Fully self-consistent calculations of magnetic structure within non-collinear Alexander–Anderson model

A. V. Ivanov[1,2], P. F. Bessarab[2,3], H. Jónsson[2,4], V. M. Uzdin[3]

[1]Saint Petersburg State University, 199034, Saint Petersburg, Russia

[2]Science Institute and Faculty of Physical Sciences, U. of Iceland, 107 Reykjavík, Iceland

[3]ITMO University, 197101 Saint Petersburg, Russia

[4]Department of Applied Physics, Aalto University, FI-00076 Espoo, Finland

ali5@hi.is

An implementation of the non-collinear Alexander–Anderson model for itinerant electrons in magnetic systems is presented where self-consistency is reached for specified directions of the magnetic moments. This is achieved by means of Lagrange multipliers and a variational principle for determining the transverse and longitudinal components of the magnetic moments as well as the average number of d-electrons using direct optimisation. Various optimisation algorithms are compared and the limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm is found to give the best performance. An application to antiferromagnetic Cr crystal is presented where spin-dynamics and curvature of the energy surface are calculated to compare results obtained with and without the constraints on the orientation of the magnetic moments.

1.   Introduction

A model for the description of a magnetic impurity in a non-magnetic metal was presented by Anderson [1] and later extended to describe the interaction between two magnetic atoms by Alexander and Anderson [2]. A multi-impurity extension of this model, allowing for arbitrary orientation of the spins and referred to as non-collinear Alexander–Anderson (NCAA) model [3–5] has been developed and used in studies of various systems of first row transition metals, for example spin-density waves [6, 7], magnetic nano-islands [8], and tip/surface interactions [9]. The calculations involve iterative self-consistency evaluations of the number of d-electrons and the magnetic moments.

An adiabatic separation of slow and fast degrees of freedom is assumed in NCAA calculations: the angles defining the orientation of the magnetic moments are treated as slow degrees of freedom, while the longitudinal components of the magnetic moments are treated as fast degrees of freedom [10–12]. This means that for a given set of angles, the length of the magnetic moments is found in a self-consistent manner. This adiabatic approximation is typically also used in density functional theory (DFT) [13] and tight binding calculations [14] of magnetic systems. It is analogous to the Born–Oppenheimer approximation in electronic structure calculations, where the electronic wave function is calculated for a fixed position of the nuclei. When stationary states are found, such as local minima corresponding to (meta)stable states or first order saddle points corresponding to transition states of magnetic transitions, some optimization algorithm is used to find orientations where the gradient of energy vanishes with respect to angles. Full self-consistency is then reached at the stationary points. However, when calculations of dynamics are carried out or the energy surface characterising the system explored in regions of non-stationary points, then the self-consistency calculation should be formulated in such a way as to keep the desired values of the angular variables fixed. This has, for example been emphasised in the context of the evaluation of exchange parameters [15].

Several methods for finding self-consistent solutions of non-stable magnetic states using tight-binding Hamiltonians or DFT have been presented [16–18]. These methods are based on the introduction of local magnetic fields to force the local magnetic moments to be pointing in the desired directions. Dederichs *et al.* [16] developed a general formalism for such constrained self-consistency calculations based on Lagrange multipliers. Various implementations of this general approach have been used [19–23] involving different algorithms for finding the optimal values of the Lagrange multipliers. A rigorous formulation based on a variational principle for such self-consistency calculations has been presented in the context of constrained DFT, primarily in the context of excited state electronic structure calculations [24–26]. A formulation in terms of a variational principle has the advantage that efficient and systematic optimisation methods can be employed to enforce the constraints and the optimisation can be carried out with respect to all degrees of freedom on an equal footing.

In this article, an implementation of the NCAA model is presented where constraints on the directions of the magnetic moments are included in the self-consistency calculations by use of a variational principle and the performance of various direct optimization algorithms tested. The article is organised as follows: In section 2, the NCAA model is reviewed and the inclusion of angular constraints presented. Numerical tests are presented in section 3 and the performance of the limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) [27], conjugate gradient and fixed point mixing methods compared. Section 4 presents applications of the method to calculations of spin dynamics and curvature of the energy surface for Cr crystal.

2. Methodology

2.1. NCAA model

The NCAA model is based on the separation of electrons into conduction 4s(p)-electrons and quasi-localised 3d-electrons. The basic Hamiltonian is [2]:

$$\hat{H} = \sum_{i\alpha} \epsilon_i^0 \hat{n}_{i\alpha} + \sum_{i \neq j, \alpha} v_{ij} \hat{d}_{i\alpha}^\dagger \hat{d}_{j\alpha} + \frac{1}{2} \sum_{i, \alpha \neq \beta} U_i \hat{n}_{i\alpha} \hat{n}_{i\beta} +$$
$$\sum_{\mathbf{k}\alpha} \epsilon_{\mathbf{k}} \hat{n}_{\mathbf{k}\alpha} + \sum_{\mathbf{k}, i, \alpha} \left(v_{i\mathbf{k}} \hat{d}_{i\alpha}^\dagger \hat{c}_{\mathbf{k}\alpha} + v_{\mathbf{k}i} \hat{c}_{\mathbf{k}\alpha}^\dagger \hat{d}_{i\alpha} \right), \tag{1}$$

where: $\hat{d}_{i\alpha}$, $\hat{d}_{i\alpha}^\dagger$ are annihilation and creation operators for d-electrons with spin α on i-th site, respectively, $\hat{c}_{\mathbf{k}\alpha}$, $\hat{c}_{\mathbf{k}\alpha}^\dagger$ are annihilation and creation operators of $s(p)$-electrons with spin α and momentum $\mathbf{k}$, respectively, $\hat{n}_{i\alpha} = \hat{d}_{i\alpha}^\dagger \hat{d}_{i\alpha}$ and $\hat{n}_{\mathbf{k}\alpha} = \hat{c}_{\mathbf{k}\alpha}^\dagger \hat{c}_{\mathbf{k}\alpha}$ are number operators, ϵ_i^0 energy level of quasi-localised d-electrons at site i, $\epsilon_{\mathbf{k}}$ energy of s(p) states with momentum $\mathbf{k}$, v_{ij} are hopping parameters for d-electrons on sites i and j, $v_{i\mathbf{k}}$ are parameters corresponding to hybridisation of d- and s(p) electrons, U_i Coulomb repulsion between quasi-localised electrons on site i. The first three terms describe the subsystem of quasi-localised electrons. A Hamiltonian consisting only of these three terms is referred to as the single-band Hubbard model [28]. The fourth term in Eq. (1) describes the conduction band, and the last term describes coupling between electrons in quasi-localised states and electrons in the conduction band.

Many magnetic systems can be modelled accurately enough within the mean field approximation. The mean field approximation of Coulomb interaction is applied to each site in the local coordinate frame chosen in such a way that the z-axis points in the same direction as the expectation value of the magnetic moment operator [3]. The procedure is illustrated in Fig. 1. Within this approximation, the Coulomb interaction is approximated according to following equation in the laboratory coordinate frame:

$$U_i \hat{n}_{i\alpha} \hat{n}_{i\beta} \rightarrow \frac{U_i N_i}{2} (\hat{n}_{i\alpha} + \hat{n}_{i\beta}) - \frac{U_i M_i}{2} \hat{\mathbf{M}}_i \cdot \mathbf{e}_i - \frac{1}{4} U_i (N_i^2 - M_i^2), \tag{2}$$

where N_i is the average number of d-electrons at site i, M_i the magnitude of the magnetic moment of spins at site i, and $\mathbf{e}_i = (\sin\theta_i \cos\phi_i, \sin\theta_i \sin\phi_i, \cos\theta_i)$ the unit vector defining the direction of the magnetic moment at site i. The magnetic moment operator is:

$$\hat{\mathbf{M}}_i = \left(\hat{M}_{ix}, \hat{M}_{iy}, \hat{M}_{iz} \right)^T, \tag{3}$$

$$\hat{M}_{iq} = \sum_{\alpha\beta} \hat{d}_{i\alpha}^\dagger \sigma_{\alpha\beta}^q \hat{d}_{i\beta}^\dagger, \quad q = x, y, z \tag{4}$$

where σ^q denotes the Pauli matrices.

