




Machine learning approach for longitudinal spin fluctuation effects in bcc Fe at T_c and under Earth-core conditions

Marian Arale Brännvall ^{*}, Davide Gambino , Rickard Armiento, and Björn Alling 

Department of Physics, Chemistry, and Biology (IFM), Linköping University, SE-581 83, Linköping, Sweden



(Received 16 November 2021; revised 10 March 2022; accepted 17 March 2022; published 13 April 2022)

We propose a machine learning approach to predict the shapes of the longitudinal spin fluctuation (LSF) energy landscapes for each local magnetic moment. This approach allows the inclusion of the effects of LSFs in, e.g., the simulation of a magnetic material with *ab initio* molecular dynamics in an effective way. This type of simulation requires knowledge of the reciprocal interaction between atoms and moments, which, in principle, would entail calculating the energy landscape of each atom at every instant in time. The machine learning approach is based on the kernel ridge regression method and developed using bcc Fe at the Curie temperature and ambient pressure as a test case. We apply the trained machine learning models in a combined atomistic spin dynamics and *ab initio* molecular dynamics (ASD-AIMD) simulation, where they are used to determine the sizes of the magnetic moments of every atom at each time step. In addition to running an ASD-AIMD simulation with the LSF machine learning approach for bcc Fe at the Curie temperature, we also simulate Fe at temperature and pressure comparable to the conditions at the Earth's inner solid core. The latter simulation serves as a critical test of the generality of the method and demonstrates the importance of the magnetic effects in Fe in the Earth's core despite its extreme temperature and pressure.

DOI: [10.1103/PhysRevB.105.144417](https://doi.org/10.1103/PhysRevB.105.144417)

I. INTRODUCTION

Theoretical simulations based on *ab initio* methods are powerful tools to further design and develop materials, with density functional theory (DFT) being the main work horse in material science. Fe and Fe-based alloys are examples of materials currently investigated with first-principles simulations [1] and are of great technological importance in several fields, from constructions to memory technologies [2]. However, the magnetic properties of Fe make *ab initio* simulations at elevated temperatures nontrivial. The complexity arises from the coupling between magnetic, electronic, and vibrational degrees of freedom. For example, Körmann *et al.* showed that the magnetic phases had to be accurately described to make correct phonon calculations [3]. It was also shown, in a study by Alling *et al.*, that the impact of lattice vibrations on magnetic and electronic properties is of importance in paramagnetic body centered cubic (bcc) and face centered cubic (fcc) Fe [4]. Consequently, methods that treat vibrations and magnetism separately are not enough for a quantitative simulation of a magnetic system at elevated temperature.

The theoretical understanding of magnetic materials has been developed through two models: the local magnetic

moments model and the itinerant electron model [5], neither of which alone is able to quantitatively describe the magnetism of Fe at finite temperature. In the local magnetic moments model (the Heisenberg model) the electrons are assumed to be localized on the atoms, producing a local magnetic moment with, typically, constant magnitude. The itinerant electron model (Stoner model) is based on the Stoner description of band theory [5]. Within this model, the magnetization is due to the spontaneous spin split of bands, and the magnetic moments are associated with the itinerant conduction electrons [6]. Moriya developed a phenomenological theory that connects the two models and turns it into a problem of spin-density fluctuations [7].

The phenomenological theory of Moriya led to several first-principles approaches for describing the paramagnetic (PM) state of magnetic materials [5]. One such method is the disordered local moment (DLM) model [8]. By combining the DLM model with *ab initio* molecular dynamics (AIMD), in the DLM-AIMD method [4,9], the vibrational degree of freedom can be accounted for while considering an ideally disordered magnetic phase, which corresponds to the assumption of adiabatically fast magnetic degrees of freedom. This approach allows for the investigation of the effects of lattice vibrations in the PM state. However, within DLM-AIMD, magnetic short-range order is neglected. Ma *et al.* introduced a spin-lattice dynamics approach for modeling magnetic bcc Fe [10], an approach which has later been extended to incorporate spin-orbit coupling [11] and to, for example, investigate the dynamic interplay between phonons and magnons in bcc Fe [12]. Inspired by the spin-lattice dynamics simulations, a first-principles based methodology was developed by Stockem *et al.*, called the combined atomistic spin

^{*}marian.brannvall@liu.se

dynamics and *ab initio* molecular dynamics (ASD-AIMD) method [13]. This method combines AIMD simulations to model the vibrations of the atoms with ASD simulations to model the change of the spin state in time through rotations of the magnetic moments, which is propagated following the Landau-Lifshitz-Gilbert (LLG) equation [14]. ASD-AIMD method was applied to CrN where they observed that coupling the spin and lattice dynamics leads to shortening of phonon lifetimes compared with using an adiabaticlike simulation like DLM-AIMD [13].

These methods have often focused only on the transverse magnetic excitations; however, excitations along the directions of the magnetic moments also occur in magnetic metals. These types of excitations are called longitudinal spin fluctuations (LSFs). The LSFs are connected to the itinerant nature of electrons and are especially important for itinerant systems where local magnetic moments are induced by entropy [15]. For localized moment systems the magnitudes do not depend as much on temperature, and the effects of LSFs become less important. LSFs are not included in the classical Heisenberg Hamiltonian, and so to include their effects, extended model Hamiltonians have been developed [7,15–24]. In the work by Uhl and Kübler [17] the effects of both transverse and longitudinal spin fluctuations are included using the semiclassical approach developed by Murata and Doniach [16] when studying Fe, Ni, and Co. In Murata and Doniach’s work the energy functional has a fourth-order dependence on the magnetic moment, applying an approach similar to Ginzburg-Landau expansion. Rosengaard and Johansson also included the effects of both longitudinal and transverse degrees of freedom in a study like Uhl and Kübler’s, however, they assume local interatomic exchange interactions, which leads to a real-space formulation [18]. Ma and Dudarev included the evolution of both transverse and longitudinal spin fluctuations in their generalized Langevin spin dynamics method and applied it to ferromagnetic Fe [20]. They also used the semiclassical approach; however, the semiclassical term was expanded up to the sixth order as it was shown to better reproduce the energy landscape for ferromagnetic Fe. The LSF energy of a certain atomic magnetic moment depends on the size of the moment itself, and a curve showing how the LSF energy changes with moment size can thus be plotted. These curves are referred to as LSF energy landscapes.

Ruban *et al.* showed that LSFs are necessary to accurately describe magnetism at high temperatures [15]. They looked at fcc Ni and bcc Fe as examples of an itinerant and a localized system, respectively. In this work, they used one type of modified Heisenberg Hamiltonian to include the effects of LSFs. In the PM state, the LSF energy landscape of bcc Fe has a deep minimum, around $2\mu_B$, which is the typical shape of the landscapes of a localized moment system. For fcc Ni, the minimum is at zero, which is representative of an itinerant system. Because of the deep minimum in the case of bcc Fe, the average size of the magnetic moment is similar in the PM and ferromagnetic (FM) states. The relatively robust magnetic moments of bcc Fe make the classical Heisenberg Hamiltonian work quite well without the need of accounting for LSFs. In the case of fcc Ni, the magnitudes of the magnetic moments are less robust, and therefore the effects of LSFs are bigger, making it especially necessary to include them. However, they

conclude that including the LSF effects is important even in the case of Fe.

In the study by Ruban *et al.*, described above, the LSF energy landscapes were calculated within the coherent potential approximation [15], which does not account for the differences in local environments due to vibrations or local magnetic configurations. One way to account for the lattice vibrations is with a supercell approach. An example of this method is found in the work by Ruban and Peil, where they investigated the effect of lattice distortions on the magnetic exchange interactions using AIMD simulations to obtain the atomic configurations [25]. Another example is the work by Gambino *et al.*, where a supercell approach was developed to calculate the LSF energy landscapes corresponding to every atomic magnetic moment and, from that, derive the finite-temperature magnitudes of the magnetic moments [26]. It was found that the shapes of the LSF energy landscapes depended on the local environment of the atomic moments, such as the lattice vibrations, with some landscapes having very shallow minima.

The impact of LSFs is especially strong in the PM state, where the magnetic moments have disordered directions, and so to accurately theoretically simulate this state, the LSF energy landscape, in principle, has to be calculated for each atom [26]. In an AIMD simulation or ASD-AIMD simulation, where the local environments (atomic configuration and, in the ASD-AIMD case, the magnetic configuration) changes with each time step, the LSF energy landscape would have to be calculated for each atom at each moment in time. In a supercell during a long AIMD run, this becomes a practically overwhelming task. One solution to this issue could be to use machine learning (ML) methods. By training ML models on relevant inputs it is possible to efficiently predict properties. Since the local environment of the magnetic moments plays an important role in how the shapes of the LSF energy landscapes turn out, ML methods are a promising tool to overcome the difficulty of determining these landscapes of multiple atoms over time. In this work, we probe this idea.

It has become increasingly popular to use ML methods in materials science and they have proven successful in predicting numerous material properties and accelerating the discovery of new materials [27,28]. Applying ML methods to magnetic materials is still in its early stages, but there are a few examples. In the work by Sanvito *et al.*, the Curie temperature was estimated using an ML model, which in turn was used to accelerate the discovery of new magnets that fulfill certain requirements [29]. Long *et al.* also developed an ML model to predict the Curie temperature, and they trained another ML model to classify between ferromagnetic and antiferromagnetic materials [30]. Recently, ML interatomic potentials were developed by Novikov *et al.* to incorporate the transverse magnetic degree of freedom, called the magnetic moment tensor potentials [31]. The work by Yu *et al.* is another recently developed application of ML to describe magnetic properties [32]. In their work, artificial neural networks were used to construct spin Hamiltonians.

