Atomic relaxation around defects in magnetically disordered materials computed by atomic spin constraints within an efficient Lagrange formalism

Omkar Hegde[1,*], Maximilian Grabowski[1,2], Xie Zhang[3], Osamu Waseda,[1] Tilmann Hickel[1],
Christoph Freysoldt[1,*], and Jörg Neugebauer[1]

[1]*Max-Planck-Institut für Eisenforschung, Max-Planck-Strasse 1, 40227 Düsseldorf, Germany*
[2]*Interdisciplinary Centre for Advanced Materials Simulation, Ruhr-Universit Ìat Bochum, 44801 Bochum, Germany*
[3]*Materials Department, University of California, Santa Barbara, California 93106-5050, USA*

Lattice and magnetic degrees of freedom are strongly coupled in magnetic materials. We propose a consistent first-principles framework to explore the joint configurational space. For this, we define atomic spin moments from the projector augmented-wave formalism of density-functional theory and control them via Lagrangian constraints. We demonstrate our approach for vacancy formation and migration in collinear paramagnetic bcc iron by implementing a relaxation scheme based on spin-space averaged forces (SSA relaxation). Based on these results we discuss the impact of the magnetic state on vacancy formation energies, migration barriers, and relaxations.

I. INTRODUCTION

The prediction of materials' properties from first-principles requires a description of their chemistry and structure at the atomic scale. If exposed to real environmental conditions, substantial deviations from the ideal crystal structure will be present: Finite temperatures result in entropic excitations of the electronic, magnetic, and lattice degrees of freedom. In addition, a large variety of crystal defects further destroys their periodic symmetry. The presence and mobility of these defects severely impact the mechanical, chemical, or optical properties of a material. Therefore, it is crucial to predict their formation energies and migration barriers, which are intimately connected to atomic relaxations that occur as a result of broken chemical bonds. For both environmental impacts, which are entropic excitations and defect relaxations, the key ingredients are atomic forces. Methodologically, their determination becomes particularly challenging in magnetic materials since the disorder of magnetic moments—present at finite temperatures—gives rise to complex coupling effects between structural and magnetic degrees of freedom.

Density-functional theory (DFT) has been and is expected to remain the workhorse for these kinds of calculations, as it offers a good compromise between computational feasibility and predictive power [1]. More sophisticated electronic structure methods such as dynamical mean-field theory (DMFT)

generally do not have easy access to atomic forces. While analytic forces in DMFT can be computed within linear-response theory [2] in principle, the extra computational effort for supercell systems (50–100 correlated atoms ⇒ 150–300 force components) seems prohibitive at present. On the other hand, more empirical approaches like the embedded atom methods fail to capture the magnetic degrees of freedom correctly. Originally designed for ground-state properties, DFT has experienced tremendous progress in its application to finite-temperature excitations [3,4] and defects [5] in recent decades.

In the context of *ab initio* thermodynamics, the coupling and mutual interplay of magnetic disorder and vibrational excitations were studied in several recent publications [6–10]. The spin-space averaging (SSA) technique [6], i.e., the super-position of magnetically disordered supercells, is one of the promising approaches to determine paramagnetic phonons. These activities rely on obtaining the total energy for spin configurations from standard DFT calculations. This makes a systematic exploration of the spin-configurational space tedious, when the spin configurations of interest involved in the calculations turn out to be unstable.

The numerical challenges are even larger for the simulation of defects in magnetically disordered materials, demonstrated by the fact that only a few works have addressed this topic so far [11–17]. In the case of extended defects, a couple of studies to calculate stacking faults in the paramagnetic limit were reported [15,16]. The disordered local magnetic moment approximation [18] has been frequently used for the paramagnetic state [16]. The employed electronic structure methods have, however, limitations in computing atomic forces. While atomic relaxations might be less important for stacking faults, they typically cannot be neglected for point defects, dislocations, or grain boundaries.

Most of the methodological developments that involve atomic relaxations next to defects in paramagnetic materials

---

*Corresponding authors; o.hegde@mpie.de; freysoldt@mpie.de

have been tested for vacancies in Fe. The strong focus on this system is triggered by the availability of experiments for this case [19,20]. They indicate that the vacancy activation energy in bcc Fe, i.e., the sum of its formation energy and migration energy, strongly depends on the magnetic state. Several DFT studies [11,21,22] have been able to quantitatively reproduce the change in activation energy between the ferromagnetic and paramagnetic states. However, in some of these cases [21,22] the paramagnetic state was treated with atomic positions obtained from relaxations in the ferromagnetic state. In other cases [23,24], the atomic relaxation results from individual snapshots of magnetically disordered structures; that is, the bulk symmetry is destroyed, and the result strongly depends on a specific local coordination. In a recent work [14] the lattice relaxations have been successfully included within the adiabatic approximation. In this case, the averaging was performed over atomic positions obtained from a dynamic simulation with changing magnetic configurations. The approach indirectly contains a convergence of forces but requires a large number of DFT calculations to achieve this goal.

Within this paper, we outline an alternative, computationally efficient way of considering atomic relaxations in magnetically disordered materials. From a physical point of view, the adiabatic limit is assumed, in which magnetic fluctuations happen on a timescale that is much faster than the motion of atoms. This allows us to employ the SSA [6] approach for the atomic motion that has already been successfully used to describe the impact of magnetic disorder on phonons in Fe [6,7]. In this approach, magnetic configurations are assumed to change so fast that an atom cannot respond to a single magnetic configuration. Therefore, the atom moves according to the averaged force over different magnetic configurations instead of the instantaneous force in each magnetic configuration. In this framework atomic relaxation can be performed by self-consistently computing spin-state averaged forces instead of instantaneous forces.

From a methodological point of view an efficient framework is required to combine electronic-structure calculations, SSA, and atomic relaxations. Most important in this context is the availability of an approach that allows for a robust computation of the atomic forces and energies for any given atomic configuration $\mathbf{R}_a$ and spin-state configuration $\mathbf{S}_a$. Here, $\mathbf{R}_a$ denotes the usual atomic coordinates in real space, and $\mathbf{S}_a$ denotes the atomic spin of each atom. To bring such an approach to success and devise algorithms for characterizing excitations in the joint space, one needs not only a suitable definition of atomic spins but also an implementation that allows for an efficient and reliable evaluation of the energy $E(\mathbf{R}_{\{a\}}, \mathbf{S}_{\{a\}})$ as well as the derivatives with respect to the coordinates, i.e., the atomic forces. Robustness in the DFT algorithm is paramount: A single spin-state averaged atomic relaxation may involve hundreds of $(\mathbf{R}_{\{a\}}, \mathbf{S}_{\{a\}})$ configurations, but excited spin configurations are notoriously difficult to converge because the spin-space energy surface can exhibit many local minima and saddle points at an energy scale that is small compared to typical self-consistency fluctuations. An unexpected failure in a single configuration can ruin the entire relaxation run. In this work, we present such an implementation of spin constraints and of the complex simulation protocol which results in a revised understanding of vacancy energetics in Fe.

The remainder of this work is organized as follows. In Secs. II A–II D, we present our spin-constraint formalism and compare it to previous ones in Sec. II E. Sections II F and II G are devoted to the SSA relaxation scheme. Section III summarizes the computational details before we discuss the application to vacancy in bcc Fe as an example in Sec. IV.

II. METHODOLOGY

The implementation of spin constraints within DFT goes back to the seminal work of Dederichs *et al.* [25], in which constraining magnetic fields were incorporated into an effective Hamiltonian. Established DFT codes such as VASP [26,27] employ this approach by adding a penalty contribution to the total energy expression [28,29]. Unless careful convergence tests ensure a vanishing penalty energy, this contribution makes the comparison of defect energies difficult. The need for convergence tests makes such calculations demanding with respect to both human and computer time. We therefore developed and implemented an alternative approach based on the projector augmented-wave (PAW) [30] formalism of DFT, which is outlined in the following.

