Eur. Phys. J. B **4**, 29–38 (1998)

THE EUROPEAN PHYSICAL JOURNAL B

The Gutzwiller approximation for degenerate bands: a formal derivation

J. Bünemann[a]

Inst. f. Physik, Universität Dortmund, 44221 Dortmund, Germany

Received: 9 December 1997 / Revised and accepted: 6 March 1998

Abstract. We present a detailed derivation of the Gutzwiller approximation for multi-band Hubbard models with density-density Coulomb interactions. For the one-band Hubbard model we introduce a mathematically well-defined formalism which is easily generalized to the multi-band case. In contrast to earlier attempts, our approach allows us to include inter-orbital hopping terms in the Hamiltonian.

PACS. 71.10.Fd Lattice fermion models (Hubbard model, etc.) – 71.27.+a Strongly correlated electron systems; heavy fermions

1 Introduction

During the last decades the theoretical investigation of correlated metals focussed on the one-band Hubbard model [1]. These intensive studies considerably deepened our understanding of the model; for a recent review, see reference [2]. However, it remains debatable whether the one-band Hubbard model actually captures the essential physics of real materials like the transition metals and their compounds. For example, it is commonly accepted that ferromagnetism in the one-band Hubbard model requires (unrealistically) large interaction strengths, or a density of states, which significantly differs from those expected in three dimensional systems with realistic hopping matrix elements. On the other hand, the variational investigations of a two-band model [3–5] indicate that intra-atomic Hund's-rule exchange terms place the ferromagnetic transition into a region of more moderate interaction strengths. Thus, the atomic structure of transition metals must be taken seriously to develop a *qualitative* understanding of the mechanisms for correlation effects in these materials. Consequently, it is necessary to study *multi-band* Hubbard models. In this work we analyze Jastrow-type many-particle trial states as approximate ground states for multi-band models.

In his seminal work Gutzwiller proposed a variational method for the investigation of the one-band Hubbard model [6]. His many-particle wave function $|\Psi\rangle$ is defined as a many-particle operator ("correlator") acting on a one-particle product state $|\Psi_0\rangle$. The correlator reduces the number of energetically unfavorable doubly occupied sites with respect to the reference state $|\Psi_0\rangle$. Gutzwiller also proposed an approximate evaluation of his trial state ("Gutzwiller approximation") which was later seen to become exact in the limit of infinite spatial dimensions [7,8]. From the results of his approximation he also concluded that ferromagnetism requires large Coulomb interactions.

After Gutzwiller's basic work there were several attempts to generalize the Gutzwiller method to more complicated situations. However, even the inclusion of an antiferromagnetic one-particle state $|\Psi_0\rangle$ in the one-band case led to some inconsistent and partly unphysical results [9,10]. Chao and Gutzwiller [11,12] generalized the Gutzwiller approximation to a degenerate two-band model but an analytical expression for the variational ground-state energy of their generalized trial state could not be derived.

The physical content of the Gutzwiller approximation became more transparent in the work by Ogawa *et al.* [9] and by Vollhardt [13], where it was derived from (semi-classical) counting arguments. This approach led to a successful application of the Gutzwiller approximation to more complicated Hamiltonians, *e.g.*, the periodic Anderson model [14–18]. Despite this success the counting approach remains unsatisfactory for two reasons. First, it is not clear to which variational wave function the final results actually apply. Second, it lacks a simple recipe that can straightforwardly be applied to all cases of interest.

For example, serious problems arise for multi-band wave functions, where the total number of electrons per orbital is not conserved. When inter-orbital hopping terms are not artificially neglected, a consistent formulation of the standard counting approximation was up to now not possible for these wave functions. In a recent work Okabe [19] found the correct results for multi-band systems with non-hybridized orbitals. However, in his formulation he could only speculate about the expectation values for inter-orbital hopping terms. Indeed, his expressions differ from our results.

---

[a] e-mail: `bueneman@grad.physik.uni-dortmund.de`

For non-degenerate orbitals additional complications arise, because in this case the generalized Gutzwiller correlator modifies the occupancies of the orbitals. As a consequence it is necessary to introduce a new one-particle state $|\Phi_0\rangle$ to derive a compact expression for the variational ground-state energy. The subtle difference between $|\Phi_0\rangle$ and $|\Psi_0\rangle$ in the periodic Anderson model was discussed in detail by Vulović and Abrahams [14]. Although they found the correct energy expression for the periodic Anderson model, their formulation could neither be generalized to multi-band systems. In this work we present a mathematical definition of the Gutzwiller approximation scheme for definite variational wave functions. The results and first applications of our generalized Gutzwiller method were given in references [3,20]. In this article we present a detailed derivation of our formalism.

An alternative route to derive the energy expression of the Gutzwiller method is the slave-boson mean-field approximation for degenerate Hubbard models [22,23]. In reference [21] our results are compared with this approach and with earlier attempts to generalize the Gutzwiller method.

Recently it became possible to evaluate analytically a more general class of Gutzwiller correlated wave functions in the limit of infinite spatial dimensions, both for a Hamiltonian with pure density-density interaction [21] (see below) and for the full atomic Hamiltonian [5]. This evaluation requires many-particle techniques and, therefore, is much more complex than the combinatorial derivation of the Gutzwiller approximation. Hence, our method should be seen as a simple alternative to the sophisticated diagrammatical formalism in reference [21]. Moreover, our approach finally clarifies the long-standing problem of a generalization of the Gutzwiller approximation to multi-band systems.

Our paper is organized as follows. In Section 2 we introduce a multi-band Hubbard model and our corresponding Gutzwiller-correlated wave functions. In Section 3 we re-derive the Gutzwiller approximation for the one-band Hubbard model. Then, in Section 4, we extend the one-band formalism to our multi-band Hamiltonian. A summary in Section 5 closes our presentation.

2 Model Hamiltonian and generalized Gutzwiller wave functions

2.1 Hamiltonian and notations

We study the multi-band Hubbard-type Hamiltonian

$$\hat{H} = \sum_{i,j,\sigma,\sigma'} t_{i,j}^{\sigma,\sigma'} \hat{c}_{i;\sigma}^{+} \hat{c}_{j;\sigma'} + \sum_{\sigma,\sigma'} U^{\sigma,\sigma'} \sum_{i} \hat{n}_{i;\sigma} \hat{n}_{i;\sigma'} . \tag{1}$$

In standard notation $\hat{c}_{i;\sigma}^{+}$ ($\hat{c}_{i;\sigma}$) creates (annihilates) an electron in the spin-orbit state σ at lattice site i, and $\hat{n}_{i;\sigma} = \hat{c}_{i;\sigma}^{+}\hat{c}_{i;\sigma}$ is the respective density operator. In the following the $2N$ spin-orbit states σ will be briefly denoted as 'orbitals'; N is the number of orbitals per site. Note that σ, σ' may also include orbitals of different atoms if it is necessary to study a basis with more than one atom per site. Furthermore, $t_{i,j}^{\sigma,\sigma'} = (t_{j,i}^{\sigma',\sigma})^{*}$ is the matrix element for an electron transfer from orbital σ' on site j to orbital σ at site i.