In order to describe magnetic properties of itinerant magnets, the s(p)-electrons are taken into account through the renormalisation of the parameters in the Hamiltonian and hybridisation of the d-level [1]. The effective Hamiltonian for non-collinear magnetism can be written as [5]:

$$\hat{H}^{NCAA} = \sum_{i\alpha} \left(E_i^0 + \frac{U_i N_i}{2} + i\Gamma \right) \hat{n}_{i\alpha} + \sum_i -\frac{U_i M_i}{2} \hat{\mathbf{M}}_i \cdot \mathbf{e}_i +$$
$$\sum_i -\frac{1}{4} U_i (N_i^2 - M_i^2) + \sum_{i \neq j, \alpha} V_{ij} \hat{d}_{i\alpha}^\dagger \hat{d}_{j\alpha}, \tag{5}$$

where:

$$E_i^0 = \epsilon_i^0 + \text{Re} \sum_{\mathbf{k}} \frac{v_{i\mathbf{k}} v_{\mathbf{k}i}}{\omega - \epsilon_{\mathbf{k}}}, \tag{6}$$

$$\Gamma = \text{Im} \sum_{\mathbf{k}} \frac{v_{i\mathbf{k}} v_{\mathbf{k}i}}{\omega - \epsilon_{\mathbf{k}}}, \tag{7}$$
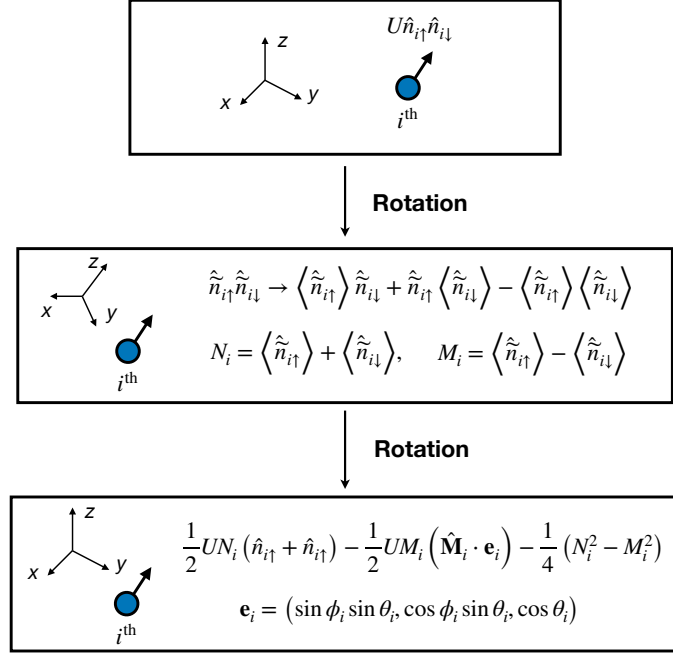
FIG. 1. Illustration of the implementation of the mean field approximation for non-collinear magnetism. Atoms i is denoted by the blue disk and the arrow indicates its average magnetic moment. First, the laboratory coordinate frame is rotated into the local coordinate frame so that the z axis points along the magnetic moment of the atom. Then, the mean field approximation for $\hat{\tilde{n}}_{i\alpha}\hat{\tilde{n}}_{i\beta}$ is applied in this local coordinate frame, and the number of d-electrons N_i and magnetic moment M_i evaluated. Then, the local coordinate frame is rotated back into the laboratory coordinate frame. This two step procedure is applied to all atoms in the system.

$$V_{ij} = v_{ij} + \sum_{\mathbf{k}} \frac{v_{i\mathbf{k}}v_{\mathbf{k}j}}{\omega - \epsilon_{\mathbf{k}}}. \tag{8}$$

This Hamiltonian is reminiscent of the Hubbard model with renormalised parameters ϵ_i^0 and v_{ij} within a mean field approximation, generalised for non-collinear alignment of the magnetic moments. For a more detailed discussion of the NCAA Hamiltonian, see Ref. [4, 5].

In order to calculate the expectation values of operators, the Green function formalism is used. As the NCAA model in the mean field approximation is quadratic with respect to fermion operators, the Green function is a resolvent of the NCAA Hamiltonian:

$$\left(\omega - \hat{H}^{NCAA}\right)\hat{G} = I. \tag{9}$$

In previous implementations, the Green function was calculated iteratively using the continued fraction approach [5]. However, we find that for small enough systems a direct solution in terms of eigenvalue decomposition of the NCAA Hamiltonian is a more efficient way to calculate the Green function. After the eigenvalues and eigenvectors has been found, the expectation values of interest can be calculated as shown in Appendix A.

The unit vectors $\{\mathbf{e_i}\}$ in the NCAA Hamiltonian should by definition point in the same direction as the magnetic moment operators. However, this will in general not be the case at a non-stationary point because there will, by definition, be a non-zero toque acting on the magnetic moments

$$\mathbf{T}_i^{NCAA} = -\mathbf{M}_i \times \mathbf{H}_i^{\mathrm{eff}} = \mathbf{e}_i \times \frac{\partial E^{NCAA}}{\partial \mathbf{e_i}} = \frac{U_i}{2}\mathbf{M}_i \times \left\langle \hat{\mathbf{M}_i} \right\rangle. \tag{10}$$

Therefore, $\mathbf{M}_i \nparallel \left\langle \hat{\mathbf{M}_i} \right\rangle$, where $\mathbf{M}_i = M_i \mathbf{e}_i$. In order to enforce the expectation value of a magnetic moment to be aligned along the direction $\mathbf{e}_i$, the following constraint needs to be added: A projection of the expectation value of the

magnetic moment on the tangent space to the direction $\mathbf{e}_i$ equals to zero, i.e.:

$$\left\langle \hat{\mathbf{M}}_i \right\rangle \cdot \mathbf{e}_{\theta_i} = 0 \tag{11}$$

$$\left\langle \hat{\mathbf{M}}_i \right\rangle \cdot \mathbf{e}_{\phi_i} = 0. \tag{12}$$

Here, $\mathbf{e}_{\phi_i} = (-\sin\phi_i, \cos\phi_i, 0)^T$, $\mathbf{e}_{\theta_i} = (\cos\theta_i \cos\phi_i, \cos\theta_i \sin\phi_i, -\sin\theta_i)^T$ are the local orthogonal unit vectors in the tangent space of $\mathbf{e}_i$. These constraints can be written formally in various forms, for example: $\left[\left\langle \hat{\mathbf{M}}_i \right\rangle, \mathbf{e_i} \right] = 0$ [19] or: $\left\langle \hat{\mathbf{M}}_i \right\rangle - \left(\left\langle \hat{\mathbf{M}}_i \right\rangle \cdot \mathbf{e}_i \right) \mathbf{e_i} = 0$ [20]. In either case, the constraint ensures that the projection of the magnetic moment on the local tangent space equals zero. We will proceed with Eqs. (11) and (12). The algorithm described in the following sections can be generalised to other formulations of the constraints in a straight-forward way.