A. Definition of atomic spins

The total spin moment

$$\mathbf{M}^{\text{tot}} = \int d^3\mathbf{r}\, \mathbf{M}(\mathbf{r}) \tag{1}$$

as obtained from the magnetization density

$$M_\alpha(\mathbf{r}) = \sum_{\sigma\sigma'} \tau^\alpha_{\sigma\sigma'} \sum_n f_n \psi^*_{n\sigma}(\mathbf{r}) \psi_{n\sigma'}(\mathbf{r}) \tag{2}$$

is a well-defined quantity in spin-polarized DFT. Here, the index α denotes a Cartesian component of the magnetization, τ^α denote the Pauli spin matrices, $\psi_{n\sigma}$ denotes the σ component ($\sigma = \uparrow, \downarrow$) of the nth Kohn-Sham two-component spinor wave function, and f_n is the corresponding occupation number. For readability, we omit the explicit Brillouin zone $\mathbf{k}$ index and the corresponding $\mathbf{k}$-point summation weights relevant for periodic boundary conditions. The sums over n shown in the following will implicitly run over $\mathbf{k}$ points, too.

In the following, we will restrict our discussion to collinear spin-density functional theory, where the spinors can be split into separate wave functions for the spin-up and spin-down channels and the magnetization reduces to the difference between them. We note, however, that the generalization to noncollinear spins is straightforward, as discussed in Sec. II D. With collinear spins, the magnetization reduces to a scalar field

$$M(\mathbf{r}) = \rho^\uparrow(\mathbf{r}) - \rho^\downarrow(\mathbf{r}), \tag{3}$$

with the spin-up $\rho^\uparrow$ and spin-down density $\rho^\downarrow$ defined as

$$\rho^\sigma(\mathbf{r}) = \sum_n f_{n\sigma} |\psi_{n\sigma}(\mathbf{r})|^2 \quad \sigma = \uparrow, \downarrow, \tag{4}$$

respectively.

The total spin moment equation (2) is not sufficient to characterize magnetically disordered materials. For sufficiently localized spins, atomic spin moments can give a viable coarse

graining of spin-excitation space. There are multiple ways of attributing a magnetic moment to atoms in the system. For instance, Bader's concept [31] of "atoms in molecule" provides atomic "basins" Ω^a based on the topology of the electron density, which is used to define an atomic moment as

$$M^a = \int_{\Omega^a} d^3\mathbf{r}\, M(\mathbf{r}). \tag{5}$$

Alternatively, if the wave functions are expanded in atom-centered orbitals (or projected to such an expansion [32]), the concept of Mulliken populations [33,34] can be generalized to atomic spins as

$$M^a = \sum_{\mu\nu}(f_{n\uparrow}c^*_{n\uparrow\mu}S_{\mu\nu}c_{n\uparrow\nu} - f_{n\downarrow}c^*_{n\downarrow\mu}S_{\mu\nu}c_{n\downarrow\nu}), \tag{6}$$

where ν and μ denote atomic basis functions, c is the corresponding expansion coefficients, S is the overlap matrix, and the sum is restricted to basis functions associated with atom a for ν (gross spins) or both indices (net spins). Last, a purely geometric ansatz is to integrate the magnetization density within a sphere for each atom, i.e.,

$$M^a = \int d^3\mathbf{r}\, M(\mathbf{r})\Theta(r^{\mathrm{cut}} - |\mathbf{r} - \mathbf{r}^a|), \tag{7}$$

where $\mathbf{r}^a$ denotes the position of the nucleus and r^{cut} is the cutoff radius [35–37]. Θ is the Heaviside step function but can be replaced by a smooth cutoff function [29].

Each of these definitions has its own advantages and disadvantages. The Bader definition [Eq. (5)] has no arbitrary parameters, but it is intricately linked to the electronic structure. Introducing spin constraints based on this definition may modify the atomic basins and lead to a complex interplay between the spin constraint, atomic basins, and the resulting spin moment. Mulliken spins [Eq. (6)] depend on the chosen basis set but have the huge advantage that they are defined in terms of the atomic orbitals used for the calculation (or the projection), leading to rather simple ways for incorporating constraints into a standard DFT implementation. Last, the spherical definition [Eq. (7)] is easy to implement, but the atomic spins do not sum up to the total spin moment in general. Moreover, an implementation in plane-wave basis sets requires a smoothing of the cutoff to avoid aliasing effects [29].

For the purpose of coarse-graining spin-excitation space, slightly different values of the atomic spin resulting from the above definitions are not critical. Instead, we will regard the resulting atomic spin moments as the signature of a specific spin configuration within the chosen definition. The main purpose of our work is to use these signatures to control to some degree the spin configuration. For instance, by locally minimizing the electronic energy with respect to the atomic spin moments we will arrive at metastable spin configurations that are independent of the chosen definition of atomic spins because the minimization condition sets the associated contributions to the spin-constrained Hamiltonian to zero (see below). But even when we refrain from minimization in order to prepare unstable or transient spin states, we can still quantify the bias on the unconstrained Hamiltonian.

Our main motivation for choosing the specific definition in the following is therefore driven not by physical elegance but mostly by practical concerns for the implementation. In return, we obtain a formalism that allows us to enforce our constraints to high accuracy and, vice versa, to obtain the derivatives with respect to the atomic spins, i.e., the equivalent to atomic forces in spin configurational space from the Lagrangian multipliers. Our DFT code SPHInX [38] adopts the PAW formalism [30]. We therefore choose a definition of atomic spin linked to the partial-wave projection of each atom, namely,

$$M^a = \sum_{ij}\left[D^{a,\uparrow}_{ij} - D^{a,\downarrow}_{ij}\right]\int d^3\mathbf{r}\,\phi_i(\mathbf{r})\phi_j(\mathbf{r})\Theta(r^{\mathrm{cut}} - |\mathbf{r} - \mathbf{r}^a|), \tag{8}$$

with the usual one-center density matrix [30] for atom a

$$D^{a\sigma}_{ij} = \sum_n f_{n\sigma}\langle\psi_{n\sigma}|p_i\rangle\langle p_j|\psi_{n\sigma}\rangle. \tag{9}$$

The ϕ_i in Eq. (8) denote the all-electron partial waves; p_i in Eq. (9) denote the corresponding PAW projectors. For the cutoff radius r^{cut} we choose the corresponding PAW-cutoff radius.

To simplify the notation in the following, we introduce the partial-wave matrix elements of the cutoff-sphere integral

$$\Omega_{ij} = \int d^3\mathbf{r}\,\phi_i(\mathbf{r})\phi_j(\mathbf{r})\Theta(r^{\mathrm{cut}} - |\mathbf{r} - \mathbf{r}^a|). \tag{10}$$

Ω_{ij} have the dimension of a (partial) volume. Introducing the spin sign

$$\chi_\sigma = \begin{cases} +1 & \text{for } \sigma = \uparrow, \\ -1 & \text{for } \sigma = \downarrow \end{cases} \tag{11}$$

allows us to write the atomic moment compactly as

$$M^a = \sum_{ij\sigma}\chi_\sigma D^{a\sigma}_{ij}\Omega_{ij}. \tag{12}$$

B. Spin constraints

To constrain the atomic spin moments defined in Eq. (8), we employ the standard Lagrange formalism [25]; that is, we search for stationary points of the functional

$$\mathcal{L} = F^{\mathrm{el}}[\psi_{\{n\sigma\}}, f_{\{n\sigma\}}]$$
$$- \sum_a \nu_a(M^a[\psi_{\{n\sigma\}}, f_{\{n\sigma\}}] - M^{a,\mathrm{target}}). \tag{13}$$

$M^{a,\mathrm{target}}$ are the target spins, and ν_a are the corresponding Lagrangian parameters. F^{el} denotes the standard electronic free-energy functional [39,40]

$$F^{\mathrm{el}}[\psi_{\{n\sigma\}}, f_{\{n\sigma\}}] = E^{\mathrm{el}}[\psi_{\{n\sigma\}}, f_{\{n\sigma\}}] - TS(f_{\{n\sigma\}}). \tag{14}$$

