Grisafi, A., Wilkins, D. M., Csányi, G. & Ceriotti, M. Symmetry-adapted machine learning for tensorial properties of atomistic systems. *Phys. Rev. Lett.* **120**, 036002 (2018).

36. Drautz, R. Atomic cluster expansion of scalar, vectorial, and tensorial properties including magnetism and charge transfer. *Phys. Rev. B* **102**, 024104 (2020).

37. Nikolov, S. et al. "Quantum-accurate magneto-elastic predictions with classical spin-lattice dynamics", *arXiv preprint arXiv:2101.07332* (2021).

38. Dusson, G. et al. "Atomic cluster expansion: completeness, efficiency and stability", *arXiv preprint arXiv:1911.03550* (2019).

39. Novikov, I. S., Gubaev, K., Podryabinkin, E. V. & Shapeev, A. V. The MLIP package: moment tensor potentials with MPI and active learning. *Mach. Learn. Sci. Technol.* **2**, 025002 (2020).

40. Rosengaard, N. & Johansson, B. Finite-temperature study of itinerant ferromagnetism in fe, co, and ni. *Phys. Rev. B* **55**, 14975 (1997).

41. Szilva, A. et al. Interatomic exchange interactions for finite-temperature magnetism and nonequilibrium spin dynamics. *Phys. Rev. Lett.* **111**, 127204 (2013).

42. Alling, B. et al. Strong impact of lattice vibrations on electronic and magnetic properties of paramagnetic fe revealed by disordered local moments molecular dynamics. *Phys. Rev. B* **93**, 224411 (2016).

43. Grabowski, B. et al. Ab initio vibrational free energies including anharmonicity for multicomponent alloys. *npj Comput. Mater.* **5**, 1 (2019).

44. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558 (1993).

45. Kresse, G. & Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251 (1994).

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

47. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).

48. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953 (1994).

49. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).

50. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **81**, 511 (1984).

AUTHOR CONTRIBUTIONS

F.K. and A.S. conceived the project. I.N. and A.S. developed the magnetic Moment Tensor Potential (mMTP). I.N. fitted mMTP, investigated prediction errors and convergence of mMTP. F.K. prepared DFT datasets, calculated, and compared the phonon spectra obtained with the magnetic MTP and DFT. B.G. conducted the disordered-local-moment MD simulations and investigated the correlation of magnetic MTP and DFT energies and spins. All the authors participated in analyzing the results and writing the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-022-00696-9.

Correspondence and requests for materials should be addressed to Ivan Novikov.

Reprints and permission information is available at http://www.nature.com/reprints

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