Interplay of Coulomb interaction and spin-orbit coupling

Jörg Bünemann, Thorben Linneweber, Ute Löw, Frithjof B. Anders, and Florian Gebhard

Phys. Rev. B **94**, 035116 — Published 7 July 2016

DOI: 10.1103/PhysRevB.94.035116

Interplay of Coulomb interaction and spin-orbit coupling

Jörg Bünemann,[1,2] Thorben Linneweber,[3] Ute Löw,[3] Frithjof B. Anders,[3] and Florian Gebhard[1]

[1] *Fachbereich Physik, Philipps-Universität Marburg, D-35032 Marburg, Germany*
[2] *Institut für Physik, BTU Cottbus-Senftenberg, D-03013 Cottbus, Germany*
[3] *Lehrstuhl für Theoretische Physik II, Technische Universität Dortmund, D-44227 Dortmund, Germany*

(Dated: Version of June 15, 2016, approved by JB, FG)

We employ the Gutzwiller variational approach to investigate the interplay of Coulomb interaction and spin-orbit coupling in a three-orbital Hubbard model. Already in the paramagnetic phase we find a substantial renormalization of the spin-orbit coupling that enters the effective single-particle Hamiltonian for the quasi-particles. Only close to half band-filling and for sizable Coulomb interaction we observe clear signatures of Hund's atomic rules for spin, orbital, and total angular momentum. For a finite local Hund's-rule exchange interaction we find a ferromagnetically ordered state. The spin-orbit coupling considerably reduces the size of the ordered moment, it generates a small ordered orbital moment, and it induces a magnetic anisotropy. To investigate the magnetic anisotropy energy, we use an external magnetic field that tilts the magnetic moment away from the easy axis $(1, 1, 1)$.

PACS numbers: 71.10.Fd,71.27.+a,71.70.Ej,75.10.Lp

I. INTRODUCTION

In atomic physics, the spin-orbit coupling (SOC) plays an important role because it determines the value of the total angular momentum in the ground state according to Hund's third rule. After maximizing the total spin s (first rule) and the total orbital moment l (second rule), the quantum number for the total angular momentum is $j = |l - s|$ ($j = l + s$) below (above) half filling (third rule).[1,2] The third rule applies in the limit where the SOC is small compared to the average Coulomb interaction of the electrons, i.e., for all 'light atoms', including transition metals. Note that the quantum numbers s and l are, in fact, well defined only in the limit of a vanishing SOC.

For atoms in a solid, the situation is obviously much more complicated because neither of the three quantum numbers s, l, or j is well defined due to a breaking of the rotational symmetry. Yet, we know that some of the basic mechanisms of Hund's rules are still relevant. For example, the maximization of the spin is a direct consequence of intra-atomic exchange correlations that are caused by the electronic Coulomb interactions. The very same Coulomb interaction is the reason for magnetic order in solids, e.g., in ferromagnets. The SOC is only a small perturbation to the dominant Coulomb interaction in transition metals and their compounds. Nevertheless, it can have profound consequences, e.g., for the direction of the magnetic moment, the so-called 'easy axis'.

From a theoretician's point of view, the analysis of the interplay and/or competition of a strong local Coulomb interaction and a (comparatively small) SOC in a solid is rather demanding. Even the study of simplifying models for the Coulomb interaction, such as multi-orbital Hubbard models, poses a tremendously difficult task. Any study of such models is possible with a limited numerical accuracy only, e.g., in determining the ground-state energy. Given the fundamental uncertainties in the treatment of the sizable Coulomb correlations it is non-trivial to come to firm conclusions on the effects of the SOC. Therefore, most theoretical studies on the interplay of Coulomb interaction and SOC focused on insulating or spin states and/or assumed a rather large SOC.[3–6] For the study of (itinerant) $4d$, $5d$ or f electron systems, the dynamical mean field theory has been used frequently in recent years, see, e.g., Refs. [7–10]. In such systems, however, the SOC tends to be significantly larger than in transition metals and their compounds that we have primarily in mind in our present model study.

In this work, we employ the Gutzwiller approach[11] to investigate approximate variational ground states for multi-orbital Hubbard models. The analytical evaluation of expectation values for Gutzwiller wave functions poses a difficult many-body problem that requires additional approximations. Most often applied in the context of multi-band models is the 'Gutzwiller approximation' which becomes exact for the Gutzwiller wave functions in the limit of infinite spatial dimensions.[11–14] It can be used to evaluate expectation values for a large set of model parameters, see Sect. II B. This allows us to study systematically the subtle interplay of Coulomb correlations and spin-orbit coupling.

We consider a Hubbard model with three degenerate t_{2g} orbitals on a three-dimensional cubic lattice. In the first part of our investigation we concentrate on the interplay of Coulomb interaction and spin-orbit coupling for paramagnetic metallic ground states. We find that the Coulomb interaction enhances the effective SOC between the quasi-particles. In addition, we investigate the significance of Hund's rules. Only Hund's first rule approximately applies in strongly correlated paramagnetic metallic systems.

It is well known that for a finite (local) exchange interaction, multi-orbital Hubbard models tend to favor ferromagnetic states for sufficiently large Coulomb interactions. In the second part of our investigation we investigate if and to what extent the ferromagnetic states

are modified by the spin-orbit coupling. We find that the SOC opposes the formation of ferromagnetic order in metals. While, in the absence of SOC, the ordered moment has no preferred direction, the SOC aligns it along the 'easy-axis', and induces a small ordered orbital moment.

Recently, the Gutzwiller method and the density functional theory (DFT) were combined in a self-consistent manner;[15,16] a formal derivation can be found in Ref. [17]. The Gutzwiller-DFT was applied to a number of materials, for example to nickel and iron, see Refs. [17,18], and references therein. From a methodological point of view, our model study in this work provides a first step towards a self-consistent treatment of the SOC within the Gutzwiller-DFT scheme.

This work is organized as follows. In Sect. II we introduce our model and summarize the Gutzwiller variational approach. In Sect. III we discuss our results for paramagnetic and ferromagnetic ground states. Summary and conclusions, Sect. IV, close our presentation. Technical details are deferred to two appendices.

II. MODELS AND METHOD

In this section, we introduce our model and explain the Gutzwiller variational approach that we use for its investigation.

A. Hamiltonian

We study a Hubbard model with three t_{2g} orbitals per site on a simple-cubic lattice in three dimensions. The Hamiltonian of this system has the form

$$\hat{H} = \hat{H}_0 + \hat{H}_C + \hat{H}_{so} \,, \tag{1}$$

where $\hat{H}_0$ denotes the kinetic energy of the electrons, $\hat{H}_C$ describes their Coulomb interaction, and $\hat{H}_{so}$ models the spin-orbit coupling.

1. *Kinetic energy and density of states*

We consider electrons that move between t_{2g} orbitals b and b' on sites i and j of our simple-cubic lattice with L sites. In second quantization the single-particle Hamiltonian reads

$$\hat{H}_0 = \sum_{i \neq j} \sum_{\sigma,\sigma'} t_{i,j}^{\sigma,\sigma'} \hat{c}_{i,\sigma}^\dagger \hat{c}_{j,\sigma'} \,, \tag{2}$$

where we introduce the combined spin-orbital index

$$\sigma \equiv (b,s) \,,\; b \in \{1,2,3\} \,,\; s \in \{\uparrow,\downarrow\} \,. \tag{3}$$

The crystal-field energies are set to zero, $t_{i,i}^{\sigma,\sigma'} = 0$.

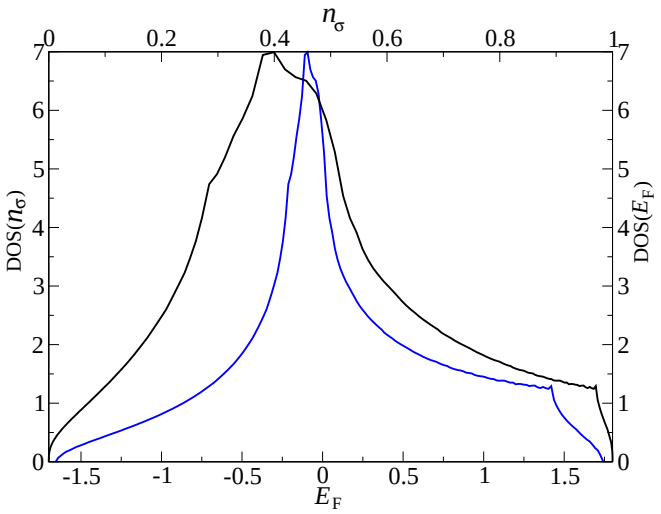

FIG. 1: Density of states at the Fermi energy E_F as a function of E_F (blue) and the orbital occupation n_σ (black).

We use the standard parameterization for the hopping amplitudes in (2) with some generic Slater-Koster parameters[19]

$$t_\pi^{(1),(2),(3)} = 0.3t, -0.1t, 0.025t \,, \tag{4}$$

$$t_\sigma^{(2),(3)} = 0.1t, 0.01t \,, \tag{5}$$

$$t_\delta^{(1),(2),(3)} = 0.1t, -0.025t, 0.02t \tag{6}$$

for the electron transfers up to 3rd nearest neighbors. By including hoppings beyond the nearest neighbors we make sure that there are no artificial features in our band structure, such as nesting vectors or particle-hole symmetry. In transition metal compounds, the value of t is of the order of 1 eV. In our pure model study in this work, we will simply set $t = 1$ as our energy unit.

The single-particle Hamiltonian (2) can be readily diagonalized in momentum space,

$$\hat{H}_0 = \sum_{\boldsymbol{k}} \sum_{\sigma,\sigma'} \varepsilon_{\boldsymbol{k};\sigma,\sigma'} \hat{c}_{\boldsymbol{k},\sigma}^\dagger \hat{c}_{\boldsymbol{k},\sigma'} \tag{7}$$

with the bare dispersion

$$\varepsilon_{\boldsymbol{k};\sigma,\sigma'} \equiv \frac{1}{L} \sum_{i \neq j} t_{i,j}^{\sigma,\sigma'} \, e^{\mathrm{i}\boldsymbol{k}(\boldsymbol{R}_i - \boldsymbol{R}_j)} \,, \tag{8}$$

and $\boldsymbol{k}$ from the first Brillouin zone. The remaining task is the diagonalization of the 6×6 matrix $\varepsilon_{\boldsymbol{k};\sigma,\sigma'}$ for each $\boldsymbol{k}$ to obtain the (bare) band structure. For non-interacting electrons, all energy levels up to the Fermi energy E_F are filled in the ground state. The corresponding density of states at the Fermi-energy E_F is shown in Fig. 1 as a function of both E_F and of the average orbital occupation $0 \leq n_\sigma \leq 1$. The total bandwidth is $W \approx 3.4$.

Apparently, the Hamiltonian for the kinetic energy is not particle-hole symmetric, as can be seen from the density of states at the Fermi energy. Fig. 1 clearly shows

that $\mathrm{DOS}(n_\sigma) \neq \mathrm{DOS}(1-n_\sigma)$. To study the influence of the spin-orbit coupling, we shall later investigate a particle-hole symmetric kinetic energy. For this case, we use the somewhat artificial Slater-Koster parameters for electron transfers between nearest neighbors only,

$$t_\pi'^{(1)} = 0.2 \quad , \quad t_\delta'^{(1)} = 0.1 \, , \tag{9}$$

which lead to a symmetric density of states of bandwidth $W' = 2$.