The constraints can be satisfied by means of Lagrange multipliers [16]:

$$\begin{aligned} \hat{H} &= \hat{H}^{NCAA} - \sum_{i=1..P} \lambda_1^i \hat{\mathbf{M}}_i \cdot \mathbf{e}_{\theta_i} - \sum_{i=1..P} \lambda_2^i \hat{\mathbf{M}}_i \cdot \mathbf{e}_{\phi_i} = \\ &= \hat{H}^{NCAA} - \sum_{i=1..P} \hat{\mathbf{M}}_i \cdot \mathbf{h}_i^c, \end{aligned} \tag{13}$$

The last term describes the interaction of a magnetic moment with a local magnetic field acting perpendicular to the magnetic moment. As a result, a revised torque acting on the magnetic moment in such a way as to satisfy the constraints is:

$$\mathbf{T}_i = -\mathbf{M}_i \times \mathbf{H}_i^{\text{eff}} = \mathbf{e}_i \times \mathbf{h}_i^c. \tag{14}$$

The energy of the system given by the NCAA model is a function of N_i, M_i, λ_1^i and λ_2^i. The optimal values of these parameters correspond to an extremum of the energy and can be determined through the implementation of a variation principle [26]. Energy gradients with respect to N_i, M_i, λ_1^i and λ_2^i, as well as optimisation algorithms for finding the optimal values are given below. First, the case of a non-stationary state where the orientation of the magnetic moments is fixed is discussed. Then, the case of a stationary state is discussed, where an optimisation of the energy with respect to angles is carried out.

2.2. Calculation of non-stationary states

In order to find a self-consistent solution at a non-stationary configuration of the magnetic moments, one needs to find the extremum of the energy with respect to N_i, M_i, λ_1^i and λ_2^i. Analysis of constrained DFT has shown that the extremum with respect to λ_1^i and λ_2^i corresponds to a maximum of the energy [26]. The same is true for the NCAA model. Also, the extremum of the energy with respect to N_i corresponds to a maximum of the energy within the mean filed approximation. The non-zero magnetic moment can be considered as an additional constraint on the occupation numbers. The partial derivatives can be evaluated analytically:

$$\frac{\partial E}{\partial N_i} = \frac{U_i}{2} \left[\langle \hat{n}_i \rangle - N_i \right], \tag{15}$$

$$\frac{\partial E}{\partial M_i} = \frac{U_i}{2} \left[M_i - \left\langle \hat{\mathbf{M}}_i \right\rangle \cdot \mathbf{e}_i \right], \tag{16}$$

$$\frac{\partial E}{\partial \lambda_1^i} = -\left\langle \hat{\mathbf{M}}_i \right\rangle \cdot \mathbf{e}_{\theta_i}, \tag{17}$$

$$\frac{\partial E}{\partial \lambda_2^i} = -\left\langle \hat{\mathbf{M}}_i \right\rangle \cdot \mathbf{e}_{\phi_i}, \quad \forall i = 1 \ldots P. \tag{18}$$

When the gradient given by the first expression is zero, a self-consistent value for the number of d-electrons has been reached. The second gradient vanishes when a self-consistent value has been obtained for the longitudinal component of the magnetic moments. The last two gradients give self-consistent equations for Lagrange multipliers that adjust the perpendicular component of magnetic moments. When they are zero, the expectation values of the magnetic moment operators point in the predefined directions, i.e. $\langle \hat{\mathbf{M}_i} \rangle \parallel \mathbf{e}_i$.

The most commonly used and simplest method for solving self-consistency equations is the fixed point iteration method with simple mixing:

$$M^{(k)} = \alpha \left\langle \hat{\mathbf{M}}^{(k-1)} \right\rangle \cdot \mathbf{e} + (1-\alpha)M^{(k-1)}, \tag{19}$$

$$N^{(k)} = \alpha \left\langle \hat{n}^{(k-1)} \right\rangle + (1-\alpha)N^{(k-1)}, \tag{20}$$

$$\lambda_1^{(k)} = \alpha \left[\lambda_1^{k-1} - \left\langle \hat{\mathbf{M}}^{(k-1)} \right\rangle \cdot \mathbf{e}_\theta \right] + (1-\alpha)\lambda_1^{(k-1)} \tag{21}$$

where k is an iteration counter and α is a mixing parameter chosen in such a way as to give a balance between stability and rate of convergence. However, by formulating the equations in terms of partial derivatives and a variational principle, more efficient optimisation methods, such as the conjugate gradient algorithms and quasi-Newton algorithms, can be used. The fixed point iteration method Eq. (19) is effectively a steepest-descent algorithm:

$$
M^{(k)} = M^{(k-1)} + \alpha \left[\left\langle \hat{\mathbf{M}}^{(k-1)} \right\rangle \cdot \mathbf{e} - M^{(k-1)} \right] =
$$
$$
= M^{(k-1)} - \alpha \frac{2}{U} \left(\frac{\partial E}{\partial M} \right)^{(k-1)}, \tag{22}
$$

$$
N^{(k)} = N^{(k-1)} + \alpha \left[\left\langle \hat{n}^{(k-1)} \right\rangle - N^{(k-1)} \right] =
$$
$$
= N^{(k-1)} + \alpha \frac{2}{U} \left(\frac{\partial E}{\partial N} \right)^{(k-1)}, \tag{23}
$$

$$
\lambda^{(k)} = \lambda^{(k-1)} + \alpha \left(\frac{\partial E}{\partial \lambda} \right)^{(k-1)} \tag{24}
$$

Note that a descent direction is used for M_i and an ascent direction for n_i, λ_i^1 and λ_i^2. This is due to the fact that the self-consistent solution for the NCAA model correspond to a saddle point in the parameter space.

In order to find the extremum of energy one can carry out optimisation with respect to all degrees of freedom. The algorithm is presented below:

Algorithm for Non-stationary States:

(1) Given initial values of the variables that need to be determined self-consistently

$$
\vec{x}^{(0)} = \left(N_1 \ldots N_P, M_1 \ldots M_P, \lambda_1^1 \ldots \lambda_1^P, \lambda_2^1 \ldots \lambda_2^P \right)^T, \tag{25}
$$

the corresponding scaled derivatives of the energy

$$
\vec{g}^{(0)} = \alpha \left(-\frac{2}{U} \frac{\partial E}{\partial N_1} \cdots, \frac{2}{U} \frac{\partial E}{\partial M_1}, -\frac{\partial E}{\partial \lambda_1^1} \cdots, -\frac{\partial E}{\partial \lambda_2^1} \cdots -\frac{\partial E}{\partial \lambda_2^P} \right)^T \tag{26}
$$

and a value of the scaling parameter α

(2) Let $\Theta = \max_i \left\{ \left| \frac{2}{U} \frac{\partial E}{\partial N_i} \right|, \left| \frac{2}{U} \frac{\partial E}{\partial M_i} \right| \right\}$, $\Delta = \max_i \left| \mathbf{T}_i^{NCAA} \right|$; ϵ, δ are convergence tolerance, l is a parameter which controls the size of the step along the search direction and $k = 0$ is iteration counter

(3) Calculate the search direction according to a given minimisation algorithm.
For example, $\vec{p}^{(k)} = -\vec{g}^{(k)}$ for gradient descent

(4) While $\Theta > \epsilon$ and $\Delta > \delta$:
  (a) $d = \|\vec{p}^{(k)}\|_{\inf}$. If $d > l$ then $\vec{p}^{(k)} \leftarrow \vec{p}^{(k)} l / d$
  (b) $\vec{x}^{(k+1)} = \vec{x}^{(k)} + \vec{p}^{(k)}$
  (c) Calculate new gradient $\vec{g}^{(k+1)}$ and new search direction $\vec{p}^{(k+1)}$
  (d) $k \leftarrow k + 1$

(5) End.

