Gutzwiller variational wave function for multiorbital Hubbard models in finite dimensions

Kevin zu Münster and Jörg Bünemann
Fachbereich Physik, Philipps Universität Marburg, D-35032 Marburg, Germany
(Received 30 October 2015; revised manuscript received 12 June 2016; published 26 July 2016)

We develop a diagrammatic method for the evaluation of general multiband Gutzwiller wave functions in finite dimensions. Our approach provides a systematic improvement of the widely used Gutzwiller approximation. As a first application, we investigate itinerant ferromagnetism and correlation-induced deformations of the Fermi surface for a two-band Hubbard model on a square lattice.

DOI: 10.1103/PhysRevB.94.045135

I. INTRODUCTION

Strongly correlated electron systems display a variety of intriguing phases, such as superconductivity, (anti)ferromagnetism, or Mott insulating phases. In order to study the fundamental properties of strongly correlated lattice systems, simplifying Hubbard-type models are often employed. Unfortunately, the calculation of ground state and dynamical properties is notoriously difficult even for these relatively simple models.

In one dimension, the density-matrix renormalization group (DMRG) method permits the numerical investigation of Hubbard-type models for fairly large chains. However, even modern variants of the DMRG such as tensor network approaches [1], are not satisfactory when applied to many-orbital models or higher-dimensional systems. In the limit of infinite dimensions, the dynamical mean-field theory (DMFT) [2] maps the problem onto a single-impurity model whose spectral function must be calculated numerically. For multiband systems, the solution of this task requires sophisticated quantum Monte-Carlo techniques and substantial computational resources. Concomitantly, it is very difficult to go beyond the mean-field limit and access multiband models in finite dimensions.

In the absence of exact analytical or quasiexact numerical methods, variational approaches have proven helpful. In this work, we employ the Gutzwiller wave function approach [3]. The evaluation of expectation values with the Gutzwiller correlated many-particle wave function poses itself a difficult many-body problem. Therefore even the Gutzwiller-correlated single-band Fermi sea can be evaluated exactly only in one dimension [4–9]. In the limit of infinite dimensions, the so-called Gutzwiller approximation (GA) becomes exact for the single-band Hubbard model [4,10]. Later, the method was extended to the multiband case [11–14]. Recently, it has been combined with the density functional theory in a self-consistent manner to describe transition metals and their compounds [15–25].

Despite many successes of the GA in improving our understanding of correlated metals, there are certain phenomena which it cannot describe properly. For example, in single-band models, the Fermi surface is independent of the local Coulomb interaction within the GA, unless a state with broken spin or translational symmetry is considered. This is obviously incorrect, as can be seen already from straightforward perturbation theory for the paramagnetic ground state [26,27]. In order to describe a Fermi-surface deformation, one needs to evaluate the Gutzwiller wave function in finite dimensions. A well established way to do this, is the "variational Monte Carlo method" in which the Gutzwiller energy functional is minimized numerically on finite lattices [28–30]. Although numerically less demanding than other techniques, such as quantum Monte Carlo, this method still has significant finite-size limitations.

An alternative approach, which has first been proposed in Refs. [10,31], constitutes a systematic improvement of the GA for Gutzwiller wave functions on finite dimensional lattices. The method has been used successfully to study Fermi-surface deformations, d-wave superconductivity, and quasiparticle band structures in single-band Hubbard models [31–34], t-J models [35], and periodic Anderson models [36].

Most transition metals and their compounds cannot be described properly by single-band models. For example, in iron pnictides, such as LaOFeAs, all five d orbitals are partially occupied and may have to be taken into account in any model study that aims to describe the superconductivity or the antiferromagnetism in these systems [37,38]. Hence, it is clearly desirable to generalize the method, developed in Ref. [31] for the single-band model, to the multiband case. It is the main purpose of this work, to formulate such a generalization. As a first application, we shall study ferromagnetism and Fermi-surface deformations in a two-orbital Hubbard model on a square lattice.

Our work is organized as follows. In Sec. II, we introduce the multiband Hubbard model and the corresponding Gutzwiller wave function. Moreover, we present a detailed derivation of the diagrammatic expansion for ground-state expectation values. In Sec. III, we discuss our results for ferromagnetism and interaction-induced deformations of the Fermi surface in a two-orbital Hubbard model on a square lattice. Finally, Sec. IV summarizes our findings and gives a brief outlook. For some technical details, we refer to three Appendices. A more detailed derivation of the results presented in this work can be found in Ref. [39].

II. GUTZWILLER WAVE FUNCTIONS

In this work, we employ Gutzwiller variational functions for multiband Hubbard models. Such wave functions start from an independent-particle picture where the electrons are distributed over all lattice sites to optimize the single-particle energy. This statistical distribution leads to atomic configurations that are energetically unfavorable for finite Hubbard interactions. In the Gutzwiller wave function, the

weight of such configurations is reduced with the help of the Gutzwiller "correlator," a product of local operators, see below. Nonlocal ("extended") Gutzwiller correlators for the single-band Hubbard model can be studied analytically [40], and numerically using variational Monte Carlo [28–30]. More recently, a two-dimensional bilayer Hubbard model was studied with the same method [41].

A. Multiband Hubbard model

In this work, we investigate Gutzwiller-correlated wave functions for *general* multiband Hubbard models. Only in the numerical applications in Sec. III, we shall be more specific by considering a two-orbital Hubbard model on a square lattice where the degenerate orbitals obey a $p_x - p_y$ symmetry. The Hubbard Hamilton operator with purely local interactions reads

$$\hat{H} = \hat{H}_0 + \hat{U}, \tag{1}$$

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

$$\hat{U} = \sum_{i,\sigma_1,\ldots,\sigma_4} U_{\sigma_1\sigma_2\sigma_3\sigma_4} \hat{c}_{i,\sigma_1}^\dagger \hat{c}_{i,\sigma_2}^\dagger \hat{c}_{i,\sigma_3} \hat{c}_{i,\sigma_4}. \tag{3}$$

Here, $\hat{c}_{j,\sigma}^\dagger$ and $\hat{c}_{j,\sigma'}$ are fermionic creation and annihilation operator, respectively. The site index is given by i and j and the combined spin-orbital index by σ. The lattice indices run over all lattice sites of the lattice Λ. Periodic boundary conditions apply. The hopping amplitudes $t_{ij}^{\sigma\sigma'}$ and the coefficients $U_{\sigma_1\sigma_2\sigma_3\sigma_4}$ of the on-site interaction energy are considered to be free model parameters.

The hopping and Coulomb parameters are restricted by symmetry. Spin conservation and rotational symmetry of the $p_x - p_y$ orbitals reduce the nearest-neighbor and next-nearest-neighbor hopping amplitudes to four independent parameters. Furthermore, the coefficients of the on-site energy can be expressed solely in terms of the Hubbard interaction U and the Hund's-rule coupling J, as shown in Appendix A. Note that the symmetry of the $p_x - p_y$ orbitals is the same as that of a pair of $d_{xz} - d_{yz}$ orbitals in our two-dimensional model. Therefore our two-band Hubbard model applies to $p_x - p_y$ orbitals and to $d_{xz} - d_{yz}$ orbitals equally well.

B. Definition of Gutzwiller variational states

The Gutzwiller correlator is given by the product of the local Gutzwiller correlators for all sites l on our lattice Λ,

$$\hat{P}_G = \prod_{l \in \Lambda} \hat{P}_l. \tag{4}$$

If the context does not lead to any ambiguities, the local index l will frequently be dropped in the following. In this work, we restrict ourselves to the homogeneous case where the variational parameters in $\hat{P}_l$ are the same for all lattice sites. The local Gutzwiller operator is given by

$$\hat{P}_l = \sum_{I_1,I_2} \lambda_{I_1,I_2} (|I_1\rangle\langle I_2|)_l, \tag{5}$$

$$\hat{P}_l^\dagger \hat{P}_l = \sum_{I_1,I_2} \bar{\lambda}_{I_1,I_2} (|I_1\rangle\langle I_2|)_l, \tag{6}$$

with

$$\bar{\lambda}_{I_1,I_2} = \sum_J \lambda_{I_2,J} \lambda_{I_1,J}, \tag{7}$$

where we already assumed that the parameters λ_{I_1,I_2} are real. The operators in (5) and (6), which act on the site i, can be written explicitly as

$$(|I_1\rangle\langle I_2|)_i = \prod_{l \in \Lambda \setminus i} \mathrm{Id}_l \otimes (|I_1\rangle\langle I_2|)_i, \tag{8}$$

where Id_l represents the identity operator on site l. In our two-band application, the local indices I_1, I_2 run over all 16 local configurations, which can contain up to four electrons.

In order to simplify the notation, we define a product of local creation or annihilation operators by the introduction of the following symbols:

$$\hat{C}_I^\dagger = \prod_{\sigma \in I} \hat{c}_\sigma^\dagger = \hat{c}_{\sigma_1}^\dagger \ldots \hat{c}_{\sigma_n}^\dagger \quad i < j \to \sigma_i < \sigma_j, \tag{9}$$

$$\hat{C}_I = \prod_{\sigma \in I} \hat{c}_\sigma = \hat{c}_{\sigma_1} \ldots \hat{c}_{\sigma_n} \quad i < j \to \sigma_i > \sigma_j, \tag{10}$$

where we introduced some arbitrary order of the spin-orbit indices σ. The multiparticle states

$$|I\rangle = \hat{C}_I^\dagger |\mathrm{vac}\rangle \tag{11}$$

are uniquely determined by the lexicographical order of their subindices σ_i in (9). Note that we consider $I = \{\sigma_1, \ldots, \sigma_n\}$ as sets of spin-orbit indices σ_i. Hence all standard mathematical set operations are well defined, e.g., $I \setminus \sigma$ or $I_1 \cup I_2$. Moreover, we define $|I|$ as the number of elements in I, i.e., the number of electrons in the state $|I\rangle$.

In this work, we will develop our diagrammtic formulism only for systems without superconductivity. Hence we can safely assume that the variational parameters λ_{I_1,I_2} in (5) are nonzero only when $|I_1| = |I_2|$. The generalization of our method to superconducting systems is straightforward, see Appendix C.