In this work, we have developed ML models to estimate the shapes of LSF energy landscapes based on the local atomic and magnetic environments of a magnetic moment. The ML method used is kernel ridge regression which has been suc-

cessfully used in earlier studies regarding formation energies of atomic crystals [33,34] and is a good starting point in investigating the use of ML in this case. The ML models are trained mainly on data for bcc Fe at the Curie temperature (1043 K) and ambient pressure. A small portion of data for bcc Fe at 6000 K and 300 GPa, corresponding to the conditions expected in the solid inner core of the Earth, were added to test the generality of the developed ML models. Since the training data only consist of data from different cases of bcc Fe, the implementation of the ML approach in this work only addresses bcc crystals. However, the same methodology can also be applied to magnetic materials with different crystal structures, with some adjustments.

This machine learned LSF approach has been implemented into the ASD-AIMD scheme (denoted as ASD-AIMD-MLLSF simulations) to accelerate the determination of sizes of the magnetic moments for each atom at each time step. With this implementation we include both the longitudinal and transverse magnetic degrees of freedom, where the ML approach accounts for the longitudinal spin fluctuations and the transverse spin fluctuations are treated within the ASD-AIMD scheme using the LLG equations. The ASD-AIMD-MLLSF simulations have been done for both bcc Fe at the Curie temperature and ambient pressure and for bcc Fe at 6000 K and 300 GPa. As a result, we find that predicting the LSF energy landscapes using ML models reduces the mean absolute error compared to other approximate methods, e.g., using a mean LSF energy landscape to describe all landscapes. The representation of the input data is also shown to have a significant impact on the performance of the ML models. We see that allowing for entropy-induced local magnetic moments increases the pressure at a given volume which becomes especially interesting in the case of estimating the density of the Earth's inner solid core.

The paper is structured as follows: Section II gives a theoretical background to longitudinal spin fluctuations (Sec. II A) and the kernel ridge regression ML method (Sec. II B). In the remaining part of Sec. II, the ASD-AIMD method combined with the ML of LSF approach is presented as well as the representation of input data in the descriptor and the computational details. Section III contains the results of applying the ML approach to ASD-AIMD simulations. Lastly, in Sec. IV, conclusions are discussed.

II. THEORETICAL METHODS

A. Theory of longitudinal spin fluctuations

To include the effects of LSFs to the Heisenberg Hamiltonian, the Heisenberg-Landau Hamiltonian is used [16–18,20]

$$H = - \sum_{i \neq j} J_{ij} \mathbf{m}_i \cdot \mathbf{m}_j + \sum_i E_i(m_i), \quad (1)$$

where J_{ij} are interatomic exchange constants between the moments at sites i and j , \mathbf{m}_i is the magnetic moment vector at site i , and $E(m_i)$ is the LSF energy depending on the magnitude of the i th moment. The LSF energy $E(m_i)$ can be expanded in

even powers,

$$E_i(m_i) = \sum_{n=0}^{\infty} a_n m_i^{2n} \approx a_i m_i^2 + b_i m_i^4, \quad (2)$$

where a fourth-order polynomial has shown to be sufficient for our purpose.

Different thermodynamic quantities can be calculated for the LSF Heisenberg Hamiltonian, which requires the partition function Z . Since the longitudinal excitations are faster than transversal, the first term in Eq. (1) can be considered constant at each time step; as a result, the partition function can be separated. Since the last term in Eq. (1) does not depend on other magnetic moments explicitly, the partition function may be written as the product of N partition functions $Z = \prod Z_i$ in the PM state. Hence,

$$Z_i = \int dm_i e^{-\frac{E_i(m_i)}{k_B T}} = \int_0^{\infty} dm_i \text{PSM} e^{-\frac{E_i(m_i)}{k_B T}}, \quad (3)$$

where the last part comes from expressing m_i in spherical coordinates and leaving out the constant from integrating the angular variables. This change of variables leads to the phase-space measure (PSM) simply being m^2 ; k_B is the Boltzmann constant and T is the temperature.

Here we should state that the issue of PSM is intimately related to the semiclassical approximations done of separating longitudinal and transverse degrees of freedom, as has been discussed heavily in the literature, e.g., in Ref. [21]. The above derivation is based on an assumption of dealing with an itinerant system, where the magnetic moments may change direction by shrinking in size and reappearing in a different direction, i.e., we can assume full coupling between longitudinal and transversal degrees of freedom. When instead dealing with a localized system, changes in the direction of the magnetic moments appear through rigid rotations which are slow compared to the fluctuations in size, and so the LSFs can be treated as restricted to the directions of the magnetic moments. In this localized moments case the magnetic moments would fluctuate along one dimension and the PSM of Eq. (3) would instead be equal to 1, which was introduced by Murata and Doniach [16]. This ambiguity in the choice of PSM can be seen in that different studies investigating bcc Fe chose different PSM. Ruban *et al.* used a PSM of m^2 when studying bcc Fe under Earth's inner core conditions [35] and Pan *et al.* used PSM = 1 when applying an extended atomistic spin model to several systems, including bcc Fe [23]. Khmelevskiy even suggested a PSM equal to m and applied it for bcc Fe and fcc Ni, which was shown to improve the estimation of the Curie temperature especially for fcc Ni [24]. We would like to point out that the ML approach described in this work is not restricted to any particular choice of PSM.

In this work, the average magnitude of every individual magnetic moment at temperature T is used, which is calculated as

$$\langle m_i(t, T) \rangle = \frac{1}{Z_i} \int_0^{\infty} dm_i \text{PSM} m_i e^{-\frac{E_i(m_i, t)}{k_B T}}, \quad (4)$$

where t is the time step and $E_i(m_i, t) = a_i(t) m_i^2 + b_i(t) m_i^4$; the a and b parameters of each LSF energy landscape changes at each time step.

B. Machine learning with kernel ridge regression

This work uses the kernel ridge regression (KRR) ML method, which is an extension of the linear regression method ridge regression. Ridge regression uses the least-square estimation method, but adds a penalizing term to the minimization of the coefficients to avoid overfitting. A function f approximates the known set of outputs $y = (y_1, y_2, \dots, y_n)^T$, corresponding to a given set of input vectors $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T$, where each $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{im})$ and x_{ij} is the j th element of the i th input vector. As a result, \mathbf{X} is an $n \times m$ matrix. The input vectors are called descriptors and their elements are referred to as features. In ridge regression, the following function approximates the output y corresponding to a descriptor \mathbf{x} for an arbitrary input:

$$f(\mathbf{x}, \boldsymbol{\alpha}) = \mathbf{k}(\mathbf{x})\boldsymbol{\alpha} = \sum_i k_i(\mathbf{x})\alpha_i, \quad (5)$$

where $\mathbf{k}(\mathbf{x}) = (k_1(\mathbf{x}), k_2(\mathbf{x}), \dots, k_n(\mathbf{x}))^T$ is a function ansatz acting on the vector \mathbf{x} , and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n)$ are the coefficients that are chosen to minimize the Euclidean norm for a given set of inputs and outputs $\|\mathbf{y} - f(\mathbf{X}, \boldsymbol{\alpha})\|_2$, which is the regular norm of an n -dimensional vector. With a penalty term λ the coefficients are then given by

$$\boldsymbol{\alpha} = [\mathbf{k}(\mathbf{X})^T \mathbf{k}(\mathbf{X}) + \lambda \mathbf{I}]^{-1} \mathbf{k}(\mathbf{X})^T \mathbf{y}, \quad (6)$$

where \mathbf{I} is the identity matrix and λ is called the regularization parameter.

The extension from ridge regression to kernel ridge regression comes from the so-called kernel trick and allows one to solve a nonlinear problem. This is possible by using a mapping $\Phi(\mathbf{x})$ from a nonlinear space to a linear space. It is not necessary to know the mapping explicitly, but simply the inner product between the mapping functions, which can be expressed as a general function $k(\mathbf{x}, \mathbf{x}')$, called a kernel, given by

$$k(\mathbf{x}, \mathbf{x}') = \langle \mathbf{x}, \mathbf{x}' \rangle = \langle \Phi(\mathbf{x}), \Phi(\mathbf{x}') \rangle. \quad (7)$$

In KRR the function to approximate the output is

$$f(\mathbf{x}, \boldsymbol{\alpha}) = \sum_i k(\mathbf{x}, \mathbf{x}_i)\alpha_i, \quad (8)$$

which is optimized using a training set of known inputs y and a set of the corresponding descriptors \mathbf{X} so that

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y}, \quad (9)$$

where \mathbf{K} is the kernel matrix with $K_{ij} = k(x_i, x_j)$ for each pair of descriptors in the training set. From Eq. (9) we see that we get one coefficient α_i for each training data point. The sum in Eq. (8) goes over all training data. Consequently, the function $f(\mathbf{x}, \boldsymbol{\alpha})$ is a sum of the the kernel function acting on the input descriptor and each of the descriptors of the training data times the corresponding coefficient.

This work uses the Laplacian kernel function

$$k(\mathbf{x}, \mathbf{x}') = e^{-\|\mathbf{x} - \mathbf{x}'\|_1 / \sigma}, \quad (10)$$

where $\|\mathbf{x} - \mathbf{x}'\|_1 = \sum_i |x_i - x'_i|$ is the Manhattan norm and σ is the kernel width. The kernel width regulates how similar two vectors (\mathbf{x} and \mathbf{x}') are categorized to be. The kernel width

σ and the regularization parameter λ are calibrated to the problem at hand.