Typical values of the parameters are $\alpha = 0.7$, $\epsilon = 10^{-7}$, $\delta = 10^{-7}\,\Gamma$, $l = 0.04$. If the algorithm does not converge, then the values of the parameters α and/or l should be reduced. As the extremum of energy that corresponds to the solution is not a minimum, a line search procedure cannot be applied in this case.

2.3. Calculations of stationary states

In order to find stationary states, i.e. configurations of the magnetic moments that correspond to a local minimum or a saddle point on the energy surface, the function that gives the energy of the system with respect to all the degrees of freedom, one can use the NCAA Hamiltonian without constraints because the Lagrange multipliers vanish at these points. Therefore, one needs to find the extremum of the energy with respect to N_i, M_i and $\mathbf{e}_i$. In order to find energy minima, the energy minimisation can be carried out in two different ways: Either the degrees of freedom are separated into fast and slow variables and the minimisation performed with an inner loop for the former and an outer loop for the latter or all degrees of freedom are treated on an equal footing. We will refer to the first option as algorithm M1, and the second option as algorithm M2. The motivation for M1 is that a deviation of N_i or M_i from the optimal values leads to larger change in the energy than a deviation in the angular variables, θ_i and ϕ_i. This, for example, leads to the separation of the dynamics into fast and slow degrees of freedom [12]: slow degrees of freedom are directions of magnetic moments obeying classical Landau–Lifshitz equations of motion while the fast degrees of freedom are

the longitudinal components of the magnetic moments found using self-consistency calculations for given directions. Similarly, one can separate the minimisation into two nested loops, where for each value of the slow degrees of freedom, the fast degrees of freedom are adjusted according to the self-consistency equations. The orthogonal spin optimisation (OSO) method [29] is used here for the slow degrees of freedom, $\mathbf{e}_i$. The inner loop finds optimal values of N_i and M_i, and the outer loop performs rotations of the magnetic moments towards the energy minimum:

Algorithm M1:

Inner loop for finding optimal values of N_i and M_i.

(1) Let
$$\vec{x} = (N_1 \dots N_P, M_1 \dots M_P)^T \tag{27}$$
be the vector consisting of variables that need to be found self-consistently,
$$\vec{g} = \alpha \left(-\frac{2}{U} \frac{\partial E}{\partial N_1} \dots, \frac{2}{U} \frac{\partial E}{\partial M_1} \right)^T \tag{28}$$
the vector consisting of the corresponding energy derivatives, and α a scaling parameter.

(2) Let $\Theta = \max_i \left\{ \left| \frac{2}{U} \frac{\partial E}{\partial N_i} \right|, \left| \frac{2}{U} \frac{\partial E}{\partial M_i} \right| \right\}$ be convergence tolerances, l a parameter controlling the size of the step along the search direction and $k_{in} = 0$ an iteration counter.

(3) Calculate the search direction according to a given minimisation algorithm. For example, $\vec{p}^{(k_{in})} = -\vec{g}^{(k_{in})}$ for gradient descent algorithm.

(4) While $\Theta > \epsilon$:
   (a) $d = \|\vec{p}^{(k_{in})}\|_{\inf}$. If $d > l$ then $\vec{p}^{(k_{in})} \leftarrow \vec{p}^{(k_{in})} l/d$.
   (b) $\vec{x}^{(k_{in}+1)} = \vec{x}^{(k_{in})} + \vec{p}^{(k_{in})}$
   (c) Calculate new gradient $\vec{g}^{(k_{in}+1)}$ and new search direction $\vec{p}^{(k_{in}+1)}$.
   (d) $k_{in} \leftarrow k_{in} + 1$.

End of inner loop.

   Outer loop:

(1) Let $k = 0$ be iteration counter. Let δ be a convergence tolerance for the maximum torque in the system. Call **Inner loop** to calculate optimal values $N_i^{(k)}$, $M_i^{(k)}$ for given initial directions of the magnetic moments: $\{\mathbf{e}_i^{(k)}\}_{i=1..N}$. Using this optimal values calculate the torques
$$\mathbf{T}_i^{(k)} = \frac{U_i M_i^{(k)}}{2} \mathbf{e}_i \times \left\langle \hat{\mathbf{M}}_{\mathbf{i}} \right\rangle^{(k)}.$$

(2) While $\max_i \left| \mathbf{T}_i^{(k)} \right| \rangle \delta$ repeat the the following steps:
   (a) Perform rotation of the magnetic moments towards the minimum of energy:
$$\mathbf{e}_i^{(k+1)} = \exp\left(-P_i\right) \mathbf{e}_i^{(k)}, \quad i = 1...N \tag{29}$$
   where P is a skew-symmetric matrix. For gradient descent algorithm it is:
$$P_i = \begin{pmatrix} 0 & -t_{iz} & t_{iy} \\ t_{iz} & 0 & -t_{ix} \\ -t_{iy} & t_{ix} & 0 \end{pmatrix}, \quad \mathbf{t}_i \equiv \mathbf{T}_i^{(k)}. \tag{30}$$

   (b) Call **Inner loop** to calculate optimal values $N_i^{(k+1)}$, $M_i^{(k+1)}$ for a new directions of magnetic moments $\{\mathbf{e}_i^{(k+1)}\}_{i=1..N}$.
   (c) Calculate the torques,
$$\mathbf{T}_i^{(k)} = \frac{U_i M_i^{(k)}}{2} \mathbf{e}_i \times \left\langle \hat{\mathbf{M}}_{\mathbf{i}} \right\rangle^{(k)}.$$
   using the optimal values $N_i^{(k+1)}$, $M_i^{(k+1)}$.
   (d) Set $k \leftarrow k + 1$.

End of outer loop

End of M1

While the expression given here for the matrices P_i in Eq. (30) are for gradient descent algorithm, it is more efficient to use the L-BFGS in the outer loop, as shown below. The relevant expressions for the P_i matrices in that case and detail on the implementation of the L-BFGS can be found in Ref. [29].

The second algorithm, M2, where all degrees of freedom are treated on an equal footing is given below.

Algorithm M2:

(1) Let

$$\vec{x} = (N_1 \ldots N_P, M_1 \ldots M_P, \theta_1, \ldots, \theta_P, \phi_1, \ldots, \phi_P,)^T \tag{31}$$

be a vector consisting of the variables that need to be found self-consistently,

$$\vec{g} = \alpha \left(-\frac{2}{U}\frac{\partial E}{\partial N_1} \cdots, \frac{2}{U}\frac{\partial E}{\partial M_1} \cdots, \frac{2}{UM_i}\frac{\partial E}{\partial \theta_1} \cdots, \frac{2}{UM_i}\frac{\partial E}{\sin\theta_i \partial\phi_1} \cdots, \right)^T \tag{32}$$

the vector consisting of corresponding derivatives of the energy, and α the scaling parameter.

(2) Let $\Theta = \max_i \left\{ \left|\frac{2}{U}\frac{\partial E}{\partial N_i}\right|, \left|\frac{2}{U}\frac{\partial E}{\partial M_i}\right| \right\}$, $\Delta = \max_i \left|\mathbf{T}_i^{NCAA}\right|$; ϵ, δ be convergence tolerances, l a parameter controling the size of steps along the search direction and $k = 0$ an iteration counter.

(3) Calculate the search direction according to the given minimisation algorithm. For example, $\vec{p}^{(k)} = -\vec{g}^{(k)}$ for a gradient descent algorithm.