A single-particle product state (SPPS) can always be cast in the form

$$|\Psi_0\rangle = \prod_{k,\gamma} \hat{h}_{k,\gamma}^\dagger |\mathrm{vac}\rangle \tag{12}$$

in some fermionic basis

$$\hat{h}_{k,\gamma}^\dagger = \sum_{i,\sigma} U_{\sigma,\gamma}^{i,k} \hat{c}_{i,\sigma}^\dagger. \tag{13}$$

We will assume that the SPPS are normalized, $\langle\Psi_0|\Psi_0\rangle = 1$, and that the canonical commutation relations hold, $\{\hat{h}_{k,\gamma}^\dagger, \hat{h}_{k',\gamma'}\} = \delta_{kk'}\delta_{\gamma\gamma'}$. Now, we define the Gutzwiller wave function as

$$|\Psi_G\rangle = \hat{P}_G |\Psi_0\rangle. \tag{14}$$

In the remaining part of this work, we optimize the Gutzwiller correlator $\hat{P}_G$ and the SPPS $|\Psi_0\rangle$ so that the approximate ground-state energy

$$E_G = \langle \hat{H} \rangle_G = \frac{\langle\Psi_G|\hat{H}|\Psi_G\rangle}{\langle\Psi_G|\Psi_G\rangle} \tag{15}$$

becomes minimal.

C. Diagrammatic expansion in finite dimensions

We consider the expectation value of some local operator $\hat{O}_i$

$$\langle \hat{O}_i \rangle_G = \frac{\langle \Psi_G | \hat{O}_i | \Psi_G \rangle}{\langle \Psi_G | \Psi_G \rangle}. \tag{16}$$

Note that the following derivation can equally be performed for expectation values of nonlocal operators such as $\langle \hat{c}^\dagger_{i,\sigma} \hat{c}_{j,\sigma'} \rangle_G$, see Appendix C.

As a first step, we follow the analysis for the single-band case derived in Refs. [10,31] and partly worked out for the multiband case in infinite dimensions in Ref. [42]. In the numerator of Eq. (16), we pull the Gutzwiller correlators with indices $l \neq i$ to the right side of $\hat{O}_i$ and denote the sandwich $\hat{P}^\dagger_i \hat{O}_i \hat{P}_i$ as $\hat{Q}_i$,

$$\langle \Psi_G | \hat{O}_i | \Psi_G \rangle = \langle \Psi_0 | \hat{Q}_i \prod_{l \in \Lambda \setminus i} \hat{P}^\dagger_l \hat{P}_l \, | \Psi_0 \rangle. \tag{17}$$

The operator $\hat{Q}_i$ and the squares of the Gutzwiller operator $\hat{P}^\dagger_l \hat{P}_l$ can be written in terms of creation and annihilation operators:

$$\hat{Q}_i = \sum_{I_1, I_2} Q'_{I_1, I_2} \hat{C}^\dagger_{i,I_1} \hat{C}_{i,I_2}, \tag{18}$$

$$\hat{P}^\dagger_i \hat{P}_i = \bar{\lambda}_{\emptyset,\emptyset} + \hat{A}'_i = \bar{\lambda}_{\emptyset,\emptyset} + \sum_{\substack{I_1, I_2 \\ |I_1|, |I_2| > 0}} X'_{I_1, I_2} \hat{C}^\dagger_{i,I_1} \hat{C}_{i,I_2}, \tag{19}$$

where the scalar contribution $\bar{\lambda}_{\emptyset,\emptyset}$ to $\hat{P}^\dagger \hat{P}$ could always be set to unity after rescaling the Gutzwiller wave function. However, we will postpone this step to a later stage of our analysis. Particle number conservation in our correlation operator (5) only terms with $|I_1| = |I_2|$ contribute in Eq. (18).

As a next step, we apply Wick's theorem

$$\langle \Psi_0 | \hat{Q}_i \prod_{l \in \Lambda \setminus i} \hat{P}^\dagger_l \hat{P}_l \, | \Psi_0 \rangle = \left\{ \hat{Q}_i \prod_{l \in \Lambda \setminus i} \hat{P}^\dagger_l \hat{P}_l \right\}_\rho, \tag{20}$$

where $\{ \ldots \}_\rho$ gives the sum over all possible contractions with respect to the density matrix ρ with the elements

$$\rho_{(i\sigma),(j\sigma')} = \rho^{\sigma\sigma'}_{ij} = \langle \Psi_0 | \hat{c}^\dagger_{i\sigma} \hat{c}_{j\sigma'} | \Psi_0 \rangle. \tag{21}$$

For example,

$$\begin{aligned} \langle \Psi_0 | \hat{c}^\dagger_{l\sigma_1} \hat{c}_{l\sigma_2} \hat{c}^\dagger_{k\sigma_3} \hat{c}_{k\sigma_4} | \Psi_0 \rangle &= \left\{ \hat{c}^\dagger_{l\sigma_1} \hat{c}_{l\sigma_2} \hat{c}^\dagger_{k\sigma_3} \hat{c}_{k\sigma_4} \right\}_\rho \\ &= \rho^{\sigma_1\sigma_2}_{ll} \rho^{\sigma_3\sigma_4}_{kk} - \rho^{\sigma_1\sigma_4}_{lk} \rho^{\sigma_3\sigma_2}_{kl}. \end{aligned} \tag{22}$$

We depict the different contributions in a diagrammatic way, as shown in Fig. 1.

Each summand of the operator $\hat{Q}$ and $\hat{P}^\dagger \hat{P}$ in Eq. (18) defines an "external node" with weight Q'_{I_1,I_2} or an "internal node" with weights X'_{I_1,I_2}, respectively. Each operator contraction can be represented by a line which is either a "self-closing line" (also denoted as "local contractions" or "Hartree bubbles") or connects two different nodes. In the following, we will simplify this diagrammatic analysis in three steps.

First, we aim to eliminate all local contractions. Therefore we map our operators to so called Hartree–Fock (HF) operators

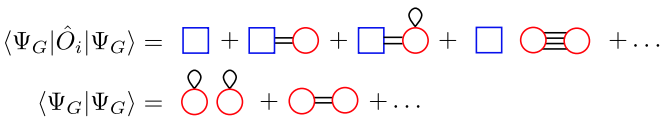

FIG. 1. Diagrammatic representation of the numerator and denominator of the expectation value of an local operator $\hat{O}_i$ on the lattice site i. The blue square and the red circles gives the external and internal nodes, respectively. Black lines correspond to the single-particle density matrix.

[10,42], which, by definition, have no Hartree bubbles in the diagrammatic expansion. For example,

$$\begin{aligned} &\langle \Psi_0 | \left(\hat{c}^\dagger_{l,\sigma_1} \hat{c}_{l,\sigma_2} \right)^{\mathrm{HF}} \left(\hat{c}^\dagger_{k,\sigma_3} \hat{c}_{k,\sigma_4} \right)^{\mathrm{HF}} | \Psi_0 \rangle \\ &\quad = \left\{ \left(\hat{c}^\dagger_{l,\sigma_1} \hat{c}_{l,\sigma_2} \right)^{\mathrm{HF}} \left(\hat{c}^\dagger_{k,\sigma_3} \hat{c}_{k,\sigma_4} \right)^{\mathrm{HF}} \right\}_\rho \\ &\quad = -\rho^{\sigma_1\sigma_4}_{lk} \rho^{\sigma_3\sigma_2}_{kl}. \end{aligned} \tag{23}$$

The mapping between the normal creation and annihilation operators and the HF-operators depends on the local density matrix $\rho^{\sigma\sigma'}_{ij}$ as can be seen from the simplest case,

$$\left(\hat{c}^\dagger_{i,\sigma} \hat{c}_{i,\sigma'} \right)^{\mathrm{HF}} = \hat{c}^\dagger_{i,\sigma} \hat{c}_{i,\sigma'} - \rho^{\sigma\sigma'}_{ii}.$$

An extension of this mapping for operator products $\hat{C}^\dagger_{i,I_1} \hat{C}_{i,I_2}$ is given in Appendix B. We write the operator $\hat{Q}$ and the square of the Gutzwiller correlator as

$$\hat{Q}_i = \sum_{I_1, I_2} Q_{I_1, I_2} \left(\hat{C}^\dagger_{i,I_1} \hat{C}_{i,I_2} \right)^{\mathrm{HF}}, \tag{24}$$

$$\hat{P}^\dagger_i \hat{P}_i = 1 + \hat{A}_i = 1 + \sum_{\substack{I_1, I_2 \\ |I_1|, |I_2| > 0}} X_{I_1, I_2} \left(\hat{C}^\dagger_{i,I_1} \hat{C}_{i,I_2} \right)^{\mathrm{HF}}, \tag{25}$$

where we set the coefficient $X_{\emptyset,\emptyset} = 1$. As mentioned above, this is equal to a rescaling of the Gutzwiller wave function by a factor which is always canceled out by the denominator in Eq. (16). Equation (24) is equivalent to Eq. (18) in the single-band derivation of Ref. [31]. If combined with (41) and (B16), see below, Eq. (25) corresponds to the single-band formula (11) in Ref. [31].