The notation $\psi_{\{n\sigma\}}$ and $f_{\{n\sigma\}}$ is used to indicate the set of *all* wave functions and *all* occupation numbers entering the calculation, respectively. E^{el} is the usual DFT electronic energy. For the sake of readability, we omit explicit orthonormalization constraints

$$\langle\psi_{n\sigma}|\psi_{n'\sigma}\rangle = \delta_{nn'} \tag{15}$$

(which require the wave functions to be eigenfunctions of the Kohn-Sham Hamiltonian $\hat{H}^{KS}$) and the electron-number constraint

$$N^{el} = \sum_{n\sigma} f_{n\sigma}. \qquad (16)$$

In the standard (unconstrained) approach, the latter defines the Fermi energy μ in the Fermi-Dirac expression for the occupation numbers [39,40]

$$f_{n\sigma} = \frac{1}{1 + \exp[(\epsilon_{n\sigma} - \mu)/k_B T]} \qquad (17)$$

from the eigenvalues

$$\epsilon_{n\sigma} = \langle \psi_{n\sigma} | \hat{H}^{KS} | \psi_{n\sigma} \rangle, \qquad (18)$$

in agreement with the electronic entropy

$$S[f_{\{n\sigma\}}] = -k_B \sum_{n\sigma} [f_{n\sigma} \ln f_{n\sigma} + (1 - f_{n\sigma}) \ln(1 - f_{n\sigma})]. \qquad (19)$$

Other occupation number schemes (with different entropies) are possible, too. We implicitly require the electronic free-energy functional to fulfill these constraints, Eqs. (15) and (16).

To make the spin-constrained functional $\mathcal{L}$ stationary, we require

$$\frac{\delta \mathcal{L}}{\delta \langle \psi_n |} = 0, \qquad (20)$$

$$\frac{\partial \mathcal{L}}{\partial f_n} = 0, \qquad (21)$$

$$\frac{\partial \mathcal{L}}{\partial v_a} = 0. \qquad (22)$$

These equations translate into the constrained-spin Kohn-Sham equation [from Eq. (20)]

$$\{\hat{H}^{KS} + \Delta H^v\} | \psi_{n\sigma} \rangle = \epsilon_{n\sigma} | \psi_{n\sigma} \rangle, \qquad (23)$$

with the constraining contribution to the Kohn-Sham Hamiltonian implicitly defined by

$$f_{n\sigma} \Delta H^v | \psi_{n\sigma} \rangle = -\sum_a v_a \frac{\delta M^a}{\delta \langle \psi_{n\sigma} |}, \qquad (24)$$

the usual occupation numbers [Eq. (17)] from Eq. (21) but with the eigenvalues taken from the spin-constraint Kohn-Sham equation (23), and, last, the desired spin constraint

$$M^a = M^{a,\text{target}} \qquad (25)$$

from Eq. (22). The solution of this last equation is discussed in Sec. II C.

By combining the implicit definition of the constraining contribution, Eq. (24), with the explicit definition of the atomic spins in Eqs. (8) and (9), we arrive at

$$\Delta H^v = -\chi_\sigma \sum_a v_a \sum_{ij} |p_i\rangle \Omega_{ij} \langle p_j|$$

$$= -\chi_\sigma \sum_{ij} |p_i\rangle \langle \phi_i | v_a \Theta(r^{cut} - r) | \phi_j \rangle \langle p_j|. \qquad (26)$$

This has the same formal structure as the one-center PAW correction to the PAW Hamiltonian [30]. It can therefore be added as an additional term in the one-center Hamiltonian. The constraining contribution comes also with a clear interpretation: It adds a uniform magnetic field of strength v_a inside the PAW cutoff radius within the one-center correction of the PAW Hamiltonian. The second line of Eq. (26) expresses this idea in the usual PAW formalism as a projection onto the partial-wave expansion (the $|p_i\rangle \cdots \langle p_j|$ term) with the partial-wave matrix element of the field within the cutoff sphere. The spin-dependent sign χ_σ plays the role of the matrix element of the unit magnetic field with the spin. The spins are thus controlled by additional per-atom magnetic fields. At the same time, the correction inherits the spatial smoothness of the PAW projectors, avoiding any sharp features near the cutoff radius.

As the spin constraints incorporate naturally into the PAW Hamiltonian, they leave the electronic minimization routine largely unaffected. However, if the total electronic energy is evaluated from the Harris-Foulkes functional [41–43], the double-counting term must be augmented by

$$E^{dc,v} = -\sum f_{n\sigma} \langle \psi_{n\sigma} | \Delta H^v | \psi_{n\sigma} \rangle$$

$$= \sum_a v_a \sum_{\sigma ij} \chi_\sigma D_{ij}^{a\sigma} \Omega_{ij}$$

$$= \sum_a v_a M^a. \qquad (27)$$

C. Determining values for the Lagrange multipliers

The key step toward a self-consistent algorithm is the determination of the magnetic fields v_a. Unfortunately, there is no viable explicit expression for v_a for a given set of $M^{a,\text{target}}$ because v_a enter the Hamiltonian and thus the eigenvalue equation (23), which modifies the one-center density matrix $D_{ij}^{a\sigma}$ via both the wave functions $\psi_{n\sigma}$ and the corresponding eigenvalues $\epsilon_{n\sigma}$. Instead, for a given density $\rho^\uparrow(\mathbf{r}), \rho^\downarrow(\mathbf{r})$ and thus for a given standard Kohn-Sham Hamiltonian $H^{KS}[\rho^\uparrow(\mathbf{r}), \rho^\downarrow(\mathbf{r})]$, we solve the set of equations (25) iteratively as explained in the following. For this, after obtaining the wave functions $\tilde{\psi}$ for a given set of atomic magnetic fields $\tilde{v}_{\{a\}}$ from standard iterative diagonalization [38], we rediagonalize the subspace Hamiltonian [44]

$$H_{mm'}^\sigma = \langle \tilde{\psi}_{m\sigma} | \hat{H}^{KS} + \Delta H^v | \tilde{\psi}_{m'\sigma} \rangle \qquad (28)$$

for a new set of atomic magnetic fields

$$v_a = \tilde{v}_a + \Delta v_a. \qquad (29)$$

The subspace Hamiltonian can be easily obtained from the original eigenvalues $\tilde{\epsilon}_{m\sigma}$ as

$$H_{mm'}^\sigma = \tilde{\epsilon}_{m\sigma} \delta_{mm'} + \sum_a -\chi_\sigma \Delta v_a \Omega_{mm'\sigma}^a, \qquad (30)$$

with the per-atom PAW partial volume subspace matrix [45]

$$\Omega_{mm'\sigma}^a = \sum_{i,j\, @a} \langle \tilde{\psi}_{m\sigma} | p_i \rangle \Omega_{ij} \langle p_j | \tilde{\psi}_{m'\sigma} \rangle. \qquad (31)$$

We then obtain the corresponding unitary transformation U_{mn}^σ consisting of the subspace eigenvectors as columns. From

these and the new occupation numbers $f_{n\sigma}$ obtained from the subspace Hamiltonian eigenvalues according to Eq. (17) and enforcing particle-number conservation by an appropriate value of the Fermi energy μ, we compute new magnetic moments

$$M^a(\Delta\nu_{\{a\}}) = \sum_{n\sigma} \chi_\sigma f_{n\sigma} \sum_{mm'} (U^\sigma_{mn})^* \Omega^a_{mm'\sigma} U^\sigma_{m'n}. \qquad (32)$$

These values will, in general, disagree with the target values. To find the set of optimal correction fields $\Delta\nu_a$ that minimize the deviation in the least-squares sense, we set up the minimization problem

$$\mathcal{D}(\Delta\nu_{\{a\}}) = \frac{1}{2}\sum_a [M^a(\Delta\nu_{\{a\}}) - M^{a,\text{target}}]^2, \qquad (33)$$

with the gradient

$$g_a(\Delta\nu_{\{a\}}) = \frac{\partial\mathcal{D}}{\partial\Delta\nu_a} = M^a(\Delta\nu_a) - M^{a,\text{target}}. \qquad (34)$$

This minimization problem can be solved iteratively with a conjugate-gradient scheme to a desired accuracy, typically

$$\frac{1}{N_a}\sum_a |M^a - M^{a,\text{target}}|^2 \leqslant 10^{-16}. \qquad (35)$$

We find a quick convergence for our systems, within a few tens of steps in the beginning of the self-consistent iterations and even fewer toward the end. This is not too surprising: We expect *a priori* the change in each atom's magnetic moment to be dominated by the change in the on-site magnetic field (weak interatomic coupling). Moreover, the on-site susceptibility $\partial^2 M^a / \partial\nu_a^2$ should be of the same order of magnitude at each constrained site. This leads to a diagonally dominant, well-conditioned minimization problem.