(4) While $\Theta > \epsilon$ and $\Delta > \delta$:
   (a) $d = \|\vec{p}^{(k)}\|_{\inf}$. If $d > l$ then $\vec{p}^{(k)} \leftarrow \vec{p}^{(k)} l/d$.
   (b) $\vec{x}^{(k+1)} = \vec{x}^{(k)} + \vec{p}^{(k)}$
   (c) Calculate new gradient $\vec{g}^{(k+1)}$ and new search direction $\vec{p}^{(k+1)}$.
   (d) $k \leftarrow k + 1$.

End of M2

Algorithm M2 is efficient when the system is far from the energy minimum and the gradients for all degrees of freedom are large. But, near a minimum it can be unstable. The reason is that it mixes slow and fast degrees of freedom. If one needs to converge N_i and M_i to a small tolerance then it is difficult to converge angles ϕ_i, θ_i to the same accuracy. Changes in M_i and N_i have large effect on the energy while changes in the angular variables do not affect the energy as strongly. In practice, it is better to use algorithm M2 in the beginning and start with relatively large tolerances for N_i and M_i, for example $\epsilon = 10^{-3}$, and then reduce ϵ to a smaller number so that the error in the torque due to deviations from self-consistency are smaller than the magnitude of the torque. Near the minimum it is usually enough to perform 1 to 3 iteration steps for N_i and M_i to reach good convergence.

Saddle points on the energy surface can be found using the geodesic nudged elastic band method [30] where a climbing image converges onto the saddle point. While constraints on the orientation of the magnetic moments should in principle be applied to images that are in between the endpoint minima and the climbing image, this turns out not to be important for convergence onto the saddle point. The intermediate images are needed to estimate the tangent along the path, and they can reveal the presence of intermediate minima, but it is not so critical that they lie exactly on the minimum energy path.

3. Numerical tests

The performance of three different optimisation algorithms: L-BFGS, Fletcher–Revere conjugate gradients (FR) and steepest-descent (SD) was tested in calculations of non-stationary states. Then, algorithms M1 and M2 as well as mixed M1+M2 algorithms were tested and compared in calculations of stationary states. The test system consists of either 400 or 196 atoms in a hexagonal two-dimensional lattice subject to periodic boundary conditions. The NCAA model parameters are $E_0 = -12\,\Gamma$, $U = 13\,\Gamma$, and $V = 0.5\,\Gamma$.

The most demanding computational part of NCAA calculations is the diagonalisation of the Hamiltonian. The comparison of the performance of the various optimisation algorithms is, therefore, measured in terms of the number of diagonalisations which in the present case, equals the number of iterations needed in order to reach convergence.

3.1. Calculations of non-stationary states

The non-stationary states are a tilted spin spirals:

$$\mathbf{e}_i = \begin{pmatrix} \sin(\psi)\cos(\mathbf{q} \cdot \mathbf{r}_i) \\ \sin(\psi)\sin(\mathbf{q} \cdot \mathbf{r}_i) \\ \cos(\psi) \end{pmatrix} \tag{33}$$
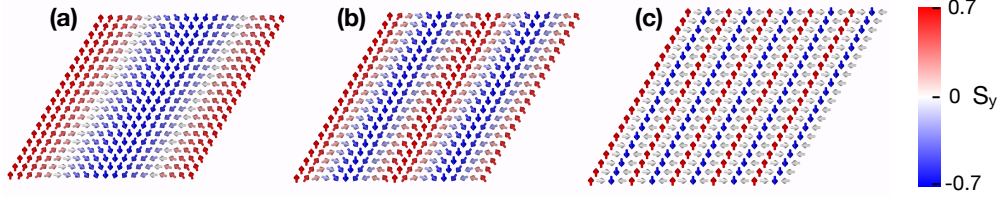
FIG. 2. Spin states corresponding to Eq. 33 for $\psi = \pi/4$ when **q** lies along ΓM direction. (a) $q = \pi/10a$, (b) $q = \pi/5a$, (c) $q = \pi/2a$. a is a lattice constant. The color corresponds to y component of the spins. The visualisation was done with OVITO software [36].

TABLE 1. Number of iterations needed in order to reach convergence using three different optimisation algorithms: L-BFGS, Fletcher–Revere conjugate gradients (FR), and steepest descent (SD). Calculations are carried out for different values of the wave number q. The convergence tolerances are $\epsilon = 10^{-8}$, ($\epsilon = 10^{-14}$ in brackets) and $\delta = 10^{-8}$. The initial values of N_i and M_i are 1.4 and 0.44, respectively, and the converged values of N_i and M_i are given in the last two columns.

q	SD	FR	L-BFGS	N_i	M_i
0	30 (58)	52 (401)	24 (29)	1.4271	0.4555
$\pi/10a$	55 (84)	127 (375)	30 (43)	1.4268	0.4556
$\pi/5a$	64 (101)	158 (206)	50 (66)	1.4259	0.4558
$3\pi/10a$	69 (110)	84 (108)	47 (66)	1.4245	0.4562
$4\pi/10a$	73 (119)	43 (536)	25 (38)	1.4225	0.4568
$\pi/2a$	75 (125)	54 (79)	20 (142)	1.4200	0.4581
$6\pi/10a$	77 (128)	372 (174)	26 (39)	1.4170	0.4599
$7\pi/10a$	77 (129)	188 (302)	28 (61)	1.4138	0.4622
$8\pi/10a$	78 (131)	244 (185)	21 (53)	1.4108	0.4646
$9\pi/10a$	78 (134)	140 (211)	32 (41)	1.4086	0.4665
π/a	77 (132)	102 (140)	26 (43)	1.4078	0.4672
average:	68 (113)	142 (247)	30 (56)		

with ψ chosen to be $\pi/4$. If $\psi = \pi/2$, the corresponding spin spirals are stationary states in the NCAA model. The wave vector **q** is chosen to lie along ΓM direction in the irreducible two-dimensional Brillouin zone. Different states correspond to different values of q. Three examples are shown in Fig. 2.

The performance of the three minimisation algorithms is summarised in Table 1. Each of the algorithms gives a different search direction during the optimization.

As can be seen from Table 1, the number of iterations needed in the SD calculations systematically grows as the wave number, q is increased. The reason is that the torque acting on the magnetic moments in the initial configuration increases with q and the expectation values of the magnetic moments deviate more from the reference direction. Overall, L-BFGS requires fewer iteration than either SD or FR conjugate gradients algorithm. On average, it requires half as many iterations as SD and 1/5 as many as FR. For one value of q, however, $q = \pi/2a$, and the tight convergence criterion, $\epsilon = 10^{-14}$, L=BFGS requires more iterations than FR and SD (142 vs. 79 and 125).

The poor performance of FR is surprising, but can probably be explained by the fact that the SCF solution does not correspond to a minimum of the NCAA energy, but rather a saddle point. The energy needs to be maximised with respect to N_i and the Lagrange multipliers while being minimised with respect to M_i. Conjugate gradient algorithms require an efficient line search procedure to find the minimum of energy along the search direction. As the extremum is a saddle point in the present case, no line search can be applied and the step taken along the search direction not optimal, thereby reducing the efficiency of the algorithm. This can be the reason why the FR algorithm performs even worse than the SD algorithm in the present case.