All operators in Eq. (20) are normal ordered because all site indices are different when we apply Wick's theorem. We can set all local entries in $\rho^{\sigma\sigma'}_{ii}$ to zero because we work with the HF operators so that all local contractions vanish automatically. Therefore we can carry out all contractions with a new density matrix

$$\bar{\rho}^{\sigma\sigma'}_{ij} = \rho^{\sigma\sigma'}_{ij} - \delta_{ij} \, \rho^{\sigma\sigma'}_{ii} \tag{26}$$

and drop the HF-operator notation at the same time,

$$\left\{ \left(\hat{C}^\dagger_{i,I_1} \hat{C}_{i,I_2} \right)^{\mathrm{HF}} \ldots \right\}_\rho \equiv \left\{ \hat{C}^\dagger_{i,I_1} \hat{C}_{i,I_2} \ldots \right\}_{\bar{\rho}}. \tag{27}$$

Without any nonzero local contraction, we get

$$\{ \hat{c}^\dagger_{i,\sigma} \hat{c}_{i,\sigma} \ldots \}_{\bar{\rho}} = -\{ \hat{c}_{i,\sigma} \hat{c}^\dagger_{i,\sigma} \ldots \}_{\bar{\rho}}. \tag{28}$$

Thus we can replace the fermionic operators $\hat{c}_{i,\sigma}$, $\hat{C}_{i,I}$ by Graßmann variables $\tilde{c}_{i,\sigma}$, $\tilde{C}_{i,I}$, respectively. These Graßmann variables are nilpotent,

$$\tilde{C}_{i,I_1} \tilde{C}_{i,I_2} = 0 \quad \text{if } I_1 \cap I_2 \neq \varnothing. \tag{29}$$

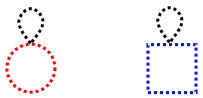

FIG. 2. All nodes with internal lines cancel because the local entries of $\bar{\rho}_{ii}^{\sigma\sigma'}$ are set to zero.

In principle, the introduction of the HF mapping is not a necessary step for the introduction of Graßmann operators as we discuss in Appendix D. All local entries $\bar{\rho}_{ii}^{\sigma\sigma'}$ of the new density matrix vanish so that the diagrammatic expansion cannot have nodes with self-closing lines, as shown in Fig. 2.

The coefficients Q_{I_1,I_2} and X_{I_1,I_2} are not affected by our mapping so that we can write

$$\tilde{Q}_i = \sum_{I_1,I_2} Q_{I_1,I_2} \tilde{C}^\dagger_{i,I_1} \tilde{C}_{i,I2}, \tag{30}$$

$$\tilde{A}_i = \sum_{\substack{I_1,I_2 \\ |I_1|,|I_2|>0}} X_{I_1,I_2} \tilde{C}^\dagger_{i,I_1} \tilde{C}_{i,I2}. \tag{31}$$

The numerator in Eq. (20) becomes

$$\left\{ \hat{Q}_i \prod_{l\in\Lambda\setminus i} (1+\hat{A}_l) \right\}_\rho = \left\{ \tilde{Q}_i \prod_{l\in\Lambda\setminus i} (1+\tilde{A}_l) \right\}_{\bar{\rho}}, \tag{32}$$

whereas the denominator reads

$$\left\{ \prod_{l\in\Lambda} (1+\hat{A}_l) \right\}_\rho = \left\{ \prod_{l\in\Lambda} (1+\tilde{A}_l) \right\}_{\bar{\rho}}. \tag{33}$$

As a second step, we merge the diagrams of the numerator and denominator with the help of the linked cluster theorem. To this end, the lattice site restrictions on the right hand site of Eq. (32) must be removed. Therefore we define

$$1+\tilde{A}_i = \exp(\tilde{G}_i), \tag{34}$$

$$\tilde{Q}_i = \tilde{M}_i \exp(\tilde{G}_i). \tag{35}$$

with

$$\tilde{G}_i = \sum_{\substack{I_1,I_2 \\ |I_1|,|I_2|>0}} Z_{I_1,I_2} \tilde{C}^\dagger_{i,I_1} \tilde{C}_{i,I_2}, \tag{36}$$

$$\tilde{M}_i = \sum_{I_1,I_2} M_{I_1,I_2} \tilde{C}^\dagger_{i,I_1} \tilde{C}_{i,I_2}. \tag{37}$$

The exponential series expansion stays finite due to the nilpotency of the Graßmann variables. Therefore, the new coefficients Z_{I_1,I_2} and M_{I_1,I_2} can be written as finite polynomials of the old coefficients Q_{I_1,I_2} and X_{I_1,I_2}. The explicit expressions are given in Appendix B. It is crucial that we perform the HF mapping before we switch to the exponential form of our correlators. Note that the explicit formulas (36) and (37) justify the definitions in Eqs. (34) and (35).

These additional redefinitions allow us to cast Eq. (32) into the form

$$\langle\Psi_G|\hat{O}_i|\Psi_G\rangle = \left\{ \tilde{M}_i \prod_{l\in\Lambda} \exp(\tilde{G}_l) \right\}_{\bar{\rho}}. \tag{38}$$

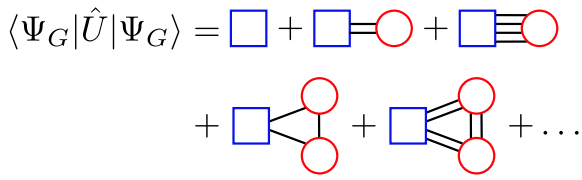

FIG. 3. The first few connected diagrams that contribute to $\langle\hat{U}\rangle_G$. The blue square represents the external node. The red circles represent the internal nodes. Black lines stand for the single-particle density matrix. The second and the forth diagrams cancel out after the introduction of a gauge in the variational parameters λ_{I_1,I_2} as shown in Sec. II D.

Note that the site index restriction $l \neq i$ disappeared. Equation (33) can be rewritten as

$$\langle\Psi_G|\Psi_G\rangle = \left\{ \prod_{l\in\Lambda} \exp(\tilde{G}_l) \right\}_{\bar{\rho}}. \tag{39}$$

Now, we are in the position to apply the linked cluster theorem (LCT), as described, e.g., in Ref. [43]. We find

$$\frac{\langle\Psi_G|\hat{O}_i|\Psi_G\rangle}{\langle\Psi_G|\Psi_G\rangle} = \sum_{L\subset\Lambda} \frac{1}{|L|!} \left\{ \tilde{M}_i \prod_{l\in L} \tilde{G}_l \right\}^{\text{conn.}}_{\bar{\rho}}, \tag{40}$$

where the summation is performed over all subsets L of the lattice Λ. The first few diagrams that are needed for the evaluation of the potential energy are shown in Fig. 3.

Some of the polynomials of $\tilde{G}_l$ in Eq. (38) vanish due to the nil potency of the Graßmann variables. In contrast to the usual application of the LCT in many-body lattice theories, we can apply our expansion for a finite lattice as well. The nil-potency property allows us to add virtually as many nodes as we need to regroup our diagrams in all orders. Note that after the application of the LCT the nodes $\tilde{G}_l$ are contracted in such a way that all nodes have to be connected to the external nodes M_i. This invalidates the nil-potency of the Graßmann variables inside the curly brackets. Therefore several nodes can be located on the same site as long as these nodes are only connected indirectly. Note that the diagrammatic rules are the same as those for the single-band model introduced, e.g., in Refs. [4,10].

D. Parameter gauge and the limit of infinite dimensions

An essential step in our expansion is now the elimination of all diagrams with internal nodes that have a single outgoing and incoming line, as illustrated in Fig. 4. This step is essential for a rapid convergence of our expansion as has been demonstrated for the single-band model in one dimension [31]. For this reason, a gauge in the variational parameters is introduced which sets the weight of these nodes to zero,

$$Z_{I_1,I_2} = 0 \quad \forall |I_1| = |I_2| = 1. \tag{41}$$

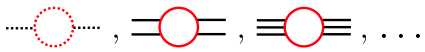

FIG. 4. The internal nodes with only two lines will be eliminated by a gauge in the variational parameters.

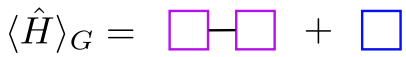

FIG. 5. Diagrammatic analysis in infinite dimensions. The external nodes that arise from the hopping operator $\hat{H}_0$ are colored in purple while the external nodes that arise from the on-site interaction operator $\hat{U}$ are colored in blue.

This constraint must be incorporated in the optimization of the variational parameters λ_{I_1,I_2}. However, we can show that this constraint will not reduce the variational freedom in our model as can be understood from an investigation of the single-band case in [31]. For a complete proof of this presumption that is based on the symmetries of our complex two-band model we refer the reader to Ref. [39].

Originally, the application of the gauge condition (41) was motivated by the simplifications of the diagrammatic expansion that arise when it is applied to a system with infinite lattice dimensions [10]. A scaling analysis of the "kinetic energy operator" $\hat{H}_0$ shows [8,10] that the lines of the density matrix scale with the lattice dimension d as

$$\rho_{ij} \sim (\sqrt{2d})^{-||i-j||_1}, \tag{42}$$

where we dropped the spin-band index for notational clarity, and $||.||_1$ gives the "one-norm" (or "Manhattan metric") of the displacement vector $i-j$. Then, with the scaling

$$t_{ij} \sim (\sqrt{2d})^{-||i-j||_1}, \tag{43}$$

of the hopping parameters, all contributions with an internal node or two external nodes that are connected by three or more independent paths scale at least as $\sim 1/d$. Due to the gauge condition (41), however, all diagrams with internal nodes vanish. Hence, in the limit $d \to \infty$, the only remaining terms are given by

$$\langle \hat{H}_0 \rangle_G = \sum_{i \neq j} \sum_{\sigma\sigma',\tau\tau'} M_{\sigma,\emptyset}(\hat{c}^\dagger_\tau) M_{\emptyset,\sigma'}(\hat{c}_{\tau'})\, t_{ij}^{\tau\tau'} \rho_{ij}^{\sigma\sigma'} \tag{44}$$

$$\langle \hat{U} \rangle_G = M_{\emptyset,\emptyset}(\hat{U}), \tag{45}$$

as illustrated in Fig. 5. Note that the symbolic operator dependence in this equations is just a reminder that the coefficients M_{I_1,I_2} in Eqs. (36) and (37) depend on the corresponding operator in Eq. (16).

In the rest of our work, we will refer to the terms in Eqs. (44) and (45), already derived in Ref. [13], as the "infinite-d limit." Furthermore, we define the order m of a diagram by the number of its internal nodes, where $m \geqslant 0$. As we have explained in this section, the constraints (41) ensure that the leading order terms of our diagrammatic expansion correspond to the exact Gutzwiller energy expectation value in infinite dimensions.