To summarize the algorithm for the spin constraints, the following steps are performed:

(1) For a given density $\{\rho^\uparrow_{\text{in}}, \rho^\downarrow_{\text{in}}\}^{(i)}$ in step i and a given set of $\nu^{(i)}_{\{a\}}$, set up the constrained-spin Hamiltonian equations (23) and (27).

(2) Solve the spin-constrained Kohn-Sham equation (23) by iterative diagonalization, yielding eigenfunctions $\tilde{\psi}^{(i)}_{n\sigma}$ and eigenvalues.

(3) Within the subspace formed by these eigenfunctions, solve the minimization problem (33) by a conjugate-gradient algorithm on $\Delta\nu_a$.

(4) For the resulting $\Delta\nu^{(i)}_a$, rotate the wave functions into eigenfunctions

$$\psi^{(i)}_{n\sigma} = \sum_m \tilde{\psi}^{(i)}_{m\sigma} U^\sigma_{mn} \qquad (36)$$

and update the initial atomic magnetic field for the next step

$$\nu^{(i+1)}_a = \nu^{(i)}_a + \Delta\nu^{(i)}_a. \qquad (37)$$

(5) Obtain the output density

$$\rho^{\sigma,(i)}_{\text{out}}(\mathbf{r}) = \sum_n f_{n\sigma} |\psi_{n\sigma}(\mathbf{r})|^2 \qquad (38)$$

from the updated wave functions and occupation numbers.

(6) Obtain a new input density $(\rho^\uparrow_{\text{in}}, \rho^\downarrow_{\text{in}})^{(i+1)}$ by a standard density-mixing scheme.

(7) Check for convergence in density and total energy; otherwise, return to step 1.

By making a full search for the atomic magnetic field in step 3 for each density, we ensure that the output density depends only on the input density and not on the initial magnetic fields $\nu^{(i)}_{\{a\}}$. This is essential for maintaining the convergence-accelerating effect of advanced mixing schemes such as direct inversion in the iterative subspace, which rely on a one-to-one mapping of input and output density and no "hidden" parameters. This approach can be seen as being in full analogy to the search for the Fermi energy μ, even though the algorithm for determining the atomic magnetic fields is conceptually and computationally more involved.

D. Generalization to noncollinear spin

We would like to briefly comment on the generalization to noncollinear spins. It has not yet been implemented here but does not require dramatic changes to the structure of the equations or the algorithm. We present these thoughts here to simplify and encourage the implementation of PAW-based spin constraints in other codes. In the noncollinear case, the electron spin is no longer a good quantum number, and the electronic states become two-component spinors. The wave functions therefore lose their explicit spin index $\psi_{n\sigma}(\mathbf{r}) \to \psi_n(\mathbf{r}, \sigma)$, while the spin-dependent on-site matrices become matrices also in spin space, e.g., $D^{a\sigma}_{ij} \to D^{a\sigma\sigma'}_{ij}$. The magnetization and local magnetic field become three-dimensional vectors; for each component the spin sign χ_σ must be replaced by the corresponding Pauli matrix.

For the eigenvalues and subspace Hamiltonians discussed in Sec. II C, a similar transformation occurs. The explicit spin index is removed, $\epsilon_{n\sigma} \to \epsilon_n$, $H^\sigma_{mm'} \to H_{mm'}$. The subspace partial volume matrix [Eq. (31)] becomes a matrix in spin space $\Omega^a_{mm'\sigma} \to \Omega^a_{mm'\sigma\sigma'}$, but it will be advantageous to store it combined with the Pauli matrices $P^\alpha_{\sigma\sigma'}$, i.e., the transformed entity

$$\chi_\sigma \Omega^a_{mm'\sigma} \to \Omega^a_{mm'\alpha} = \sum_{\sigma\sigma'} P^\alpha_{\sigma\sigma'} \Omega^a_{mm'\sigma\sigma'}. $$

E. Comparison to previous approaches

The most commonly employed alternative to our collinear spin constraints within the PAW formalism is the constrained DFT formalism for noncollinear magnetism by Ma and Dudarev [29]. This formalism uses noncollinear spins and aims at constraining the direction (but not the magnitude) of magnetic moments. It therefore can be used to set up a collinear spin configuration, too. By carefully comparing the working equations, we see that the converged result of the Ma and Dudarev approach corresponds to a constraint solution that is fully consistent with our approach. However, the actual moments from the Ma-Dudarev approach do not exactly correspond to the prescribed targets. This can be most easily seen from the structure of their constraining magnetic field (see Eq. (7) in Ref. [29], adapted in part to our notation)

$$\mathbf{b}^a(\mathbf{r}) = -\lambda\left(\frac{\mathbf{M}^a}{|\mathbf{M}^a|} - \mathbf{e}^a\right)F^a(\mathbf{r} - \mathbf{r}^a), \qquad (39)$$

where F^a is their equivalent of our cutoff sphere for atom a, $\mathbf{e}^a$ is the intended direction of the magnetic moment, and λ is a tunable input parameter a bit misleadingly denoted as the "Lagrange parameter." If the magnetic moment pointed exactly in the desired direction, the constraining field would vanish. In general, this is not the case, and a small deviation is needed to produce the field. In our formalism, we would arrive at exactly the same field

$$\nu_a = -\lambda_a\left(\frac{\mathbf{M}^a}{|\mathbf{M}^a|} - \mathbf{e}^a\right) \tag{40}$$

for constraining the moments at their *actual, off-target* values $\mathbf{M}^a \neq |\mathbf{M}^a|\mathbf{e}^a$. In practice, the Ma-Dudarev approach as implemented can be seen as biasing the direction of the magnetic moment toward a desired target direction with a bias strength proportional to λ. Alternatively, it may be seen as a constrained DFT formalism that does not allow direct control over the direction. To arrive at a specific moment, one would have to set the target direction slightly away. This was recognized by the original authors, and they showed that the deviation can be made arbitrarily small by increasing the λ parameter. However, this comes at the expense of stability in the convergence properties because small fluctuations in direction can then introduce huge fields. Our approach does not have this uncertainty, but we pay for this by having to explicitly search for the desired constraining field strengths ν_a in each step of the self-consistent iteration. In terms of stability, however, this makes a huge difference: Our exact constraint formalism projects out the spin degrees of freedom from the outer loop and treats them entirely in the (comparatively cheap) inner loop to search for the constraining fields, while the Ma-Dudarev approach hardens the spin-directional degrees of freedom within the outer loop. To arrive at a well-conditioned self-consistent problem, the latter approach must be carefully balanced, and one must compromise between numerical stability and the accuracy of the final result.

More recently, Cuadrado *et al.* introduced a fully Lagrangian approach with per-atom Lagrangian multipliers to constrain the direction of magnetization in noncollinear relativistic DFT [34] in an atomic-orbital based code (SIESTA). In their implementation, they use the Mulliken population based definition of atomic spins. In contrast to our procedures, however, the Lagrangian multipliers are optimized in an outer loop, while the inner loop involves the self-consistency iterations at fixed Lagrangian multipliers. We estimate that the optimization algorithms are interchangeable, while a suitable definition of atomic spins is (as explained above) mostly driven by practical considerations for the specific basis set used. Direct comparison of the two implementations is beyond the scope of the present paper.

F. Vacancies in paramagnetic materials

In order to consider a magnetically fully disordered material, i.e., the paramagnetic limit, the spin constraints are applied for a random distribution of magnetic moments that sum up to a total magnetic moment of zero. To achieve an efficient sampling of spin configurational space that mimics true random orientations in a limited supercell, we employ special quasirandom structures (SQSs) [46,47] for an equal number of collinear [48] spin-up and spin-down atoms. The use of collinear SQSs to represent magnetic disorder is well established: Gyorffy *et al.* [48] showed that the local moments can be assumed to be collinear by assuming complete disorder between local moments and neglecting the spin-orbit coupling. The SQS approach has been extended for magnetic disorder [47,49] and has been successfully employed in numerous previous works [6,47,49,50].