In our ferromagnetic calculations we focus on the filling $n_\sigma \approx 0.4$ where the (paramagnetic) density of states has a maximum at the Fermi energy. At such a maximum we can expect a stronger tendency towards ferromagnetic order according to the Stoner criterion.[20]

2. Local interactions

The Coulomb and spin-orbit interaction are assumed to be purely local,

$$\hat{H}_\mathrm{C} = \sum_i \hat{H}_{i;\mathrm{C}} \quad , \quad \hat{H}_\mathrm{so} = \sum_i \hat{H}_{i;\mathrm{so}} \, . \tag{10}$$

The local Coulomb interaction for a model with three degenerate $t_{2\mathrm{g}}$ orbitals reads[21]

$$\begin{aligned} 2\hat{H}_{i;\mathrm{C}} \;=\; & U \sum_{b,s} \hat{n}_{i,b,s}\hat{n}_{i,b,\bar{s}} + \sum_{\substack{b(\neq)b' \\ s,s'}} (U' - \delta_{s,s'}J)\hat{n}_{i,b,s}\hat{n}_{i,b',s'} \\ & + J \sum_{b(\neq)b'} \Big[\left(\hat{c}^\dagger_{i,b,\uparrow}\hat{c}^\dagger_{i,b,\downarrow}\hat{c}_{i,b',\downarrow}\hat{c}_{i,b',\uparrow} + \mathrm{h.c.} \right) \\ & + \sum_s \hat{c}^\dagger_{i,b,s}\hat{c}^\dagger_{i,b',\bar{s}}\hat{c}_{i,b,\bar{s}}\hat{c}_{i,b',s} \Big] \, , \end{aligned} \tag{11}$$

where we use the convention $\bar{\uparrow} = \downarrow$, $\bar{\downarrow} = \uparrow$, and $\hat{n}_{i,b,s} = \hat{c}^\dagger_{i,b,s}\hat{c}_{i,b,s}$ counts the electrons with spin s in orbital b on site i. Note that for $t_{2\mathrm{g}}$-orbitals the three parameters in (11) are not independent because they obey the symmetry relation $U' = U - 2J$.[21]

For the SOC we use

$$\hat{H}_{i;\mathrm{so}} = \sum_{\sigma,\sigma'} \epsilon^\mathrm{so}_{i;\sigma,\sigma'} \hat{c}^\dagger_{i,\sigma}\hat{c}_{i,\sigma'} \, . \tag{12}$$

When we work with the following order for our local basis $|\sigma\rangle$,

$$|1\rangle = |yz,\uparrow\rangle,\ |2\rangle = |yz,\downarrow\rangle,\ |3\rangle = |xz,\uparrow\rangle, \ldots, |6\rangle = |xy,\downarrow\rangle, \tag{13}$$

the six-dimensional SOC matrix $\tilde{\epsilon}^\mathrm{so}$ in (12) has the well-known form

$$\tilde{\epsilon}^\mathrm{so} = -\mathrm{i}\frac{\zeta}{2} \begin{pmatrix} 0 & -\tilde{\sigma}_3 & \tilde{\sigma}_2 \\ \tilde{\sigma}_3 & 0 & -\tilde{\sigma}_1 \\ \tilde{\sigma}_2 & \tilde{\sigma}_1 & 0 \end{pmatrix} \equiv \zeta\tilde{\Sigma} \tag{14}$$

with the standard two-dimensional Pauli matrices $\tilde{\sigma}_1$, $\tilde{\sigma}_2$, $\tilde{\sigma}_3$, and the SOC constant ζ.

The local Hamiltonian

$$\hat{H}_{i;\mathrm{loc}} = \hat{H}_{i;\mathrm{C}} + \hat{H}_{i;\mathrm{so}} \tag{15}$$

in the 64-dimensional local Hilbert space is readily diagonalized,

$$\hat{H}_{i;\mathrm{loc}}|\Gamma\rangle_i = E_\Gamma|\Gamma\rangle_i \, . \tag{16}$$

For parameter values that are typical for transition metals, $\zeta/J = 0.2 \ldots 1.0$ and $J/U = 0.2$, the atomic spectrum has a generic form. In table I we list the degenerate eigenspaces of $\hat{H}_{i;\mathrm{loc}}$, ordered by increasing energy for given particle number $0 \leq n_\mathrm{loc} \leq 6$. We give the degeneracy g of each level, its total spin s, orbital moment l, and total 'angular momentum' j.

#	n_loc	g	s	l	j
1	0	1	0	0	0
1	1	4	1/2	1	3/2
2	1	2	1/2	1	1/2
1	2	5	1	1	2
2	2	3	1	1	1
3	2	1	1	1	0
4	2	5	0	2	2
5	2	1	0	0	0

#	n_loc	g	s	l	j
1	6	1	0	0	0
1	5	2	1/2	1	1/2
2	5	4	1/2	1	3/2
1	4	1	1	1	0
2	4	3	1	1	1
3	4	5	1	1	2
4	4	5	0	2	2
5	4	1	0	0	0

#	n_loc	g	s	l	j
1	3	4	3/2	0	3/2
2	3	4	1/2	2	3/2
3	3	6	1/2	2	5/2
4	3	2	1/2	1	1/2
5	3	4	1/2	1	3/2

TABLE I: Degenerate eigenspaces of $\hat{H}_{i;\mathrm{loc}}$, ordered by energy for a given particle number $0 \leq n_\mathrm{loc} \leq 6$ with a specification of the degeneracy g, total spin s, orbital moment l, and total 'angular momentum' j.

Since the rotational symmetry is broken in our cubic environment, the quantum numbers l and j do, in fact, not label eigenstates of the total 'angular momentum' operator. It is well known, however, that in the $t_{2\mathrm{g}}$ subspace we have

$$\tilde{\boldsymbol{l}}^2 = \sum_{i\in\{x,y,z\}} \tilde{l}_i^2 = 2\mathbb{1} \tag{17}$$

for the vector $\tilde{\boldsymbol{l}}$ of the three matrices

$$\tilde{l}_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \mathrm{i} \\ 0 & -\mathrm{i} & 0 \end{pmatrix} ,$$
$$\tilde{l}_y = \begin{pmatrix} 0 & 0 & -\mathrm{i} \\ 0 & 0 & 0 \\ \mathrm{i} & 0 & 0 \end{pmatrix} ,$$
$$\tilde{l}_z = \begin{pmatrix} 0 & \mathrm{i} & 0 \\ -\mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \quad (18)$$

Hence, the orbital moment behaves like that of $l = 1$ states ('T-P equivalence'),[21] because $\langle \hat{\boldsymbol{l}}^2 \rangle = 1(1+1) = 2$. To be more precise, one finds

$$\tilde{l}_i = -\tilde{l}_i^{(l=1)} \quad (19)$$

where on the r.h.s. we introduced the representation of the orbital momentum for $(l = 1)$ p-orbitals. Due to the T-P equivalence we can label the multiplet states $|\Gamma\rangle$ by a quantum number j that formally corresponds to a total angular momentum of $l = 1$ orbitals. Table I shows that Hund's rules are still valid for the ground states of all particle numbers if we make the replacement $l \to -l$ in Hund's third rule, as a consequence of eq. (19). In particular, as seen from table I, the local spectrum is not particle-hole symmetric. As we will show in Sect. III A 1, the particle-hole asymmetry induced by the SOC is visible in our itinerant three-band lattice model even when we work with a symmetric density of states.

B. Gutzwiller wave functions and energy functional

1. Wave functions

For the variational investigation of the Hamiltonian (1) we use the Gutzwiller wave functions

$$|\Psi_{\mathrm{G}}\rangle = \prod_i \hat{P}_i |\Psi_0\rangle , \quad (20)$$

where $|\Psi_0\rangle$ is a normalized single-particle product state (Slater determinant) and the local Gutzwiller correlator is defined as

$$\hat{P}_i = \sum_{\Gamma,\Gamma'} \lambda_{i;\Gamma,\Gamma'} |\Gamma\rangle_{ii} \langle\Gamma'| \equiv \sum_{\Gamma_{\mathrm{d}}} \lambda_{i;\Gamma_{\mathrm{d}}} |\Gamma_{\mathrm{d}}\rangle_{ii} \langle\Gamma_{\mathrm{d}}| . \quad (21)$$

Here, we introduce the matrix $\tilde{\lambda}_i$ of (complex) variational parameters $\lambda_{i;\Gamma,\Gamma'}$ which allows us to optimize the occupation and the form of the eigenstates $|\Gamma_{\mathrm{d}}\rangle_i$ of $\hat{P}_i$.

We assume that the matrix $\tilde{\lambda}_i$ is Hermitian which ensures that the eigenstates $|\Gamma_{\mathrm{d}}\rangle_i$ exist and form a basis of the local Hilbert space. Without SOC it is usually a sensible approximation to work with a diagonal (and hence real) matrix $\tilde{\lambda}_i$. For a finite SOC, however, it is essential to include at least some non-diagonal elements in $\tilde{\lambda}_i$. In this work, we will take into account all non-diagonal parameters in $\lambda_{i;\Gamma,\Gamma'}$ with states $|\Gamma\rangle_i$ and $|\Gamma'\rangle_i$ that have the same particle number.

The evaluation of expectations values with respect to the wave function (20) poses a difficult many-particle problem that cannot be solved in general. As shown in Refs. [11,22], it is possible to derive analytical expressions for the variational ground-state energy in the limit of infinite spatial dimensions ($D \to \infty$). An application of this energy functional to finite-dimensional systems is usually termed 'Gutzwiller approximation'. It will also be used in this work. One should keep in mind, however, that the Gutzwiller approximation has its limitations, and the study of some phenomena requires an evaluation of expectation values in finite dimensions.[23,24]

Since the energy functional of the Gutzwiller approximation has been derived in detail in previous work, we will only summarize the main results in this section. In the following we are only interested in systems and wave functions that are translationally invariant. Hence, we shall drop lattice-site indices whenever this does not lead to ambiguities.

2. Constraints

As shown in Refs. [11,22] it is most convenient for the evaluation of Gutzwiller wave functions in infinite spatial dimensions to impose the following (local) constraints

$$\langle \hat{P}^\dagger \hat{P} \rangle_{\Psi_0} - 1 \equiv g_1^{\mathrm{c}}(\tilde{\lambda}, |\Psi_0\rangle) = 0 , \quad (22)$$
$$\langle \hat{c}_\sigma^\dagger \hat{P}^\dagger \hat{P}\ \hat{c}_{\sigma'} \rangle_{\Psi_0} - C_{\sigma',\sigma} \equiv g_{\sigma,\sigma'}^{\mathrm{c}}(\tilde{\lambda}, |\Psi_0\rangle) = 0 \quad (23)$$

for the local correlation operators $\hat{P} \equiv \hat{P}_i$. Here, we introduce the local density matrix $\tilde{C} \equiv \tilde{C}_i$ with the elements

$$C_{i;\sigma,\sigma'} = \langle \hat{c}_{i,\sigma'}^\dagger \hat{c}_{i,\sigma} \rangle_{\Psi_0} . \quad (24)$$

Note that the order of indices in (24) has been chosen deliberately because it slightly simplifies the analytical results in Sect. II B 5.

The constraints can be evaluated by means of Wick's theorem; explicit expressions are given in Appendix A. In systems with a high symmetry, the matrix $\tilde{C}$ is often diagonal, e.g., for d orbitals in a cubic environment. In such a case, one usually has to take into account only the diagonal constraints (23), because the l.h.s. of (23) for $\sigma \neq \sigma'$ is automatically zero for all values of $\tilde{\lambda}_{\mathrm{i}}$ that are included in the variational Ansatz. Here, the matrix $\tilde{C}$ is non-diagonal in our system with a finite SOC. Even if one introduces a local basis which has a diagonal local density matrix with respect to $|\Psi_0\rangle$, see Appendix A, one still has to take into account some non-diagonal constraints.

3. Expectation values

Each local operator $\hat{O}_i$, e.g., the operator $\hat{H}_{i;\mathrm{so}}$, can be written as

$$\hat{O}_i = \sum_{\Gamma,\Gamma'} O_{\Gamma,\Gamma'} \hat{m}_{i;\Gamma,\Gamma'} , \tag{25}$$

$$\hat{m}_{i;\Gamma,\Gamma'} \equiv |\Gamma\rangle_{ii}\langle\Gamma'| . \tag{26}$$

In infinite dimensions the expectation value of $\hat{O}_i$ has the form

$$\langle\hat{O}_i\rangle_{\Psi_\mathrm{G}} = \sum_{\Gamma_1,\Gamma_2,\Gamma_3,\Gamma_4} O_{\Gamma_2,\Gamma_3} \lambda^*_{\Gamma_2,\Gamma_1} \lambda_{\Gamma_3,\Gamma_4} \langle\hat{m}_{i;\Gamma_1,\Gamma_4}\rangle_{\Psi_0} , \tag{27}$$

where the remaining expectation values

$$m^0_{i;\Gamma,\Gamma'} \equiv \langle\hat{m}_{i;\Gamma,\Gamma'}\rangle_{\Psi_0} \tag{28}$$

can readily be evaluated using Wicks theorem, see Appendix A.

The expectation value of a hopping operator in infinite dimensions reads ($i \neq j$)

$$\left\langle \hat{c}^\dagger_{i,\sigma_1} \hat{c}_{j,\sigma_2} \right\rangle_{\Psi_\mathrm{G}} = \sum_{\sigma_1',\sigma_2'} q^{\sigma_1'}_{\sigma_1} \left(q^{\sigma_2'}_{\sigma_2}\right)^* \left\langle \hat{c}^\dagger_{i,\sigma_1'} \hat{c}_{j,\sigma_2'} \right\rangle_{\Psi_0} , \tag{29}$$

where an analytical expression for the (local) renormalization matrix $q^{\sigma'}_{\sigma}$ is also given in Appendix A. Note that the matrix $q^{\sigma'}_{\sigma}$ is, in general, neither real nor Hermitian. Any symmetries among its elements are caused by those of the orbital basis states $|\sigma\rangle$ and the form of the Gutzwiller wave function. For example, if we have no SOC and no magnetic or orbital order in our degenerate three-band system, the renormalization matrix has the simple form $q^{\sigma'}_{\sigma} = \delta_{\sigma,\sigma'}\sqrt{q}$ with only one renormalization factor for all orbitals.