E. Optimization of $|\Psi_0\rangle$

In this work, we use the optimization algorithm which was introduced in Ref. [31]. The energy (15) depends on the variational parameters λ_{I_1,I_2} and the state $|\Psi_0\rangle$, where the latter enters the functional only through the single-particle density matrix (21),

$$E_G = E_G\big(\lambda_{I_1,I_2}, \rho_{ij}^{\sigma\sigma'}\big). \tag{46}$$

As shown, e.g., in Appendix A of Ref. [25], the minimization of E_G with respect to ρ leads to the following effective single-particle Hamiltonian:

$$\hat{H}_0^{\text{eff}} = \sum_{i \neq j} \sum_{\sigma\sigma'} t_{ij}^{\text{eff};\sigma\sigma'} \hat{c}^\dagger_{i\sigma} \hat{c}_{j\sigma'}, \tag{47}$$

$$t_{ij}^{\text{eff};\sigma\sigma'} = \partial_{\rho_{ij}^{\sigma\sigma'}} E_G, \tag{48}$$

which has $|\Psi_0\rangle$ as a ground state. Hence, for the minimization of (46), we need to solve

$$\hat{H}_{\text{eff}} |\Psi_0\rangle = E_S |\Psi_0\rangle \tag{49}$$

and minimize E_G with respect to λ_{I_1,I_2},

$$\partial_{\lambda_{I_1,I_2}} E_G = 0. \tag{50}$$

In order to solve these equations self-consistently, we usually start with the ground state $|\Psi_0\rangle^{(0)}$ of the free system, i.e., we set $t_{ij}^{\text{eff};\sigma\sigma'} = t_{ij}^{\sigma\sigma'}$ in (48) and (49). Then we compute the density matrix, solve (50), and determine new parameters (48). The optimization terminates if the change of the effective hopping parameters between two cycles drops below some threshold. In order to test the stability of the algorithm, we can start from a different initial state. This initial state may be constructed from a perturbed kinetic energy operator $\hat{H}_0 + \delta\hat{h}_0$. Usually the optimization algorithm remains stable against these perturbations but in some cases the fix point of this map does not need to be unique, as shown in Ref. [31] where a symmetry breaking of the Fermi surface (Pomeranchuk phase) has been investigated.

III. RESULTS

A. Magnetism

The occurrence of a ferromagnetic phase is favored by two conditions. The local Hamiltonian favors the formation of local magnetic moments for positive values of the Hund's-rule coupling J. Then, the two-particle eigenstates of the on-site energy $\hat{U}$ with maximal local spin $S = 1$ are lowest in energy, in accordance with Hund's first rule. Therefore, for large values of J, the ground state of the lattice system may show global ferromagnetism if the preformed local moments align. In contrast to that, the Stoner picture gives a different explanation for the origin of ferromagnetism. In this picture, a splitting between majority and minority bands reduces their mutual Coulomb repulsion due to the Pauli principle. This effect becomes significant when the density of states $D(E_F)$ at the Fermi energy is large. In the Gutzwiller variational approach, the number of energetically costly multiple occupancies is reduced by an adjustment of the variational parameters. Therefore we can expect that the Gutzwiller wave function predicts ferromagnetism at much larger interaction strengths than the uncorrelated SPPS. The Gutzwiller variational description leaves room both for the Stoner band splitting and the local moment formation as a source for itinerant ferromagnetism.

Throughout this section, we focus on the kinetic energy operator with some particular amplitudes

$$t_x^{11} = -1.0, \quad t_y^{11} = -0.6, \quad t_{xy}^{11} = 0.2, \quad t_{xy}^{12} = -0.4, \tag{51}$$

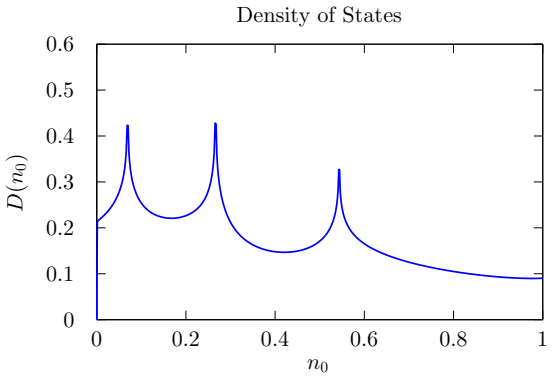

FIG. 6. The density of states $D(n_0)$ as a function of the density n_0. The hopping amplitudes are $t_x^{11} = -1.0$, $t_y^{11} = -0.6$, $t_{xy}^{11} = 0.2$, $t_{xy}^{12} = -0.4$.

where the coefficients t_x^{11} and t_y^{11} give the hopping amplitudes for a transition process between two p_x orbitals to its horizontal and vertical neighbors, respectively. The coefficients t_{xy}^{11} and t_{xy}^{12} give the hopping amplitude for the transition process to the next-nearest neighbors with and without an inter-orbital transition, see Fig. 13 in Appendix A. A detailed analysis of the lattice symmetries and all remaining hopping coefficients can be found in that Appendix. Note that our model parameters (51) do not aim to reflect the situation in any real material where, e.g., the signs of t_x^{11} and t_y^{11} are normally different.

The density of states $D(n_0) = D(E)|_{E=E_F(n_0)}$ in Fig. 6 shows three peaks at $n_0 \approx 0.069$, 0.266, and 0.543. Below we investigate the ferromagnetic transition at $n_0 = 0.2$, 0.265, 0.275, 0.3. The densities are located near the second peak in the density of states. We define the quantity M as

$$\begin{aligned} \langle \hat{n}_{i\uparrow} \rangle &= n_\uparrow = (1+M)n_0, \\ \langle \hat{n}_{i\downarrow} \rangle &= n_\downarrow = (1-M)n_0 \end{aligned} \tag{52}$$

with $i \in \{1,2\}$, so that $0 \leqslant M \leqslant 1$, and the total density remains constant. The total magnetization will be given by $M_{\text{tot}} = 2(n_\uparrow - n_\downarrow)$ when both bands are considered. In order to obtain the optimal magnetization, we will perform a scan in the M-U plane, while we keep the ratio of J and U fixed to $J/U = 8/30$. Then, we optimize the band symmetric Gutzwiller variational parameters λ_{I_1,I_2} for each magnetization. Our diagrammatic expansion includes all lines $\rho_{ij}^{\sigma\sigma'}$ with $||i-j||_1 \leqslant 4$.

As seen in Fig. 7, in the Hartree-Fock approximation the magnetization jumps to a finite value at $U_{\text{HF}} \approx 3.3$. The magnetization then increases monotonically until the ground state is fully polarized at $U_{\text{HF}}^{\text{sat.}} \approx 3.6$. A detailed analysis of the ground state energy shows that the nature of the jumps can be understood as a first-order phase transition. The Gutzwiller approach reveals a different picture. In second order, the magnetization shows a finite magnetization for $U_{\text{G}} \approx 6.4$ and becomes fully magnetized for $U_{\text{G}}^{\text{sat.}} \approx 6.68$. This shows that the magnetization is shifted to much larger interaction strengths in the Gutzwiller wave function. The infinite-d approximation becomes magnetized at $U_{\text{G}}^{\infty} \approx 6.4$ and becomes fully magnetized for $U_{\text{G}}^{\infty,\text{sat.}} \approx 7.35$. Therefore

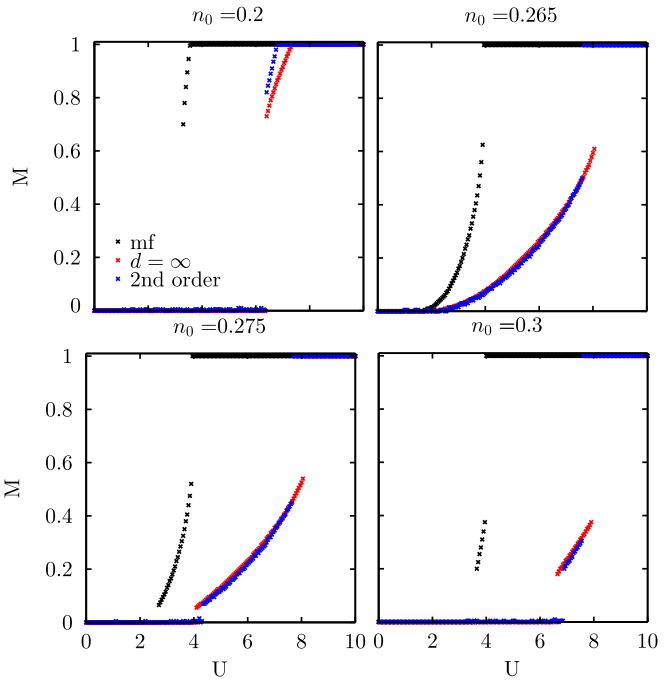

FIG. 7. Magnetization of the Gutzwiller wave function. The black, red, and blue crosses give the HF, the infinite-d, and the second-order approximation, respectively.

the second-order diagrams in our diagrammatic expansion do not change the results on ferromagnetism significantly.

Next, we analyze the density $n_0 = 0.265$ that lies very close to the second peak in the density of states, see Fig. 6. In the Stoner picture, the large density of states causes a finite magnetization at much smaller interaction strength. In our second-order Gutzwiller approach, the ground state becomes already magnetized at $U_{\text{G}} \gtrsim 2$ although a precise evaluation of the threshold is hindered by numerical difficulties. The magnetization in the second-order approximation jumps to the fully magnetized state at $U_{\text{G}}^{\text{sat.}} = 7.6$. The infinite-$d$ approximation lies almost on top of the second-order expansion except at the transition to the fully magnetized state which occurs at $U_{\text{G}}^{\infty,\text{sat.}} \approx 8.1$. The HF result shows the same qualitative behavior but the onset of ferromagnetism is at $U_{\text{HF}} < U_{\text{G}}$. Moreover, the magnetization increases more rapidly as function of the interaction strengths and saturates already at $U_{\text{HF}}^{\text{sat.}} \approx 3.95$.