The Hamiltonian (1) contains only density-density interactions. Hence, all local terms of the form

$$\hat{c}_{i;\sigma_1}^{+} \hat{c}_{i;\sigma_2}^{+} \hat{c}_{i;\sigma_3} \hat{c}_{i;\sigma_4} \quad \text{with} \quad \{\sigma_1, \sigma_2\} \neq \{\sigma_3, \sigma_4\} , \tag{2}$$

which cannot be written as products of density operators, are neglected. Nevertheless, Hund's rule exchange terms are partly included in the Hamiltonian (1). To see this we divide the spin-orbit index σ in an orbital (b) and a spin (s) part, and write our coefficients as

$$U^{(b,s),(b',s')} = U_{\mathrm{C}}^{b,b'} - \delta_{s,s'} J^{b,b'} . \tag{3}$$

$U_{\mathrm{C}}^{b,b'}$ and $J^{b,b'}$ are the standard Coulomb and exchange integrals between the two orbitals b, b' ($J^{b,b} \equiv 0$). Thus, the Hamiltonian (1) favors atomic configurations with a maximum total spin into the z direction. However, Hund's first rule is not fulfilled, since these configurations are not degenerate with the other states of the same spin multiplet. To obtain the correct spin multiplicity and ligand field symmetry we have to include all relevant exchange terms (2); for more details and an example, see reference [5]. Within the Gutzwiller approximation scheme it seems to be necessary to restrict ourself to the incomplete Hamiltonian (1). In the single-band case ($N = 1$) we recover the well-known Hubbard model,

$$\hat{H} = \sum_{i,j} \sum_{\sigma=1}^{2} t_{i,j} \hat{c}_{i;\sigma}^{+} \hat{c}_{j;\sigma} + U \sum_{i} \hat{n}_{i;1} \hat{n}_{i;2}, \tag{4}$$

where the two numbers $\sigma = 1, 2$ represent the two spin directions and $U \equiv 2U^{1,2}$.

To properly define the wave functions for the variational investigation of our Hamiltonian (1) it is convenient to introduce the following notations.

(i) At each lattice site 2^{2N} different site occupancies are possible. For example, in a two-band model ($N = 2$) the 16 occupancies are: one empty, four single, six double, four triple, and one quadruple occupancy. In general, we will denote these occupancies by multiple-orbital indices

$$\begin{aligned} I \in \{ & \emptyset; (1), \ldots, (2N); (1,2), (1,3) \ldots, (2,3), (2,4), \ldots ; \\ & (1,2,3), \ldots ; \ldots ; (1, \ldots, 2N) \} , \end{aligned} \tag{5}$$

i.e., the index I contains all occupied orbitals, which are arbitrarily numbered as $\sigma = 1, \ldots, 2N$. The symbol $\emptyset$ means an empty site. We will interpret the indices as sets in the usual sense. Thus, the standard mathematical set operations apply. For example, in the configuration $I \backslash I'$ only those orbitals in I are occupied which are not in I'. The complement of I is $\overline{I} = (1, 2, \ldots, 2N) \backslash I$, *i.e.*, in the configuration $\overline{I}$ all orbitals but those in I are occupied.

The absolute value $|I|$ of a configuration is the number of elements in it, *i.e.*,

$$|(\emptyset)| = 0; |(\sigma)| = 1; |(\sigma,\sigma')| = 2; \dots ;\\ |(1,\dots,2N)| = 2N. \tag{6}$$

(ii) The operator which projects onto a specific configuration I on site i is given by

$$\hat{m}_{i;I} = \prod_{\sigma\in I} \hat{n}_{i;\sigma} \prod_{\sigma'\in\overline{I}} (1-\hat{n}_{i;\sigma'}) \,. \tag{7}$$

Thus, the operator

$$\hat{M}_I = \sum_i \hat{m}_{i;I} \tag{8}$$

counts the number of sites with an occupancy I. These "net" occupancy operators differ from the respective "gross" occupancy operators

$$\hat{n}_{i;I} = \prod_{\sigma\in I} \hat{n}_{i;\sigma}, \tag{9a}$$

$$\hat{N}_I = \sum_i \hat{n}_{i;I}. \tag{9b}$$

$\hat{N}_I$ measures the number of sites where all orbitals $\sigma \in I$ are occupied, and all sites with an occupancy $I' \supseteq I$ are also included. Each gross (net) operator can be written as a sum of net (gross) operators:

$$\hat{N}_I = \sum_{I'(\supseteq I)} \hat{M}_{I'}, \tag{10a}$$

$$\hat{M}_I = \sum_{I'(\supseteq I)} (-1)^{|I'\setminus I|}\, \hat{N}_{I'}. \tag{10b}$$

With the help of the above definitions we may rewrite our Hamiltonian (1) as

$$\hat{H} = \sum_{i,j,\sigma,\sigma'} t_{i,j}^{\sigma,\sigma'}\, \hat{c}_{i;\sigma}^{+} \hat{c}_{j;\sigma'} + \sum_I U_I\, \hat{M}_I, \tag{11a}$$

with the configuration energies

$$U_I = \sum_{\sigma,\sigma'\in I} U^{\sigma,\sigma'}. \tag{11b}$$

2.2 Variational wave functions

To examine the one-band Hubbard model,

$$\hat{H} = \sum_{i,j} \sum_{\sigma=1}^{2} t_{i,j} \hat{c}_{i;\sigma}^{+} \hat{c}_{j;\sigma} + U\, \hat{M}_{12}, \tag{12}$$

Gutzwiller [6] proposed the following variational wave function,

$$|\Psi\rangle = g^{\hat{M}_{12}} |\Psi_0\rangle. \tag{13}$$

Here, $|\Psi_0\rangle$ is any normalized one-particle product state and g is a real variational parameter which controls the number of doubly occupied sites. A straightforward generalization of this wave function for a multi-orbital system is given by [20]

$$|\Psi\rangle = \prod_{I\,(|I|\geq 2)} g_I^{\hat{M}_I} |\Psi_0\rangle. \tag{14}$$

Besides the $2^{2N} - (2N+1)$ real numbers g_I, the one-particle wave function $|\Psi_0\rangle$ may also depend on further variational parameters. For example, they may be used to control the gross occupancies $\langle n_{i,\sigma}\rangle_\Psi$ when non-degenerate orbitals are involved, or magnetic order must be taken into account. Due to the relations (10) we could also use the gross operators (9b) in the ansatz (14) without changing our variational space. However, as we will see below, the net operators are more useful for the formulation of our approximation.