4. Structure of the energy functional

In a translationally invariant system, the expectation values that we introduced in the previous section lead to the following variational energy functional (per lattice site)

$$\begin{aligned} E_\mathrm{G}\big(\tilde{\lambda}, |\Psi_0\rangle\big) &= \sum_{\substack{\sigma_1,\sigma_2\\ \sigma_1',\sigma_2'}} q^{\sigma_1'}_{\sigma_1} \left(q^{\sigma_2'}_{\sigma_2}\right)^* E_{\sigma_1,\sigma_2,\sigma_1',\sigma_2'} \\ &\quad + \sum_{\Gamma,\Gamma_1,\Gamma_2} E_\Gamma \lambda^*_{\Gamma,\Gamma_1} \lambda_{\Gamma,\Gamma_2} m^0_{\Gamma_1,\Gamma_2} . \end{aligned} \tag{30}$$

Here, we introduce the tensor

$$\begin{aligned} E_{\sigma_1,\sigma_2,\sigma_1',\sigma_2'} &\equiv \frac{1}{L} \sum_{i\neq j} t^{\sigma_1,\sigma_2}_{i,j} \langle \hat{c}^\dagger_{i,\sigma_1'} \hat{c}_{j,\sigma_2'} \rangle_{\Psi_0} \\ &= \frac{1}{L} \sum_{\boldsymbol{k}} \varepsilon_{\boldsymbol{k};\sigma_1,\sigma_2} \big\langle \hat{c}^\dagger_{\boldsymbol{k},\sigma_1'} \hat{c}_{\boldsymbol{k},\sigma_2'} \big\rangle_{\Psi_0} \end{aligned} \tag{31}$$

with the bare dispersion $\varepsilon_{\boldsymbol{k};\sigma,\sigma'}$ from eq. (8).

The energy (30) is a function of $\lambda_{\Gamma,\Gamma'}$ and $|\Psi_0\rangle$ where $|\Psi_0\rangle$ enters (30), (31) solely through the (non-interacting) density matrix $\tilde{\rho}$ with the elements

$$\rho_{(i\sigma),(j\sigma')} \equiv \langle \hat{c}^\dagger_{j,\sigma'} \hat{c}_{i,\sigma} \rangle_{\Psi_0} . \tag{32}$$

Note that the non-local elements of $\tilde{\rho}$ ($i \neq j$) determine the tensor (31) while its local elements

$$\rho_{(i\sigma),(i\sigma')} = C_{i;\sigma,\sigma'} , \tag{33}$$

as introduced in eq. (24), enter the elements $q^{\sigma'}_{\sigma}$ of the renormalization matrix, the expectation value (28), and the constraints (22), (23).

The energy

$$E_\mathrm{G} = E_\mathrm{G}(\tilde{\lambda}, \tilde{\rho}, \tilde{C}) \tag{34}$$

has to be minimized with respect to the variational parameters $\lambda_{\Gamma,\Gamma'}$ and the density matrices $\tilde{\rho}$ and $\tilde{C}$ obeying the constraints (22), (23), (33), and

$$\tilde{\rho}^2 = \tilde{\rho} . \tag{35}$$

This additional constraint ensures that $\tilde{\rho}$ corresponds to a Slater determinant $|\Psi_0\rangle$. Note that introducing the local density matrix $\tilde{C}$ as an independent variational object in (34), at the expense of the additional constraint (33), is actually not necessary. Instead one could consider the energy solely as a function of $\tilde{\lambda}, \tilde{\rho}$. Our form of the energy functional, however, will turn out to be slightly more convenient because, in the Gutzwiller approximation, $\tilde{\rho}$ enters the energy in a non-linear way only through its local elements.

5. Minimization of the energy functional

We introduce the real and the imaginary parts of the variational parameters

$$\lambda_{\Gamma,\Gamma'} = \lambda^{(\mathrm{r})}_{\Gamma,\Gamma'} + \mathrm{i}\lambda^{(\mathrm{i})}_{\Gamma,\Gamma'} . \tag{36}$$

Due to the Hermiticity of $\tilde{\lambda}$ we have

$$\lambda^{(\mathrm{r})}_{\Gamma,\Gamma'} = \lambda^{(\mathrm{r})}_{\Gamma',\Gamma} , \tag{37}$$

$$\lambda^{(\mathrm{i})}_{\Gamma,\Gamma'} = -\lambda^{(\mathrm{i})}_{\Gamma,\Gamma'} \to \lambda^{(\mathrm{i})}_{\Gamma,\Gamma} = 0 , \tag{38}$$

which leads to a number n_v of independent (and real) variational parameters $\lambda^{(\mathrm{r/i})}_{\Gamma',\Gamma}$. They will be considered as the components v_z of the n_v-dimensional vector

$$\boldsymbol{v} = (v_1, \ldots, v_{n_v})^\mathrm{T} . \tag{39}$$

The (in general) complex constraints (22), (23) are not all independent, e.g., because of the Hermiticity of $\tilde{g}^\mathrm{c}$.

We denote the set of all independent real and imaginary parts of (22), (23) by the n_c real constraints

$$g_l(\boldsymbol{v}, \tilde{C}) = 0 \qquad (l = 1, \ldots, n_c) \; . \tag{40}$$

The constraints (33), (35), and (40) are implemented via Lagrange parameters $\eta_{\sigma,\sigma'}$, $\Omega_{(i\sigma),(j\sigma')}$, and Λ_l. This leads to the Lagrange functional

$$\begin{aligned} L_{\mathrm{G}} \; \equiv \; & E_{\mathrm{G}}(\boldsymbol{v}, \tilde{\rho}, \tilde{C}) - \sum_l \Lambda_l g_l(\boldsymbol{v}, \tilde{C}) \\ & - \sum_{\sigma,\sigma'} \eta_{\sigma,\sigma'} \sum_i (C_{\sigma',\sigma} - \rho_{(i\sigma'),(i\sigma)}) \\ & - \sum_{i,j} \sum_{\sigma,\sigma'} \Omega_{(i\sigma),(j\sigma')} [\tilde{\rho}^2 - \tilde{\rho}]_{(j\sigma'),(i\sigma)} \end{aligned} \tag{41}$$

which provides the basis of our minimization.

As shown, e.g., in Ref. [25], the minimization of (41) with respect to $\tilde{\rho}$ leads to the effective single-particle Hamiltonian

$$\hat{H}_0^{\mathrm{eff}} = \sum_{i,j} \sum_{\sigma,\sigma'} (\bar{t}_{i,j}^{\sigma,\sigma'} + \delta_{i,j} \eta_{\sigma,\sigma'}) \hat{c}_{i,\sigma}^\dagger \hat{c}_{j,\sigma'} \tag{42}$$

with the renormalized hopping parameters

$$\bar{t}_{i,j}^{\sigma_1,\sigma_2}(\boldsymbol{v}, \tilde{C}) = \sum_{\sigma_1',\sigma_2'} q_{\sigma_1'}^{\sigma_1}(\boldsymbol{v}, \tilde{C}) \left(q_{\sigma_2'}^{\sigma_2}(\boldsymbol{v}, \tilde{C}) \right)^* t_{i,j}^{\sigma_1',\sigma_2'} \; . \tag{43}$$

The optimal Slater determinant $|\Psi_0\rangle$ is the ground state of $\hat{H}_0^{\mathrm{eff}}$,

$$\hat{H}_0^{\mathrm{eff}} |\Psi_0\rangle = E_0^{\mathrm{eff}} |\Psi_0\rangle \; . \tag{44}$$

From the minimization of (41) with respect to $\tilde{C}$ we obtain an equation for $\eta_{\sigma,\sigma'}$ in (42),

$$\eta_{\sigma,\sigma'} = \frac{\partial}{\partial C_{\sigma,\sigma'}} E_{\mathrm{G}} - \sum_l \Lambda_l \frac{\partial}{\partial C_{\sigma,\sigma'}} g_l \; . \tag{45}$$

Finally, the minimization with respect to $\boldsymbol{v}$

$$\frac{\partial}{\partial v_Z} E_{\mathrm{G}} - \sum_l \Lambda_l \frac{\partial}{\partial v_Z} g_l = 0 \tag{46}$$

determines the Lagrange parameters Λ_l and the optimal value of $\boldsymbol{v}$. Equations (42)–(46) need to be solved self-consistently. In Appendix B we explain in more detail how we solve this problem numerically. Note that our minimization algorithm does not require the constraints $g_l(\boldsymbol{v}, \tilde{C})$ to be independent. This is a major advantage over the method that we had proposed in the earlier work [25].

III. RESULTS

In the following we discuss the paramagnetic and the ferromagnetic cases separately.

A. Paramagnetic ground states

1. *Effective spin-orbit coupling*

Without any breaking of spin or orbital symmetries, the minimization of the Gutzwiller energy functional leads to effective on-site energies (45) that have the same form as the SOC (14) but with the coupling constant ζ replaced by ζ^{eff}. This change from the bare to an effective coupling constant also changes the quasi-particle dispersion of $\hat{H}_0^{\mathrm{eff}}$. Therefore, the energy splittings at certain high-symmetry points as seen in ARPES experiments are a measure for the effective, not the bare spin-orbit coupling. Note that extracting the quasi-particle dispersion from our Gutzwiller approach relies on a Fermi-liquid interpretation.[26] However, all changes of the effective single-particle Hamiltonian, e.g., energy shifts, are related to changes of certain ground-state expectation values. Since the latter are variationally controlled, it is very likely that the exact single-particle spectrum reflects the same trends.

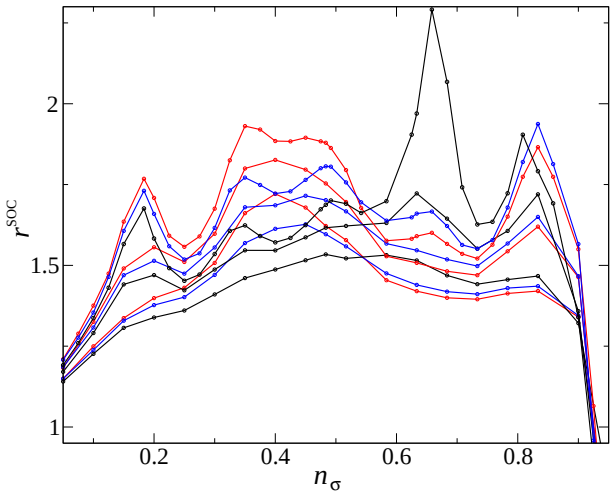

FIG. 2: SOC-renormalization r^{SOC} as a function of the orbital occupation n_σ for $J = 0$, $\zeta = 0.05$ (red), $\zeta = 0.1$ (blue), $\zeta = 0.2$ (black) and $U = 2, 3, 4$ (in ascending order).

In Fig. 2 we show the renormalization of ζ,

$$r^{\mathrm{SOC}} \equiv \zeta^{\mathrm{eff}} / \zeta \tag{47}$$

as a function of the orbital occupation n_σ for the three bare values $\zeta = 0.05, 0.1, 0.2$ and interaction parameters $U = 2, 3, 4$ and $J = 0$. Apparently, the effective spin-orbit coupling increases as a function of U, apart from a small region of an almost filled shell where $\zeta^{\mathrm{eff}}(U) < \zeta$. For $U = 4$, which is approximately equal to the band width, the spin-orbit couping can be renormalized by a factor two or more, $r^{\mathrm{SOC}}(U = 4, n_\sigma = 2/3, \zeta = 0.2) \approx 2.3$. This substantial increase is clearly visible in the quasi-particle band structure, see Fig. 3, where we show the bare ($U = J = 0$) and renormalized band structures ($U = 4, J/U = 0.2$) for $n_\sigma = 2/3$ and $\zeta = 0.2$. For

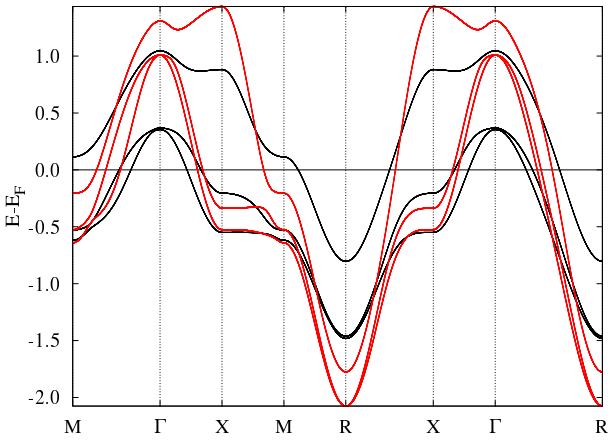

FIG. 3: Quasi-particle bands along high-symmetry lines for $\zeta = 0.2$, $n_\sigma = 2/3$, $U = 0$, $J = 0$ (red), and $U = 4$, $J/U = 0.2$ (black).

example, the splitting of the bands at the Γ-point and the R-point is noticeably enhanced in presence of the Coulomb interaction.