To evaluate the performance of the ML models, a subset of the available data is left as a test set and the ML models are trained on the remainder of the data (the training set). The errors of the ML models predictions of previously unseen data are given by the mean absolute error (MAE)

$$\text{MAE}_{\text{test}} = \frac{1}{n} \sum_{i=1}^n |y_i - f(\mathbf{x}_i, \boldsymbol{\alpha})|, \quad (11)$$

where n is the size of the test set. To more accurately estimate the MAE for predictions, 10-fold cross validation is used, where the data are divided into 10 subsets and in 10 iterations the subsets alternate in acting as the test set. A mean of all MAEs is then taken as the estimate of the MAE for predictions on data outside the training set.

C. Descriptors

The accuracy of the KRR method depends on the precise choice of representation of the input data in the descriptor, i.e., the vector \mathbf{x} given as input to $f(\mathbf{x}, \boldsymbol{\alpha})$ in Eq. (8). There are certain requirements for a good descriptor: for instance, it should contain all relevant features, but have a minimal number of redundant features, and it should require as little computational effort as possible. The descriptor should also be able to detect similarities in the input data of different instances by representing instances that are close with descriptors that are close, i.e., give a small $\|\mathbf{x}_1 - \mathbf{x}_2\|$ [33].

Since the LSF energy landscapes depend on the structural and magnetic local environments of the atomic magnetic moments, the descriptors need to contain information which describes the local environment. One such feature that has proven important is the local Voronoi volume [26]. The positions and the magnetic moments of neighboring atoms should also be of interest when describing the local environment of a specific atomic magnetic moment. Four different descriptors are proposed and tested in this work, each containing the Voronoi volume and the positions of neighboring atoms in the form of spherical coordinates with respect to the central atom. The content of the different proposed descriptors is presented in Table I, and they differ in the following way: descriptor 1 includes the explicit directions of the magnetic moments, i.e., e_x , e_y , and e_z . The other descriptors, instead, contain the scalar products between the magnetic moment of the central atom and of each neighboring atom. Descriptor 2 contains each scalar product and descriptor 3 contains the sum of the scalar products. Descriptor 4 only differs from descriptor 2 by one element: it adds the standard DFT result of the size of the magnetic moment of the central atom, denoted as m_0 . Adding m_0 requires running an additional DFT simulation, where only the directions of the magnetic moments are constrained, giving the size of the moments corresponding to a temperature of 0 K for this degree of freedom. For an itinerant system, with its LSF energy landscape minimum at zero, the solution at 0 K would be nonmagnetic in this type of DFT calculation.

Aside from *which* features to include in the descriptor, another important aspect is *how many* features to include. The number of features couples to the amount of available training

TABLE I. The four descriptors tested in this work. V is the Voronoi volume, \mathbf{e}_i is the direction of the i th magnetic moment (with components e_{x_i} , e_{y_i} , and e_{z_i}), r , θ , ϕ are the spherical coordinates of the neighboring atoms with respect to the central atom, and m_i is the size of the i th magnetic moment obtained from a DFT simulation without longitudinal constraints. Subscript “0” indicates the central atom and “1, 2, ...” indicate neighboring atoms.

Descriptor no.	Descriptor, \mathbf{x}
1	$[V, e_{x_0}, e_{y_0}, e_{z_0}, r_1, \theta_1, \phi_1, e_{x_1}, e_{y_1}, e_{z_1}, \dots, r_N, \theta_N, \phi_N, e_{x_N}, e_{y_N}, e_{z_N}]$
2	$[V, r_1, \theta_1, \phi_1, \mathbf{e}_0 \cdot \mathbf{e}_1, \dots, r_N, \theta_N, \phi_N, \mathbf{e}_0 \cdot \mathbf{e}_N]$
3	$[V, r_1, \theta_1, \phi_1, \dots, r_N, \theta_N, \phi_N, \sum_{n=1}^N \mathbf{e}_0 \cdot \mathbf{e}_n]$
4	$[V, m_0, r_1, \theta_1, \phi_1, \mathbf{e}_0 \cdot \mathbf{e}_1, \dots, r_N, \theta_N, \phi_N, \mathbf{e}_0 \cdot \mathbf{e}_N]$

data since a small amount of data will not work well with a too big feature space. In the present case, the number of shells of neighboring atoms included in the descriptor impacts the size of the feature space. We determine the number of shells to take into account by comparing the MAEs when including different numbers of shells.

In Fig. 1 the MAE is shown as a function of the training set size for each descriptor and LSF energy landscape parameter, when including up to four coordination shells. We see that, especially for descriptors 2 and 3, the MAE converges quite rapidly, already at about 200 data points. For descriptors 1 and 4, there is still a slight tilt, indicating that more data points in the training data set should further decrease their MAEs. In the implementations of the ML models presented in this paper, we use descriptor 3 with only one coordination shell included since it gives the lowest error without having to do an extra DFT calculation to get the m_0 feature. Our training data set contains about 680 data points (90% of the complete data set) which, based on Fig. 1, should be quite enough to get the lowest possible MAE with this descriptor.

In Table II the lowest MAE of each descriptor is presented with the corresponding number of coordination shells included. Including more than one coordination shell in the descriptor increases the error in most cases. Using the scalar

products of the directions of the magnetic moments is clearly an advantage compared to the explicit directions; this is briefly discussed below.

A final aspect of the representation of input data in the descriptor is the ordering of the features. The neighboring atoms are sorted based on their ideal lattice positions as described schematically in Fig. 2. From the ideal positions, the neighboring atoms are sorted based on which coordination shell they belong to (i.e., distance from the central atom) and by their angle, θ and ϕ , to the central atom. Although the indexing is based on the ideal crystal points, the actual positions, including thermal displacements, are used in the descriptors. As previously mentioned, a requirement on the descriptor is that similar LSF energy landscapes (the output) should be represented by similar descriptors, i.e., with a small Manhattan distance in the case of a Laplacian kernel. When constructing the descriptor in the manner described previously, with the explicit atomic positions with regards to the central atom ordered in this way, some detections of similarities will most likely be lost. As seen in the right-hand illustration of Fig. 2, a rotation of the coordinate system can shift the order of the atoms in the descriptor. We would therefore get two different descriptors for two LSF energy landscapes that refer to the same physical system (a rigid rotation of the crystal does not

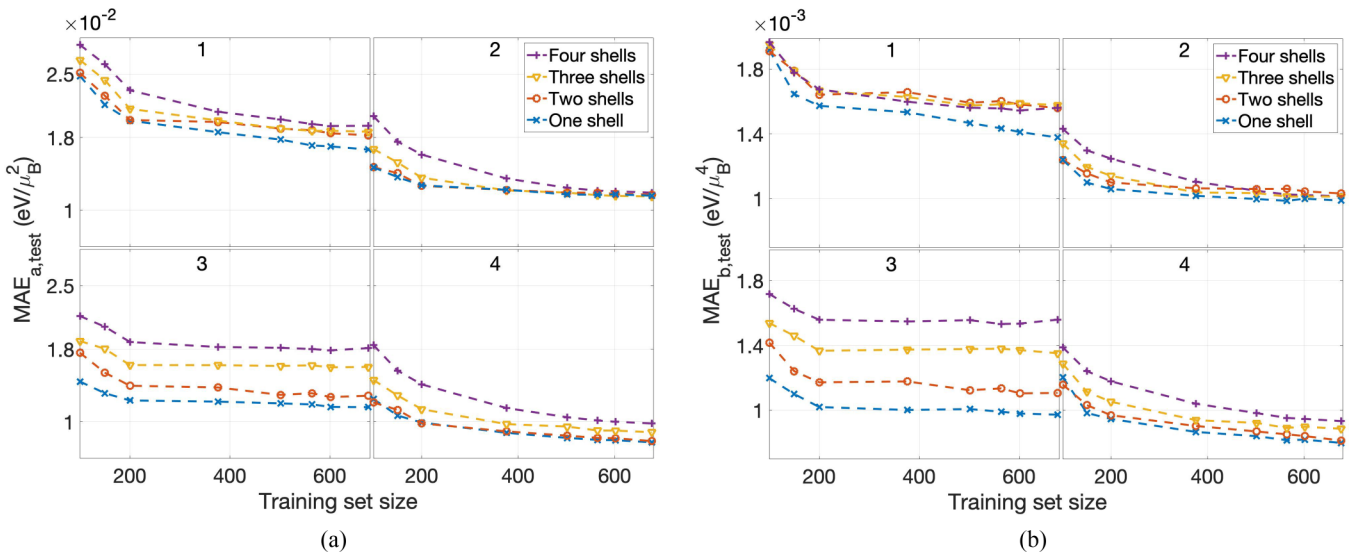


FIG. 1. Mean absolute error (MAE) as a function of the training set size for each of the four descriptors proposed in this work (listed in Table I), represented by the numbers 1–4 in the figures. Graphs for up to four coordination shells in the descriptor are shown. (a), (b) Show, respectively, the parameters a and b of the LSF energy landscape parametrization.

TABLE II. The lowest mean absolute error (MAE) for predictions of the validation data set for parameters a and b for each descriptor. The number of included coordination shells is shown in parentheses.