In the course of our derivation it will also become clear that an analytical expression for the energy of our wave function (14) becomes feasible because *every* multiple occupancy ($|I| \geq 2$) is controlled by its own variational parameter. The earlier attempts of Chao and Gutzwiller [11, 12] to extend the Gutzwiller approximation were bound to fail just at this point. Their variational wave function,

$$|\Psi'\rangle = \prod_{I\,(|I|=2)} g_I^{\hat{N}_I} |\Psi_0\rangle, \tag{15}$$

contains only $N(2N-1)$ independent variational parameters, and, thus, is too restricted to allow for a successful generalization of the Gutzwiller approximation to the multi-band case.

3 Gutzwiller approximation for the one-band Hubbard model

In this section we re-derive the Gutzwiller approximation for the one-band Hubbard model. Our formulation will enable us to treat the multi-band case as a straightforward generalization of the one-band formalism.

As required in any variational examination of a quantum mechanical system, we have to determine the expectation value for the Hamiltonian (12) in our trial state (13)

$$\langle \hat{H}\rangle_\Psi \equiv \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi\mid\Psi\rangle} = \sum_{i,j} t_{i,j} \sum_{\sigma=1}^{2} \langle \hat{c}_{i;\sigma}^{+} \hat{c}_{j;\sigma}\rangle_\Psi \\ + U \sum_i \langle \hat{m}_{i,12}\rangle_\Psi \;. \tag{16}$$

Without further approximations the analytical evaluation of this energy expression was possible so far only in the special cases of one [7] and infinite spatial dimensions [8]. Gutzwiller himself proposed an approximate evaluation of equation (16), the so-called "Gutzwiller approximation", which was later motivated in a more transparent manner by Ogawa *et al.* [9] and Vollhardt [13]. Our formulation is similar to that of Vulović and Abrahams [14].

To formulate a well-defined mathematical description of the Gutzwiller approximation it is convenient to introduce a local basis for our many-particle Hilbert space,

$$|\Phi_{\mathcal{N}_1,\mathcal{N}_2}\rangle \equiv \prod_{i_1\in\mathcal{N}_1}\prod_{i_2\in\mathcal{N}_2}\hat{c}^+_{i_1;1}\hat{c}^+_{i_2;2}|vacuum\rangle. \tag{17}$$

$\mathcal{N}_1$ and $\mathcal{N}_2$ are sets of lattice sites with $N_1 = |\mathcal{N}_1|$ and $N_2 = |\mathcal{N}_2|$ elements. As the basis (17) is complete, we may expand the one-particle wave function in the form

$$|\Psi_0\rangle = \sum_{\mathcal{N}_1,\mathcal{N}_2} T(\mathcal{N}_1,\mathcal{N}_2)|\Phi_{\mathcal{N}_1,\mathcal{N}_2}\rangle, \tag{18}$$

with expansion coefficients $T(\mathcal{N}_1,\mathcal{N}_2)$. The sum over $\mathcal{N}_1,\mathcal{N}_2$ has to include all subsets of the whole set of lattice sites $\mathcal{L}$ ($|\mathcal{L}| = L$) with N_1, N_2 elements, respectively. In this way we also fix the number of particles in $|\Psi_0\rangle$. This is a reasonable restriction since the Hamiltonian (12) commutes with the gross operators $\hat{N}_1$ and $\hat{N}_2$. Then, our correlated trial state (13) becomes

$$|\Psi\rangle = \sum_{\mathcal{N}_1,\mathcal{N}_2} g^{M_{12}(\mathcal{N}_1,\mathcal{N}_2)}T(\mathcal{N}_1,\mathcal{N}_2)|\Phi_{\mathcal{N}_1,\mathcal{N}_2}\rangle, \tag{19}$$

where the function $M_{12}(\mathcal{N}_1,\mathcal{N}_2)$ counts the number of doubly occupied sites, provided that the up (down) electrons are placed at the sites belonging to $\mathcal{N}_1$ ($\mathcal{N}_2$). We use the fact that the states (17) are orthogonal and normalized, and write the expectation values in (16) as

$$\langle\Psi \mid \Psi\rangle = \sum_{\mathcal{N}_1,\mathcal{N}_2} g^{2M_{12}(\mathcal{N}_1,\mathcal{N}_2)}\left|T(\mathcal{N}_1,\mathcal{N}_2)\right|^2, \tag{20a}$$

$$\begin{aligned}\langle\Psi|\hat{c}^+_{i;1}\hat{c}_{j;1}|\Psi\rangle = &\sum_{\mathcal{N}_1',\mathcal{N}_2} g^{M_{12}(\mathcal{N}_1'\cup i,\mathcal{N}_2)}g^{M_{12}(\mathcal{N}_1'\cup j,\mathcal{N}_2)}\\ &\times T^*(\mathcal{N}_1'\cup i,\mathcal{N}_2)\,T(\mathcal{N}_1'\cup j,\mathcal{N}_2),\end{aligned} \tag{20b}$$

$$\begin{aligned}\langle\Psi|\hat{m}_{i,12}|\Psi\rangle = &\sum_{\mathcal{N}_1',\mathcal{N}_2'} g^{2M_{12}(\mathcal{N}_1'\cup i,\mathcal{N}_2'\cup i)}\\ &\times\left|T(\mathcal{N}_1'\cup i,\mathcal{N}_2'\cup i)\right|^2.\end{aligned} \tag{20c}$$

Here, the following restrictions for the sum over $\mathcal{N}_1'$ and $\mathcal{N}_2'$ hold,

$$i,j\notin\mathcal{N}_1' \quad\text{and}\quad |\mathcal{N}_1'| = N_1-1 \quad\text{in (20b)}, \tag{21a}$$

$$\begin{aligned} i\notin\mathcal{N}_1',\mathcal{N}_2' \quad\text{and}\quad &|\mathcal{N}_1'| = N_1-1,\\ &|\mathcal{N}_2'| = N_2-1 \quad\text{in (20c)}.\end{aligned} \tag{21b}$$

To make further progress we need expressions for the two types of products between the coefficients T in (20). They can be obtained approximately if we consider the corresponding expectation values in the *uncorrelated* wave function $|\Psi_0\rangle$. As a first step we address the norm

$$\langle\Psi_0 \mid \Psi_0\rangle = \sum_{\mathcal{N}_1,\mathcal{N}_2}\left|T(\mathcal{N}_1,\mathcal{N}_2)\right|^2 = 1. \tag{22}$$