TABLE 2. Number of iterations needed in order to obtain the ferromagnetic ground state starting from various initial spiral states with wave vector q and $\psi = \pi/4$. A magnetic field is applied along the y-direction, $H = 10^{-4}\,\Gamma$. The simulation cell consists of 196 atoms. The first two columns correspond to algorithm M1 involving inner and outer loop iterations, with tolerances $\epsilon = 10^{-3}$ and $\epsilon = 10^{-7}$. In both cases $\delta = 10^{-7}$. The third column corresponds to calculations using a hybrid algorithm where M2 is first used to obtain convergence to a tolerance of $\delta = 10^{-4}$, and then algorithm M1 is used until with tolerances $\epsilon = 10^{-7}$ and $\delta = 10^{-7}$

q	M1($\epsilon = 10^{-3}$)	M1($\epsilon = 10^{-7}$)	M2+M1 ($\epsilon = 10^{-7}$)
0	46	135	110
$\pi/7a$	165	1353	186
$4\pi/14a$	108	1456	153
$6\pi/14a$	197	1459	108
$8\pi/14a$	160	1041	112
$10\pi/14a$	169	894	100
$12\pi/14a$	138	1067	105
π/a	122	750	91
average:	138	1200	120

3.2. Calculations of stationary states

In calculations of energy minima or saddle points on the energy surface, the constraint fields can be set to zero at each iteration during the optimisation as these fields vanish at stationary points. Performance tests were carried out for a 196 atom system where the initial state corresponds to one of the titled spiral states with $\psi = \pi/4$ described in the previous subsection, and the final state corresponds to the uniform, ferromagnetic ground state. A magnetic field of $H = 10^{-4}\,\Gamma$ is applied along the y-direction. The search direction is calculated using the L-BFGS algorithm.

The results are shown in Table 2. The performance of the M1 algorithm where fast and slow degrees of freedom are treated separately in an inner and outer loop is given for two values of the inner loop convergence tolerance, ϵ. For the outer loop, the tolerance is $\delta = 10^{-7}$ in both cases. In the calculations corresponding to the first column $\epsilon = 10^{-3}$, but since at least one inner loop iteration is taken for each outer loop iteration, the convergence of N and M is actually better than 10^{-3}.

The results in the second column correspond to inner loop tolerance of $\epsilon = 10^{-7}$ and the outer loop tolerance is again $\delta = 10^{-7}$. The number of iterations is significantly larger compared with the results in the first column. The difference in the energy of the system when $\epsilon = 10^{-3}$ and $\epsilon = 10^{-7}$ is around $10^{-4}\,\Gamma$.

The third column gives results for a hybrid algorithm, where all degrees of freedom are first treated on the same footing using algorithm M2, until convergence corresponding to $\delta = 10^{-4}$ has been reached and then algorithm M1 is used with a tolerance of $\epsilon = 10^{-7}$ in the inner loop and $\delta = 10^{-7}$ in the outer loop. This turns out to be the optimal algorithm, especially when the initial state is far from the minimum energy configuration and the initial values of the gradients are large.

4. Application

The NCAA model has previously been applied in studies of spin density waves and the antiferromagnetic state of BCC Cr [6] and the same approached is used here. A pair of neighboring atoms are included explicitly in order to represent the antiferromagnetic state, but the effect of neighboring atoms in the crystal is included implicitly. The NCAA model parameter values are: $E_0 = -3.37\,\Gamma, U = 6.745\,\Gamma$ and $v = 0.9\,\Gamma$. The effective hopping parameter is $V = \sqrt{8}v$ corresponding to the d-d hopping to the eight nearest neighbors in the BCC lattice. The self-consistent NCAA calculations for the antiferromagnetic ground state give magnetic moment of $0.6\mu_B$ and 5.0 d-electrons.

The angle between the prescribed directions of the magnetic moments is denoted by θ, as illustrated in Fig. 3. The deviation from the antiferromagnetic order is $\phi = \pi - \theta$. The angle between the prescribed direction and the calculated magnetic moment and the angle obtained from the self-consistency calculation without constraints is denoted by α. This represents the error due to the lack of constraints on the transverse component of the magnetic moments. A comparison is made between fully self-consistent calculations where constraints are applied and calculations where
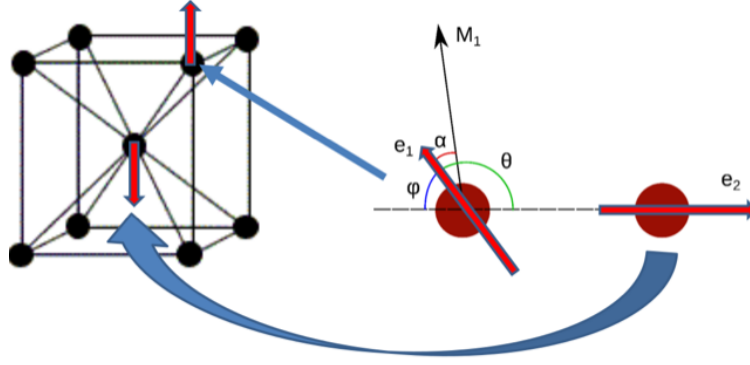
FIG. 3. Illustration of the simulation of a BCC crystal of Cr. Two atoms are included explicitly in the calculations in order to represent the antiferromagnetic order, but the effect of neighbors is included implicitly. The prescribed directions of the magnetic moments are $\mathbf{e_1}$ and $\mathbf{e_2}$. $\mathbf{M_1}$ is a calculated magnetic moment of atom 1. α is the angle between $\mathbf{M_1}$ and $\mathbf{e_1}$ and represents the error in the direction when constraints on the transverse components of the magnetic moments are not included.
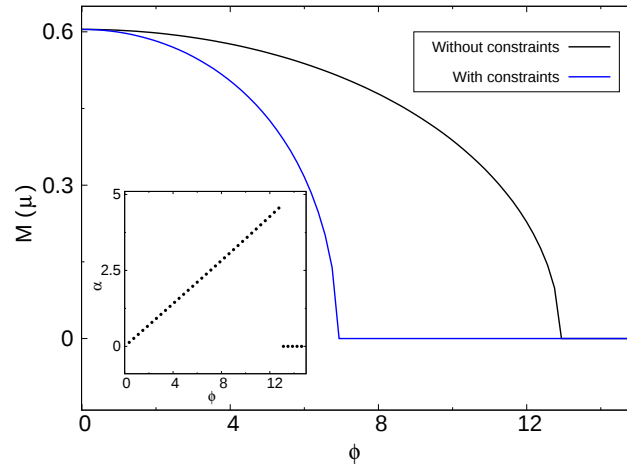

FIG. 4. Magnitude of the magnetic moments as a function of the angle ϕ (see Fig. 3). The blue curve represents calculations including the constraints and the black curve represents the calculations without constraints. Inset: The deviation of the direction of calculated magnetic moment from prescribed directions as functions of angle ϕ when constraints are not included.

the self-consistency calculations only include the average number of d-electrons and the longitudinal component of magnetic moments.

As one of the magnetic moments is rotated with respect to the other, the magnitude of the magnetic moments is reduced as shown in Fig. 4. At a certain angle the magnetic moment vanishes. When the angular constraints are included, the magnitude drops to zero at $\phi \approx 7°$, while it drops to zero at $\phi \approx 13°$. when the constraints are not included. The error that occurs when the contstraints are neglected is, therefore, large in this case. The qualitative shape of the curve is similar, but the magnetic moment vanishes at an angle that is nearly twice as large when the constraints are not included.

The inset in Fig. 4 shows the maximum deviation of the direction of the calculated magnetic moments from the prescribed direction when constraints are not included, i.e. the dependence of α on ϕ. The maximum deviation is about $5°$ obtained right when the magnetic moment vanishes.