For $n_0 = 0.275$, the second-order magnetization result jumps to a finite value at $U_{\text{G}} \approx 4.3$ and becomes fully spin polarized at $U_{\text{G}}^{\text{sat.}} \approx 7.6$. The transition points of the infinite-d (HF) approximation lie at $U_{\text{G}}^{\infty} \approx 4.1$ ($U_{\text{HF}} \approx 2.7$) and $U_{\text{G}}^{\infty,\text{sat.}} \approx 8.05$ ($U_{\text{HF}}^{\text{sat.}} \approx 3.95$), respectively. The magnetization curve shows the same qualitative behavior in all three approximations: The magnetization jumps to a small but finite value. Then the magnetization increases gradually as a function of U whereby the slope is much steeper in HF than in Gutzwiller theory. Lastly, the magnetization jumps to full saturation at $U^{\text{sat.}}$. In general, the critical values are much larger in Gutzwiller theory than in the Hartree-Fock approach. Note that the second-order terms to the result in $d = \infty$ lead to fairly small quantitative corrections. The magnetization

onset requires a larger interaction strength U for $n_0 = 0.275$ because the density of states is lower for $n_0 = 0.275$ than for $n_0 = 0.265$. Furthermore, we can see that the transitions to the fully magnetized state occur at almost the same interaction strength as for $n_0 = 0.265$ in all approximations. This shows that the transition to the fully polarized state depends on the density but not on the density of states.

For $n_0 = 0.3$, we still recover qualitatively the same behavior as for $n_0 = 0.275$ but the region between the onset of ferromagnetism and the transition to the fully polarized phase becomes smaller. Again, the critical values in Gutzwiller theory are about a factor of two larger than in Hartree-Fock theory.

In summary, we can state that a large density of states at the Fermi energy promotes ferromagnetism. The Gutzwiller approach shows, however, that ferromagnetism, in general, requires large Coulomb interactions. Moreover, the Gutzwiller approach leaves room in parameter space for nonsaturated ferromagnetism. For the system parameters, considered in this work, phases with long-range magnetic order are already well described within the GA. This supports the use of this approximation in many earlier works, see, e.g., Refs. [25,37,38]. Note that the magnetizations calculated in this section are based on the optimization of our powerful but restricted class of variational wave functions. As usual in a variational approach, it is difficult to assess the quality of such a wave function and its predictions for physical quantities as long as the true ground state is unknown. The same uncertainty remains in our calculations of the Fermi surface that are performed in the next section.

B. Fermi surface deformations

In this section, we show that the optimization of the SPPS can lead to a deformed Fermi surface. In some cases, the Fermi surface even changes its topology. Note that Fermi surface deformations within the Gutzwiller variational approach have already been studied in Ref. [31] for the single-band Hubbard model.

For our degenerate Hubbard model in the infinite-d limit, neither for the Gutzwiller wave function nor within a more sophisticated DMFT calculation, we would find any correlation-induced changes of the Fermi surface. Hence all results in this section can be understood as an effect of the finite-dimensional evaluation provided by our higher-order expansion.

We examine the Fermi-surface deformations for the following parameter set:

$$t_x^{11} = -1.0, \quad t_y^{11} = -0.5, \quad t_{xy}^{11} = 0.4, \\ t_{xy}^{12} = -0.2, \quad U = 6.0, \quad J = 0.8. \tag{53}$$

The amplitudes are chosen in such a way that the topology of the Fermi surface changes near half-filling where the effect of the Gutzwiller correlator is strongest. The density of states is shown in Fig. 8. The peak near $n_0 = 0.47$ is caused by the change in the topology of the Fermi surface.

Our diagrammatic expansion includes all lines $\rho_{ij}^{\sigma\sigma'}$ with $||i - j||_1 \leqslant 5$. In some cases, the optimization algorithm

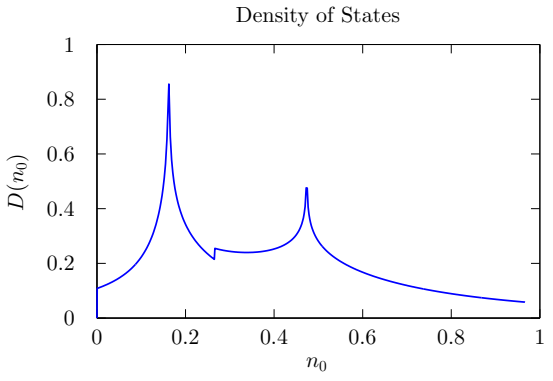

FIG. 8. The density of states $D(n_0)$ as a function of the density n_0. The hopping amplitudes are $t_x^{11} = -1.0$, $t_y^{11} = -0.5$, $t_{xy}^{11} = 0.4$, and $t_{xy}^{12} = -0.2$.

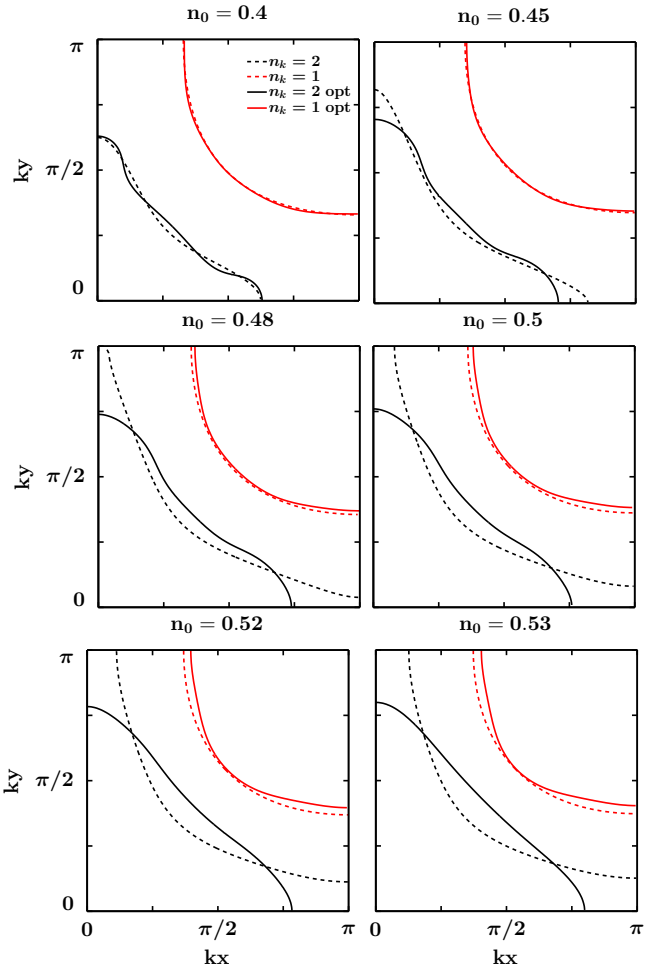

FIG. 9. Fermi surface deformations for densities $n_0 = 0.4$, 0.45, 0.48, 0.50, 0.52, and 0.53. The local interaction strengths are set to $U = 6.0$ and $J = 0.8$. The dashed lines give the initial Fermi edge and the solid lines give the optimized Fermi surface. Both bands are occupied in the region between the origin and the solid (dashed) black line. In the region between the black and the red lines, only the lower band is occupied. For densities $n_0 > 0.48$, the optimized inner Fermi edge has a closed topology while the initial Fermi surface topology is open.

alternates between two fix points which are energetically very close. However, the Fermi surface of these fix points may differ significantly. In these cases, it is useful to introduce some damping for the effective hopping parameters (48) in the self-consistency cycle. In our calculations, we usually find a fix point after $n < 15$ iteration steps.

The single-particle states $|\Psi_0\rangle^n$ are calculated by a diagonalization of the effective Hamiltonian on the left hand side of Eq. (49). In momentum space, each k point for a spin σ can be empty, singly or doubly occupied. We denote the spin-independent occupation number of the momentum k as $n_k = n_\sigma$. In Fig. 9, the Fermi edges of the initial and optimized SPPS are shown for the densities $n_0 = 0.4$, 0.45, 0.48, 0.50, 0.52, and 0.53. The Hubbard/Hund parameters are set to $U = 6.0$ and $J = 0.8$. Although the SPPS $|\Psi_0\rangle$ can have a small but finite magnetization in this parameter regime, we restrict ourselves to a paramagnetic wave function. The deformation of the inner Fermi surface between $n_k = 2$ and 1 start for densities $n_0 > 0.4$. The outer Fermi edge between $n_k = 1$ and 0 is more robust. For densities $n_0 > 0.48$, the optimized inner Fermi edge still has a closed topology while the initial Fermi surface topology is open. The optimization becomes difficult for densities larger than $n_0 = 0.53$. Alternatively, a particle-hole symmetry can be used to determine the optimal Gutzwiller wave function [39]. In this way, we can show that the deviations in the Fermi surface are small for $n_0 \gtrsim 0.6$ where the topology of the optimized Fermi surface becomes open.

The dependence of the Fermi surface deformations on the Hund's-rule coupling J is shown in Fig. 10. The Fermi edge for $n_k = 2$ remains open for vanishing J. An increase of the Hund's-rule interaction strength to $J = 0.4$ leads to the appearance of small islands in which both bands are filled. These islands collapse when we further increase the interaction strength to $J = 0.8$ so that the Fermi surface becomes closed. The left panel of Fig. 12 shows that the energy gain ΔE increases linearly in J and becomes vanishingly small for

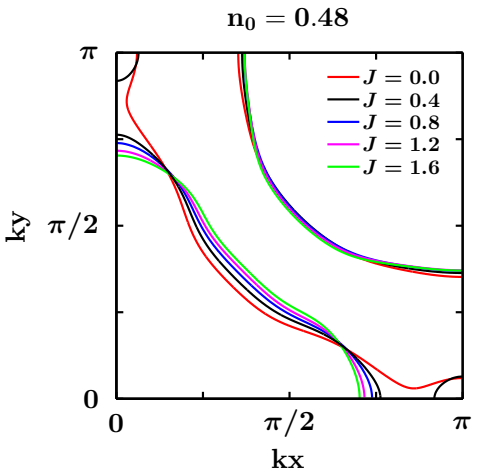

FIG. 10. Fermi surface deformations for different interaction strength J. The density and the interaction strength are set to $n_0 = 0.48$ and $U = 6.0$, respectively. The deformations increase for larger values of J. For $J = 0$, the Fermi surface topology (for $n_k = 2$) is still open.