Generating a vacancy in such a bulk SQS is connected to a specific magnetic configuration for the adjacent atoms that cannot represent the full disorder. Therefore, a superposition of different magnetic structures in the vicinity of the vacancy needs to be considered. For this purpose, one can average over different SQSs having the vacancy always at the same position, or one can stick to the same SQS and remove the atom at all structurally equivalent positions. The latter approach ensures the completeness of the set of local spin structures and therefore the translational symmetry for a given supercell, similar to the spin-wave method [11,21]. In addition, previous studies [50] confirm that different magnetic SQSs with different pair correlation functions yield similar energies and equilibrium volumes in bcc Fe. Hence, we do not expect a significant dependence of vacancy energies on the choice of SQS. Further, forces on atoms are symmetrized in each structure using the point group symmetries of the crystal structure.

When a vacancy is created by removing an atom in the bulk SQS, a residual moment arises which is equal in magnitude and opposite in direction to the spin of the removed atom. Such a residual moment can be attributed as the moment of the vacancy. However, within our formalism, since the average over many magnetic configurations with either positive or negative residual spins is considered, the vacancy moment is zero.

G. Simulation protocol

The DFT calculations are performed for each vacancy configuration in the given SQS using the electronic-structure code SPHInX [38]. This allows us to benefit from the method for spin constraints outlined in Sec. II B using the algorithm given in Sec. II C. After completing the electronic minimization, the Hellmann-Feynman forces are spin-space averaged following the SSA scheme [6] originally developed for paramagnetic phonons. The SSA forces are defined as [6]

$$\boldsymbol{F}^i_{\text{SSA}} = \sum_{K=1,N} p_K \boldsymbol{F}^i_K, \tag{41}$$

where $\boldsymbol{F}^i_K$ is the Hellmann-Feynman force on the ith atom in the Kth magnetic configuration, N is the total number of magnetic configurations, and $\boldsymbol{F}^i_{\text{SSA}}$ is the spin-space averaged force on the ith atom. p_K denotes the Boltzmann weight for the Kth magnetic configuration, given as $p_K = \exp[-E^{BO}_K/k_BT]/Z$, where E^{BO}_K is the magnetic Born-Oppenheimer energy for the Kth magnetic configuration, k_B is the Boltzmann constant, and Z is the magnetic partition sum. In the present work, we are interested in the fully paramagnetic limit of $T \to \infty$, in which all configurations have the same weight $p_K = 1/N$.

These SSA forces are afterwards symmetrized according to the point group symmetry of the defect structure. This force symmetrization is an efficient way to ensure that with each spin configuration also all symmetrically equivalent configurations are taken into account without additional DFT calculations. In general, defects may show a symmetry-lowering structural instability and symmetrization will suppress the required relaxation. For the present case of the vacancy in bcc Fe, there is no indication of such an instability.

In the next step, the atoms are relaxed according to these spin averaged forces that properly conserve the symmetry. This ensures a relaxation in the adiabatic limit of quickly fluctuating spins and slowly moving atoms. Accordingly, the atomic positions are updated until a convergence with respect to the SSA forces is achieved.

For structural relaxations, the "external structure optimizer" (SxExtOpt) [51] is used for the following reasons: (i) Since the average of forces is performed outside the DFT code, a structure optimizer independent of the potential energy surface calculation is required. (ii) SxExtOpt shows a very fast convergence to the structure of lowest energy as it models the Hessian based on internal coordinates and parametrizes it on the fly. The fast convergence is important because the averaging over different spin configurations makes the calculation significantly more expensive than a standard DFT calculation.

Using the new positions obtained from SxExtOpt as input for the next ionic step, DFT calculations are again performed on each magnetic structure. These steps are repeated until the forces converge with desired accuracy. Henceforth, such a relaxation scheme will be referred to as "SSA relaxation." On achieving convergence, the energies of all structures are averaged and combined to obtain the vacancy formation energy [5]

$$E_f = E_{N-1}^{\text{vac}} - \left(\frac{N-1}{N}\right)E_N^{\text{bulk}} \qquad (42)$$

and the migration energy

$$E_m = E_{N-1}^{\text{trans}} - E_{N-1}^{\text{vac}}. \qquad (43)$$

Here, E_{N-1}^{vac}, E_N^{bulk}, E_{N-1}^{trans} are the energies of a vacancy supercell, a bulk system, and a supercell in the transition state, respectively. The migration energies are determined using the activation-relaxation technique [52]. The activation energy E_a is defined as the sum of the vacancy formation energy E_f and the vacancy migration energy E_m.

The algorithm outlined above therefore combines three tools: The DFT code SPHInX, the structure optimizer SxExtOpt, and the force averaging SSA (Fig. 1). To develop and disseminate such a complex simulation protocol, we have used the PYTHON-based framework pyiron [53]. Like an integrated development environment for programming languages, pyiron allows one to interactively implement and test simulation protocols and to upscale them for high-throughput simulations on large computer clusters. One of the advantages of this framework is that the different software tools can run interactively; that is, the communication between them occurs via named pipes on the fly, without terminating one of the codes before starting the next atomic relaxation
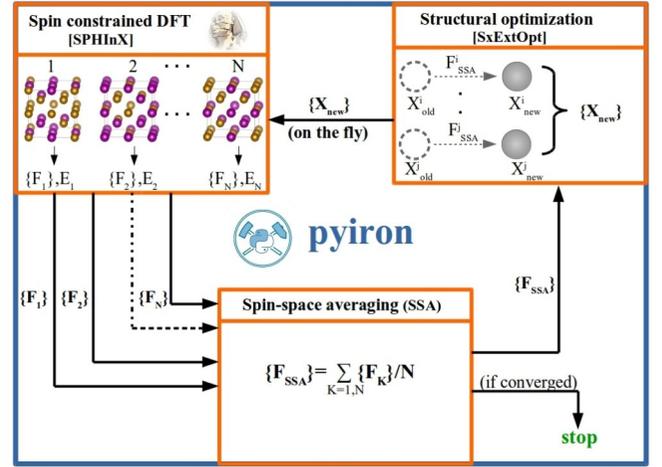

FIG. 1. Schematic representation of the self-consistent simulation protocol for a spin-space averaged relaxation. The DFT code SPHInX, the structure optimizer SxExtOpt, and the SSA force averaging scheme are orchestrated by the p-based framework pyiron [53].

step. This yields a substantial improvement of computational efficiency.

III. COMPUTATIONAL DETAILS

The DFT calculations with the SPHInX package [38] are performed within the framework of the PAW [30,54] formalism, using the Perdew-Burke-Ernzerhof flavor of the generalized gradient approximation [55,56]. A supercell of 54 atoms ($3\times3\times3$ cell) is used for the ferromagnetic (FM) and for the paramagnetic (PM) SQS bulk calculations. An energy cutoff of 600 eV and a Monkhorst-Pack **k**-point mesh of $6\times6\times6$ are used. All calculations are performed with Fermi-Dirac smearing, with a width of 0.1 eV (electronic temperature close to the Curie temperature of 1041 K). We verified that halving the smearing width changed the vacancy formation energy by less than 2 meV. The convergence criterion for the total energy in the electronic relaxation loop is 10^{-5} eV. The convergence criterion for forces in the ionic relaxation loop is 0.015 eV/Å. The lattice parameter obtained from an energy-volume optimization is 2.83 Å for both the FM and the SQS structures. To understand the effect of lattice expansion, a set of calculations is performed for a lattice constant of 2.87 Å in the PM state. This value was deduced using the relative lattice expansion coefficient [57] near the Curie temperature from experiment.