The Gutzwiller approximation will now be defined as follows. We assume that each term in (22) is independent of the special choice of the sets $\mathcal{N}_1,\mathcal{N}_2$. This means that all local distributions of electrons have the same probability in the uncorrelated state $|\Psi_0\rangle$. Consequently, equation (22) gives

$$\begin{aligned}\left|T(\mathcal{N}_1,\mathcal{N}_2)\right|^2 &= \binom{L}{N_1}^{-1}\binom{L}{N_2}^{-1}\\ &\equiv P(L,N_1)P(L,N_2).\end{aligned} \tag{23}$$

As a second step we make an equivalent assumption about an expectation value which involves an electron transfer,

$$\begin{aligned}\langle\Psi_0|\hat{c}^+_{i;1}\hat{c}_{j;1}|\Psi_0\rangle &\equiv \langle\hat{c}^+_{i;1}\hat{c}_{j;1}\rangle_0\\ &= \sum_{\mathcal{N}_1',\mathcal{N}_2} T^*(\mathcal{N}_1'\cup i,\mathcal{N}_2)T(\mathcal{N}_1'\cup j,\mathcal{N}_2).\end{aligned} \tag{24}$$

Again, we require that the individual terms in (24) do not depend on the sets $\mathcal{N}_1'$, $\mathcal{N}_2$. Then, equation (24) leads to

$$\begin{aligned}&T^*(\mathcal{N}_1'\cup i,\mathcal{N}_2)T(\mathcal{N}_1'\cup j,\mathcal{N}_2) =\\ &P(L-2,N_1-1)P(L,N_2)\langle\hat{c}^+_{i;1}\hat{c}_{j;1}\rangle_0.\end{aligned} \tag{25}$$

Equations (23, 25) constitute the Gutzwiller approximation for the single-band case.

Now we are in the position to evaluate equations (20). First, the norm in the *interacting* case becomes

$$\langle\Psi \mid \Psi\rangle = P(L,N_1)P(L,N_2)\sum_{\mathcal{N}_1,\mathcal{N}_2} g^{2M_{12}(\mathcal{N}_1,\mathcal{N}_2)}. \tag{26}$$

The sum over the sets $\mathcal{N}_1,\mathcal{N}_2$ can be replaced by a sum over the number of doubly occupied sites M_{12} if we introduce the combinatorial factor

$$\begin{aligned}&C_{M_{12}}(L,N_1,N_2)\\ &= \frac{L!}{M_{12}!(N_1-M_{12})!(N_2-M_{12})!(L-N_1-N_2+M_{12})!}\end{aligned} \tag{27a}$$

$$= \frac{L!}{M_{12}!M_1!M_2!M_\emptyset!}\,. \tag{27b}$$

This factor counts the number of possibilities to distribute N_1, N_2 electrons with spin up and down on the lattice such that M_{12} sites are doubly occupied. It allows us to write equation (26) as

$$\begin{aligned}\langle\Psi \mid \Psi\rangle = &P(L,N_1)P(L,N_2)\\ &\times\sum_{M_{12}} g^{2M_{12}}C_{M_{12}}(L,N_1,N_2)\,.\end{aligned} \tag{28}$$

With the help of equation (23) the expectation value for a double occupancy (20c) becomes

$$\begin{aligned}\langle\Psi|\hat{m}_{i,12}|\Psi\rangle &= P(L,N_1)P(L,N_2)\\ &\quad\times\sum_{\mathcal{N}_1',\mathcal{N}_2'} g^{2M_{12}(\mathcal{N}_1'\cup i,\mathcal{N}_2'\cup i)} \qquad (29a)\\ &= P(L,N_1)P(L,N_2)\\ &\quad\times\sum_{M_{12}} g^{2M_{12}+2}C_{M_{12}}(L-1,N_1-1,N_2-1). \qquad (29b)\end{aligned}$$

To relate this expression to (28) for the norm we recast equation (29b) into the form

$$\begin{aligned}\langle\Psi|\hat{m}_{i,12}|\Psi\rangle &= P(L,N_1)P(L,N_2)\\ &\quad\times\sum_{M_{12}} g^2\frac{(N_1-M_{12})(N_2-M_{12})}{L(L-N_1-N_2+M_{12})}\\ &\quad\times\left[g^{2M_{12}}C_{M_{12}}(L,N_1,N_2)\right]. \qquad (30)\end{aligned}$$

Here, we used the fact that we may set

$$L-N_1-N_2+M_{12}+1\simeq L-N_1-N_2+M_{12} \qquad (31)$$

in the quotient (30) when we perform the thermodynamic limit ($L\to\infty$, $N_1\to\infty$, $N_2\to\infty$, $M_{12}\to\infty$; $N_{1,2}/L$ and M_{12}/L finite). Note that the factors $P(L,N_1)^{-1}$, $P(L,N_2)^{-1}$, and $C_{M_{12}}(L,N_1,N_2)$ are macroscopic quantities of the order of $\mathcal{O}(\exp(L))$. Below we will use this fact to replace the sum in (30) by its maximum term when we divide $\langle\Psi|\hat{m}_{i;12}|\Psi\rangle$ by the norm (28).

A similar form as equation (30) can be found for the expectation value for a hopping term. With the help of (25) we find

$$\begin{aligned}\langle\Psi|\hat{c}_{i;1}^{+}\hat{c}_{j;1}|\Psi\rangle &= \langle\hat{c}_{i;1}^{+}\hat{c}_{j;1}\rangle_0 P(L-2,N_1-1)P(L,N_2)\\ &\quad\times\sum_{\mathcal{N}_1',\mathcal{N}_2} g^{M_{12}(\mathcal{N}_1'\cup i,\mathcal{N}_2)}g^{M_{12}(\mathcal{N}_1'\cup j,\mathcal{N}_2)}. \qquad (32)\end{aligned}$$

Again, the sum over the sets $\mathcal{N}_1',\mathcal{N}_2$ can be replaced by a sum over M_{12}. To this end we have to distinguish four different cases in (32) since the sites i and j can either be occupied with the opposite spin or empty when the hopping process takes place. Hence, we have to address (i) $i,j\notin\mathcal{N}_2$; (ii) $i\in\mathcal{N}_2, j\notin\mathcal{N}_2$; (iii) $j\in\mathcal{N}_2, i\notin\mathcal{N}_2$; (iv) $i,j\in\mathcal{N}_2$. Altogether, (32) can be written as

$$\begin{aligned}\langle\Psi|\hat{c}_{i;1}^{+}\hat{c}_{j;1}|\Psi\rangle &= \langle\hat{c}_{i;1}^{+}\hat{c}_{j;1}\rangle_0 P(L-2,N_1-1)P(L,N_2)\\ &\quad\times\sum_{M_{12}} g^{2M_{12}}\Big[C_{M_{12}}(L-2,N_1-1,N_2)\\ &\quad+2gC_{M_{12}}(L-2,N_1-1,N_2-1)\\ &\quad+g^2C_{M_{12}}(L-2,N_1-1,N_2-2)\Big]. \qquad (33)\end{aligned}$$