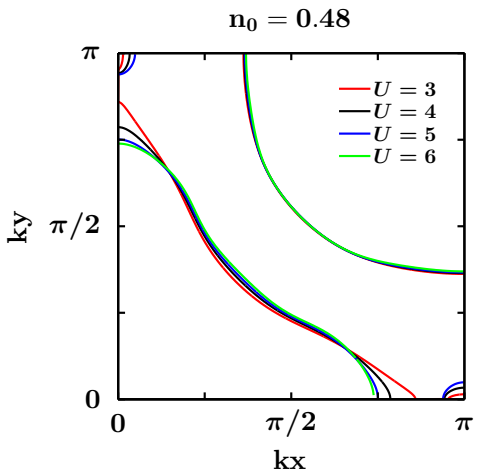

FIG. 11. Fermi surface deformation for different interaction strength U. The density and the Hund's-rule coupling are set to $n_0 = 0.48$ and $J = 0.8$, respectively. Small islands appear for $U = 3$ in which $n_k = 2$.

$J = 0$. From an energetic point of view, the transition from an open to a closed inner Fermi surface is gradual as a function of J. The existence of intermediate islands also shows that the hopping matrix elements change gradually.

The change in the Fermi-surface topology as a function of U for $n_0 = 0.48$ and $J = 0.8$ is shown in Fig. 11. For an interaction strength $U = 3$, the Fermi surface starts to deform from an open to a closed topology and small islands appear. The islands at the border of the Brillouin zone vanish for $U = 6$ again. The energy gain ΔE increases linearly in U as shown in the right panel of Fig. 12.

In this section, we showed that the interaction-induced Fermi surface deformations are clearly visible and may even change the Fermi-surface topology. This is a warning against a naive application of Fermi liquid theory if this works with the Fermi surface of the noninteracting system. The contributions beyond our second-order approximation still affect the Fermi surface and the density matrix. However, a higher-order expansion of the single-band Hubbard model [31,44] showed that the Fermi surface deformations are true features of the Gutzwiller wave function. Therefore we can assume that the qualitative findings are valid in all orders of the approximation.

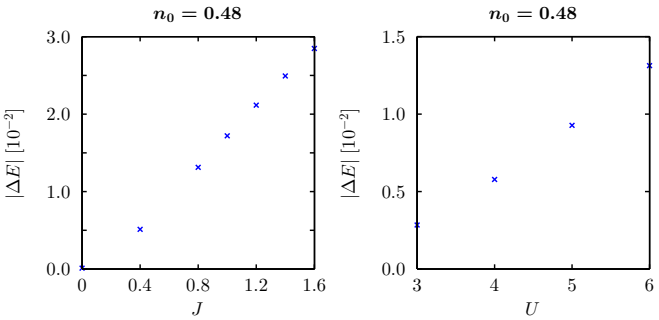

FIG. 12. (Left) The energy gain ΔE increases linearly in J. (Right) Energy gain ΔE as a function of U.

IV. CONCLUSION AND OUTLOOK

In this work, we have given a comprehensive derivation of a diagrammatic method that allows us to evaluate multiband Gutzwiller wave functions in finite dimensions. Our approach constitutes a systematic improvement of the widely used Gutzwiller approximation, which corresponds to the zeroth order of our expansion.

As our first application, we studied the ferromagnetic phase transition in a two-band Hubbard model on a square lattice as a function of the model parameters for various band fillings. In general, a large density of states and a strong Hund's-rule exchange favor ferromagnetism. In the Gutzwiller wave function, the ferromagnetic order is strongly suppressed so that much larger interaction strength are needed than predicted by the Hartree-Fock solution. Moreover, the regions in parameter space where nonsaturated ferromagnetism occurs are much broader in Gutzwiller theory. As shown in earlier studies, this gives room for the experimental observations of nonsaturated ferromagnetism, e.g., in transition metals such as nickel and iron. It turned out that long-range ferromagnetic order in our model is already well described within the Gutzwiller approximation.

As a second application, we investigated the interaction-induced deformation of the Fermi surface. These effects occur for large interaction strength, when the potential energy of the system is twice as large as the kinetic energy. For weaker interactions and small densities, the deformations of the Fermi surface can be neglected. Close to half band-filling and for special choices of the electron transfer parameters, the interactions can induce a change in the Fermi-surface topology from open to closed constant-energy contours. These effects are a result of the finite-order diagrams and cannot be seen in the Gutzwiller approximation.

It will be an interesting question for future work whether the Fermi surface deformation can lead to symmetry broken phases (Pomeranchuk phase) as in the single-band Hubbard model. In our two-band model, such a broken rotational spatial symmetry leads to different orbital densities, which are energetically unfavorable. Hence, it is an open question if the ground state of our two-band model can have an asymmetric Fermi surface. Another open questions concerns the appearance of superconductivity in our model, as seen in earlier work on two-orbital Hubbard models [45–47].

ACKNOWLEDGMENTS

We like to thank Florian Gebhard for many valuable discussions in all stages of this project. Furthermore we thank Jan Kaczmarczyk for a valuable discussion of his higher-oder study of a single-band Hubbard model.

APPENDIX A: LATTICE SYMMETRIES

Figure 13 shows the hopping processes to nearest and next-nearest neighbors. The amplitude t_x^{11} for the transition from the p_x ($\sigma = 1$) orbitals on site i to the site $i \pm \mathrm{d}x$ equals the amplitude t_y^{22} for the transition from the p_y orbitals ($\sigma = 2$) on site i to the site $i \pm \mathrm{d}y$. The same holds for the amplitudes t_{xy}^{11} and t_{xy}^{22} for the transitions between the p_x orbitals on i and $i \pm \mathrm{d}y$ and the p_y orbitals on sites i and $i \pm \mathrm{d}x$, respectively.

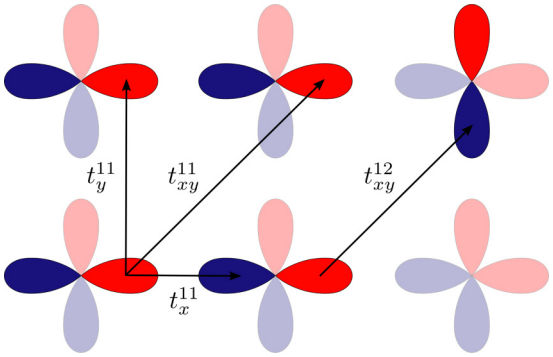

FIG. 13. Hopping amplitudes of our two-orbital system with $p_x - p_y$ symmetry. The arrows mark some hopping parameters between p-orbitals. All others can be obtained via symmetry considerations.

The amplitudes for the hopping processes from i to $i \pm \mathrm{d}x \pm \mathrm{d}y$ between the p_x orbitals is the same as between the p_y orbitals. The symmetry of the orbitals does not allow any $p_x - p_y$ transition to nearest neighbors. For transitions between next-nearest neighbors, the sign of the amplitudes t_{xy}^{12} will change after a rotation of $\pi/2$ so that $t_{xy}^{12} = -t_{yx}^{12}$. The xy symmetry of the inter-orbital hopping processes leads to a diagonal local density matrix $\rho_{\sigma,\sigma'} = n_\sigma^0 \delta_{\sigma,\sigma'}$. Furthermore, the rotational symmetry of the hopping processes within the same orbitals guarantees that all diagonal entries of the local density matrix are the same, $\rho_{\sigma,\sigma'} = n_0\,\mathrm{Id}$.

The calculation of the on-site Coulomb interaction (3) requires the evaluation of two-particle expectation values of the Coulomb energy

$$U_{\sigma_1\sigma_2\sigma_3\sigma_4} = \left\langle \phi_{\sigma_1}\phi_{\sigma_2} \middle| \hat{V}_{\mathrm{Coul.}} \middle| \phi_{\sigma_3}\phi_{\sigma_4} \right\rangle, \tag{A1}$$

with $\sigma_i \in \{p_x, p_y\}$. These coefficients can be simplified after a decomposition of the $p_x - p_y$ orbitals in terms of Laguerre and Legendre polynomials, respectively. The explicit derivation of the coefficients can be found in many text books (e.g., in Ref. [48]) so that we simply state the result. The matrix representation of the two-particle sector of $\hat{U}_{\mathrm{int}}$ is given by

$$U_{\mathrm{int}} = U\,\mathrm{Id} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & J \\ 0 & -3J & 0 & 0 & 0 & 0 \\ 0 & 0 & -2J & -J & 0 & 0 \\ 0 & 0 & -J & -2J & 0 & 0 \\ 0 & 0 & 0 & 0 & -3J & 0 \\ J & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{A2}$$

where the standard ordering $|\uparrow\downarrow,0\rangle, |\uparrow,\uparrow\rangle, |\downarrow,\uparrow\rangle, |\uparrow,\downarrow\rangle, |\downarrow,\downarrow\rangle, |0,\uparrow\downarrow\rangle$ of the two-particle states has been used. Obviously, there are no terms where only one electron switches from a p_x to a p_y orbital or that violates spin conversation.

The variational coefficients λ_{I_1,I_2} in the Gutzwiller correlator (5) have the same structure (nonvanishing elements) as the matrix elements of the on-site Coulomb interaction in Eq. (A2). In the paramagnetic case, the variational coefficients obey a spin-band symmetry. In the ferromagnetic case, all parameters are symmetric under an interchange of the band index.

APPENDIX B: HF-OPERATORS AND DIAGRAMMATIC WEIGHTS

In this section, we give explicit results for the mapping of an arbitrary operator product to its corresponding HF operator. We use this mapping to define the coefficients X_{I_1,I_2}, Q_{I_1,I_2}, Z_{I_1,I_2}, and M_{I_1,I_2} that give the weights of the internal and external nodes in our diagrammatic expansion. A detailed derivation of this mappings, its inversion, and a derivation of all coefficients can be found in Ref. [39].