The spin constraints for the PM state (the FM state is performed without constraints) use an atomic moment of $2.1\mu_{\text{B}}$ as obtained from a moment-energy optimization at 0 K. Fluctuations in the magnitudes of magnetic moments due to temperature effects are therefore neglected. The change in the magnitude of magnetic moments for the first nearest-neighbor atoms of the vacancy for each magnetic configuration is taken into account: A number of spin constrained calculations, each with a different moment magnitude for the nearest-neighbor atoms of the vacancy, are carried out, and the ideal magnitude

of $2.3\mu_B$ is determined from the moment-energy optimization. Constraining the nearest neighbors to the bulk value $(2.1\mu_B)$ instead increases the vacancy formation energy by 0.05 eV compared to setting them to the optimum value $(2.3\mu_B)$.

The energy associated with constraining the fields to fixed values can be estimated from the Lagrange multipliers $\nu_a = dF/dM^a$. From the observation that the energy varies quadratically near the optimum value, the energy can be expressed as $\Delta F = 1/2(dF/dM)^2/S$, with the spin stiffness $S = d^2F/dM^2 = d\nu/dM$. A typical rms for the constraining field strength is $\mathrm{rms}(\nu) = 30$ meV$/\mu_B$ with rather little variation for different setups. With a spin stiffness of $S = 550$ meV$/\mu_B^2$ per atom extracted from bulk calculations, we estimate the constraints to raise the energy by about 1 meV per atom. Energy differences will profit from error cancellation if the rms values are the same. Therefore, constraining all the magnitudes to a fixed value is not a major limitation, while it dramatically improves the robustness of the calculations.

A total of eight magnetic structures are taken for the SSA after ensuring that the difference in vacancy formation energies calculated between six and eight magnetic structures is less than 20 meV, while the difference in formation energies calculated with seven and eight structures is less than 5 meV. For consistent comparisons, the same eight magnetic configurations are considered for different relaxation schemes. Since the point-group symmetries of the crystal structure are considered for the symmetrization of forces, the total number of magnetic snapshots is larger (for example, $8 \times 48 = 384$ in the case of the vacancy state). In the case of systems with very few symmetries, a larger number of magnetic configurations will be needed for a sufficient sampling of the paramagnetic state.

IV. RESULTS AND DISCUSSION

Applying the proposed approach to compute atomic forces and perform atomic relaxations in the PM state, we now present the vacancy formation and migration energies in pure bcc Fe summarized in Table I.

A. Impact of magnetic disorder

First, we discuss the impact of magnetic disorder on the overall activation energy. To this end, our *ab initio* values for the FM state are included as a reference in Table I and Fig. 2. They are in good agreement with experiments ($E_a^{\mathrm{FM}} = 3.01$ eV, 2.75 eV) [19,59] and previous theoretical results [21,22]. As highlighted in Fig. 2, the activation energy for vacancy migration turns out to be systematically reduced in the PM state. This is in particular the case for the SSA relaxation proposed in the present paper (last bar in Fig. 2), for which the impact of magnetic disorder is about 1 eV. The strong contribution of the vacancy formation energy to this reduction has previously been investigated and has been connected with the well-known kinks at the Curie temperature in the Arrhenius plots for the diffusion rate in bcc Fe. Here, we note that the absolute reduction in the vacancy migration energy E_m is almost identical to that of the formation energy E_f (~ 0.5 eV, respectively), while the relative change in E_m

TABLE I. *Ab initio* determined energy contributions to the activation energy E_a for self-diffusion in bcc Fe: The vacancy formation energy E_f and the migration barrier E_m for the nearest-neighbor jumps. Different relaxation schemes for the PM state are compared. Unless indicated differently, the calculations have been performed for the theoretical equilibrium lattice constant $a = 2.83$ Å.

	E_f (eV)	E_m (eV)	E_a (eV)
This work			
FM	2.15	0.74	2.89
PM (SSA relaxed)	1.62	0.30	1.92
PM (unrelaxed)	1.94	0.63	2.57
PM (FM relaxed)	1.71	0.50	2.21
PM (individually relaxed)	0.99	0.45	1.44
PM (SSA, $a = 2.87$ Å)	1.89	0.35	2.24
Experimental values for PM			
Matter *et al.* [58]	1.60		
De Schepper *et al.* [20]	1.79		
Iijima *et al.* [19]			2.61
De Schepper *et al.* [59]			2.37
Other theoretical values for PM			
Gambino and Alling [14] ($a = 2.84$ Å)	1.61		
Ruban and Razumovskiy [11] ($a = 2.90$ Å)	1.77		
Sandberg *et al.* [22]	1.54	0.40	1.97
Ding *et al.* [21] ($a = 2.90$ Å)	1.98	0.43	2.41

is more than 50% and the magnetic effect is, therefore, much more pronounced for this contribution. Hence, the impact of magnetic disorder on the reduction of activation energy is significant.

We note that the magnetic disorder as provided by the SQS structure is, indeed, required to properly represent the PM
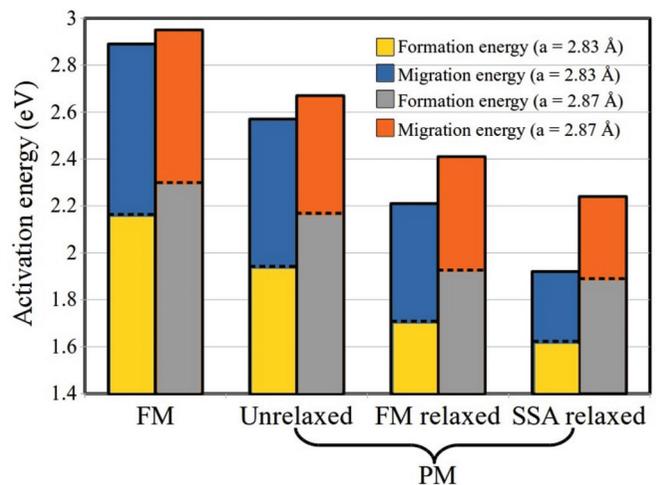

FIG. 2. Comparison of activation energies, which consist of formation energies (lower parts) and migration energies (upper parts) for different relaxation schemes. Calculations performed at the equilibrium lattice constant (yellow and blue bars) and at a high-temperature lattice constant close to the Curie temperature (gray and orange bars) are compared. The same eight magnetic configurations are considered for different relaxation schemes in the PM state.

state. An ordered magnetic configuration with vanishing total magnetization as given by the antiferromagnetic state (AFM) state is not sufficient. We performed vacancy calculations in the AFM state and obtained a negative vacancy formation energy. The negative value could be a result of a more favorable ferromagnetic interaction of the eight nearest-neighbor atoms around the vacancy, once an atom is removed to create the vacancy.

B. Effect of relaxation schemes

The novelty of the present calculations is the SSA relaxation, i.e., the relaxation of atoms according to SSA forces that are determined by employing an efficient algorithm for spin constraints. In the following, we discuss the impact of such a relaxation scheme and compare it with alternative approximations with respect to the formation energy, the migration energy, and the structural displacements.

We first focus on the values for equilibrium lattice constant ($a = 2.83$ Å, the left in each pair of bars in Fig. 2). In the unrelaxed case, each magnetic configuration is frozen in the perfect bcc bulk geometry with an atom in the bulk replaced by a vacancy. Thus, if the relaxation is completely omitted, then the vacancy activation energy is 2.57 eV (second bar in Fig. 2). In this case, a PM average over different spin configurations is still employed but is evaluated only for the vacancy energies and not for the forces. Compared to the FM state, such an approach yields only a reduction of the energy by 0.32 eV, approximately 1/3 of the full impact of magnetic disorder (1 eV) in our approach. This indicates already the importance of relaxations in the PM state.

To properly evaluate the impact of PM fluctuations on the relaxation energy, we compare the SSA relaxation with a calculation in which the relaxation in the FM state has been conserved and the spin-state averaging is thus limited to a single structure (third bar in Fig. 2). In other words, each magnetic configuration is frozen in the geometry that is obtained as a result of relaxations in the FM state. The vacancy activation energy is, in this case, 2.21 eV. We can conclude that 3/4 of the full energy reduction (1 eV) is due to a direct impact of magnetic disorder, whereas the remaining 1/4 can be attributed to a proper treatment of magnetism for relaxations. This contribution is not equally distributed: It is only 1/6 for the vacancy formation energy, while the change in relaxation due to the PM state is responsible for 1/2 of the corrections for the migration energy. We note that a relaxation in the FM state has often been employed in the literature [21,22] as an approximation for the energetics in the PM state.