We extract the factors equivalent to the norm, and arrive at

$$\begin{aligned}\langle\Psi|\hat{c}_{i;1}^{+}\hat{c}_{j;1}|\Psi\rangle &= \langle\hat{c}_{i;1}^{+}\hat{c}_{j;1}\rangle_0 P(L,N_1)P(L,N_2)\frac{L^2}{N_1(L-N_1)}\\ &\quad\times\sum_{M_{12}}\Big[\frac{(N_1-M_{12})(L-N_1-N_2+M_{12})}{L^2}\\ &\quad+2g\frac{(N_1-M_{12})(N_2-M_{12})}{L^2} \qquad (34)\\ &\quad+g^2\frac{(N_1-M_{12})(N_2-M_{12})^2}{L^2(L-N_1-N_2+M_{12})}\Big]g^{2M_{12}}C_{M_{12}}(L,N_1,N_2),\end{aligned}$$

where we again neglected numbers of the order unity in the quotients, similar to equation (31).

Now we are in the position to take advantage of the simplifications in the thermodynamic limit when we divide by the norm. The sum in equation (28) will have a macroscopic maximum for a certain value $\overline{M}_{12}$ (and respective values $\overline{M}_1$, $\overline{M}_2$, $\overline{M}_\emptyset$). The same holds for the expectation values (30, 34) since the factors of order one do not affect the macroscopic peak position. Hence, we may replace the sum by its maximum term in all these cases. After the application of Stirling's formula, $X!\approx X^X$, we see that this maximization problem has the general form

$$\frac{\partial}{\partial x}\left(p^x\prod_i f_i(x)^{\alpha_i f_i(x)}\right)=0, \qquad (35a)$$

where the linear functions $f_i(x)$ obey the additional condition

$$\sum_i\alpha_i\frac{\partial}{\partial x}f_i(x)=0. \qquad (35b)$$

After a short calculation we find

$$p\prod_i f_i(x)^{\alpha_i\frac{\partial}{\partial x}f_i(x)}=1. \qquad (36)$$

For example, for the norm we obtain from (27a, 28) that $f_1(x)=M_{12}=x$, $f_2(x)=N_1-x$, $f_3(x)=N_2-x$, $f_4(x)=L-N_1-N_2+x$, $\alpha_i=-1$, and $p=g^2$. Hence, we immediately find from (36) the well-known "maximum-term condition" of the Gutzwiller approximation,

$$g^2=\frac{m_\emptyset m_{12}}{m_1m_2} \qquad \left(m_I\equiv\overline{M}_I/L\right). \qquad (37)$$

Equation (37) allows us to replace the original variational parameter g by the new and more transparent parameter m_{12}.

When we replace the expressions (30, 34) by their maximum terms and divide them by the norm we obtain the expectation values in the Hamiltonian (16). Equation (30) gives the expected result for the mean number of double occupancies,

$$\langle\hat{m}_{i,12}\rangle_\Psi=g^2\frac{m_1m_2}{m_\emptyset}=m_{12}, \qquad (38)$$

where we used the maximum term condition (37) in the second step. Analogously, we derive from (34)

$$\langle \hat{c}^{+}_{i;\sigma}\hat{c}_{j;\sigma}\rangle_{\Psi} = q_\sigma \langle \hat{c}^{+}_{i;\sigma}\hat{c}_{j;\sigma}\rangle_0 . \tag{39}$$

Here, we introduced Gutzwiller's "loss-factors"

$$\begin{aligned} q_1 &= m_1 m_\emptyset + 2g m_1 m_2 + g^2 \frac{m_1 m_2^2}{m_\emptyset} \\ &= \left(\sqrt{m_1 m_\emptyset} + \sqrt{m_2 m_{12}}\,\right)^2 , \\ q_2 &= \left(\sqrt{m_2 m_\emptyset} + \sqrt{m_1 m_{12}}\,\right)^2 . \end{aligned} \tag{40}$$

Thus, we finally obtain

$$\begin{aligned} \langle \hat{H}\rangle_{\Psi} = \sum_{i(\neq)j} t_{i,j} \sum_{\sigma=1}^{2} q_\sigma \langle \hat{c}^{+}_{i;\sigma}\hat{c}_{j;\sigma}\rangle_0 \\ + L\epsilon \sum_{\sigma=1}^{2} n_\sigma + LU m_{12} \end{aligned} \tag{41}$$

for the variational ground-state energy of the one-band Hubbard model within the Gutzwiller approximation. Here, we included the site-independent on-site energies $\epsilon \equiv t_{i,i}$. Equation (41) has to be minimized with respect to the variational parameter m_{12} and (perhaps) additional parameters which determine the one-particle product state $|\Psi_0\rangle$.

4 Gutzwiller approximation for multi-band Hubbard models

In principle, the derivation of the Gutzwiller approximation for multi-band Hubbard models is a straightforward generalization of our one-band formalism presented in the last section. There is only one obstacle which complicates the calculation. As first discussed by Vulović and Abrahams [14] in the context of the periodic Anderson model, the expectation values for the gross one-particle occupancies are different for the uncorrelated state $|\Psi_0\rangle$ and the many-particle wave function $|\Psi\rangle$. This can most easily be seen for a system with non-degenerate orbitals. In this case it is obvious that the correlator in (14) tends to reduce the number of electrons in the same orbitals. Consequently, the occupancies $\langle \hat{N}_\sigma\rangle_\Psi$ are modified by the Coulomb interaction and, in general, they will not coincide with the respective expectation values in $|\Psi_0\rangle$.

To overcome the resulting technical problems in the derivation of the Gutzwiller approximation we replace $|\Psi_0\rangle$ by another (normalized) one-particle product wave function $|\Phi_0\rangle$,

$$|\Psi_0\rangle = \prod_\sigma \eta_\sigma^{\hat{N}_\sigma} |\Phi_0\rangle . \tag{42}$$

The parameters η_σ are real numbers, and will be chosen such that the gross one-particle occupancies are the same in $|\Phi_0\rangle$ and $|\Psi\rangle$; see below. Since $|\Psi_0\rangle$ and $|\Phi_0\rangle$ both cover the whole space of one-particle product wave functions, we do not modify the space of correlated variational wave functions by this procedure. In principle, we could also work with the original form of $|\Psi\rangle$, equation (14), to derive the same final energy functional. However, the transformation (42) considerably simplifies our calculations and the explicit form of the energy functional.