	Descriptor 1	Descriptor 2	Descriptor 3	Descriptor 4
$MAE_{a,\text{test}}$	0.0167 (1 shell)	0.0115 (3 shells)	0.0116 (1 shell)	0.00767 (1 shell)
$MAE_{b,\text{test}}$	0.00138 (1 shell)	0.000989 (1 shell)	0.000973 (1 shell)	0.000798 (1 shell)

change the energy landscapes). It was previously seen that including the scalar products instead of the explicit magnetic moment directions decreases the errors, and a reason for this could be this issue of rotations changing the order of the elements in the descriptor. When including the scalar products, and especially the sum of the scalar products as in descriptor 3, there are fewer elements that are ordered differently after a rotation. If we have two systems that are the same except rotated differently, their descriptors will be slightly more similar when using scalar products compared to using the explicit directions of the moments.

In this work we focus on two different cases of bcc Fe and the form of the descriptor is, consequently, limited to the bcc structure. For other crystal structures, we will need to modify the current descriptor.

D. Atomistic spin dynamics and *ab initio* molecular dynamics with a machine learning approach for longitudinal spin fluctuations

In ASD-AIMD simulations, the coupling between spin dynamics and lattice vibrations is accounted for [13]. In the AIMD part, the atomic positions change over time and are updated based on the forces acting on each atom. The forces are calculated using noncollinear DFT, where the magnetic moments are constrained to a certain direction that changes at every time step. The direction of the moment is dictated by the ASD part, which run in parallel, and where the orientations are updated according to the Landau-Lifshitz-Gilbert equation for atomistic spin dynamics [36] and the interatomic exchange constants $J_{ij}(R_{ij})$ are assigned to each pair of atoms and depend on the local geometry. The exchange constants are parametrized prior to the simulation. In the ASD-AIMD simulations of Ref. [13], CrN did not include the effects of LSFs, which were motivated by the robust local magnetic

moments of Cr, almost $3\mu_B$ [37], in CrN. When studying Fe, however, it is important to include the effects of LSFs, as concluded in both Refs. [15,26]. This work includes the LSF effects in ASD-AIMD simulations in an efficient way using the ML approach. Figure 3 is an illustration of the resulting ASD-AIMD-MLLSF scheme. The ML models are used to predict the sizes of the magnetic moments and, therefore, in the AIMD step, the magnetic moments are constrained both in direction and size. The ML models predict the a and b parameters of the LSF energy landscapes of each magnetic moment, and this is then used to calculate the size of the magnetic moment as described by Eq. (4). The expectation value in Eq. (4) is therefore used as the instantaneous magnetic moment at time t . Since the descriptor consists of information about the positions of neighboring atoms $\{R_i\}$ and the directions of their magnetic moments $\{e_j\}$ this information is fed into the ML step. The ML step, together with LSF thermodynamics, gives the sizes of the magnetic moments which are used in the subsequent ASD and AIMD steps.

In the atomistic spin dynamics simulation, the adiabatic approximation is used [36,38,39]. Antropov *et al.* showed that the orientations of the local magnetic moments can be considered as slowly varying relative to their magnitudes [40,41] since the transversal degree of freedom is related to collective electronic dynamics and the longitudinal degree of freedom to individual electron hopping. There is approximately a two orders of magnitude difference between the timescales of the longitudinal ($\sim 10^{-15}$ s) and the transverse spin fluctuations ($\sim 10^{-13}$ s) [38]. This adiabatic approximation is valid for systems with well-defined local magnetic moments which is true for, e.g., Fe [40].

E. Computational details

All calculations are employed using density functional theory as implemented in the Vienna *ab initio* simulation package (VASP) [42–45] with projector augmented-wave (PAW) potentials [46,47] and the Perdew-Burke-Ernzerhof (PBE) generalized gradient approximation (GGA) [48] for approximating the exchange-correlation functional. Two systems are used: bcc Fe at the Curie temperature T_c (1043 K) and ambient pressure (T_c system), and bcc Fe at the temperature 6000 K and volume $7 \text{ \AA}^3/\text{atom}$ [high-temperature high-pressure (HTHP) system]. For both systems the calculations are carried out on a supercell of 54 atoms ($3 \times 3 \times 3$ bcc unit cells). Fermi-Dirac smearing and noncollinear magnetism are applied in all calculations.

When calculating the LSF energy landscapes, a $3 \times 3 \times 3$ and $5 \times 5 \times 5$ Monkhorst-Pack k -point grid [49] is used for the T_c system and the HTHP system, respectively. For the HTHP case, the $3s$ and $3p$ electrons are considered valence

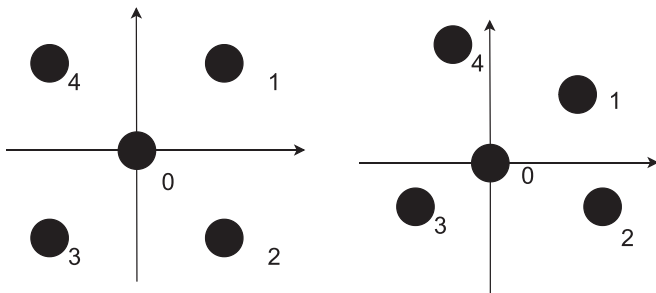


FIG. 2. 2D illustration of how the neighboring atoms are sorted in the descriptor: (left) ideal crystal and (right) vibrating crystal. A rigid rotation of, e.g., 90° of the crystal would make neighbor 4 become neighbor 1 and so on, changing the order of elements in the descriptor.

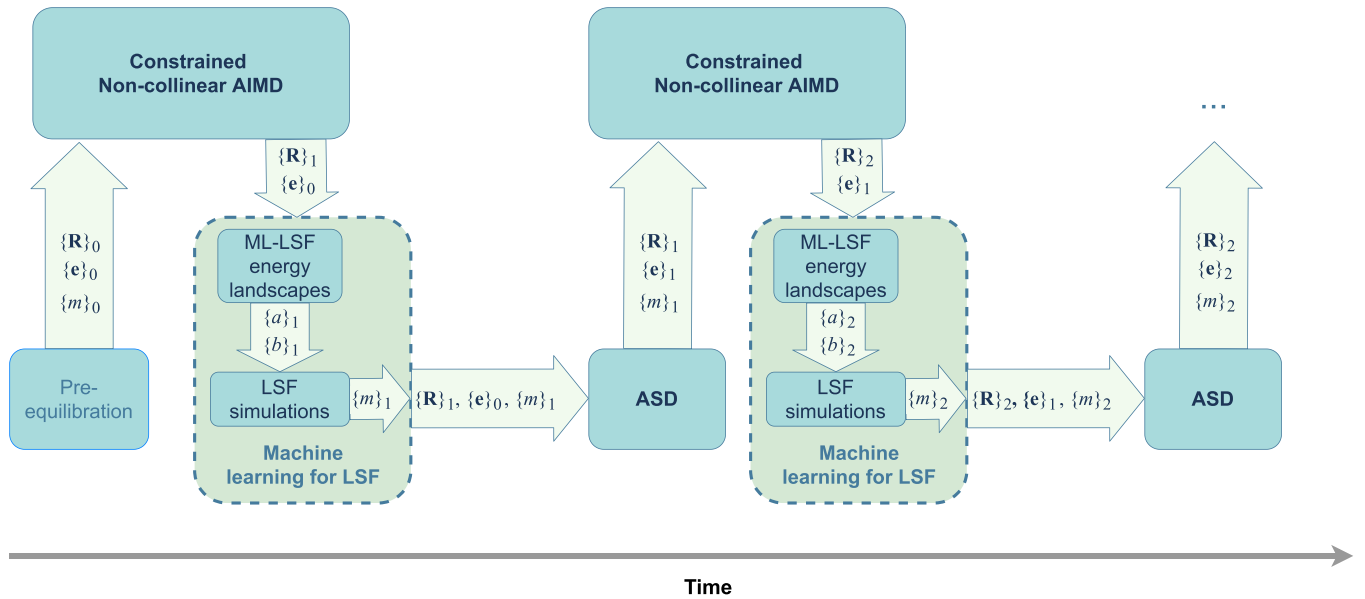


FIG. 3. Combined atomistic spin dynamics and *ab initio* molecular dynamics including our machine learning approach for longitudinal spin fluctuations (ASD-AIMD-MLLSF). The ASD-AIMD scheme [13] is followed but with an additional step where the trained machine learning models use the atomic positions $\{\mathbf{R}\}_n$ and directions of the magnetic moments $\{\mathbf{e}\}_n$ to predict the LSF energy landscapes, in terms of coefficients a_i and b_i , for each atom i in the cell. The LSF energy landscapes are subsequently used in thermodynamical simulations, at the simulated temperature, where the magnitudes of all the individual magnetic moments ($\{m\}_n$) are calculated numerically according to Eq. (4). The $\{\mathbf{R}\}_n$, $\{\mathbf{e}\}_n$, and $\{m\}_n$ are then passed on to the following ASD step. The subscripts n indicate a certain iteration or time step. In this work, time steps of 1 fs are used.

electrons because of the high-pressure conditions. The LSF energy landscapes of the T_c system are obtained from several snapshots out of an ASD-AIMD run performed at T_c and ambient pressure. The LSF energy landscapes of the HTHP system are obtained from a snapshot of a $3 \times 3 \times 3$ bcc supercell from a nonmagnetic MD run at 6000 K and a volume of $7 \text{ \AA}^3/\text{atom}$. This snapshot is taken from the work by Belonoshko *et al.* [50]. To calculate the LSF energy landscapes, the supercell scheme developed in Ref. [26] is employed.