Alternative algorithms suggested in the literature [23,24] correspond to a relaxation of the atomic positions for individual disordered magnetic structures (fifth row in Table I). Though the energies are, again, averaged over spin configurations, this relaxation scheme does not decouple the timescales of atomic and spin motions adiabatically. Instead, it assumes that any given magnetic configuration is sufficiently long-lived that the atomic coordinates can adapt to it. In this formalism, the resulting vacancy energies are obtained by averaging over the total energies of these relaxed supercells. Since each calculation corresponds to a full structural mini-
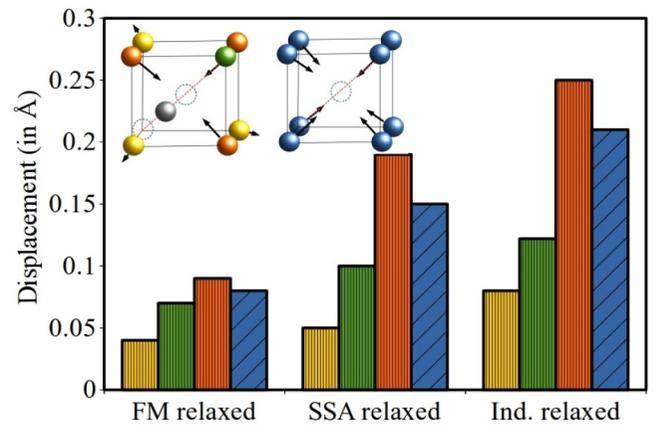

FIG. 3. Displacements of atoms in the nearest-neighbor shell of the vacancy in the vacancy state (blue bars) and in the transition state (yellow, green, and orange bars). Three different relaxation schemes for the PM state of bcc Fe are compared. Colors of atoms shown in the schematic are linked to their displacements in the bar diagram.

mization, we find that the activation energy value (1.44 eV) is significantly lower than in previous cases. The large decrease in activation energy is mainly caused by the reduced vacancy formation energy, which is about 0.6 eV lower than the corresponding energy obtained with the SSA relaxation scheme. The energy reduction is expected to be particularly large because employing individual relaxations for a highly symmetric defect configuration such as a single vacancy state causes symmetry breaking, leading to a larger relaxation of atoms into the vacancy. In contrast to this, the configuration obtained from the SSA relaxation is minimized with respect to the averaged forces over different spin configurations. This means that before an atom can respond to the forces of a single magnetic state, a new state appears and modifies these forces. Thus, in the adiabatic limit an atom "sees" only the averaged force over different magnetic configurations. Since this effective force is a compromise of all individual forces, the obtained minimum structure does not represent the minimization of any of them and is thus higher in energy.

The difference in activation energies underlines the substantial difference of the timescales for the magnetic and atomic degrees of freedom, thereby justifying the adiabatic decoupling between them. Furthermore, the formation energy (1.62 eV) obtained with our method at the equilibrium lattice constant (2.83 Å) is in good agreement with that reported by Gambino and Alling [14] (1.61 eV, lattice constant = 2.84 Å), where the position averaging was carried out in the adiabatic limit.

It is important to note that although the vacancy formation energy is lowest for an average of individual relaxations, the migration energy is the lowest for the SSA relaxation in the adiabatic limit, as the relaxations in both the vacancy state and the transition state affect the migration energy calculation.

The resulting displacements of atoms next to the defect for the vacancy and the migration state are visualized in Fig. 3. It displays the displacement of the first-nearest-neighbor atoms around the vacancy for each relaxation scheme. In the

low-symmetry transition state, three classes of neighboring atoms exist: (i) the atoms marked in yellow experience a net repulsion away from the vacancies since they are closest to the transition site, (ii) the atom marked in green is the nearest neighbor of the vacancy opposite to the transition site, and (iii) the orange-marked atoms experience an attractive force from two vacancies and simultaneously benefit from the reduced displacements of the other atoms and are further shifted toward the vacancy.

It is interesting to note that in both the initial and barrier configurations, the SSA relaxation values lie between the FM relaxed scheme and the individual relaxation scheme. The observation that the paramagnetic (or SSA) relaxations for vacancy activation are larger than the ones in the FM state agrees with observations from tracer diffusion experiments by Iijima *et al.* [19].

C. Inclusion of volume expansion

The results discussed in the previous section were obtained for supercells with the $T = 0$ K equilibrium lattice constant, 2.83 Å. However, at temperatures where the system becomes paramagnetic a sizable volume expansion occurs. To include these effects and understand how they change vacancy energetics, we repeated calculations and analysis for an expanded lattice constant of 2.87 Å for the PM state as obtained using the experimental relative lattice expansion coefficient [57] near the Curie temperature.

In Fig. 2 we compare the activation energies for the two lattice constants. It can be clearly seen that the activation energies in the case of the expanded lattice constant are substantially higher than their equilibrium lattice constant counterparts. This is largely caused by the increased vacancy formation energy, resulting from a smaller vacancy formation volume. Applying the lattice expansion to the vacancy calculations performed with the SSA relaxation yields an activation energy of 2.24 eV. This value is very similar to the activation energy obtained in the FM relaxation scheme (2.21 eV). This similarity should be regarded as mere coincidence; in a proper treatment the adiabatic decoupling of spin and lattice degrees of freedom as well as volume expansion need to be taken into account.

Table I lists the calculated vacancy formation energy, migration energy, and literature values. Vacancy formation energy values for bcc Fe obtained from different positron annihilation experiments [20,58,60] and activation energies obtained from tracer diffusion experiments [19,59] are included. Both experimental and theoretical values show a large scatter. The vacancy formation energy including volume expansion and calculated from our method (1.89 eV) is within the error range of the experimental data reported by De Schepper [20] (1.79 ± 0.10 eV). In these experiments a sample with higher chemical purity compared to previous positron annihilation experiments was considered. On the other hand, the activation energy as calculated from our SSA relaxation (2.24 eV) is lower than the values obtained from the tracer diffusion experiments in the literature (2.37–2.61 eV).

In the present work we did not consider magnetic short-range order and lattice vibrations since including them in a computationally efficient way would require substantial new developments. Magnetic short-range order effects are known to be small [19,21]. The impact of lattice vibrations will be addressed in a future study.

V. CONCLUSIONS

In the present work, we introduced and described the SSA relaxation scheme. Such a scheme is numerically efficient and robust and has strong physical implications to study atomic relaxations induced by crystal defects in PM materials.

To achieve this goal, a numerically efficient approach to enforce spin constraints that is free of penalty energies had to be developed. The method utilizes projectors in the PAW formalism and uses a Lagrangian scheme that avoids penalty forces. The implementation of this approach allows us to consider large numbers of spin configurations, even if many of them are unstable.

Further, to explore the potential energy surfaces in the joint atomic and spin configurational space with a reasonable computational effort, an effective simulation protocol had to developed and implemented. To this end, we have used the interactive development environment pyiron. Within this environment, we specifically employed the concept of interactively linking simultaneously running atomistic software tools. This allowed us to run, interrupt, and subsequently continue the electronic self-consistency cycle in simultaneously performed DFT calculations. Each DFT calculation treats an individual spin-state configuration. The SSA and the update of positions are performed externally by the pyiron framework.

To benchmark and analyze the SSA relaxation scheme we applied it to compute formation and migration energies of a vacancy in PM bcc Fe. We then compared the results of the SSA approach in the adiabatic spin limit with different approximative relaxation schemes that are commonly employed in the literature. Both considered alternatives, relaxing atomic positions in the FM state and relaxing them independently for various individual magnetic configurations, yield activation energies that show substantial differences compared to the SSA relaxation scheme. This finding clearly underlines the significance of the adiabatic de-coupling of magnetic and atomic degrees of freedom at high temperatures. Furthermore, it is shown that for a comparison between calculated and experimentally measured energies the impact of thermal expansion on the lattice constant has to be included.