As already known from the one-band case it is convenient to evaluate $|\Phi_0\rangle$ in a local basis

$$|\Phi_0\rangle = \sum_{\widetilde{\mathcal{N}}} T(\widetilde{\mathcal{N}}) |\Phi_{\widetilde{\mathcal{N}}}\rangle , \tag{43a}$$

$$|\Phi_{\widetilde{\mathcal{N}}}\rangle \equiv \prod_\sigma \prod_{i\in\mathcal{N}_\sigma} \hat{c}^{+}_{i,\sigma} |vacuum\rangle . \tag{43b}$$

Here and in the following we use the abbreviations

$$\widetilde{\mathcal{N}} \equiv \{\mathcal{N}_1, \dots , \mathcal{N}_{2N}\} , \tag{44a}$$

$$\widetilde{N} \equiv \{N_1, \dots , N_{2N}\} \tag{44b}$$

for all sub-sets $\mathcal{N}_\sigma$ of lattice sites where the orbitals σ are occupied, and the corresponding set of numbers $N_\sigma = |\mathcal{N}_\sigma|$. The hybridization terms in (1) lead to a major difference between equations (18, 43a). In (43a) the sum over $\mathcal{N}_1, \dots , \mathcal{N}_{2N}$ will include all possible number distributions $N_1, \dots , N_{2N}$ of electrons on the lattice, since, in general, the gross occupancy operators $\hat{N}_\sigma$ do not commute with the Hamiltonian (1).

In close analogy to the one-band case we need to derive expressions for certain products of the coefficients T in (43a). Again, we consider the one-particle state $|\Phi_0\rangle$ to set up the Gutzwiller approximation for the multi-band case. First, the norm becomes

$$\langle \Phi_0 | \Phi_0 \rangle = \sum_{\widetilde{\mathcal{N}}} \left| T(\widetilde{\mathcal{N}}) \right|^2 = \sum_{\widetilde{N}} \Omega(\widetilde{N}) , \tag{45a}$$

$$\Omega(\widetilde{N}) \equiv \sum_{\widetilde{\mathcal{N}}} \left| T(\widetilde{\mathcal{N}}) \right|^2 \prod_\sigma \delta_{|\mathcal{N}_\sigma|, N_\sigma} , \tag{45b}$$

where $\Omega(\widetilde{N})$ is the probability to find exactly $N_1, \dots , N_{2N}$ electrons in the orbitals $\sigma = 1, \dots , 2N$. The Gutzwiller approximation is defined by

$$\left| T(\widetilde{\mathcal{N}}) \right|^2 = \Omega(\widetilde{N}) \prod_\sigma P(L, N_\sigma) ; \tag{46}$$

compare equation (23) for the one-band case.

In principle, we could fix certain occupation densities in the wave functions $|\Psi\rangle$ and $|\Phi_0\rangle$. For example, the operator for the total number of spin-up (spin-down) electrons commutes with the Hamiltonian (1), as the latter usually includes no spin-flip terms. Thus, $|\Psi\rangle$ and $|\Phi_0\rangle$ could be chosen as eigenstates of these operators, *i.e.*, we could work with wave functions from a "canonical ensemble". Additional conserved quantities arise if some orbitals are not mixed by the hybridization term in (1). Such "conservation laws" could be implemented by imposing certain boundary conditions on the occupancies N_σ. In this case, the number distribution $\Omega(\widetilde{N})$ would be non-zero only in a d-dimensional subspace ($d < 2N$) of the whole space of occupancies $\{N_1, \dots, N_{2N}\}$, where $2N - d$ is the number of boundary conditions.

Here, we use a more elegant way to derive our results. We work with "grand-canonical" instead of "canonical" wave functions, *i.e.*, $|\Psi\rangle$ and $|\Phi_0\rangle$ are allowed to include components with any number of occupancies $\{N_1, \dots, N_{2N}\}$. In this way even the total number of electrons $\sum_\sigma N_\sigma$ is not a good quantum number for these wave functions. However, in the thermodynamic limit only expectation values are relevant and, thus, the canonical and the grand-canonical derivation become equivalent. In fact, a straightforward evaluation of the canonical wave functions leads to the same final results.

In our grand-canonical point of view we may consider the number distribution $\Omega(\widetilde{N})$ as a continuous function in the whole $2N$-dimensional space $\{N_1, \dots, N_{2N}\}$. As a macroscopic probability distribution, $\Omega(\widetilde{N})$ will have a sharp maximum for the expectation values

$$\overline{N}^0_\sigma \equiv \langle \Phi_0 | \hat{N}_\sigma | \Phi_0 \rangle. \tag{47}$$

At this position

$$\frac{\partial}{\partial N_\sigma} \Omega(\widetilde{N}) \bigg|_{\widetilde{N} = \widetilde{\overline{N}}^0} = 0, \tag{48a}$$

where we set $\widetilde{\overline{N}}^0 \equiv \{\overline{N}^0_1, \dots, \overline{N}^0_{2N}\}$. Moreover,

$$\frac{\Omega(\{\overline{N}^0_1 + \Delta N_1, \dots, \overline{N}^0_{2N} + \Delta N_{2N}\})}{\Omega(\widetilde{\overline{N}}^0)} \approx 1 \tag{48b}$$

is fulfilled in the thermodynamic limit if the numbers $\Delta N_1, \dots, \Delta N_{2N}$ are of the order of one. For the following calculation we do not need the explicit form of the function $\Omega(\widetilde{N})$ but only the simple relations (48).

Apart from equation (46) we need expressions for the products which stem from the evaluation of hopping contributions in (1). Again, we consider the respective uncorrelated expectation values

$$\begin{aligned} \langle \Phi_0 | \hat{c}^+_{i,\sigma} \hat{c}_{j,\sigma'} | \Phi_0 \rangle &= \sum_{\widetilde{\mathcal{N}}'} T^*(\widetilde{\mathcal{N}}'_{i,\sigma}) T(\widetilde{\mathcal{N}}'_{j,\sigma'}) \\ &= \sum_{\widetilde{N}} \Omega^{\sigma,\sigma'}_{i,j}(\widetilde{N}), \end{aligned} \tag{49a}$$

$$\begin{aligned} \Omega^{\sigma,\sigma'}_{i,j}(\widetilde{N}) \equiv & \sum_{\widetilde{\mathcal{N}}'} T^*(\widetilde{\mathcal{N}}'_{i,\sigma}) T(\widetilde{\mathcal{N}}'_{j,\sigma'}) \prod_{\sigma''} \delta_{|\mathcal{N}'_{\sigma''}|, N_{\sigma''}} \\ & (i \notin \mathcal{N}'_\sigma \,,\; j \notin \mathcal{N}'_{\sigma'}) \,. \end{aligned} \tag{49b}$$