For the T_c system, we run a DFT calculation, constraining only the direction of the magnetic moments. To do this type of constrained DFT calculation, the method developed by Ma and Dudarev [51], with constraining parameter λ set to 10, is used. This calculation gives the magnitude of the magnetic moments corresponding to a temperature of 0 K for the longitudinal degree of freedom. The magnetic moments are then constrained in both direction and magnitude for the following calculations, with $\lambda = 10$. The LSF energy landscape of one atomic magnetic moment is now obtained by scaling the size of the moment in steps while constraining the other moments to the sizes corresponding to a temperature of 0 K, given in the previous step. The magnetic moment is set to about six or seven different sizes, generating data points in a graph of energy as a function of the magnetic moment. A fourth-order polynomial, as in Eq. (2), is fitted to these data points and this gives us the LSF energy landscape. For the HTHP system, a similar scheme is employed. Except, in the HTHP case, the initial step of constraining only the directions of the magnetic moments is not performed. Instead, the stepwise change of the size of the magnetic moment is initially done in a nonmagnetic background [26]. This different approach means that the surrounding magnetic moments are not set to the sizes

corresponding to a temperature of 0 K, instead they are the sizes corresponding to 6000 K in this case.

The ASD-AIMD runs are carried out with Langevin dynamics on supercells of 54 atoms on both the T_c system and the HTHP system. The ASD part of the simulation is performed with the UPPASD code [36,39]. The k -point grid is reduced to a Γ -centered $2 \times 2 \times 2$ grid and the energy cutoff for the plane-wave basis set is set to 500 eV for both systems. In the T_c system the lattice parameter is set to 2.877 \AA which is the equilibrium lattice parameter at 0 K expanded with thermal expansion to the Curie temperature. For the HTHP system the lattice parameter is 2.410 \AA (volume V is $7.0 \text{ \AA}^3/\text{atom}$). Both the direction and sizes of the magnetic moments are constrained in these calculations since the sizes are fixed at the values given by Eq. (4) after the ML models have predicted the a and b parameters. The damping parameter in the LLG equation is set to 0.05 and the time step for evolution of the magnetic moment direction is set to 0.01 fs.

For comparison, a collinear ferromagnetic AIMD simulation at T_c and ambient pressure, and a nonmagnetic AIMD simulation at the temperature 6000 K and the volume $7 \text{ \AA}^3/\text{atom}$, are performed. In both these simulations, a $3 \times 3 \times 3$ bcc supercell is used and a Γ -centered $2 \times 2 \times 2$ k -point grid.

The exchange interactions J_{ij} in Eq. (1) are parametrized for the first two coordination shells from the results of the study by Ruban and Peil [25]. The interactions for shells 3, 4, and 5 are considered constant at 0.023, 0.002, and -0.100 meV, respectively.

The ML method KRR is implemented using the Python toolkit for quantum machine learning (QML) [52]. The QML

TABLE III. Kernel widths σ corresponding to the lowest mean absolute error for the a/b parameter when including up to four coordination shells.

	Descriptor 1	Descriptor 2	Descriptor 3	Descriptor 4
1 shell	10/10	10/10	10/10	100/10
2 shells	100/100	10/10	100/100	100/100
3 shells	100/100	100/100	100/100	1000/100
4 shells	100/100	100/100	100/100	1000/100

library offers a kernel module from which the Laplacian kernel can be calculated. The regularization parameter λ is set to 10^{-4} and the kernel width σ is set according to Table III when including up to four coordination shells for each descriptor. The cross-validation scheme is carried out using the Python machine learning library SCIKIT-LEARN [53].

III. RESULTS

A. Training data set

The training data set consists of the inputs (the descriptors) and corresponding outputs (the a and b parameters described in Sec. II A). One ML model is trained to predict the a parameters and one the b parameters, i.e., the two parameters are predicted independently. In practice, this means that the function $f(x, \alpha)$ in Eq. (8) differs for the two parameters of the LSF energy landscape. Specifically, the two models will have different sets of coefficients derived from Eq. (9).

The training data consist of data from the T_c system and the HTHP system (6000 K and 300 GPa). The main part of the training data set comes from the former system, with 752 LSF energy landscapes, and a smaller part comes from the latter system, with 54 landscapes. The focus of this work is to train the ML models on the T_c system, and the high-temperature-high-pressure data are used as a test of the generality of the trained ML models. In Fig. 4 some examples of LSF energy landscapes are shown both from the T_c and the HTHP systems [26]. Most of the LSF energy landscapes at T_c show finite minima but the different local environments lead to some moments having a more itinerant behavior, i.e., a shallower LSF energy landscape [26]. The LSF energy landscapes at 6000 K and 300 GPa all have their minima at a magnitude of the magnetic moment of $0\mu_B$. The finite magnitudes of magnetic moments in the HTHP case are therefore entropy induced whereas straightforward DFT calculations predict their sizes to be zero [35]. There is also less of a spread around the mean LSF energy landscape in the HTHP case, i.e., the local environment does not have the same impact as in the T_c case.

B. Other approximate methods for determining LSF energy landscapes

The performance of the machine learning models is compared to other approximate methods for determining the shapes of the LSF energy landscapes. In this work, three such approximate methods are investigated: (i) the Voronoi volume is used as the single parameter in determining the shape of the LSF energy landscape, (ii) a mean landscape of a set of LSF energy landscapes for a certain system is used as an approximation of all landscapes for that particular

system, and (iii) the magnitudes of the magnetic moments obtained by performing DFT calculations using VASP where the magnitudes are unconstrained is used as an approximation of

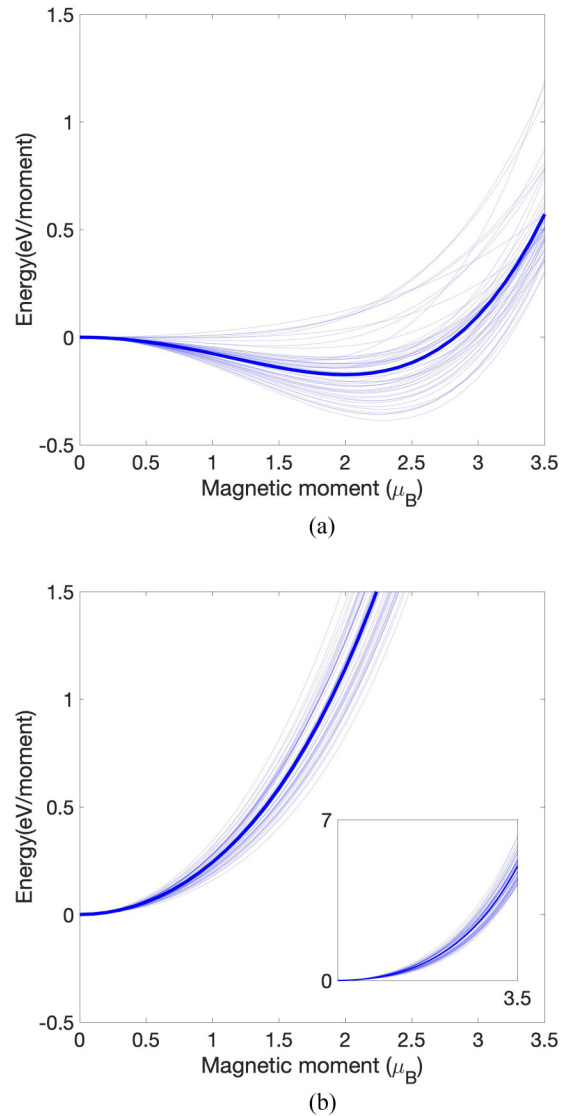


FIG. 4. Examples of longitudinal spin fluctuation energy landscapes from (a) the T_c system (ambient pressure, 1043 K) and (b) the Earth-core-like HTHP system (300 GPa, 6000 K). For the T_c system, only a selection of the complete set of 752 LSF energy landscapes are plotted. For the HTHP system, all 54 LSF energy landscapes are shown. Thick blue line represents the mean energy landscape. The inset in (b) simply shows the HTHP LSF energy landscapes in a larger energy range.

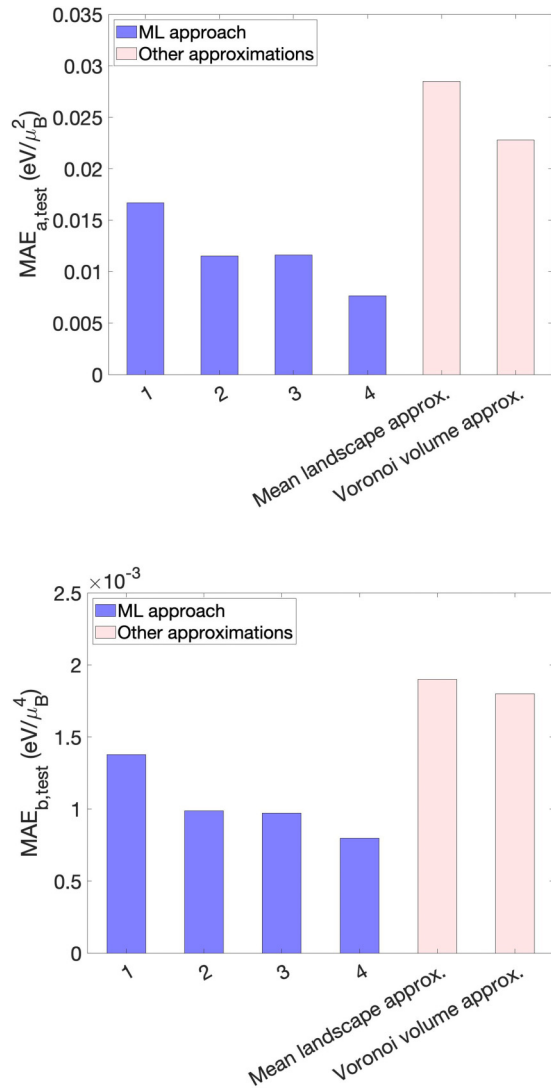


FIG. 5. Mean absolute errors (MAE) when predicting (a) the a parameters and (b) the b parameters using any of the four types of descriptors or two of the other approximation methods described in the text.

the magnitude at the minimum of the LSF energy landscape. These DFT calculations with unconstrained magnetic moment magnitudes essentially give the magnitude of each moment at the LSF temperature of 0 K. Based on the small spread seen in Fig. 4(b), we expect that approximation (ii), the mean landscape approximation, should work quite well in the HTHP case but less so for the T_c case. Approximation (i) should capture some of the differences in local environments for the magnetic moments because of the variation in the Voronoi volume, and therefore be better than approximation (ii). One may also note that approximation (iii) would, for example, not be useful for the HTHP system since the minima are all at zero size of the magnetic moment in this case.