The renormalization $r^{\rm SOC}$ is not monotonous as a function of the bare coupling ζ. Moreover, it is *not* particle-hole symmetric, i.e., it is not invariant under the transformation $n_\sigma \to 1 - n_\sigma$. This is only partly caused by the particle-hole asymmetry of the bare density of states in Fig. 1. As discussed already in Sect. II A 2, the SOC inherently breaks the particle-hole symmetry. To illustrate this point, we show the results for a symmetric density of states that results from the nearest-neighbor electron transfers (9), displayed in the inset of Fig. 4. As seen from the figure, the SOC alone induces a particle-hole asymmetry in the renormalization of the effective spin-orbit

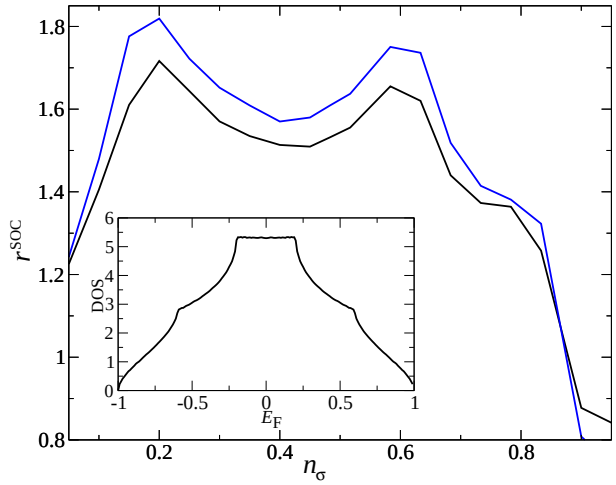

FIG. 4: SOC-renormalization $r^{\rm SOC}$ for the symmetric density of states from the nearest-neighbor electron transfers (9), as a function of the orbital occupation n_σ for $\zeta = 0.1$, $J = 0$, $U = 3$ (black) and $U = 4$ (blue); inset: density of states at the Fermi energy.

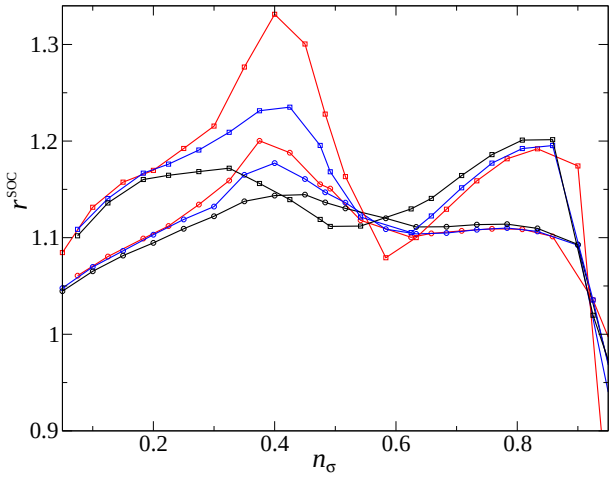

FIG. 5: SOC-renormalization $r^{\rm SOC}$ as a function of the orbital occupation n_σ for $J/U = 0.2$, $\zeta = 0.05$ (red), $\zeta = 0.1$ (blue), $\zeta = 0.2$ (black), and $U = 1$ (circles), $U = 2$ (squares).

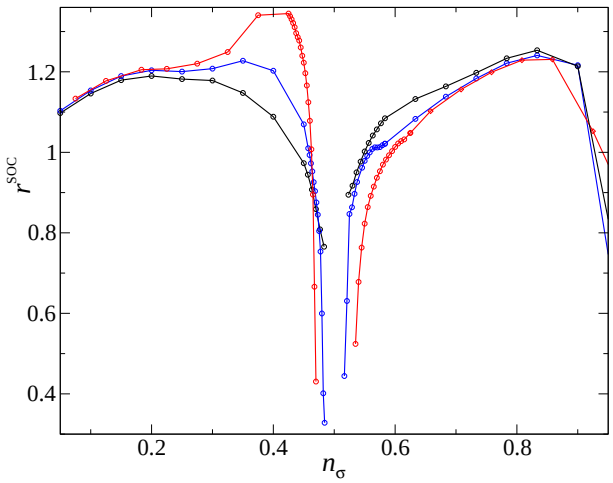

FIG. 6: SOC-renormalization $r^{\rm SOC}$ as a function of the orbital occupation n_σ for $J/U = 0.2$, $\zeta = 0.05$ (red), $\zeta = 0.1$ (blue), $\zeta = 0.2$ (black), and $U = 2.5$.

coupling. We not in passing, that band structures with a fairly similar density of states may, nevertheless, display a very different n_σ dependence of $r^{\rm SOC}$. Apparently, the full momentum dependence of the band structure determines the details of the $r^{\rm SOC}$-curves.

For finite values of the exchange interaction J, the effective coupling constants are smaller than for $J = 0$, in general. This can be seen in Figs. 5 and 6 where we show the renormalization for $J/U = 0.2$, and $U = 1, 2$ (Fig. 5), $U = 2.5$ (Fig. 6). Note that for $U = U_{\rm c} \lesssim 2.5$ there appears a Brinkmann-Rice type of insulating phase[27] at half filling where the renormalization matrix $\tilde{q}$ is zero. Therefore we could perform our calculations shown in Fig. 6 only away from half filling.

The dependence of the renormalization on the band-filling n_σ appears to be even more complicated for fi-

nite J, in particular in the region around half filling. One must keep in mind, however, that there is a 'trivial' contribution to the renormalization of ζ which simply stems from the band-width renormalization induced by the renormalization matrix $q_\sigma^{\sigma'}$. To understand this effect, we consider, for the sake of argument, a renormalization matrix of the simplest form $q_\sigma^{\sigma'} = \delta_{\sigma,\sigma'}\sqrt{q}$. In that case, the effective hopping parameters in (42) are given by $\bar{t}_{i,j}^{\sigma,\sigma'} = q t_{i,j}^{\sigma,\sigma'}$. Hence, in order to obtain the same expectation values of $|\Psi_0\rangle$ as in the non-interacting limit, we must introduce a scaling $\zeta \to q\zeta < \zeta$. The effect of the enhancement of $\zeta^{\rm eff}$ is therefore amplified by the renormalization of the hopping parameters.

For a more quantitative analysis, we define an average value $\bar{q}$ of the bandwidth renormalization through

$$\bar{q} = \langle \hat{H}_0 \rangle_{\rm G} / \langle \hat{H}_0 \rangle_0 \ , \tag{48}$$

i.e., $\bar{q}$ quantifies the reduction of the average kinetic energy in presence of the Coulomb interaction. The relative SOC-renormalization is then plotted in Fig. 7 for the same parameters as in Fig. 5. It shows that the non-trivial renormalization is, in fact, largest in the region around half filling. Moreover, it is actually fairly independent of the bare SOC, a feature that cannot be seen in the original representation of the data in Fig. 5.

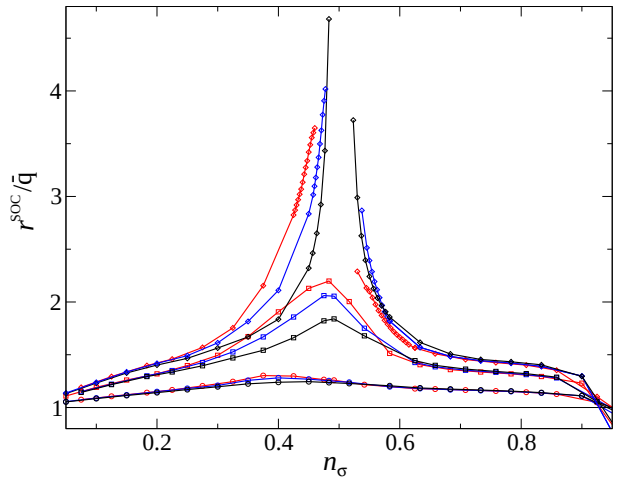

FIG. 7: Relative SOC-renormalization $r^{\rm SOC}/\bar{q}$ as a function of the orbital occupation n_σ for $J/U = 0.2$, $\zeta = 0.05$ (red), $\zeta = 0.1$ (blue), $\zeta = 0.2$ (black) and $\zeta = 0.1$ (blue), $\zeta = 0.2$ (black), and $U = 1$ (circles), $U = 2$ (squares), $U = 2.5$ (diamonds).

2. *Hund's rules in a solid?*

In the introduction we raised the question if, and to what extent, Hund's rules are still discernible in a solid. To clarify this issue, we define the three 'quantum numbers' s, l, j via the local expectation values

$$\begin{aligned} \langle \hat{\boldsymbol{S}}_i^2 \rangle_{\rm G} &= s(s+1) \ , \\ \langle \hat{\boldsymbol{L}}_i^2 \rangle_{\rm G} &= l(l+1) \ , \\ \langle (\hat{\boldsymbol{S}}_i + \hat{\boldsymbol{L}}_i)^2 \rangle_{\rm G} &= j(j+1) \ . \end{aligned} \tag{49}$$

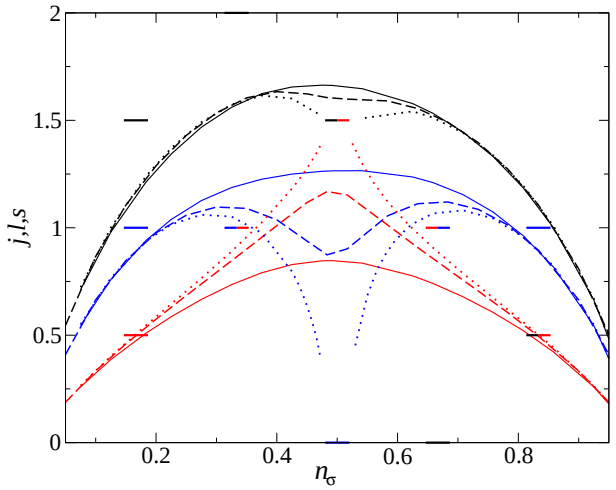

FIG. 8: Quantum numbers j (black), l (blue), s (red) as a function of the orbital occupation n_σ for $J/U = 0.2$, $\zeta = 0.05$, and $U = 1$ (solid), $U = 2$ (dashed), $U = 2.5$ (dotted).

Figure 8 shows these three numbers for $\zeta = 0.05$, $J/U = 0.2$ and $U = 1, 2, 2.5$. The bars give the values in the atomic limit, as extracted from the ground states in table I. As expected, all quantum numbers move towards their atomic values when we increase the Coulomb interaction parameters. This is best visible near half-filling when the system is close to the metal-insulator transition that appears at half filling.

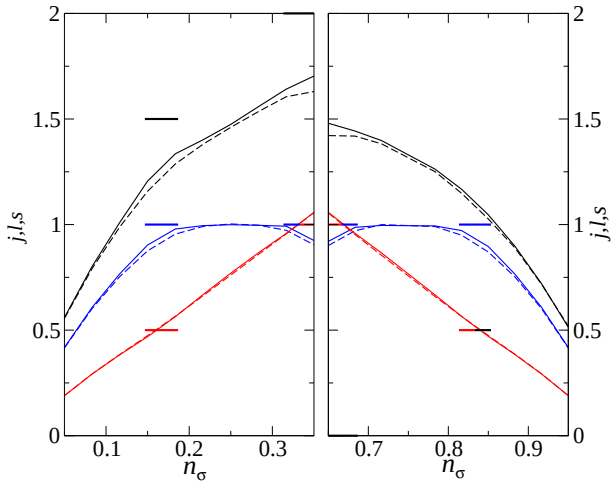

FIG. 9: Quantum numbers j (black), l (blue), s (red) as a function of the orbital occupation n_σ for $J/U = 0.2$, $\zeta = 0.05$, and $U = 6$ (solid), $U = 9$ (dashed).