Here, we introduced the notation

$$\widetilde{\mathcal{N}}'_{i(j),\sigma(\sigma')} \equiv \left\{ \mathcal{N}'_1, \dots, \mathcal{N}'_{\sigma(\sigma')} \cup i(j), \dots, \mathcal{N}'_{2N} \right\}. \tag{50}$$

$\Omega^{\sigma,\sigma'}_{i,j}(\widetilde{N})$ is the probability for the hopping transfer of an electron from the orbital σ' on site j into the orbital σ on site i under the restriction that the number distribution of the other electrons is given by $\widetilde{N} = \{N_1, \dots, N_{2N}\}$. We can surely write (49b) as

$$\Omega^{\sigma,\sigma'}_{i,j}(\widetilde{N}) = \omega^{\sigma,\sigma'}_{i,j}(\widetilde{N})\, \Omega(\widetilde{N}), \tag{51a}$$

where $\omega^{\sigma,\sigma'}_{i,j}(\widetilde{N})$ is a function of the order of unity. In particular, for the expectation values $\widetilde{\overline{N}}^0$ the following relation holds

$$\omega^{\sigma,\sigma'}_{i,j}(\widetilde{\overline{N}}^0) = \langle \Phi_0 | \hat{c}^+_{i,\sigma} \hat{c}_{j,\sigma'} | \Phi_0 \rangle. \tag{51b}$$

We use the fact that the resulting probability factors for both cases $\sigma' = \sigma$ and $\sigma \neq \sigma'$ in (49a) have the same form,

$$\begin{aligned} & P(L-1, N_\sigma) P(L-1, N_{\sigma'}) \\ & = \frac{L^2}{(L - N_\sigma)\,(L - N_{\sigma'})} P(L, N_\sigma) P(L, N_{\sigma'}), \end{aligned} \tag{52a}$$

$$P(L-2, N_\sigma) = \frac{L^2}{(L - N_\sigma)^2} P(L, N_\sigma). \tag{52b}$$

Then, the Gutzwiller approximation reads

$$\begin{aligned} & T^*(\widetilde{\mathcal{N}}'_{i,\sigma}) T(\widetilde{\mathcal{N}}'_{j,\sigma'}) \\ & = \Omega^{\sigma,\sigma'}_{i,j}(\widetilde{N}) \frac{L^2}{(L - N_\sigma)\,(L - N_{\sigma'})} \prod_{\sigma''} P(L, N_{\sigma''}). \end{aligned} \tag{53}$$

Equations (46, 53) define the Gutzwiller approximation for the multi-band case.

Now we are in the position to evaluate the correlated expectation values. First, equation (46) allows us to write the norm as

$$\langle \Psi | \Psi \rangle = \sum_{\widetilde{\mathcal{N}}} \left| T(\widetilde{\mathcal{N}}) \right|^2 \prod_\sigma \eta_\sigma^{2N_\sigma} \prod_{I(|I| \geq 2)} g_I^{2 Z_I^{\mathcal{L}}(\widetilde{\mathcal{N}})} \tag{54a}$$

$$\begin{aligned} & = \sum_{\widetilde{N}} \Omega(\widetilde{N}) \prod_\sigma P(L, N_\sigma) \eta_\sigma^{2N_\sigma} \\ & \times \sum_{\widetilde{\mathcal{N}}(|\mathcal{N}_\sigma| = N_\sigma)} \prod_{I(|I| \geq 2)} g_I^{2 Z_I^{\mathcal{L}}(\widetilde{\mathcal{N}})}, \end{aligned} \tag{54b}$$

where $Z_I^{\mathcal{X}}(\widetilde{\mathcal{N}})$ is the number of sites (included in $\mathcal{X}$) with a multiple-occupancy I, provided that the distribution of all electrons is given by $\widetilde{\mathcal{N}}$; in equations (54) we have $\mathcal{X} = \mathcal{L}$, the set of all lattice sites. The sum over $\widetilde{\mathcal{N}}$ in (54b) can be replaced by a sum over all possible multiple-occupancies $\widetilde{M} \equiv \{M_{12}, \dots, M_{1,\dots,2N}\}$,

$$\begin{aligned} \langle \Psi | \Psi \rangle = & \sum_{\widetilde{N}} \Omega(\widetilde{N}) \prod_\sigma P(L, N_\sigma) \eta_\sigma^{2N_\sigma} \\ & \times \sum_{\widetilde{M}} C_{\widetilde{M}}(L, \widetilde{N}) \prod_{I(|I| \geq 2)} g_I^{2M_I}, \end{aligned} \tag{55}$$

where we introduced the factor

$$C_{\widetilde{M}}(L, \widetilde{N}) \equiv L! \Big/ \Big(\prod_I M_I! \Big). \tag{56a}$$

This factor counts the number of possibilities to put $N_1, \ldots, N_{2N}$ electrons onto the lattice for a given number of sites with multiple-occupancies $(12), \ldots, (1 \ldots 2N)$. Here, the numbers of empty and single occupied sites are determined by the completeness relations

$$M_{\emptyset} = L - \sum_{\sigma} N_{\sigma} + \sum_{I(|I| \geq 2)} (|I|-1) M_I, \tag{56b}$$

$$M_{\sigma} = N_{\sigma} - \sum_{I(|I| \geq 2, \sigma \in I)} M_I. \tag{56c}$$

As in the one-band case, we will later see that all relevant expectation values have the same form as the norm, apart from factors of the order of unity in the respective sums. Hence, we may replace all these expressions by their maximum term.

As seen from (55) we have to find the maximum term with respect to $\widetilde{M}$ and $\widetilde{N}$. Let us first consider the sum over $\widetilde{M}$ in (55). The maximum-term condition for one of these multiple occupancies has the simple form of equations (35) if we use Stirling's formula and equations (56). Thus we directly find as a generalization of equations (37)

$$\begin{aligned} g_{12}^2 &= \frac{m_{\emptyset} m_{12}}{m_1 m_2}, \\ &\vdots \qquad\qquad\qquad\qquad (m_I \equiv \overline{M}_I / L) \\ g_{1\ldots 2N}^2 &= \frac{(m_{\emptyset})^{2N-1} m_{1\ldots 2N}}{m_1 \ldots m_{2N}}. \end{aligned} \tag{57}$$

The maximum-term conditions (57) also allow us to replace the original variational parameters g_I by their physical counterparts m_I.