In Fig. 5 the mean absolute error of approximation (i) and (ii) are compared with each descriptor when predicting the parameters a and b . In Fig. 6 the predictions of the magnitude of the magnetic moment at the LSF energy landscape minimum are compared between the ML approach and the

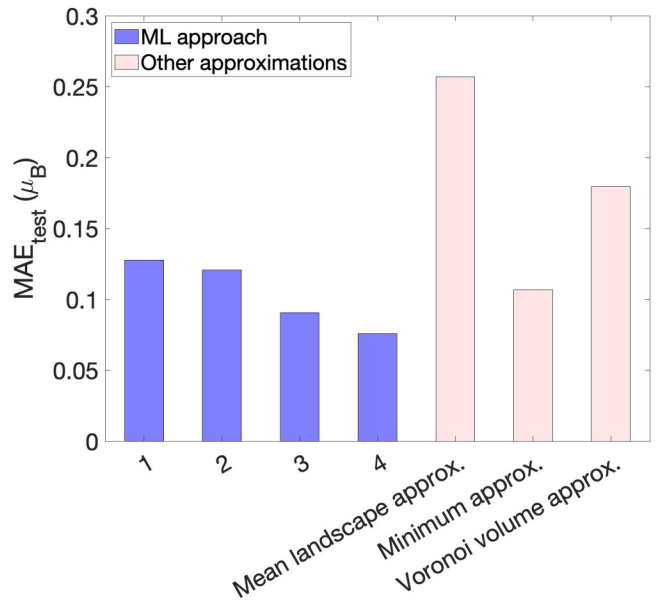


FIG. 6. Mean absolute errors (MAE) when predicting the magnetic moment at the LSF energy landscape minima using any of the four types of descriptors or any of the three other approximation methods described in the text.

three other approximate methods. The predictions of the a and b parameters are improved using the ML models regardless of descriptor. With descriptors 2 to 4, the errors are half (or less) of the errors of both approximations (i) and (ii). In the case of predicting the sizes of the magnetic moments at the minima of the LSF energy landscapes, approximation (iii) slightly outperforms descriptors 1 and 2 whereas 3 and 4 give a marginally smaller error. Descriptor 4 contains the information of the size of the magnetic moment at the minimum along with additional knowledge of the local environment and it is therefore not surprising that it gives a smaller error than approximation (iii). The fact that descriptor 3 also gives a slightly lower error compared to approximation (iii), with the only difference compared to descriptor 2 being summing the scalar products or not, illustrates the importance of how the input data are represented in the descriptor. Summing the scalar products should lead to more similarities being caught by the descriptor, and the result of this improvement is most likely what we see when comparing descriptors 2 and 3.

C. Predictions at Earth's inner core conditions

The ML models are trained on T_c data and, considering that we are using the KRR method, the models should struggle with making predictions of data that are far from the training data. To test the generalizability of the present ML models, we apply them to predict the LSF energy landscapes in Fe under very different conditions from T_c , namely, at Earth's inner core conditions. Studies have indeed shown that magnetic effects could have an important impact even at these conditions [26,35,54], which provides a perfect test system for the ML models developed. In particular, in the following, we will compare models not trained on any HTHP data with

models where we use a transfer learning approach to fine tune the model using a small set of such data. Here, descriptor 3 in Table I is used. One should note that descriptor 4 cannot be expected to give the same improvement in performance in the HTHP case as observed for the T_c system since the magnitudes of the magnetic moments vanish under such conditions in straightforward DFT calculations [35]. The bcc structure is one of several proposed alternatives for Fe at Earth's inner core (other alternatives are hexagonal close packed and fcc), which has recently been predicted to be stable under this high temperature and pressure [50,55,56]. Belonoshko *et al.* demonstrated that it is complicated to accurately study the bcc phase under these conditions because of its superionic behavior and, as a consequence, it is required to use large supercells for the bcc structure to be thermodynamically stable under the conditions of Earth's inner core [50,55]. Our $3 \times 3 \times 3$ supercell with 54 atoms does not satisfy this requirement; we will therefore restrain ourselves from making any quantitative comparison except regarding the change in total pressure, a quantity that should be relatively robust with respect to changes in the supercell size.

In Fig. 7, scatter plots to compare predicted and DFT calculated a and b parameter values for four different combinations of training data and test data content are shown. As expected, when trained only on the T_c data, the ML models struggle to make correct predictions for both the a and b parameter values of the HTHP system. The local environments of the atomic moments in the HTHP system are clearly too different compared to the T_c system for the ML models to accurately predict the a and b parameter values in this case. On the other hand, when including the small amount of HTHP data into the training set, there is a significant improvement in the predictions of the ML models. This notable reduction in MAE suggests a high level of generalizability of the ML models. In Fig. 7, there is a clear separation between two regions of a parameters of the HTHP (high values) and T_c conditions (low values), which is the signature of the transition from an energy-driven localized moment situation to an itinerant situation at high temperature and high pressure where the moments are entropy driven. A similar separation can be seen for the b values, where the ones corresponding to the HTHP system are grouped together with the higher b values of the T_c domain. Because of the small amount of HTHP data, the MAEs are, however, still worse (about 2.5 times larger) for the HTHP system compared to the T_c system even after including a small part of this type of data in the training set.

D. Applying machine learning LSF to ASD-AIMD simulations

The ASD-AIMD-MLLSF simulations are conducted for both the T_c system and the HTHP system. Again, descriptor 3 is used. For the ASD-AIMD-MLLSF simulation of the HTHP system, the ML models have been trained on the complete training set, i.e., including the HTHP data. In these simulations, the ML models predict the a and b parameter values of the LSF energy landscape of each atomic magnetic moment. From the energy landscapes, the magnitudes of the magnetic moments can be calculated (see Sec. II A) following the scheme described in Fig. 3.

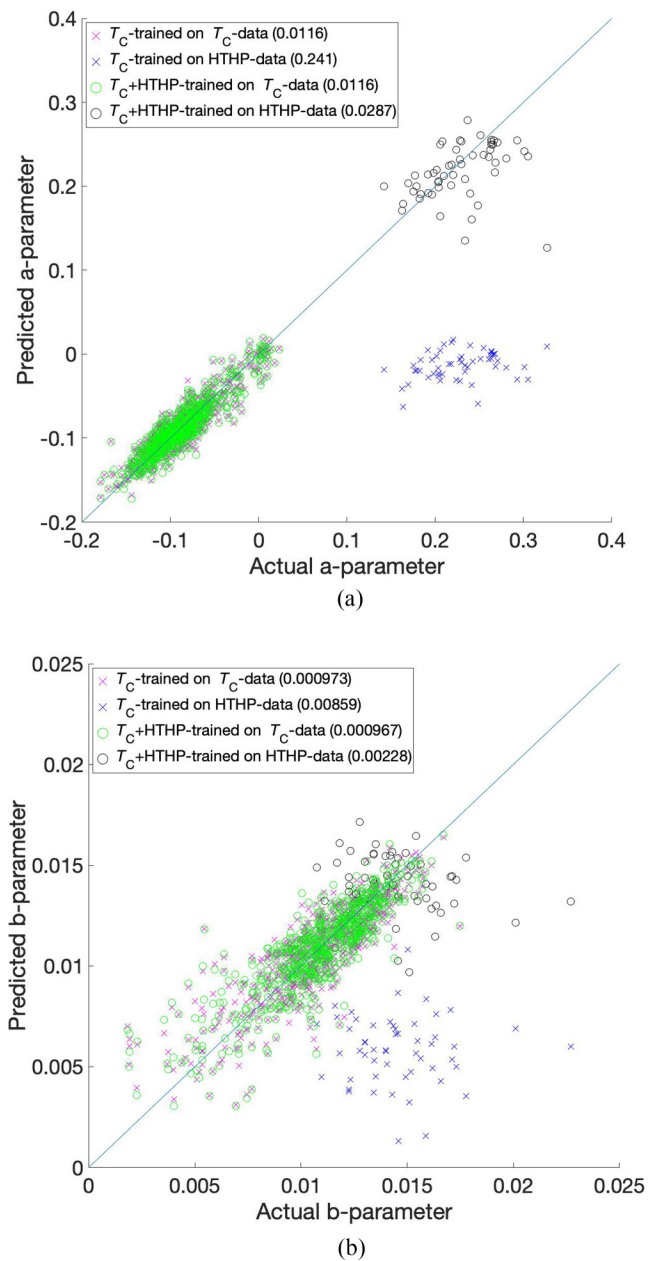


FIG. 7. Comparison between predicted and DFT calculated values of (a) the a parameter and (b) the b parameter. Four different cases are shown: (1) the models are trained only on T_c data and tested on T_c data (magenta crosses); (2) the models are trained only on T_c data and tested on HTHP data (blue crosses); (3) the models are trained on both T_c and HTHP data and tested on T_c data (green circles); (4) the models are trained on both T_c and HTHP data and tested on HTHP data (black circles). In parentheses are the mean absolute errors of the corresponding case with units eV/μ_B^2 and eV/μ_B^4 for the a and b parameter, respectively.