As shown in previous work,[11,28] this transition is of first-order where in the Gutzwiller insulating state all atoms are in their ground state. This means that at U_c all three quantum numbers will jump to their atomic values at half filling. For all other (integer) fillings, the system is still rather itinerant and some of the quantum numbers, in particular j, deviate significantly from their atomic values. This is best visible at a filling of $n_\sigma = 2/3$ where the value of j is far off its atomic value $j_{\rm atomic} = 0$. The results change only slightly when we increase the values of U (and J) as can be seen from Fig. 9 where we display j, l, and s for larger values of U away from half filling.

The difference between the behavior close to half filling and the other integer fillings can be understood from the atomic spectra. The high-spin ground state at half filling is only slightly changed by a small SOC and, most importantly, its degeneracy is not lifted. Hence, the energy difference between the Hund's-rule ground state and the first excited state is of the order of J. In contrast, at all other integer fillings, the ground states are created by a splitting of the (degenerate) ground states at $\zeta = 0$, caused by the SOC. Therefore, the energy difference between the Hund's-rule ground state and the first excited states is much smaller away from half filling. As a consequence, it is energetically not favorable to lose a lot of kinetic energy by only occupying the Hund's-rule ground state. Unlike in the half-filled case, the Hund's-rule ground state does not dominate the quantum numbers in the metallic phase at or around other integer fillings. As seen from Figs. 8 and 9, only Hund's first rule is seen to be obeyed in strongly correlated paramagnetic metals close to integer fillings.

B. Ferromagnetic ground states

Without the spin-orbit coupling, the Hamiltonian commutes with the total spin operator. Hence, the energy of a ferromagnetic ground state cannot depend on the direction of the magnetic moment. For finite SOC, there is a preferred direction of the moment, the so-called 'easy axis'. In order to find this axis, we minimize the energy functional with respect to $|\Psi_0\rangle$ without any bias on the magnetic-moment direction using a completely general matrix $\eta_{\sigma,\sigma'}$. It turns out that in our system and for the parameters considered in this section, the magnetic moment always points into the $(1,1,1)$-direction.

1. *Ordered moment*

In Fig. 10 we display the total spin $S \equiv |\langle \hat{\boldsymbol{S}}_i \rangle|$ for seven different values of ζ ($0 \le \zeta \le 0.3$) as a function of U for $J/U = 0.2$. As seen from the figure, the SOC destabilizes the ferromagnetic order, i.e., the value U_c for noticeable ferromagnetic order ($S > 0.1$) substantially increases as a function of ζ. Concomitantly, the ordered magnetic moment $m = 2S$ strongly depends on the SOC as long as the magnetic order is weak, $S < 1/2$. The SOC becomes a small perturbation only in the saturation region, $S > 1$.

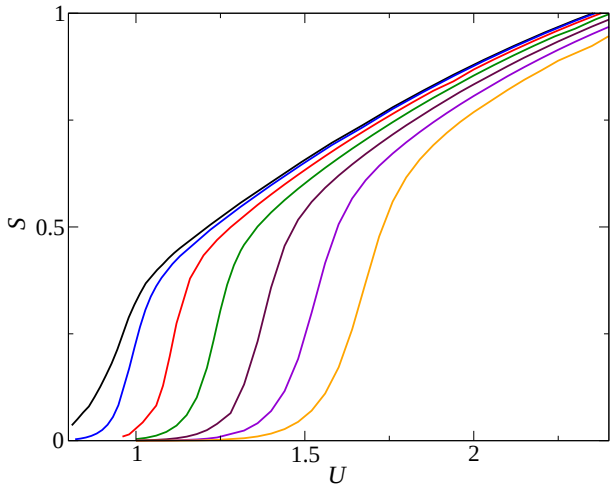

FIG. 10: Spin S in $(1,1,1)$-direction as a function of U with $J/U = 0.2$ for $n_\sigma = 0.4$ and $\zeta = 0.05$ (blue), 0.1 (red), 0.15 (green), 0.2 (maroon), 0.255 (violet), 0.3 (orange).

The SOC not only reduces the ordered spin moment, it also induces an orbital moment, i.e., $L \equiv |\langle \hat{\boldsymbol{L}}_i \rangle|$ is nonzero. This is shown in Fig. 11 where we display L for the same parameters as in Fig. 10, apart from $\zeta = 0$ where $L = 0$. The orbital contribution to the magnetic moment, however, remains rather small, of the order of 10% of the spin moment, especially for values of $\zeta < 0.1$ that are realistic for transition metals. Therefore, the gain in orbital moment does not compensate the loss in the ordered spin moment induced by the spin-orbit coupling.

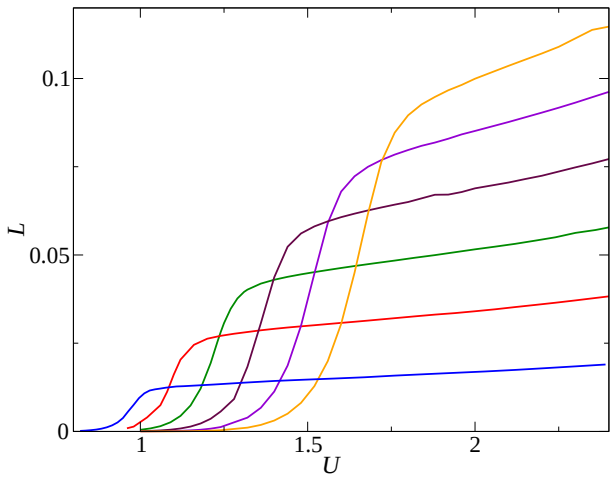

FIG. 11: Orbital moment L in (1,1,1)-direction as a function of U with $J/U = 0.2$ for $n_\sigma = 0.4$ and $\zeta = 0.05$ (blue), 0.1 (red), 0.15 (green), 0.2 (maroon), 0.255 (violet), 0.3 (orange).

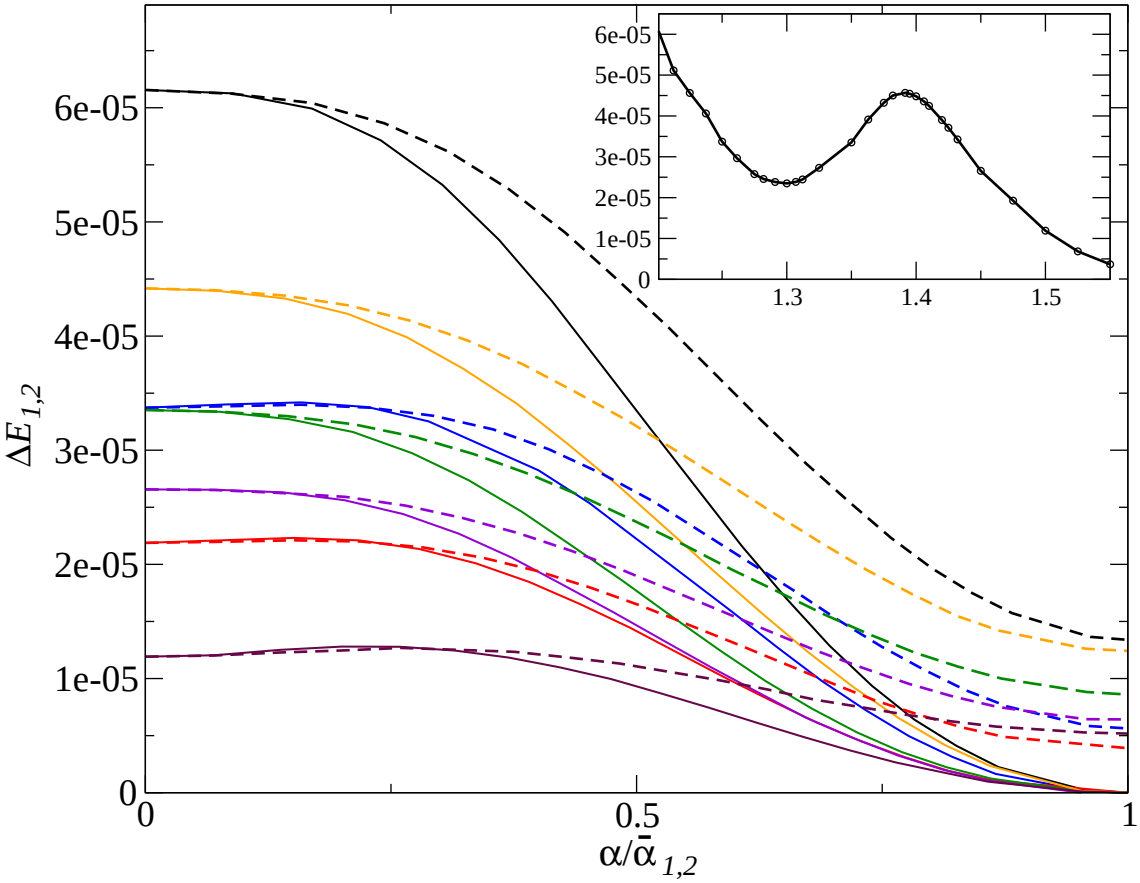

FIG. 12: Anisotropy energy $\Delta E_{1,2}$ as a function of the magnetic-moment direction that is rotated, (1), from $(1,0,0)$ to $(1,1,1)$ (solid lines) and, (2), from $(1,0,0)$ to $(1,1,0)$ (dashed lines) with maximum rotation angles $\bar{\alpha}_1 = \arccos 1/\sqrt{3}$ and $\bar{\alpha}_2 = \pi/4$; parameters: $U = 1.2$ (black), $U = 1.25$ (blue), $U = 1.3$ (red), $U = 1.35$ (green), $U = 1.4$ (orange), $U = 1.45$ (violet), $U = 1.5$ (maroon), $J/U = 0.2$, $\zeta = 0.1$; inset: maximal anisotropy energy as a function of U for $J/U = 0.2$, $\zeta = 0.1$.

2. *Anisotropy energy*

Finally, we take a look at the 'anisotropy energy', i.e., the dependence of the energy on the magnetic-moment direction. To this end, we could introduce additional constraints that fix the moment direction during the minimization. However, this would require additional programming work that we prefer to avoid. Therefore, we apply an external magnetic field that allows us to change the magnetic-moment direction. In fact, this is how the anisotropy energy would actually be measured.

Since our field is just a technical tool, we couple it to the spin only, i.e., we add

$$\hat{H}_B = -B \sum_i \boldsymbol{e}_B \cdot \boldsymbol{S}_i \tag{50}$$

to the Hamiltonian of our system. Here, $\boldsymbol{e}_B$ is the direction of the magnetic field that we adjust in our calculations. The size of the field amplitude B must be chosen with care to obtain meaningful results. On the one hand, it must be large enough to force the magnetic moment into all directions that we aim to investigate, i.e., in the ground state we must approximately find $\langle \boldsymbol{S}_i \rangle_{\mathrm{G}} || \boldsymbol{e}_B$. On the other hand, the variation in the field contribution to the energy must be small compared to the variation of the system's energy that we actually want to determine. Meeting these criteria becomes difficult, in particular, in the region of small magnetic moments. In all calculations that we are going to present below, we found that a field amplitude of $B = 0.002$ leads to meaningful results for the anisotropy energy.

In the following we consider rotations of the magnetic moment from the $(1,0,0)$ direction, (1), into the $(1,1,1)$ direction, and, (2), into the $(1,1,0)$ direction. The corresponding maximal rotation angles are $\bar{\alpha}_1 = \arccos(1/\sqrt{3})$ and $\bar{\alpha}_2 = \pi/4$, respectively. From our minimization we obtain the two energies $E_{1,2}(\alpha)$ as a function of the angle α. Since, as mentioned before, the easy axis always points into the $(1,1,1)$-direction, we define the anisotropy energy as $\Delta E_i(\alpha) \equiv E_i(\alpha) - E_1(\bar{\alpha}_1)$. This quantity is displayed in Fig. 12 for several values of U (and consequently also different values of the magnetic moment) for $J/U = 0.2$ and $\zeta = 0.1$. The figure shows that, although the anisotropy energy is quite small, of the order of several ten μeV per site, our approach is

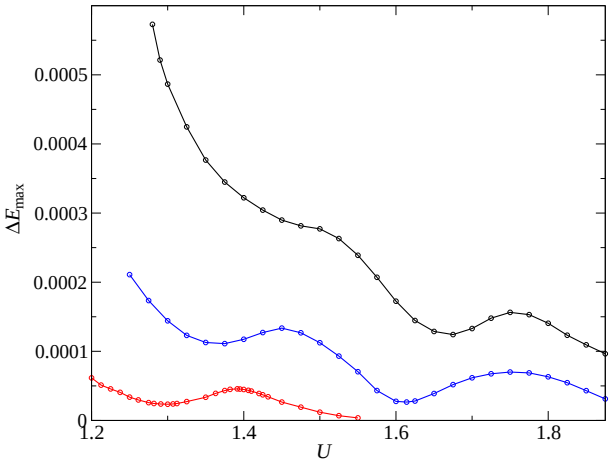

FIG. 13: Maximal anisotropy energy as a function of U for $J/U = 0.2$, $\zeta = 0.2$ (black), $\zeta = 0.15$ (blue), $\zeta = 0.1$ (red).