The spin-constraint method and the SSA relaxation scheme introduced in this paper pave a numerically efficient way to study the impact of atomic relaxations on thermodynamic and kinetic defect properties in magnetically disordered materials. The proposed approach can be applied to materials science problems, where forces on magnetic atoms are required and where the magnetic degrees of freedoms are much faster than the atomic ones.

[1] R. J. Maurer, C. Freysoldt, A. M. Reilly, J. G. Brandenburg, O. T. Hofmann, T. Björkman, S. Lebègue, and A. Tkatchenko, Annu. Rev. Mater. Res. **49**, 1 (2019).

[2] I. Leonov, V. I. Anisimov, and D. Vollhardt, Phys. Rev. Lett. **112**, 146401 (2014).

[3] T. Hickel, B. Grabowski, F. Körmann, and J. Neugebauer, J. Phys.: Condens. Matter **24**, 053202 (2011).

[4] L.-F. Zhu, B. Grabowski, and J. Neugebauer, Phys. Rev. B **96**, 224202 (2017).

[5] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, Rev. Mod. Phys. **86**, 253 (2014).

[6] F. Körmann, A. Dick, B. Grabowski, T. Hickel, and J. Neugebauer, Phys. Rev. B **85**, 125104 (2012).

[7] F. Körmann, B. Grabowski, B. Dutta, T. Hickel, L. Mauger, B. Fultz, and J. Neugebauer, Phys. Rev. Lett. **113**, 165503 (2014).

[8] P. Steneteg, B. Alling, and I. A. Abrikosov, Phys. Rev. B **85**, 144404 (2012).

[9] I. Stockem, A. Bergman, A. Glensk, T. Hickel, F. Körmann, B. Grabowski, J. Neugebauer, and B. Alling, Phys. Rev. Lett. **121**, 125902 (2018).

[10] B. Dutta, F. Körmann, S. Ghosh, B. Sanyal, J. Neugebauer, and T. Hickel, Phys. Rev. B **101**, 094201 (2020).

[11] A. V. Ruban and V. I. Razumovskiy, Phys. Rev. B **85**, 174407 (2012).

[12] P. Delange, T. Ayral, S. I. Simak, M. Ferrero, O. Parcollet, S. Biermann, and L. Pourovskii, Phys. Rev. B **94**, 100102(R) (2016).

[13] I. Abrikosov, A. Ponomareva, P. Steneteg, S. Barannikova, and B. Alling, Curr. Opin. Solid State Mater. Sci. **20**, 85 (2016).

[14] D. Gambino and B. Alling, Phys. Rev. B **98**, 064105 (2018).

[15] I. Bleskov, T. Hickel, J. Neugebauer, and A. Ruban, Phys. Rev. B **93**, 214115 (2016).

[16] S. Lu, Q.-M. Hu, B. Johansson, and L. Vitos, Acta Mater. **59**, 5728 (2011).

[17] A. Schneider, C.-C. Fu, F. Soisson, and C. Barreteau, Phys. Rev. Lett. **124**, 215901 (2020).

[18] F. J. Pinski, J. Staunton, B. L. Gyorffy, D. D. Johnson, and G. M. Stocks, Phys. Rev. Lett. **56**, 2096 (1986).

[19] Y. Iijima, K. Kimura, and K. Hirano, Acta Metall. **36**, 2811 (1988).

[20] L. De Schepper, D. Segers, L. Dorikens-Vanpraet, M. Dorikens, G. Knuyt, L. M. Stals, and P. Moser, Phys. Rev. B **27**, 5257 (1983).

[21] H. Ding, V. I. Razumovskiy, and M. Asta, Acta Mater. **70**, 130 (2014).

[22] N. Sandberg, Z. Chang, L. Messina, P. Olsson, and P. Korzhavyi, Phys. Rev. B **92**, 184102 (2015).

[23] A. V. Ponomareva, Y. N. Gornostyrev, and I. A. Abrikosov, Phys. Rev. B **90**, 014439 (2014).

[24] A. V. Ponomareva, B. O. Mukhamedov, and I. A. Abrikosov, Phys. Rev. Mater. **4**, 024401 (2020).

[25] P. H. Dederichs, S. Blügel, R. Zeller, and H. Akai, Phys. Rev. Lett. **53**, 2512 (1984).

[26] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).

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

[28] D. Hobbs, G. Kresse, and J. Hafner, Phys. Rev. B **62**, 11556 (2000).

[29] P.-W. Ma and S. L. Dudarev, Phys. Rev. B **91**, 054420 (2015).

[30] P. E. Blochl, Phys. Rev. B **50**, 17953 (1994).

[31] R. W. F. Bader, *Atoms in Molecules: A Quantum Theory* (Clarendon Press, Oxford, 1990).

[32] B. Lange, C. Freysoldt, and J. Neugebauer, Phys. Rev. B **84**, 085101 (2011).

[33] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).

[34] R. Cuadrado, M. Pruneda, A. García, and P. Ordejón, J. Phys. Mater. **1**, 015010 (2018).

[35] G. M. Stocks, B. Ujfalussy, X. Wang, D. M. C. Nicholson, W. A. Shelton, Y. Wang, A. Canning, and B. L. Györffy, Philos. Mag. B **78**, 665 (1998).

[36] M. Eisenbach, B. L. Györffy, G. M. Stocks, and B. Újfalussy, Phys. Rev. B **65**, 144424 (2002).

[37] P. Kurz, F. Forster, L. Nordstrom, G. Bihlmayer, and S. Blugel, Phys. Rev. B **69**, 024415 (2004).

[38] S. Boeck, C. Freysoldt, A. Dick, L. Ismer, and J. Neugebauer, Comput. Phys. Commun. **182**, 543 (2011).

[39] M. J. Gillan, J. Phys.: Condens. Matter **1**, 689 (1989).

[40] C. Freysoldt, S. Boeck, and J. Neugebauer, Phys. Rev. B **79**, 241103(R) (2009).

[41] J. Harris, Phys. Rev. B **31**, 1770 (1985).

[42] W. M. C. Foulkes, Ph.D. thesis, University of Cambridge, 1988.

[43] W. M. C. Foulkes and R. Haydock, Phys. Rev. B **39**, 12520 (1989).

[44] For periodic boundary conditions with implicit inclusion of **k**-point indices in n, the formal all-state subspace Hamiltonian separates into $N_\mathbf{k}$ **k**-dependent subspace Hamiltonians.

[45] We note in passing that $\Omega^a_{mm'\sigma}$ can be efficiently formed on the fly: The computational effort is $\sim N_\mathbf{k} N_n^2 N_p^2$, with N_n bands for each **k** point ($N_\mathbf{k}$ many of them) and N_p projectors per atom.

[46] A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, Phys. Rev. Lett. **65**, 353 (1990).

[47] B. Alling, T. Marten, and I. A. Abrikosov, Phys. Rev. B **82**, 184430 (2010).

[48] B. Gyorffy, A. Pindor, J. Staunton, G. Stocks, and H. Winter, J. Phy. F **15**, 1337 (1985).

[49] A. Dick, T. Hickel, and J. Neugebauer, Steel Res. Int. **80**, 603 (2009).

[50] Y. Ikeda, A. Seko, A. Togo, and I. Tanaka, Phys. Rev. B **90**, 134106 (2014).

[51] C. Freysoldt, Comput. Mater. Sci. **133**, 71 (2017).

[52] N. Mousseau and G. T. Barkema, Phys. Rev. E **57**, 2419 (1998).

[53] J. Janssen, S. Surendralal, Y. Lysogorskiy, M. Todorova, T. Hickel, R. Drautz, and J. Neugebauer, Comput. Mater. Sci. **163**, 24 (2019).

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

[55] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).

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

[57] Z. S. Basinski, W. Hume-Rothery, and A. Sutton, Proc. R. Soc. London, Ser. A **229**, 459 (1955).

[58] H. Matter, J. Winter, and W. Triftshäuser, Appl. Phys. **20**, 135 (1979).

[59] L. De Schepper, G. Knuyt, and L. Stals, J. Phys. Chem. Solids **44**, 171 (1983).

[60] K. Maier, H. Metz, D. Herlach, and H.-E. Schaefer, J. Nucl. Mater. **69**, 589 (1978).