In Fig. 8, the mean-square displacement (MSD) over time is plotted for the T_c system both from the ASD-AIMD-MLLSF simulation and a ferromagnetic AIMD simulation of bcc Fe at T_c . The MSD is more than twice as large in the ASD-AIMD-MLLSF simulation. As concluded in Ref. [3], magnetic disorder has a large impact on the vibrational properties of bcc Fe. The shift upwards is most likely due to

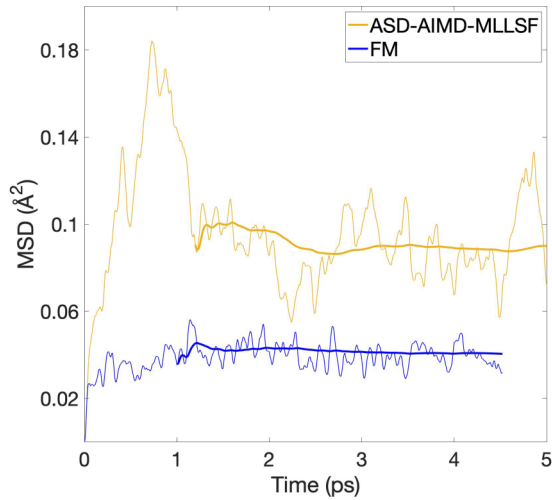


FIG. 8. Mean-square displacement (MSD) over time for bcc Fe at T_c in a ferromagnetic (FM) AIMD simulation and in an ASD-AIMD simulation with the added machine learning step for calculating the longitudinal spin fluctuation energy landscapes (ASD-AIMD-MLLSF) in the paramagnetic phase. Cumulative averages are represented by the thicker lines.

the disordered magnetic moments in the ASD-AIMD-MLLSF simulation and their impact on the interatomic forces above T_c , which allows for larger fluctuations in atomic positions. This demonstrates our ASD-AIMD-MLLSFs possibility to capture the feedback effects these higher amplitude vibrations give on the magnetic moments and structure.

In Fig. 9, the pressure is plotted over time for the HTHP system, both from the ASD-AIMD-MLLSF simulation and a nonmagnetic (NM) AIMD simulation. In the NM-AIMD simulation the pressure is ≈ 320 GPa when the volume is

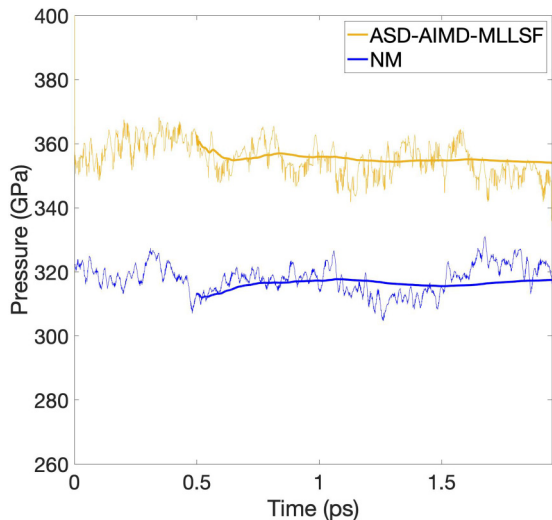


FIG. 9. Pressure over time for bcc Fe at 6000 K and a volume of $7.0 \text{ \AA}^3/\text{atom}$ in a nonmagnetic (NM) AIMD simulation and in an ASD-AIMD simulation with the added machine learning step for calculating the longitudinal spin fluctuation energy landscapes (ASD-AIMD-MLLSF) in the paramagnetic phase. Cumulative averages are shown by the thicker lines.

$7.0 \text{ \AA}^3/\text{atom}$. When adding magnetic moments in the ASD-AIMD simulation (which is made possible by our MLLSF method) the pressure is increased to ≈ 360 GPa, supporting the results from Refs. [26,35]. In Ref. [35], lattice vibrations were not included and a pressure increase of ≈ 50 GPa was observed. In Ref. [26], a pressure increase of ≈ 30 GPa was found when LSF moments were present and lattice vibrations were included by using a single snapshot from a nonmagnetic MD run. One may speculate on the impact of these results on the density of Earth's inner core. The density of pure Fe and Fe-Ni alloys at Earth's core conditions has been found to be higher compared to the preliminary reference Earth model (PREM) [57]. To adjust for this discrepancy, it has been proposed that some additional light elements might be present [58]. The increase in pressure we find when including magnetic moments indicates a lowering in density if the volume was allowed to expand to match the external pressure of 320 GPa. The presence of magnetic moments even when subject to high-temperature vibrations should therefore reduce the density deficit between previous nonmagnetic models for bcc Fe under Earth's inner core conditions and the PREM data.

At this point, we should comment on the efficiency of including the ML models in the ASD-AIMD scheme. If we imagine the calculations needed to determine all LSF energy landscapes of a configuration of 54 atoms being parallelized to the fullest, it would take about 40 minutes to calculate these 54 LSF energy landscapes in the T_c system. For our ASD-AIMD simulation at T_c of approximately 5000 fs we would need more than 3300 h of DFT calculations to determine the LSF energy landscapes. It would take about 10 h to generate the entire training data set from the T_c system used in this work, if we assume the same level of parallelizability as before. At each time step in the ASD-AIMD simulations, predicting the complete set of LSF energy landscapes takes approximately 30 s.

We can do the same comparison for the HTHP system. In this case, it takes a lot longer to generate the individual LSF energy landscapes. We imagine, as before, the calculations being parallelized as much as possible. All in all, it would take about 10.5 h to generate the 54 LSF energy landscapes of the HTHP system used in this work, and it would take 8000 h to calculate all LSF energy landscapes of the ASD-AIMD simulation of 2000 fs at Earth's inner core conditions.

In conclusion, it is clear that the hours of computational time needed to generate the training data are quickly compensated for by using the ML models to predict the LSF energy landscapes.

IV. CONCLUSIONS

We have developed a kernel ridge regression ML approach for determining the shape of the longitudinal spin fluctuation energy landscapes of atomic magnetic moments based on their local environments. Our results show that using an ML approach improves the predictions of longitudinal spin fluctuation energy landscapes compared to other more approximate methods such as, e.g., approximating the energy landscapes to a mean landscape. Our method enables efficient calculations of temperature-dependent magnitudes of the magnetic

moments in dynamic simulations such as AIMD where each magnetic moment at each time step needs to be determined. Our approach thus avoids explicit calculations of thousands of LSF energy landscapes, which would be unfeasible using DFT or other computationally expensive *ab initio* methods.

The approach has been applied both to bcc Fe at T_c and at Earth-core-like conditions of 6000 K and approximately 300 GPa. Our tests show that a model trained only on data at T_c does not extrapolate well to other conditions, which is the expected behavior of this type of ML method. Nevertheless, a transfer-learning approach where a small portion of data (less than 10%) is introduced to cover the HTHP conditions significantly improves the accuracy of the predictions in the HTHP domain. This behavior suggests that the model has a high level of generalizability. However, even when introducing the HTHP data into the training data set, the errors of the HTHP-data predictions are ≈ 2.5 times larger than for the T_c data, which is as expected because of the comparatively scarce training data from the HTHP domain.

In this work, the only crystal structure looked at is bcc. The bcc Fe at Earth's inner core conditions is quite far from perfect bcc crystal but still the ML models are able to train on this data and make reasonable predictions. However, in principle, for the ML models to be able to handle chemical disorder, transition between crystal structures, or even melting, a generalization of how the descriptor is chosen is needed, this is, nonetheless, outside the scope of this work. Since hexagonal close-packed Fe is another suggested crystal structure of Fe under Earth's inner core conditions, an interesting future study would be to do similar simulations for this crystal structure as well.

Including this ML-LSF approach into ASD-AIMD simulations has shown its potential in overcoming difficulties with incorporating both longitudinal and transverse magnetic degrees of freedom and vibrations in theoretical simulations. As pointed out by the example of the ASD-AIMD-MLLSF simulation at the condition of Earth's inner core, allowing for entropy-induced magnetic moments can be of importance when calculating properties such as the density of a mate-

rial. The ASD-AIMD-MLLSF simulation gives an increase of ≈ 40 GPa or 12.5% compared to the nonmagnetic simulation under these HTHP conditions. The ASD-AIMD-MLLSF simulation at T_c demonstrates the difference in vibrational behavior of the atoms between ferromagnetic and paramagnetic Fe at this temperature, where the mean-square displacement is determined to be much larger in the ASD-AIMD-MLLSF simulation. The MSD in the FM simulation is $\approx 0.04 \text{ \AA}^2$ compared to $\approx 0.09 \text{ \AA}^2$ in the ASD-AIMD-MLLSF simulation.