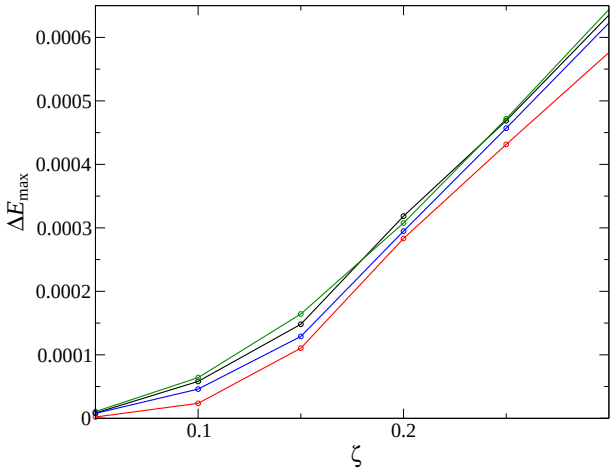

FIG. 14: Maximal anisotropy energy as a function of ζ for values of U with the same ordered spin moment of $S = 0.4$ (black), $S = 0.45$ (blue), $S = 0.5$ (red), $S = 0.6$ (green), and $J/U = 0.2$.

perfectly capable to resolve it. The maximal anisotropy energy $\Delta E_{\rm max}$, i.e., its value for $\alpha = 0$ is a non-trivial function of U. This can be seen from the inset of Fig. 12 where we display $\Delta E_{\rm max}$.

When we increase the SOC, the anisotropy energies change significantly, see Fig. 13 where we show $\Delta E_{\rm max}$ as a function of U for $\zeta = 0.1, 0.15, 0.2$. The non-monotonic behavior of $\Delta E_{\rm max}$ has its cause in the band structure. For example, the maxima in the red and blue curves and the corresponding structure in the black curve correspond to almost the same magnetization, cf. Fig. 10.

To extract the genuine ζ dependence of $\Delta E_{\rm max}$, it is best to consider states with the same moment. This is done in Fig. 14 where we display $\Delta E_{\rm max}$ as a function of ζ for values of U which lead to the same ordered spin moments. These curves reveal that the anisotropy depends very sensitively on ζ for small values of ζ whereas it becomes linear for sizable ζ.

IV. SUMMARY AND CONCLUSIONS

In this work we investigated the interplay of local Coulomb interactions and the spin-orbit coupling in a three-orbital Hubbard model in three dimensions. Based on the Gutzwiller approximation to general multi-band Gutzwiller wave functions, we find that the Coulomb interaction leads to a considerable renormalization of the effective SOC in paramagnetic metals; the spin-orbit couplings can be enhanced over their atomic values by a factor of more than two. This effect could be seen in experiment as enhanced band splittings in the quasi-particle dispersion.

Hund's rules determine spin and orbital moments of an atom. In metallic systems, signatures of Hund's rules are visible only close to half band-filling. For all other (integer) fillings, the local Hund'-rule ground states cannot dominate over states with other quantum numbers because this would be very unfavorable for the electrons' kinetic energy. At best, Hund's first rule applies in strongly correlated metallic systems close to integer fillings.

For ferromagnetic ground states, we find magnetization curves that are significantly influenced by the spin-orbit coupling. Overall, the SOC tends to destabilize the ferromagnetic order. For example, it shifts the onset of ferromagnetism to higher values of the Coulomb parameters. In the presence of an ordered spin moment, the SOC has two main effects: (i), the magnetic spin moment points into a preferred direction (easy axis), and, (ii), it generates a small but finite orbital moment in the same direction as the spin moment.

We analyzed the magnetic anisotropy by applying an external magnetic field with constant strength and varying direction. Our method is capable to resolve the anisotropy energy which can be rather small for spin-orbit couplings that are realistic for transition metals. As a function of the Coulomb interaction, the anisotropy energy shows a non-monotonic behavior which we could trace back to details of the electronic band structure.

In this study we worked with the most general Ansatz for a Gutzwiller wave function. For the calculation of anisotropy energies, it is mandatory to avoid the often used approximation of a diagonal variational-parameter matrix because this approximation results in anisotropy energies that can be off by several orders of magnitude.

For our three-band model, it is possible to include all elements of the variational-parameter matrix. Of course, this cannot be done for five d-bands. Therefore, strategies must be developed to include only the most significant matrix elements. In a separate, more technical work, we analyze in detail the importance of non-diagonal variational parameters, and show how to obtain accurate results with a properly chosen subset of such parameters.[29]

Our method can directly be applied to materials that

can be described by effective three-band models, e,g,, Sr_2RuO_4. It will be interesting to see the consequences of the substantial spin-orbit coupling on the ground-state phase diagram and other electronic properties of these systems.[8,30]

Acknowledgements

We thank R. Schade for valuable discussions on optimization algorithms.

This work was supported in part by the Priority Programme 1458 of the Deutsche Forschungsgemeinschaft (DFG) under GE 746/10-1. T.L., U.L., and F.B.A. acknowledge the financial support by the Deutsche Forschungsgemeinschaft and the Russian Foundation of Basic Research through the Transregio TRR 160.

The authors gratefully acknowledge the computing time granted by the John-von-Neumann Institute for Computing (NIC), and provided on the supercomputer JURECA at Jülich Supercomputing Centre (JSC) under project no. HDO08.

Appendix A: Energy functional and its derivatives

1. Local basis

The local density matrix (24) is non-diagonal when we include the spin-orbit coupling. For a fixed state $|\Psi_0\rangle$, however, we can always find a local basis, described by operators

$$\hat{d}^\dagger_{i,\gamma} = \sum_\sigma u_{i;\sigma,\gamma} \hat{c}^\dagger_{i,\sigma}, \quad , \quad \hat{d}_{i,\gamma} = \sum_\sigma u^*_{i;\sigma,\gamma} \hat{c}_{i,\sigma} \tag{A1}$$

and a unitary matrix $\tilde{u}_i$, so that the local density matrix $\tilde{D}_i$ is diagonal,

$$D_{i;\gamma',\gamma} \equiv \langle \hat{d}^\dagger_{i,\gamma} \hat{d}_{i,\gamma'} \rangle_{\Psi_0} = \delta_{\gamma,\gamma'} n_{i,\gamma} \ . \tag{A2}$$

Working with this new orbital basis $|\gamma\rangle$ is quite useful because the energy functional (30) as well as the constraints (22), (23) have a much simpler form, see Sect. A 3.

In general, the basis $|\gamma\rangle$ is not uniquely defined. For instance, in our three-band model without any charge or magnetic order, our local density matrix has the form

$$\tilde{C} = n_0 \mathbb{1} - \Delta n_0^{\mathrm{so}} \tilde{\Sigma} \tag{A3}$$

with $\tilde{\Sigma}$ as defined in (14). The diagonalization of (A3) leads to a two-fold and a four-fold degenerate set of states $|\gamma\rangle$ with the occupation numbers $n_0 - 2\Delta n_0^{\mathrm{so}}$ and $n_0 + \Delta n_0^{\mathrm{so}}$, respectively. Therefore, the states $|\gamma\rangle$ are defined only up to an arbitrary unitary transformation within these two degenerate sub-spaces. Even for a system with three non-degenerate orbitals there would be a remaining two-fold (Kramer's) degeneracy in the spectrum of $\tilde{C}$.

2. Atomic spectrum

We introduce the configuration basis $|I\rangle$ of the local Hilbert space,

$$|I\rangle \equiv \prod_{\sigma \in I} \hat{c}^\dagger_\sigma |0\rangle \equiv \hat{c}^\dagger_{\sigma_1} \dots \hat{c}^\dagger_{\sigma_{|I|}} |0\rangle \ , \tag{A4}$$

where the operators $\hat{c}^\dagger_\sigma$ are in ascending order, i.e., we have $\sigma_1 < \sigma_2 \dots < \sigma_{|I|}$ where $|I|$ is the number of particles in state $|I\rangle$. Using the standard mathematical notations for set operators, we frequently encounter the states $|I \cup \sigma\rangle$ or $|I \backslash \sigma\rangle$ which result from the local creation/annihilation of an electron. Since we work with fermions, we define the minus-sign function

$$\mathrm{fsgn}(\sigma, I) \equiv \langle I \cup \sigma | \hat{c}^\dagger_\sigma | I \rangle \ . \tag{A5}$$

With the basis (A4), we can readily set up the local Hamilton matrix

$$H^{\mathrm{loc}}_{I,I'} = \langle I | \hat{H}_{\mathrm{loc}} | I' \rangle \tag{A6}$$

and determine its eigenstates

$$|\Gamma\rangle = \sum_I T_{I,\Gamma} \, |I\rangle \tag{A7}$$

by standard numerical techniques. For the numerical minimization of the Gutzwiller energy functional, however, we prefer to work with the orbital states $|\gamma\rangle$ and its corresponding configuration basis

$$|J\rangle \equiv \prod_{\gamma \in J} \hat{d}^\dagger_\gamma |0\rangle \equiv \hat{d}^\dagger_{\gamma_1} \dots \hat{d}^\dagger_{\gamma_{|I|}} |0\rangle \ . \tag{A8}$$

One way to determine the expansion of $|\Gamma\rangle$ with respect to this basis,

$$|\Gamma\rangle = \sum_J A_{J,\Gamma} \, |J\rangle \ , \tag{A9}$$

would be to transform the local Hamiltonian $\hat{H}_{\mathrm{loc}}$ to the basis $|\gamma\rangle$ and to set up and diagonalize the Hamilton matrix $H^{\mathrm{loc}}_{J,J'}$. Alternatively, one may determine the eigenstates (A7) and calculate the coefficients $A_{J,\Gamma}$ in (A9) from the formula

$$\begin{aligned} A_{J,\Gamma} &= \sum_I T_{I,\Gamma} \langle J | I \rangle \ , \\ \langle J | I \rangle &= \mathrm{Det}(u^*_{\sigma_i,\gamma_j}) \quad , \quad (\sigma_i \in I, \gamma_j \in J) \ . \end{aligned} \tag{A10}$$

3. Energy functional

For a (still general) orbital basis $|\gamma\rangle$, we find the following expression for the constraints (22), (23),

$$\sum_{\Gamma,\Gamma_1,\Gamma_2} \lambda^*_{\Gamma,\Gamma_1} \lambda_{\Gamma,\Gamma_2} m^0_{\Gamma_1,\Gamma_2} = 1 \ , \tag{A11}$$

$$\sum_{\Gamma,\Gamma_1,\Gamma_2} \lambda^*_{\Gamma,\Gamma_1} \lambda_{\Gamma,\Gamma_2} m^0_{\Gamma_1 \cup \gamma, \Gamma_2 \cup \gamma'} = \delta_{\gamma,\gamma'} n_\gamma \ , \tag{A12}$$

where

$$|\Gamma \cup \gamma\rangle \;\equiv\; \hat{d}^\dagger_\gamma |\Gamma\rangle = \sum_{J(\gamma\notin J)} \mathrm{fsgn}(\gamma, J) A_{J,\Gamma} |J \cup \gamma\rangle \;, \tag{A13}$$

$$m^0_{\Gamma,\Gamma'} \;\equiv\; \langle \hat{m}_{\Gamma,\Gamma'}\rangle_{\Psi_0} \;. \tag{A14}$$

Since $|J\rangle$ is a basis of the local Hilbert space, all expectation values of the form (A14) are determined by the determinants

$$m^0_{J,J'} \equiv \langle \hat{m}_{J,J'}\rangle_{\Psi_0} = \begin{vmatrix} \Omega^{J,J'} & -\Omega^{J,\bar{J}} \\ \Omega^{\bar{J},J'} & \bar{\Omega}^{\bar{J},\bar{J}} \end{vmatrix} \;. \tag{A15}$$

Here, $\Omega_{J,J'}$ are the matrices

$$\Omega_{J,J'} = \begin{pmatrix} D_{\gamma'_1,\gamma_1} & D_{\gamma'_2,\gamma_1} & \dots & D_{\gamma'_{|J'|},\gamma_1} \\ D_{\gamma'_1,\gamma_2} & D_{\gamma'_2,\gamma_2} & \dots & D_{\gamma'_{|J'|},\gamma_2} \\ \vdots & \vdots & \ddots & \vdots \\ D_{\gamma'_1,\gamma_{|J|}} & D_{\gamma'_2,\gamma_{|J|}} & \dots & D_{\gamma'_{|J'|},\gamma_{|J|}} \end{pmatrix} \;, \tag{A16}$$

in which the entries are the elements of the uncorrelated local density matrix (A2) that belong to the configurations $J = (\gamma_1, \dots, \gamma_{|J|})$ and $J' = (\gamma'_1, \dots, \gamma'_{|J'|})$. The matrix $\bar{\Omega}^{\bar{J},\bar{J}}$ in (A15) is defined by

$$\bar{\Omega}_{\bar{J},\bar{J}} = \begin{pmatrix} 1 - D_{\gamma_1,\gamma_1} & -D_{\gamma_1,\gamma_2} & \dots & -D_{\gamma_{|\bar{J}|},\gamma_1} \\ -D_{\sigma_2,\sigma_1} & 1 - D_{\sigma_2,\sigma_2} & \dots & -D_{\gamma_{|\bar{J}|},\gamma_2} \\ \vdots & \vdots & \ddots & \vdots \\ -D_{\gamma_1,\gamma_{|\bar{J}|}} & -D_{\gamma_2,\gamma_{|\bar{J}|}} & \dots & 1 - D_{\gamma_{|\bar{J}|},\gamma_{|\bar{J}|}} \end{pmatrix} \;, \tag{A17}$$

with $\gamma_i \in \bar{J} \equiv (1, \dots, N)\backslash(J \cup J')$.