Consider a product of fermionic creation/annihilation operators $\hat{a}_1\hat{a}_2\dots\hat{a}_n$. For the corresponding HF-operator $(\hat{a}_1\hat{a}_2\dots\hat{a}_n)^{\mathrm{HF}}$, the evaluation of

$$\{\hat{O}(\hat{a}_1\hat{a}_2\dots\hat{a}_n)^{\mathrm{HF}}\} \tag{B1}$$

shall, by definition, not include contractions between any pairs of operators $\hat{a}_1,\dots,\hat{a}_n$. We use the notation

$$\{\hat{a}_1\hat{a}_2\dots\hat{a}_n\}_m, \tag{B2}$$

where m denotes the number of internal contractions, e.g., for $m=1$,

$$\{\hat{a}_1\hat{a}_2\dots\hat{a}_n\}_1 = \sum_{j<k}(-1)^{j+k+1}\{\hat{a}_j\hat{a}_k\}\hat{a}_1\dots\hat{a}_{j-1}\hat{a}_{j+1} \dots\dots\hat{a}_{k-1}\hat{a}_{k+1}\dots\hat{a}_n. \tag{B3}$$

Each internal contraction reduces the number of operators by two. With the abbreviation (B2), we can give the following closed expression for the HF operator of an arbitrary operator:

$$(\hat{a}_1\hat{a}_2\dots\hat{a}_n)^{\mathrm{HF}} = \sum_{k=0}^{[n]}(-1)^k\{a_1a_2\dots a_n\}_k, \tag{B4}$$

where $[n]$ denotes the next smallest even number. This result agrees with the expressions for the definition of the HF operators in Ref. [42].

In our diagrammatic expansion, we must express all operators $\hat{C}^\dagger_{K_1}\hat{C}_{K_2}$ in terms of HF-operators:

$$\hat{C}^\dagger_{K_1}\hat{C}_{K_2} = \sum_{I_1,I_2} X^{K_1,K_2}_{I_1,I_2}\big(\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\big)^{\mathrm{HF}}. \tag{B5}$$

Let us contract both sides with an arbitrary operator $\hat{O}$. In order to determine the value of the coefficient $X^{K_1K_2}_{I_1,I_2}$, we have to evaluate the term where the operators $\hat{C}^\dagger_{I_1}$ and $\hat{C}_{I_2}$ with $I_1\subset K_1$, $I_2\subset K_2$ form the external contractions with $\hat{O}$. We need to shift the operators that are reserved for the external contractions to the front of the operator $\hat{O}$, and contract all remaining operators internally. The operator $\hat{O}$ can be chosen without any restrictions and just indicates which operators are reserved for an external contraction. The contraction with $\hat{O}$ can be carried out symbolically by a replacement of $\hat{C}^\dagger_{I_1}\hat{C}_{I_2}$ with the HF operators $(\hat{C}^\dagger_{I_1}\hat{C}_{I_2})^{\mathrm{HF}}$. Thus we get

$$X^{K_1,K_2}_{I_1,I_2} = [\overleftarrow{I_1,K_1}]^+_\Sigma[\overrightarrow{K_2,I_2}]^-_\Sigma\big\{\hat{C}^\dagger_{K_1\setminus I_1}\hat{C}_{K_2\setminus I_2}\big\}, \tag{B6}$$

where we introduced the following symbols to indicate that we consider a sign change after a reordering of creation or annihilation operators:

$$\begin{aligned}
[\overrightarrow{I,J}]^+_\Sigma &\quad \text{sign after splitting } \hat{C}^\dagger_{I\cup J}\to\hat{C}^\dagger_I\hat{C}^\dagger_J,\\
[\overrightarrow{I,J}]^-_\Sigma &\quad \text{sign after splitting } \hat{C}_{I\cup J}\to\hat{C}_I\hat{C}_J,\\
[\overleftarrow{J,I}]^+_\Sigma &\quad \text{sign after splitting } \hat{C}^\dagger_{I\cup J}\to\hat{C}^\dagger_J\hat{C}^\dagger_I,\\
[\overleftarrow{J,I}]^-_\Sigma &\quad \text{sign after splitting } \hat{C}_{I\cup J}\to\hat{C}_J\hat{C}_I,\\
[\overrightarrow{I}\,\overleftarrow{J}\,]^+_\Sigma &\quad \text{sign after merging } \hat{C}^\dagger_I\hat{C}^\dagger_J\to\hat{C}^\dagger_{I\cup J},\\
[\overrightarrow{I}\,\overleftarrow{J}\,]^-_\Sigma &\quad \text{sign after merging } \hat{C}_I\hat{C}_J\to\hat{C}_{I\cup J}.
\end{aligned} \tag{B7}$$

All operators are assumed to be normally ordered before and after the process. The reversed ordering of the annihilation and creation operators ensures that

$$[\overrightarrow{I,J}]^+_\Sigma = [\overleftarrow{J,I}]^-_\Sigma = [\overrightarrow{I\setminus J}\,\overleftarrow{J}\,]^+_\Sigma = [\overrightarrow{J}\,\overleftarrow{I\setminus J}]^-_\Sigma. \tag{B8}$$

As a next step we transform the square of the local Gutzwiller operator into a sum of HF operators:

$$P^\dagger P = \sum_{I_1,I_2} X_{I_1,I_2}\big(\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\big)^{\mathrm{HF}}. \tag{B9}$$

An application of the mapping (B6) gives

$$\begin{aligned}
X_{I_1,I_2} = &\sum_{\substack{J_1\subset\bar{I}_1,J_2\subset\bar{I}_2\\|J_1|=|J_2|}}\big\{\hat{C}^\dagger_{J_1}\hat{C}_{J_2}\big\}[\overrightarrow{I_1}\,\overleftarrow{J_1}\,]^+_\Sigma[\overrightarrow{J_2}\,\overleftarrow{I_2}\,]^-_\Sigma\\
&\times\sum_{\substack{J_3\subset\\(I_1\cup J_1)\cap(I_2\cup J_2)}}(-1)^{|J_3|}\bar{\lambda}^{(I_2\cup J_2)\setminus J_3}_{(I_1\cup J_1)\setminus J_3}\\
&\times[\overrightarrow{\overrightarrow{I_1\cup J_1,J_3}}]^+_\Sigma[\overleftarrow{\overleftarrow{J_3,I_2\cup J_2}}]^-_\Sigma.
\end{aligned} \tag{B10}$$

The external nodes defined in Eq. (17) can be written as

$$\hat{Q}(\hat{O}) = \sum_{I_1,I_2} K_{I_1,I_2}(\hat{O})|I_1\rangle\langle I_2| \tag{B11}$$

with

$$K_{I_1,I_2}(\hat{O}) = \sum_{I_3,I_4}\lambda_{I_3,I_1}\lambda_{I_4,I_2}\langle I_3|\hat{O}|I_4\rangle. \tag{B12}$$

Then we can use the HF mapping to find

$$\hat{Q}(\hat{O}) = \sum_{I_1,I_2} Q_{I_1,I_2}(\hat{O})\big(\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\big)^{\mathrm{HF}}, \tag{B13}$$

which still depends on the operator $\hat{O}$. Here,

$$\begin{aligned}
Q_{I_1,I_2}(\hat{O}) = &\sum_{\substack{J_1\subset\bar{I}_1,J_2\subset I_2\\|J_1|=|J_2|}}\{\hat{C}^\dagger_{J_1}\hat{C}_{J_2}\}[\overrightarrow{I_1}\,\overleftarrow{J_1}\,]^+_\Sigma[\overrightarrow{J_2}\,\overleftarrow{I_2}\,]^-_\Sigma\\
&\times\sum_{\substack{J_3\subset\\(I_1\cup J_1)\cap(I_2\cup J_2)}}(-1)^{|J_3|}K^{(I_2\cup J_2)\setminus J_3}_{(I_1\cup J_1)\setminus J_3}(\hat{O})\\
&\times[\overrightarrow{\overrightarrow{I_1\cup J_1,J_3}}]^+_\Sigma[\overleftarrow{\overleftarrow{J_3,I_2\cup J_2}}]^-_\Sigma,
\end{aligned} \tag{B14}$$

where the coefficients K_{I_1,I_2} play the role of the coefficients $\bar{\lambda}_{I_1,I_2}$ in Eq. (B10).

The internal nodes defined in Eq. (34) are given by

$$\tilde{G} = \sum_{\substack{I_1,I_2 \\ |I_1|,|I_2|>0}} Z_{I_1,I_2} \tilde{C}^\dagger_{I_1} \tilde{C}_{I_2}, \tag{B15}$$

with

$$Z_{I_1,I_2} = \sum_{m>0} \frac{(-1)^{m+1}}{m} \sum_{\substack{\{(J_1^s,J_2^s)\} \\ s=1,\ldots,m}} \Sigma\big[J_1^s,J_2^s\big] \prod_{s=1}^{m} X_{J_1^s,J_2^s}, \tag{B16}$$

where the sum in Eq. (B16) runs over all (disjunct) partitions $\{(J_1^s,J_2^s)\}$ of the set (I_1,I_2) such that

$$\bigcup_s J_1^s = I_1 \text{ and } \bigcup_s J_2^s = I_2, \tag{B17}$$

and $\Sigma[J_1^s,J_2^s]$ gives the sign, which is necessary to convert the operator product $\prod_s \hat{C}^\dagger_{J_1^s} \hat{C}_{J_2^s}$ into normal order again. The weight of the external nodes in Eq. (35) can be written as

$$\tilde{M} = \sum_{I_1,I_2} M_{I_1,I_2} \tilde{C}^\dagger_{I_1} \tilde{C}_{I_2}, \tag{B18}$$

with

$$M_{I_1,I_2} = Q_{I_1,I_2} - \sum_{m=1} (-1)^{m+1} \sum_{\substack{\{(J_1^s,J_2^s)\} \\ s=0,\ldots,m}} \Sigma\big[J_1^s,J_2^s\big] \times Q_{J_1^0,J_2^0} \prod_{s=1}^{m} X_{J_1^s,J_2^s}. \tag{B19}$$