An important possible improvement of the ML models explored in this work is the design of the descriptor. Descriptor 3 in Table I gives the best results without having to make an extra DFT simulation out of the four tested here. However, as mentioned previously, the representation of the input data in the descriptor could be improved to reduce the representation of redundant structural information, which may reduce the errors further. Furthermore, the success of this work motivates investigating if other types of models, e.g., neural-network-based methods, can further increase the performance.

ACKNOWLEDGMENTS

The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) located at National Super Computer Centre (NSC) in Linköping, partially funded by the Swedish Research Council through Grant Agreement No. 2018-05973. B.A. acknowledges financial support from the Swedish Research Council (VR) through Grant No. 2019-05403 from the Swedish Government Strategic Research Area in Materials Science on Functional Materials at Linköping University (Faculty Grant SFOMatLiU No. 2009-00971), and from the Knut Alice Wallenberg Foundation (Wallenberg Scholar Grant No. KAW-2018.0194), as well as support from the Swedish Foundation for Strategic Research (SSF) through the Future Research Leaders 6 program, FFL 15-0290. R.A. acknowledges financial support from the Swedish Research Council (VR) Grant No. 2020-05402 and the Swedish e-Science Research Centre (SeRC).

-
- [1] T. Hickel, A. Dick, B. Grabowski, F. Körmann, and J. Neugebauer, *Steel Res. Int.* **80**, 4 (2009).
 - [2] S. Bhatti, R. Sbiaa, A. Hirohata, H. Ohno, S. Fukami, and S. N. Piramanayagam, *Mater. Today* **20**, 530 (2017).
 - [3] F. Körmann, A. Dick, B. Grabowski, T. Hickel, and J. Neugebauer, *Phys. Rev. B* **85**, 125104 (2012).
 - [4] B. Alling, F. Körmann, B. Grabowski, A. Glensk, I. A. Abrikosov, and J. Neugebauer, *Phys. Rev. B* **93**, 224411 (2016).
 - [5] I. A. Abrikosov, A. Ponomareva, P. Steneteg, S. Barannikova, and B. Alling, *Curr. Opin. Solid State Mater. Sci.* **20**, 85 (2016).
 - [6] S. Blundell, *Magnetism in Condensed Matter* (Oxford University Press, Oxford, 2001)
 - [7] T. Moriya, *Spin Fluctuations in Itinerant Electron Magnetism*, Springer Series in Solid-State Sciences (Springer, Berlin, 1985).
 - [8] B. L. Gyorffy, A. J. Pindor, J. Staunton, G. M. Stocks, and H. Winter, *J. Phys. F: Met. Phys.* **15**, 1337 (1985).
 - [9] P. Steneteg, B. Alling, and I. A. Abrikosov, *Phys. Rev. B* **85**, 144404 (2012).
 - [10] P.-W. Ma, C. H. Woo, and S. L. Dudarev, *Phys. Rev. B* **78**, 024434 (2008).
 - [11] D. Perera, M. Eisenbach, D. M. Nicholson, G. M. Stocks, and D. P. Landau, *Phys. Rev. B* **93**, 060402(R) (2016).
 - [12] D. Perera, D. M. Nicholson, M. Eisenbach, G. M. Stocks, and D. P. Landau, *Phys. Rev. B* **95**, 014431 (2017).
 - [13] I. Stockem, A. Bergman, A. Glensk, T. Hickel, F. Körmann, B. Grabowski, J. Neugebauer, and B. Alling, *Phys. Rev. Lett.* **121**, 125902 (2018).
 - [14] T. Gilbert, *IEEE Trans. Magn.* **40**, 3443 (2004).
 - [15] A. V. Ruban, S. Khmelevskiy, P. Mohn, and B. Johansson, *Phys. Rev. B* **75**, 054402 (2007).
 - [16] K. K. Murata and S. Doniach, *Phys. Rev. Lett.* **29**, 285 (1972).
 - [17] M. Uhl and J. Kübler, *Phys. Rev. Lett.* **77**, 334 (1996).

- [18] N. M. Rosengard and B. Johansson, *Phys. Rev. B* **55**, 14975 (1997).
- [19] L. M. Sandratskii, *Phys. Rev. B* **78**, 094425 (2008).
- [20] P.-W. Ma and S. L. Dudarev, *Phys. Rev. B* **86**, 054416 (2012).
- [21] A. L. Wysocki, J. K. Glasbrenner, and K. D. Belashchenko, *Phys. Rev. B* **78**, 184419 (2008).
- [22] Z. Dong, W. Li, D. Chen, S. Schönecker, M. Long, and L. Vitos, *Phys. Rev. B* **95**, 054426 (2017).
- [23] F. Pan, J. Chico, A. Delin, A. Bergman, and L. Bergqvist, *Phys. Rev. B* **95**, 184432 (2017).
- [24] S. Khmelevskiy, *J. Magn. Magn. Mater.* **461**, 14 (2018).
- [25] A. V. Ruban and O. E. Peil, *Phys. Rev. B* **97**, 174426 (2018).
- [26] D. Gambino, M. Arale Brännvall, A. Ehn, Y. Hedström, and B. Alling, *Phys. Rev. B* **102**, 014402 (2020).
- [27] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, *npj Comput. Mater.* **5**, 83 (2019).
- [28] G. R. Schleder, A. C. M. Padilha, C. M. Acosta, M. Costa, and A. Fazzio, *J. Phys.: Mater.* **2**, 032001 (2019).
- [29] S. Sanvito, C. Oses, J. Xue, A. Tiwari, M. Zic, T. Archer, P. Tozman, M. Venkatesan, M. Coey, and S. Curtarolo, *Sci. Adv.* **3**, e1602241 (2017).
- [30] T. Long, N. M. Fortunato, Y. Zhang, O. Gutfleisch, and H. Zhang, *Mater. Res. Lett.* **9**, 169 (2021).
- [31] I. Novikov, B. Grabowski, F. Körmann, and A. Shapeev, *npj Comput. Mater.* **8**, 13, 13 (2022).
- [32] H. Yu, C. Xu, F. Lou, L. Bellaiche, Z. Hu, X. Gong, and H. Xiang, *arXiv:2110.00724*.
- [33] F. Faber, A. Lindmaa, O. A. von Lilienfeld, and R. Armiento, *Int. J. Quantum Chem.* **115**, 1094 (2015).
- [34] F. A. Faber, A. Lindmaa, O. A. von Lilienfeld, and R. Armiento, *Phys. Rev. Lett.* **117**, 135502 (2016).
- [35] A. V. Ruban, A. B. Belonoshko, and N. V. Skorodumova, *Phys. Rev. B* **87**, 014405 (2013).
- [36] O. Eriksson, A. Bergman, L. Bergqvist, and J. Hellsvik, *Atomistic Spin Dynamics: Foundations and Applications*, 1st ed. (Oxford University Press, Oxford, 2017).
- [37] B. Alling, *Phys. Rev. B* **82**, 054408 (2010).
- [38] A. V. Ruban and I. A. Abrikosov, *Rep. Prog. Phys.* **71**, 046501 (2008).
- [39] B. Skubic, J. Hellsvik, L. Nordström, and O. Eriksson, *J. Phys.: Condens. Matter* **20**, 315203 (2008).
- [40] V. P. Antropov, M. I. Katsnelson, B. N. Harmon, M. van Schilfhaarde, and D. Kusnezov, *Phys. Rev. B* **54**, 1019 (1996).
- [41] V. P. Antropov, M. I. Katsnelson, M. van Schilfhaarde, and B. N. Harmon, *Phys. Rev. Lett.* **75**, 729 (1995).
- [42] G. Kresse and J. Hafner, *Phys. Rev. B* **47**, 558 (1993).
- [43] G. Kresse and J. Hafner, *Phys. Rev. B* **49**, 14251 (1994).
- [44] G. Kresse and J. Furthmüller, *Comput. Mater. Sci.* **6**, 15 (1996).
- [45] G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
- [46] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- [47] G. Kresse and D. Joubert, *Phys. Rev. B* **59**, 1758 (1999).
- [48] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [49] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- [50] A. B. Belonoshko, T. Lukinov, J. Fu, J. Zhao, S. Davis, and S. I. Simak, *Nat. Geosci.* **10**, 312 (2017).
- [51] P.-W. Ma and S. L. Dudarev, *Phys. Rev. B* **91**, 054420 (2015).
- [52] A. S. Christensen, L. A. Bratholm, J. C. Kromann, S. Amabilino, F. A. Faber, B. Huang, A. Tkatchenko, K.-R. Müller, D. R. Glowacki, and O. A. von Lilienfeld, *QML: A Python Toolkit for Quantum Machine Learning* (2017), <https://github.com/qmlcode/qml>.
- [53] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, *J. Machine Learning Res.* **12**, 2825 (2011).
- [54] L. V. Pourovskii, T. Miyake, S. I. Simak, A. V. Ruban, L. Dubrovinsky, and I. A. Abrikosov, *Phys. Rev. B* **87**, 115130 (2013).
- [55] A. B. Belonoshko, J. Fu, and G. Smirnov, *Phys. Rev. B* **104**, 104103 (2021).
- [56] S. Chatterjee, S. Ghosh, and T. Saha-Dasgupta, *Minerals* **11**, 258 (2021).
- [57] A. M. Dziewonski and D. L. Anderson, *Phys. Earth Planet. Inter.*, **25** 297 (1981).
- [58] D. J. Stevenson, *Science* **214**, 611 (1981).