The Fig. 5(a) shows the variation of the torques acting on magnetic moments as a function of the angle ϕ. The torques are evaluated using Eq. (14) when the constraints are included, but from Eq. (10) when the constraints are neglected. The torques differ by up to a factor of three in magnitude and the maximum is obtained at different values
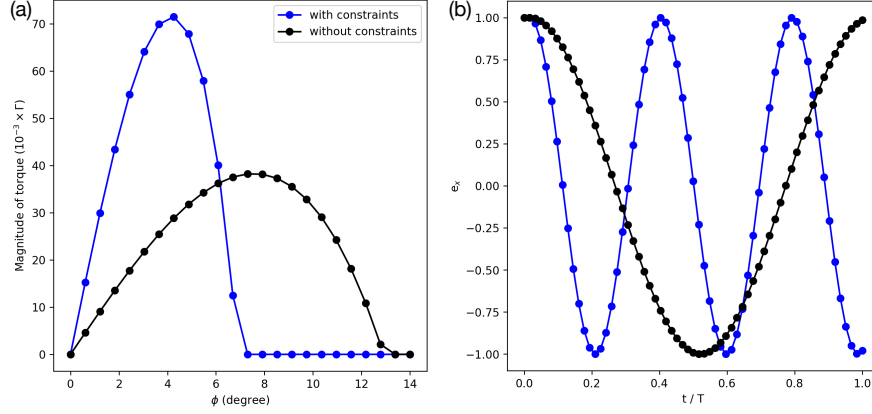
FIG. 5. Comparison of results obtained for Cr crystal with (blue curves) and without (black curves) constraints on the transverse components of the magnetic moments. (a) Torque as a function of ϕ. (b) Projection of the magnetic moment on the x-axis as a function of time in units of the oscillation period obtained without constraints. Here, $\phi = 4°$.

of the angle ϕ. When constraints are not included, the torque rises and falls more gradually than when constraints are included. Both curves level out at zero, corresponding to a non-magnetic solution.

Spin dynamics were simulated by solving numerically the classical Landau-Lifshitz equation of motion for the slow degrees of freedom

$$\frac{d\mathbf{e}_i}{dt} = -\gamma\, \mathbf{e}_i \times \mathbf{H}_i^{\text{eff}}. \tag{34}$$

The Eq. 14 for torques coincides with what is used in tight-binding and DFT simulations of the spin-dynamics with constraints [17, 19]. Indeed, substitution of Eq. 14 into Eq. 34 gives:

$$\frac{d\mathbf{e}_i}{dt} = \gamma\, \mathbf{e}_i \times \mathbf{h}_\perp^{\mathbf{c}}(\mathbf{i}). \tag{35}$$

We have used the two step numerical method of Mentink *et al.* [31] to numerically integrate the equation of motion. As the calculated torques are so different, the spin dynamics are also very different. The two sublattices oscillate around a common axis with a frequency that depends directly on the magnitude of the torques. Since the torques differ by almost a factor of three depending on whether the constraints are applied or not, the frequencies in the dynamical motion differ by the same amount, as can be seen from Fig. 5. When the constraints are applied, the frequency is about three times larger.

Effective exchange coupling parameters [15, 32, 33] and pre-exponential factors for estimates of the lifetime of magnetic states [34, 35] can be obtained from the curvature of the energy surface at the minimum. When the curvature is calculated by finite differences, the energy needs to be evaluated at non-stationary points corresponding to small tilts of the magnetic moments away from the ground state, local minimum configurations. Constraints need to be included in order for the angular deviations to be well defined. The second derivative of the energy, $\partial^2 E/\partial\theta_0 \partial\theta_1$, at $\phi = 0$ is calculated with the NCAA model to be $290 \times 10^{-3}\,\Gamma$ for the Cr crystal when the constraints are included. When the constraints are neglected a significantly smaller curvature is deduced, $87 \times 10^{-3}\,\Gamma$. This shows the importance of the constraints when the pre-exponential factors and effective exchange parameters are calculated using self-consistency calculations.

Acknowledgements

Appendix A. Calculation of expectation values

The Green function method is used to calculate expectation values of the number of d-electrons, magnetic moments and the energy of the system. Since the Hamiltonian is quadratic with respect to creation and annihilation operators, the Green function is the resolvent:

$$\left(\omega - \hat{H}^{NCAA}\right)\hat{G} = I \tag{36}$$

In the case of non-collinear magnetism, the density of states is a matrix:

$$\rho_{ij}(\epsilon) = \begin{pmatrix} \rho_{ij}^{\alpha\alpha} & \rho_{ij}^{\alpha\beta} \\ \rho_{ij}^{\alpha\beta*} & \rho_{ij}^{\beta\beta} \end{pmatrix} \tag{37}$$

which can be calculated using the imaginary part of the Green function

$$\rho_{ij}^{\alpha\beta}(\epsilon) = \frac{1}{\pi}\lim_{s\to 0}\operatorname{Im}G_{ij}^{\alpha\beta}(\epsilon - is) \tag{38}$$

where i, j refer to the atomic sites and α, β refer the spin projection.

The expectation values can be calculated as:

$$\langle \hat{n}_i \rangle = \int_{-\infty}^{\epsilon_f} d\epsilon\,\operatorname{Tr}_s \rho_{ii}(\epsilon) \tag{39}$$

$$\left\langle \hat{\mathbf{M}}_i \right\rangle = \int_{-\infty}^{\epsilon_f} d\epsilon\,\operatorname{Tr}_s \sigma\,\rho_{ii},(\epsilon), \tag{40}$$

where the trace operation is carried out with respect to spin degrees of freedom, and $\sigma = (\sigma^x, \sigma^y, \sigma^z)^T$ are the Pauli matrices. To calculate matrix elements of the Green function, an eigendecomposition of the Hamiltonian is used. Let $\{\epsilon_i\}, \{\vec{y}_i\}$ be the eigenvalues and eigenvectors of the NCAA Hamiltonian. Using the eigenvectors as a basis, the Green function is diagonal:

$$G_{km} = \frac{\delta_{km}}{\omega - \epsilon_k - i\Gamma}, \quad k, m = 1..2P \tag{41}$$

where P number of atoms. The Green function in the laboratory basis set can be obtained using the similarity transformation:

$$G_{ij}^{\alpha\beta} = \sum_{k=1}^{2P}\langle x_i^\alpha|y_k\rangle G_{kk}^d\langle y_k|x_j^\beta\rangle \quad i, j = 1..P, \tag{42}$$

where $|x_i^\alpha\rangle$ corresponds to the state on site i with the spin projection α. Using the equation above, one can find expectation values of the operators in terms of eigenvalues and eigenvectors:

$$\langle \hat{n}_i \rangle = \sum_{k=1}^{2P}\left(|\langle x_i^\uparrow|y_k\rangle|^2 + |\langle x_i^\downarrow|y_k\rangle|^2\right)R_{kk} \tag{43}$$

$$\left\langle \hat{\mathbf{M}}_i \right\rangle_x = \sum_{k=1}^{2P} 2\operatorname{Re}\left(\langle x_i^\uparrow|y_k\rangle\langle y_k|x_i^\downarrow\rangle\right)R_{kk} \tag{44}$$

$$\left\langle \hat{\mathbf{M}}_i \right\rangle_y = \sum_{k=1}^{2P} -2\operatorname{Im}\left(\langle x_i^\uparrow|y_k\rangle\langle y_k|x_i^\downarrow\rangle\right)R_{kk} \tag{45}$$

$$\left\langle \hat{\mathbf{M}}_i \right\rangle_z = \sum_{k=1}^{2P}\left(|\langle x_i^\uparrow|y_k\rangle|^2 - |\langle x_i^\downarrow|y_k\rangle|^2\right)R_{kk}, \tag{46}$$

where

$$R_{kk} = \frac{1}{\pi}\lim_{s\to 0}\operatorname{Im}G_{kk}^d = \frac{1}{\pi}\operatorname{arccot}\frac{\epsilon_k}{\Gamma}. \tag{47}$$

and the energy is [5]:

$$E = \frac{1}{\pi}\sum_i \epsilon_i\operatorname{arccot}\left(\frac{\epsilon_i}{\Gamma}\right) + \frac{\Gamma}{2}\ln\left(1 + \frac{\epsilon_i^2}{\Gamma^2}\right) - \frac{U_i}{4}\left(N_i^2 - M_i^2\right). \tag{48}$$

References

[1] Anderson P.W. Localized Magnetic States in Metals. *Phys. Rev.*, 1961, **124**, P. 41.