APPENDIX C: NONLOCAL OPERATORS AND SUPERCONDUCTING SYSTEMS

The generalization of our derivation in Sec. II to a nonlocal operator is straightforward. The numerator of the expectation value of a nonlocal operator $\hat{O}_{2,i}\hat{O}_{1,j}$ with respect to the Gutzwiller wave function reads

$$\langle \Psi_G | \hat{O}_{2,i} \hat{O}_{1,j} | \Psi_G \rangle = \langle \Psi_0 | \hat{Q}_{1,i} \hat{Q}_{2,j} \prod_{l \in \Lambda \setminus \{i,j\}} \hat{P}_l^\dagger \hat{P}_l \, | \Psi_0 \rangle. \tag{C1}$$

The Gutzwiller correlators for the sites i and j are combined with the operators $\hat{O}_{2,i}$ and $\hat{O}_{1,j}$ to the operators $\hat{Q}_{1,i}$ and $\hat{Q}_{2,j}$. Then, Eq. (32) generalizes to

$$\left\{ \hat{Q}_{1,i} \hat{Q}_{2,j} \prod_{l \in \Lambda \setminus \{i,j\}} (1 + \hat{A}_l) \right\}_\rho = \left\{ \tilde{Q}_{1,i} \tilde{Q}_{2,i} \prod_{l \in \Lambda \setminus \{i,j\}} (1 + \tilde{A}_l) \right\}_{\tilde{\rho}}. \tag{C2}$$

Finally, the definitions (34) and (35) allow us to write

$$\langle \Psi_G | \hat{O}_{1,i} \hat{O}_{2,j} | \Psi_G \rangle = \left\{ \tilde{M}_i \tilde{M}_j \prod_{l \in \Lambda} \exp(\tilde{G}_l) \right\}_{\tilde{\rho}}. \tag{C3}$$

Now, we are in the position to apply the LCT again.

The derivation in Sec. II is given for an arbitrary local Hilbert space and does not rely on translational invariance. Furthermore, we can generalize our derivation to systems without particle-number conservation in superconducting systems as treated, e.g., in Ref. [32]. To this end, we need to incorporate contractions of the form $\{\hat{C}^\dagger_{I_1} \hat{C}_{I_2}\}_\rho$, with $|I_1| \neq |I_2|$. These additional contractions lead to an adjustment of some of the index restrictions used in this work. For example, we need to allow even differences between the absolute value of the (set) indices I_1 and I_2 in Eqs. (18), (24), and (36) and between the indices $|J_1|,|J_2|$ in Eqs. (B10), (B14).

APPENDIX D: PREVIOUS APPROACHES

The diagrammatic analysis of the single-band case which includes the HF operators was first worked out in Ref. [10]. There, the correlator $\tilde{A} = x_d \tilde{n}_\uparrow \tilde{n}_\downarrow$ is employed with x_d being the only nonvanishing X_{I_1,I_2} coefficient. In this case, we can set

$$1 + x_d \tilde{C}^\dagger_{\uparrow,\downarrow} \tilde{C}_{\uparrow,\downarrow} = \exp(x_d \tilde{C}^\dagger_{\uparrow,\downarrow} \tilde{C}_{\uparrow,\downarrow}). \tag{D1}$$

When we work with multiple bands or if we allow for terms such as $x_\uparrow \tilde{c}^\dagger_\uparrow \tilde{c}_\uparrow$ in our correlator, an equation of the form (D1) does not hold. Hence we need to "re-exponentiate" our correlators as in Eq. (34). This has been overlooked in the application of the LCT in the multiband case in Ref. [42] although the problem was already noticed in Ref. [13].

In principle, the transformation of the ladder operators to HF-operators is not a necessary step for the transformation to Graßmann variables. After all operators have been brought to normal ordering every operator in the numerator (denominator) will appear only once. The operators can be mapped to Graßmann variables $\hat{c}^{(\dagger)}_{i\sigma} \to \tilde{c}^{(\dagger)}_{i\sigma}$ with vanishing anticommutator $\{\tilde{c}^\dagger_{i\sigma}, \tilde{c}_{i\sigma}\} = 0$. In contrast to the Graßmann variables defined in Eq. (28), the local contractions are still finite. The external and internal nodes can be defined as in Eq. (18).

In this approach, the diagrams will also include local lines. That means that we need to include the summation of an arbitrary number of (directly connected) nodes sitting on the same site, in order to sum up all local contractions. For the single-band model, the diagrammatic expansion of this case is derived in Refs. [4,6,8,9], where the cases of one and infinite dimensions are treated analytically. A similar approach for a three-flavor system with an Gutzwiller correlator of the form $1 + \alpha \hat{n}_1 \hat{n}_2 \hat{n}_3$ can be found in Ref. [49], where the local contractions are still present. The transformation to an exponential function is again trivial.

[1] U. Schollwöck, Ann. Phys. **326**, 96 (2011).
[2] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006).
[3] M. C. Gutzwiller, Phys. Rev. **134**, A923 (1964).
[4] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **59**, 121 (1987).
[5] F. Gebhard and D. Vollhardt, Phys. Rev. Lett. **59**, 1472 (1987).

[6] W. Metzner and D. Vollhardt, Phys. Rev. B **37**, 7382 (1988).
[7] F. Gebhard and D. Vollhardt, Phys. Rev. B **38**, 6911 (1988).
[8] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989).
[9] M. Kollar and D. Vollhardt, Phys. Rev. B **65**, 155121 (2002).
[10] F. Gebhard, Phys. Rev. B **41**, 9452 (1990).
[11] J. Bünemann and W. Weber, Physica B **230-232**, 412 (1997).
[12] J. Bünemann and W. Weber, Phys. Rev. B **55**, 4011 (1997).
[13] J. Bünemann, W. Weber, and F. Gebhard, Phys. Rev. B **57**, 6896 (1998).
[14] J. Bünemann, F. Gebhard, and W. Weber, in *Frontiers in Magnetic Materials*, edited by A. Narlikar (Springer, Berlin, 2005).
[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] X. Y. Deng, L. Wang, X. Dai, and Z. Fang, Phys. Rev. B **79**, 075114 (2009).
[18] G. T. Wang, X. Dai, and Z. Fang, Phys. Rev. Lett. **101**, 066403 (2008).
[19] G. T. Wang, Y. Qian, G. Xu, X. Dai, and Z. Fang, Phys. Rev. Lett. **104**, 047002 (2010).
[20] H. Weng, G. Xu, H. Zhang, S.-C. Zhang, X. Dai, and Z. Fang, Phys. Rev. B **84**, 060408 (2011).
[21] Y. X. Yao, J. Schmalian, C. Z. Wang, K. M. Ho, and G. Kotliar, Phys. Rev. B **84**, 245112 (2011).
[22] M.-F. Tian, X. Deng, Z. Fang, and X. Dai, Phys. Rev. B **84**, 205124 (2011).
[23] N. Lanatà, Y.-X. Yao, C.-Z. Wang, K.-M. Ho, J. Schmalian, K. Haule, and G. Kotliar, Phys. Rev. Lett. **111**, 196801 (2013).
[24] R. Dong, X. Wan, X. Dai, and S. Y. Savrasov, Phys. Rev. B **89**, 165122 (2014).
[25] T. Schickling, J. Bünemann, F. Gebhard, and W. Weber, New J. Phys. **16**, 093034 (2014).
[26] H. Schweitzer and G. Czycholl, Z. Phys. B **83**, 93 (1991).
[27] C. J. Halboth and W. Metzner, Z. Phys. B **102**, 501 (1997).
[28] L. F. Tocchio, F. Becca, A. Parola, and S. Sorella, Phys. Rev. B **78**, 041101 (2008).
[29] L. F. Tocchio, F. Becca, and C. Gros, Phys. Rev. B **83**, 195138 (2011).
[30] L. F. Tocchio, F. Becca, and C. Gros, Phys. Rev. B **86**, 035102 (2012).
[31] J. Bünemann, T. Schickling, and F. Gebhard, Europhys. Lett. **98**, 27006 (2012).
[32] J. Kaczmarczyk, J. Spałek, T. Schickling, and J. Bünemann, Phys. Rev. B **88**, 115127 (2013).
[33] J. Kaczmarczyk, Phil. Mag. **95**, 563 (2015).
[34] J. Kaczmarczyk, T. Schickling, and J. Bünemann, Phys. Status Solidi B **252**, 2059 (2015).
[35] J. Kaczmarczyk, J. Bünemann, and J. Spałek, New J. Phys. **16**, 073018 (2014).
[36] M. M. Wysokiński, J. Kaczmarczyk, and J. Spałek, Phys. Rev. B **92**, 125135 (2015).
[37] T. Schickling, F. Gebhard, and J. Bünemann, Phys. Rev. Lett. **106**, 146402 (2011).
[38] T. Schickling, F. Gebhard, J. Bünemann, L. Boeri, O. K. Andersen, and W. Weber, Phys. Rev. Lett. **108**, 036406 (2012).
[39] K. zu Münster, Ph.D. thesis, Philipps-Universität Marburg, 2015.
[40] D. Baeriswyl, D. Eichenberger, and M. Menteshashvili, New J. Phys. **11**, 075010 (2009).
[41] R. Rüger, L. F. Tocchio, R. Valentí, and C. Gros, New J. Phys. **16**, 033010 (2014).
[42] J. Bünemann (unpublished).
[43] P. Horsch and P. Fulde, Zeitschrift für Physik B Condensed Matter **36**, 23 (1979).
[44] J. Kaczmarczyk (private communication).
[45] M. Zegrodnik and J. Spałek, Phys. Rev. B **86**, 014505 (2012).
[46] M. Zegrodnik, J. Spałek, and J. Bünemann, New J. Phys. **15**, 073050 (2013).
[47] M. Zegrodnik, J. Bünemann, and J. Spałek, New J. Phys. **16**, 033001 (2014).
[48] S. Sugano, T. Yukito, and H. Kamimura, *Multiplets of Transition Metal Ions in Crystals* (Academic Press, New York, London, 1970).
[49] A. Rapp, W. Hofstetter, and G. Zaránd, Phys. Rev. B **77**, 144520 (2008).