So far we have not used yet the defining condition (A2) of the $|\gamma\rangle$-basis. By applying it, the expectation values (A15) have the much simpler form

$$\begin{aligned} m^0_{J,J'} \;&=\; \delta_{J,J'} m^0_J \;, \\ m^0_J \;&=\; \prod_{\gamma\in J} n_\gamma \prod_{\gamma\notin J} (1 - n_\gamma) \;. \end{aligned} \tag{A18}$$

It is this simplification that makes the use of the $|\gamma\rangle$-basis particularly convenient in the evaluation of ground-state expectation values. For the calculation of derivatives with respect to $D_{\gamma,\gamma'}$, however, we have to start from the general expression (A15), see Sect. A 4.

With the above results, eqs. (A14)–(A17) , we can calculate the local energy as

$$E_{\mathrm{loc}} = \sum_{\Gamma,\Gamma_1,\Gamma_2} E_\Gamma \lambda^*_{\Gamma,\Gamma_1} \lambda_{\Gamma,\Gamma_2} m^0_{\Gamma_1,\Gamma_2} \;. \tag{A19}$$

In the $|\gamma\rangle$-basis we have explicitly

$$m^0_{\Gamma_1,\Gamma_2} = \sum_{J_1,J_2} A_{J_1,\Gamma_1} A^*_{J_2,\Gamma_2} m^0_{J_1,J_2} \;. \tag{A20}$$

For a ground-state calculation, this expression can be simplified further using eq. (A18).

Finally, the renormalization matrix has the form

$$\begin{aligned} q^{\gamma'}_\gamma \;=\; &\sum_{\Gamma_1,\dots,\Gamma_4} \lambda^*_{\Gamma_2,\Gamma_1} \lambda_{\Gamma_3,\Gamma_4} \langle \Gamma_2 | \hat{d}^\dagger_\gamma | \Gamma_3 \rangle \\ &\times \sum_{J_1,J_4} A_{J_1,\Gamma_1} A^*_{J_4,\Gamma_4} H^{\gamma'}_{J_1,J_4} \end{aligned} \tag{A21}$$

with

$$\begin{aligned} H^{\gamma'}_{J_1,J_4} \;\equiv\; &(1 - f_{\gamma',J_1}) \langle J_4 | \hat{d}_{\gamma'} | J_4 \cup \gamma' \rangle m^0_{J_1,J_4\cup\gamma'} \\ &+ \left(f_{\gamma',J_4} m^0_{J_1\backslash\gamma',J_4} + (1 - f_{\gamma',J_4}) m^{0;\gamma'}_{J_1\backslash\gamma',J_4} \right) \\ &\times \langle J_1\backslash\sigma' | \hat{d}_{\gamma'} | I_1 \rangle \;, \end{aligned} \tag{A22}$$

and

$$f_{\gamma,J} \equiv \langle J | \hat{d}^\dagger_\gamma \hat{d}_\gamma | J \rangle \tag{A23}$$

is either zero or unity. Here, the expectation value $m^{0;\gamma'}_{J_1\backslash\gamma',J_4}$ has the same form as the one in (A15), except that the index $\bar{J}$ has to be replaced by $\bar{J}\backslash\gamma'$. We need the general result (A21) for the renormalization matrix for the calculation of derivatives with respect to non-diagonal elements of $D_{\gamma,\gamma'}$, see Sect. A 4. For a ground-state calculation one can use eq. (A2) and obtain the simpler expression

$$q^{\gamma'}_\gamma = \frac{1}{n_{\sigma'}} \sum_{\Gamma_1\dots\Gamma_4} \lambda^*_{\Gamma_2,\Gamma_1} \lambda_{\Gamma_3,\Gamma_4} \langle \Gamma_2 | \hat{d}^\dagger_\gamma | \Gamma_3 \rangle m^0_{\Gamma_1,\Gamma_4\cup\gamma'} \tag{A24}$$

which may also be written in the form[31]

$$q^{\gamma'}_\gamma = \frac{1}{n_{\sigma'}} \langle \hat{P}^\dagger \hat{d}^\dagger_\gamma \hat{P} \hat{d}_{\gamma'} \rangle_{\Psi_0} \;. \tag{A25}$$

In summary, the Gutzwiller energy functional in the $|\gamma\rangle$-basis is given as

$$\begin{aligned} E_{\mathrm{G}}(\boldsymbol{v}, \tilde{\rho}, \tilde{D}) \;=\; &\sum_{\substack{\gamma_1,\gamma_2 \\ \gamma'_1,\gamma'_2}} q^{\gamma'_1}_{\gamma_1} \left(q^{\gamma'_2}_{\gamma_2} \right)^* E_{\gamma_1,\gamma_2,\gamma'_1,\gamma'_2} \\ &+ \sum_{\Gamma,\Gamma_1,\Gamma_2} E_\Gamma \lambda^*_{\Gamma,\Gamma_1} \lambda_{\Gamma,\Gamma_2} m^0_{\Gamma_1,\Gamma_2} \;. \end{aligned} \tag{A26}$$

Here, we applied the transformation to the $|\gamma\rangle$-basis

$$E_{\gamma_1,\gamma_2,\gamma'_1,\gamma'_2} = \sum_{\substack{\sigma_1,\sigma_2 \\ \sigma'_1,\sigma'_2}} u^*_{\sigma_1,\gamma_1} u_{\sigma_2,\gamma_2} u_{\sigma'_1,\gamma'_1} u^*_{\sigma'_2,\gamma'_2} E_{\sigma_1,\sigma_2,\sigma'_1,\sigma'_2} \tag{A27}$$

and

$$q^{\gamma'}_\gamma = \sum_{\sigma,\sigma'} u_{\sigma,\gamma} u^*_{\sigma',\gamma'} q^{\sigma'}_\sigma \;. \tag{A28}$$

4. Derivatives

The minimization algorithm which we explain in this section requires the calculation of derivatives of the energy and of the constraints with respect to the variational parameters v_z and the local density matrix $\tilde{C}$ or $\tilde{D}$.

a. Derivatives with respect to v_Z

The constraints, the local energy, and the renormalization factors are all quadratic functions of the variational parameters v_z, i.e., they are of the form

$$f(\boldsymbol{v}) = \sum_{Z,Z'} f_{Z,Z'} v_{Z'} v_Z \ . \tag{A29}$$

The fast calculation of derivatives

$$\partial_{v_Z} f(\boldsymbol{v}) = \sum_{Z'} (f_{Z,Z'} + f_{Z,Z'}) v_{Z'} \tag{A30}$$

is then possible if all coefficients $f_{Z,Z'}$ are stored in the main memory. In our calculations we observe that the number of contributing coefficients $f_{Z,Z'}$ in the expansion is particularly large in the renormalization factors when we include non-diagonal elements in the variational parameter matrix $\lambda_{\Gamma,\Gamma'}$. Hence, our minimization for the three-orbital model that includes all $n_v = 924$ non-diagonal variational parameters is numerically much more demanding than the minimization, e.g., for a five-orbital model with only diagonal parameters ($n_v = 1024$).

b. Derivatives with respect to $C_{\sigma,\sigma'}$

For the calculation of the effective on-site energies (45), we need to determine the derivatives of the energy and of the constraints with respect to $C_{\sigma,\sigma'}$. Again, it is easier to calculate the derivatives first in the γ-basis and then transform them via

$$\frac{\partial}{\partial C_{\sigma,\sigma'}} = \sum_{\gamma,\gamma'} u^*_{\sigma,\gamma} u_{\sigma',\gamma'} \frac{\partial}{\partial D_{\gamma,\gamma'}} \ . \tag{A31}$$

For the derivatives of the constraints and of the local energy, we just need to determine the derivative of (A15). This gives

$$\frac{\partial}{\partial D_{\gamma,\gamma}} m^0_{J,J'} = \delta_{J,J'} m^0_{J,J} \begin{cases} 1/n_\gamma & \text{for } \gamma \in J \\ -1/(1-n_\gamma) & \text{for } \gamma \notin J \end{cases} \tag{A32}$$

for $\gamma = \gamma'$, and

$$\frac{\partial}{\partial D_{\gamma',\gamma}} m^0_{J,J'} = \delta_{\bar{I},I\setminus\gamma} \delta_{\bar{I},I'\setminus\gamma'} \frac{m^0_{\bar{I},\bar{I}}}{(1-n_\gamma)(1-n_{\gamma'})} \tag{A33}$$

for $\gamma \neq \gamma'$, where $\gamma \in J$ and $\gamma' \in J'$. The only remaining problem is to calculate derivatives of the object $m^{0;\bar{\gamma}}_{J,J'}$ that appears in the definition of the renormalization matrix, eqs. (A21), (A22) with respect to $D_{\gamma',\gamma}$. It contributes only when $\gamma \neq \bar{\gamma}$ and $\gamma' \neq \bar{\gamma}$. Then we can use the simple relationship

$$\frac{\partial}{\partial D_{\gamma',\gamma}} m^{0;\bar{\gamma}}_{J,J'} = \frac{1}{1-n_{\bar{\gamma}}} \frac{\partial}{\partial D_{\gamma',\gamma}} m^0_{J,J'} \ . \tag{A34}$$

Appendix B: Minimization algorithm

1. Inner minimization

For a given single-particle state $|\Psi_0\rangle$, or, equivalently, a given single-particle density matrix $\tilde{\rho}$, we have to minimize the energy functional (34) obeying the constraints (40). In Ref. [25] we introduced a very efficient method for this minimization which was used in a number of previous studies, for example on elementary iron and nickel.[17,18] This method, however, is only applicable if the gradients

$$\boldsymbol{F}^l \equiv \partial_{\boldsymbol{v}} g_l(\boldsymbol{v}) \tag{B1}$$

of the constraints (40) are linearly independent because it requires a matrix $W_{l,l'} \equiv \boldsymbol{F}^l \cdot \boldsymbol{F}^{l'}$ to be regular.

In principle, this problem can be overcome by a group-theoretical analysis that identifies the maximum set of independent constraints. Such a solution, however, is rather cumbersome and it runs into difficulties if one aims to study the transition between minima with different point-group symmetries. Even if we ensure that the gradients $\boldsymbol{F}^l$ are linearly independent, however, we observe that the algorithm introduced in Ref. [25] becomes prohibitively slow when we aim to minimize the energy functional for a general (complex) variational parameter matrix $\lambda_{\Gamma,\Gamma'}$.