[2] Alexander S., Anderson P.W. Interaction Between Localized States in Metals. *Phys. Rev.*, 1964, **133**, P. A1594.

[3] Uzdin V.M., Yartseva N.S. Periodic Anderson model for the description of noncollinear magnetic structure in low-dimensional 3d-systems. *Comput. Mater. Sci.*, 1998, **10**, P. 211.

[4] Bessarab P.F., Uzdin V.M., Jónsson H. Calculations of magnetic states and minimum energy paths of transitions using a non-collinear extension of the Alexander-Anderson model and a magnetic force theorem. *Phys. Rev. B*, 2014, **89**, P. 214424.

[5] Bessarab P.F., Skorodumov A., Uzdin V.M., Jónsson H. Navigation on the Energy Surface of the Noncollinear Alexander-Anderson model. *Nanosystems: physics, chemistry, mathematics*, 2014, **6**, P. 757.

[6] Uzdin V.M., Demangeat C. Spin-density wave in Cr without the nesting property of the Fermi surface. *J. Phys.: Condens. Matter*, 2006, **18**, P. 2717.

[7] Uzdin V.M., Vega A. Magnetization reversal process at atomic scale in systems with itinerant electrons. *J. Phys.: Condens. Matter*, 2012, **24**, P. 176002.

[8] Bessarab P.F., Uzdin V.M., Jónsson H. Navigation on the energy surface of the non-collinear Alexander-Anderson model. *Nanosystems: Physics, Chemistry, Mathematics*, 2014, **5**, P. 757.

[9] Ivanov A., Bessarab P.F., Uzdin V.M., Jónsson H. Magnetic exchange force microscopy: Theoretical analysis of induced magnetization reversals. *Nanoscale*, 2017, **9**, P. 13320.

[10] You M.V., Heine V. Magnetism in transition metals at finite temperatures I. Computational model. *J. Phys. F: Met. Phys.*, 1982, **12**, P. 177–94.

[11] Gyorffy B.L., Pindors A.J., Staunton J., Stocks G.M. and Winter H. A first-principles theory of ferromagnetic phase transitions in metals. *J. Phys. F: Met. Phys.*, 1985, **15**, P. 1337–1386.

[12] Antropov V.P., Katsnelson M.I., Harmon B.N., M. van Schilfgaarde, Kusnezov D. Spin dynamics in magnets: Equation of motion and finite temperature effects. *Phys. Rev. B*, 1996, **54**, P. 1019.

[13] Kohn W. Nobel Lecture: Electronic structure of matterwave functions and density functionals. *Rev. Mod. Phys.*, 1999, **71**, P. 1253.

[14] Harrison W.A. *Elementary Electronic Structure*. Singapore, World Scientific, 2009.

[15] Bruno P. Exchange Interaction Parameters and Adiabatic Spin-Wave Spectra of Ferromagnets: A Renormalized Magnetic Force Theorem. *Phys. Rev. Lett.*, 2003, **90**, P. 087205.

[16] Dederichs P.H., Blügel S., Zeller R. and Akai H. Ground States of Constrained Systems: Application to Cerium Impurities. *Phys. Rev. Lett.*, 1984, **53**, P. 2512.

[17] Small L.M., Heine V. A couple method for calculating interatomic interactions in itinerant electron magnetic systems. *J. Phys. F: Met. Phys.*, 1984, **14**, P. 3041–3052.

[18] Luchini M.U., Heine V., McMulla G.J. Calculation of longitudinal magnetic fluctuations in iron by a Landau energy and force method. *J. Phys.: Condens. Matter*, 1991, **3**, P. 8647–8664.

[19] Stocks G.M., Ujfalussy B., Xindong Wang, Nicholson D.M.C., Shelton W.A., Yang Wang, Canning A., Gyorffy B.L. Towards a constrained local moment model for first principles spin dynamics. *Philos. Mag.*, 1998, **78**, Nos. 5/6 665-673.

[20] Kurz Ph., Föster F., Nordström L., Bihlmayer G., Blügel S. Ab initio treatment of noncollinear magnets with the full-potential linearized augmented plane wave method. *Phys. Rev. B*, 2004, **69**, P. 024415.

[21] Singer R., Fähnle M. and Bihlmayer G. Constrained spin-density functional theory for excited magnetic configurations in an adiabatic approximation. *Phys. Rev. B*, 2005, **71**, P. 214435.

[22] Stocks G.M., Eisenbach M., Újfalussy B., Lazarovits B., Szunyogh L., Weinberger P. On calculating the magnetic state of nanostructures. *Prog. Mater. Sci*, 2007, **52**, P. 371–387.

[23] Ma P.W., Dudarev S.L. Constrained density functional for noncollinear magnetism. *Phys. Rev. B*, 2015, **91**, P. 054420.

[24] Wu Q., and T. Van Voorhis. Direct optimization method to study constrained systems within density-functional theory. *Phys. Rev. A*, 2005, **72**, P. 024502.

[25] Wu Q., and T. Van Voorhis. Constrained Density Functional Theory and Its Application in Long-Range Electron Transfer. *J. Chem. Theory Comput.*, 2006, **2**, P. 765.

[26] ORegan D.D., Teobaldi G. Optimization of constrained density functional theory. *Phys. Rev. B*, 2016, **94**, P. 035159.

[27] Nocedal J., Wright S.J. *Numerical Optimization*, 2nd ed. Springer, New York, NY, USA, 2006.

[28] Hubbard J. Electron correlations in narrow energy bands. *Proc. Roy. Soc. A*, 1963, **276**, P. 238.

[29] Ivanov A.V., Uzdin V.M., Jónsson H. Fast and Robust Algorithm for the Energy Minimization of Spin Systems Applied in an Analysis of High Temperature Spin Configurations in Terms of Skyrmion Density. arXiv:1904.02669v2 [physics.comp-ph] 2019.

[30] Bessarab P.F., Uzdin V.M., Jónsson H. Method for finding mechanism and activation energy of magnetic transitions, applied to skyrmion and antivortex annihilation. *Computer Physics Communications*, 2015, **196**, P. 335–347.

[31] Mentink J.H., Tretyakov M.V., Fasolino A., Katsnelson M.I. and Rasing Th. Stable and fast semi-implicit integration of the stochastic LandauLifshitz equation. *J. Phys.: Condens. Matter*, 2010, **22**, P. 176001.

[32] Liechtenstein A.I., Katsnelson M.I., Antropov V.P., Gubanov V.A. *J. Magn. Magn. Mater.*, 1987, **67**, P. 65–74.

[33] Katsnelson M.I., Lichtenstein A.I. Magnetic susceptibility, exchange interactions and spin-wave spectra in the local spin density approximation. *J. Phys.: Condens. Matter*, 2004, **16**, P. 7439–7446.

[34] Bessarab P.F., Uzdin V.M., Jónsson H. Harmonic transition-state theory of thermal spin transitions. *Phys. Rev. B.*, 2012, **85**, P. 184409.

[35] Bessarab P.F., Uzdin V.M., Jónsson H. Size and shape dependence of thermal spin transitions in nanoislands. *Phys. Rev. Lett.*, 2013, **110**, P. 020604.

[36] Stukowski A. Visualization and analysis of atomistic simulation data with OVITO the Open Visualization Tool. *Modelling Simul. Mater. Sci. Eng.*, 2010, **18**, P. 015012.