For this reason we tested a couple of alternative minimization algorithms that are discussed in textbooks on numerical optimization.[32] We found the 'Penalty and Augmented Lagrangian Method' (PALM) to be most useful in our context when combined with an unconstrained Broyden-Fletcher-Goldfarb-Shanno (BFGS) minimization. We shall briefly summarize these methods in the following.

a. PALM

In the PALM one studies the functional

$$L^{\mathrm{PALM}}_{\mathrm{G}}(\boldsymbol{v}, \{\Lambda_l\}, \mu) \equiv E_{\mathrm{G}}(\boldsymbol{v}) - \sum_l \Lambda_l g_l(\boldsymbol{v}) + \frac{\mu}{2} \sum_l [g_l(\boldsymbol{v})]^2 \ , \tag{B2}$$

which contains Lagrange parameter terms ($\sim \Lambda_l$) and penalty terms ($\sim \mu$). In a pure 'penalty method' one would set $\Lambda_l = 0$ and minimize (B2) for a given value of

$\mu > 0$. If, in the minimum $\boldsymbol{v} = \boldsymbol{v}_0$, the constraints are sufficiently well fulfilled, i.e.,

$$\sum_l g_l(\boldsymbol{v}_0)^2 < g_c^2 \tag{B3}$$

with some properly chosen value of g_c, we may consider $E_0 = E_G(\boldsymbol{v}_0)$ as a decent approximation for the Gutzwiller ground-state energy. Otherwise, we increase μ and start another minimization.

For our Gutzwiller energy functional it turns out that the convergence to the minimum is much faster when we use a full PALM algorithm with Lagrange parameters $\Lambda_l \neq 0$. This method works as follows.[32]

(i) Start from some initial values $\Lambda_l = \Lambda_{l;0}$ and $\mu = \mu_0$, e.g., $\Lambda_{l;0} = 1$ and $\mu_0 = 50|E_G(\boldsymbol{v}^{nc})|$ where v_l^{nc} are the variational parameters in the non-interacting limit, i.e., with $\lambda_{\Gamma,\Gamma'} = \delta_{\Gamma,\Gamma'}$.

(ii) Minimize

$$L_{G;0}^{PALM}(\boldsymbol{v}) \equiv L_G^{PALM}(\boldsymbol{v}, \{\Lambda_{l;0}\}, \mu_0) \tag{B4}$$

with respect to $\boldsymbol{v}$. For this step we use the method of steepest descent combined with the BFGS method, see Sect. B 1 b. We denote the minimum found in step (ii) by $\boldsymbol{v}_0$.

(iii) Set

$$\Lambda_{l;k+1} = \Lambda_{l;k} - \mu_k g_l(\boldsymbol{v}_0) , \tag{B5}$$
$$\mu_{k+1} = \beta\mu_k \tag{B6}$$

with some properly chosen number $\beta > 1$. In our calculations we worked with $\beta = 2$.

(iv) Go back to step (ii) until eq. (B3) is satisfied.

b. *Steepest decent and BFGS method*

We still have to choose a method for the unconstrained minimization in step (ii) in the PALM. It is a major advantage of our Gutzwiller minimization that calculating gradients of the energy or of the constraints works just as fast as the calculation of these objects themselves. Of course, this is only the case when we use eq. (A30) and do not try to calculate the gradients numerically from the difference quotient.

Let $E(\boldsymbol{v})$ be our functional and

$$\boldsymbol{F}_0 = \partial_{\boldsymbol{v}} E(\boldsymbol{v})|_{\boldsymbol{v}=\boldsymbol{v}_0} \tag{B7}$$

its gradient at the point $\boldsymbol{v}_0$. Then the simplest way of minimizing $E(\boldsymbol{v})$ is the 'method of steepest descent' where the one-dimensional function

$$\Delta E(\alpha) = E(\boldsymbol{v}_0 + \alpha \boldsymbol{F}_0) \tag{B8}$$

is minimized with respect to α. Instead of the optimal value $\alpha = \alpha_0$, in practical numerics we use a value $\tilde{\alpha}_0$ that reduces the value of our functional $E(\boldsymbol{v})$. We calculate a new point $\boldsymbol{v}_0 \to \boldsymbol{v}_0 + \tilde{\alpha}_0 \boldsymbol{F}_0$ and reiterate the procedure until $|\boldsymbol{F}_0|$ is below a pre-determined threshold. It is the decisive advantage of this method that it always converges towards a (potentially local) minimum as long as the functional is well-behaved, which we can take for granted in physics. The main disadvantage of the method is its rather slow convergence. Therefore, we found it necessary to combine it with a faster algorithm, the BFGS method, which, however, works reliably only in the vicinity of the minimum.

The starting point of the BFGS method is a second-order expansion of the functional

$$E(\boldsymbol{v}_0 + \delta\boldsymbol{v}) \approx E(\boldsymbol{v}_0) + \boldsymbol{F}_0 \cdot \delta\boldsymbol{v} + \frac{1}{2}\delta\boldsymbol{v}^{\mathrm{T}} \cdot \tilde{H}_0 \cdot \delta\boldsymbol{v} , \tag{B9}$$

where $\tilde{H}_0$ is the Hessian matrix of second derivatives at the point $\boldsymbol{v}_0$. Provided that $\tilde{H}_0$ is positive definite, the right-hand site is minimized for

$$\delta\boldsymbol{v} = -\tilde{B}_0 \cdot \boldsymbol{F}_0 , \tag{B10}$$

where $\tilde{B}_0 = \tilde{H}_0^{-1}$. Making iterative steps in the variational parameter space by means of eq. (B10) is a multi-dimensional version of the Newton method.

The main obstacle of the Newton method is the numerical calculation of $\tilde{H}_0$ and the solution of eq. (B10). Therefore, it is better to use a so-called 'quasi Newton method' of which BFGS is one example. This method employs eq. (B10) without calculating $\tilde{B}_0$ (or $\tilde{H}_0$) exactly. It works as follows.[32]

(i) Start at some point $\boldsymbol{v}_k$ and calculate the gradient $\boldsymbol{F}_k$ and the inverse $\tilde{B}_k$ of the Hessian matrix. Due to the benign structure of our functional we can afford this initial calculation of $\tilde{B}_k$ because it is done only once.

(ii) Calculate the new point

$$\boldsymbol{v}_{k+1} = \boldsymbol{v}_k - \tilde{B}_k \cdot \boldsymbol{F}_k . \tag{B11}$$

(iii) Calculate $\boldsymbol{F}_{k+1}$ from

$$\boldsymbol{F}_{k+1} = \partial_{\boldsymbol{v}} E(\boldsymbol{v})|_{\boldsymbol{v}=\boldsymbol{v}_{k+1}} \tag{B12}$$

and an approximate update of $\tilde{B}_k$ from

$$\tilde{B}_{k+1} = (\tilde{1} - \alpha_k \boldsymbol{s}_k \boldsymbol{y}_k^{\mathrm{T}})\tilde{B}_{k+1}(\tilde{1} - \alpha_k \boldsymbol{y}_k \boldsymbol{s}_k^{\mathrm{T}}) + \alpha_k \boldsymbol{s}_k \boldsymbol{s}_k^{\mathrm{T}} , \tag{B13}$$

where

$$\boldsymbol{s}_k \equiv \boldsymbol{v}_{k+1} - \boldsymbol{v}_k ,$$
$$\boldsymbol{y}_k \equiv \boldsymbol{F}_{k+1} - \boldsymbol{F}_k ,$$
$$\alpha_k \equiv \boldsymbol{y}_k^{\mathrm{T}} \boldsymbol{s}_k . \tag{B14}$$

(iv) Go back to step (ii) until $|\boldsymbol{F}_k|$ is below some predefined threshold.

Within the BFGS method it is not ensured that going from $\boldsymbol{v}_k$ to $\boldsymbol{v}_{k+1}$ always leads to a decrease of our functional. Therefore, we need the method of steepest decent as a backup to reach a region in the variational parameter space where the BFGS method converges.

2. Outer minimization

Given the optimum variational parameters $\boldsymbol{v}_0$ from the inner minimization we need to determine a new single-particle state by means of eqs. (42)–(46). All derivatives in eqs. (45)–(46) are calculated with the formulae given in eq. (B7). Then, the remaining problem is the calculation of the Lagrange parameters Λ_l from eqs. (46). The number n_v of these linear equations is usually much larger than the number of Lagrange parameters n_c. Due to a possible inter-dependence of the constraints, the solution of the equations may not be unique. Hence, we cannot use the trick of Ref. [25] (see Sec. 4.2.1 of that work), which led to a number of n_c linear equations.

Here, we choose to determine one of the infinitely many possible sets of Lagrange parameters by minimizing the functional

$$Y(\{\Lambda_l\}) = \sum_Z \left(\left.\frac{\partial E_\mathrm{G}}{\partial v_Z}\right|_{\boldsymbol{v}=\boldsymbol{v}_0} - \sum_l \Lambda_l \left.\frac{\partial g_l}{\partial v_Z}\right|_{\boldsymbol{v}=\boldsymbol{v}_0} \right)^2 \tag{B15}$$

with respect to Λ_l. Note that the lack of uniqueness for the Lagrange parameters Λ_l has no consequences for the fields (45). The latter are always uniquely defined, apart from a total energy shift that can be absorbed in the chemical potential.

With the fields (45) and the renormalization matrix determined, we diagonalize (42) and determine $|\Psi_0\rangle$ by means of the standard tetrahedron method.

[1] F. Hund. *Zeitschrift für Physik*, 40:742, 1927.
[2] F. Hund. *Zeitschrift für Physik*, 42:93, 1927.
[3] J.-P. Julien, J.-X. Zhu, and R. C. Albers. *Phys. Rev. B*, 77:195123, 2008.
[4] L. Du, L. Huang, and X. Dai. *Eur. Phys. J. B*, 86:94, 2013.
[5] L. Du, X. Sheng, H. Weng, and X. Dai. *Europhys. Lett.*, 101:27003, 2013.
[6] A. Farrell and T. Pereg-Barnea. *Phys. Rev. B*, 89:035112, 2014.
[7] H. Shinaoka, S. Hoshino, M. Troyer, and P. Werner. *Phys. Rev. Lett.*, 115:156401, 2015.
[8] G. Zhang, E. Gorelov, E. Sarvestani, and E. Pavarini. *Phys. Rev. Lett.*, 116:106402, 2016.
[9] Z. Y. Meng, Y. B. Kim, and H.-Y. Kee. *Phys. Rev. Lett.*, 113:177003, 2014.
[10] A. O. Shorikov and V. I. Anisimov. arxiv:1412.0140.
[11] J. Bünemann, W. Weber, and F. Gebhard. *Phys. Rev. B*, 57:6896, 1998.
[12] W. Metzner and D. Vollhardt. *Phys. Rev. Lett.*, 59:121, 1987.
[13] F. Gebhard. *Phys. Rev. B*, 41:9452, 1990.
[14] K. zu Münster and J. Bünemann. *Phys. Rev. B*, 2016.
[15] K. M. Ho, J. Schmalian, and C. Z. Wang. *Phys. Rev. B*, 77:073101, 2008.
[16] X. Deng, X. Dai, and Z. Fang. *Europhys. Lett.*, 83:37008, 2008.
[17] T. Schickling, J. Bünemann, F. Gebhard, and W. Weber. *New Journal of Physics*, 16:93034, 2014.
[18] T. Schickling, J. Bünemann, L. Boeri, and F. Gebhard. *Phys. Rev. B*, 2016.
[19] J. C. Slater and G. F. Koster. *Phys. Rev.*, 94:1498, 1954.
[20] E. C. Stoner. *Proc. of the Royal Soc. London A*, 165:372, 1938.
[21] S. Sugano, Y. Tanabe, and H. Kamimura. *Multiplets of Transition-Metal Ions in Crystals*. Pure and Applied Physics 33, Academic Press, New York, 1970.
[22] J. Bünemann, F. Gebhard, and W. Weber. In A. Narlikar, editor, *Frontiers in Magnetic Materials*. Springer, Berlin, 2005.
[23] J. Büneman, T. Schickling, and F. Gebhard. *Europhys. Lett.*, 98:27006, 2012.
[24] J. Kaczmarczyk, J. Spałek, T. Schickling, and J. Bünemann. *Phys. Rev. B*, 88:115127, 2013.
[25] J. Bünemann, F. Gebhard, T. Schickling, and W. Weber. *physica status solidi (b)*, 249:1282, 2012.
[26] J. Bünemann, F. Gebhard, and R. Thul. *Phys. Rev. B*, 67:75103, 2003.
[27] W. F. Brinkman and T. M. Rice. *Phys. Rev. B*, 2:4302, 1970.
[28] J. Bünemann and W. Weber. *Phys. Rev. B*, 55:4011, 1997.
[29] J. Bünemann, T. Linneweber, and F. Gebhard. *physica status solidi*, 2016.
[30] S. Hoshino and P. Werner. *Phys. Rev. Lett.*, 115:247001, 2015.
[31] M. Fabrizio. *Phys. Rev. B*, 76:165110, 2007.
[32] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, New York, 2nd edition, 2006.