The evaluation of the sum over the gross occupancies $\widetilde{N}$ in (55) would lead to considerable complications if we worked with the original one-particle wave function $|\Psi_0\rangle$. In this case ($\eta_{\sigma} = 1$) we would have to examine the explicit form of the probability distribution $\Omega(\widetilde{N})$ since, in general, the maximum term in the sum in (55) will not occur for the same set $\overline{\widetilde{N}}$ as for $\Omega(\widetilde{N})$, *i.e.*, for the set $\overline{\widetilde{N}}^0$. However, the parameters η_{σ} may be chosen at our convenience and, hence, we are allowed to demand the identity of these peak positions, *i.e.*,

$$\overline{N}_{\sigma} / L \equiv n_{\sigma} \overset{!}{=} n_{\sigma}^0 \equiv \overline{N}_{\sigma}^0 / L. \tag{58}$$

Then, the maximum term condition in (55) with respect to N_{σ} reads

$$0 = \frac{\partial}{\partial N_{\sigma}} \left(\Omega(\widetilde{N}) \eta_{\sigma}^{2N_{\sigma}} \frac{P(L, N_{\sigma})}{M_{\sigma}! M_{\emptyset}!} \right) \Bigg|_{\widetilde{N} = \overline{\widetilde{N}}^0} \tag{59a}$$

$$= \frac{\partial}{\partial N_{\sigma}} \left(\eta_{\sigma}^{2N_{\sigma}} \frac{(L-N_{\sigma})^{(L-N_{\sigma})} N_{\sigma}^{N_{\sigma}}}{M_{\sigma}^{M_{\sigma}} M_{\emptyset}^{M_{\emptyset}}} \right) \Bigg|_{\widetilde{N} = \overline{\widetilde{N}}^0}, \tag{59b}$$

where we used Stirling's formula and equation (48a). This problem has again the form of equation (35). Hence, equation (59b) fixes the parameters η_{σ} as a function of the gross occupancies n_{σ},

$$\eta_{\sigma}^2 = \frac{1-n_{\sigma}}{n_{\sigma}} \frac{m_{\sigma}}{m_{\emptyset}}. \tag{60}$$

In addition, the η_{σ} depend on the multiple-occupancies m_I *via* equations (56b, 56c) .

So far we did not really prove that the numbers $\overline{N}_{\sigma}$ and $\overline{M}_I$ which determine the maximum term in (55) are indeed the expectation values for the respective operators. However, this statement apparently holds, and it is an simple exercise to prove that the relations

$$\overline{M}_I = L m_I = \frac{\langle \Psi | \hat{M}_I | \Psi \rangle}{\langle \Psi | \Psi \rangle}, \tag{61a}$$

$$\overline{N}_{\sigma} = L n_{\sigma} = \frac{\langle \Psi | \hat{N}_{\sigma} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{61b}$$

are fulfilled within the Gutzwiller approximation scheme.

As a last step we have to evaluate the expectation value for an electron transfer. With the help of equation (53) we find

$$\begin{aligned} \langle \Psi | \hat{c}_{i,\sigma}^{+} \hat{c}_{j,\sigma'} | \Psi \rangle = \sum_{\widetilde{N}} H_{i,j}^{\sigma,\sigma'}(\widetilde{N})\, \Omega_{i,j}^{\sigma,\sigma'}(\widetilde{N})\, \eta_{\sigma} \eta_{\sigma'} \\ \times \frac{L^2}{(L-N_{\sigma})(L-N_{\sigma'})} \prod_{\sigma''} P(L, N_{\sigma''}) \eta_{\sigma''}^{2N_{\sigma''}}, \end{aligned} \tag{62a}$$

$$\begin{aligned} H_{i,j}^{\sigma,\sigma'}(\widetilde{N}) = \sum_{\substack{\widetilde{\mathcal{N}}' \\ (|\mathcal{N}_{\sigma}'| = N_{\sigma})}} \prod_{I\,(|I| \geq 2)} \exp\Big[\ln(g_I) \Big(2 Z_I^{\mathcal{L} \setminus \{i,j\}}(\widetilde{\mathcal{N}}') \\ + Z_I^{\{i,j\}}(\widetilde{\mathcal{N}}_{i,\sigma}') + Z_I^{\{i,j\}}(\widetilde{\mathcal{N}}_{j,\sigma'}') \Big) \Big]. \end{aligned} \tag{62b}$$

The sum over $\widetilde{\mathcal{N}}'$ in (62b) may again be replaced by a sum over the multiple occupancies $\widetilde{M}$. As in the one-band case we must take all possible occupancies of the sites i and j into account. Depending on these occupancies different arguments occur in the respective combinatorial factors of (56a). We set $g_I \equiv 1$ for $|I| \leq 1$ and use the abbreviation $\delta_{\sigma}^I \equiv |\sigma \cap I|$. Then, equation (62b) can be written as

$$\begin{aligned} H_{i,j}^{\sigma,\sigma'}(\widetilde{N}) = \sum_{\widetilde{M}} \left(\prod_{I} g_I^{2M_I} \right) \sum_{I_i (\sigma \notin I_i)} \sum_{I_j (\sigma' \notin I_j)} g_{I_i \cup \sigma} g_{I_i} g_{I_j \cup \sigma'} g_{I_j} \\ \times C_{\widetilde{M}} \Big(L-2, \big\{ N_1 - \delta_1^{I_i} - \delta_1^{I_j}, \ldots, N_{\sigma} - \delta_{\sigma}^{I_j}, \ldots, \\ N_{\sigma'} - \delta_{\sigma'}^{I_i}, \ldots, N_{2N} - \delta_{2N}^{I_i} - \delta_{2N}^{I_j} \big\} \Big). \end{aligned} \tag{63}$$

In the case $\sigma = \sigma'$ the terms $\delta_{\sigma}^{I_j}$ and $\delta_{\sigma'}^{I_i}$ obey $\delta_{\sigma}^{I_j} = \delta_{\sigma}^{I_i} = 0$ because of the summation restriction in (63). In the thermodynamic limit it follows that

$$
\begin{aligned}
&C_{\widetilde{M}}\Big(L-2,\{N_1-\delta_1^{I_i}-\delta_1^{I_j},\dots,N_\sigma-\delta_\sigma^{I_j},\dots\\
&\qquad N_{\sigma'}-\delta_{\sigma'}^{I_i},\dots,N_{2N}-\delta_{2N}^{I_i}-\delta_{2N}^{I_j}\}\Big)\\
&\qquad = C_{\widetilde{M}}(L,\widetilde{N})\left(\frac{M_\sigma}{M_\emptyset}\right)^{\delta_\sigma^{I_j}}\left(\frac{M_{\sigma'}}{M_\emptyset}\right)^{\delta_{\sigma'}^{I_i}}\frac{M_\emptyset^2}{L^2}\\
&\qquad\times\prod_{\sigma''\,(\neq\sigma,\sigma')}\left(\frac{M_{\sigma''}}{M_\emptyset}\right)^{\delta_{\sigma''}^{I_i}+\delta_{\sigma''}^{I_j}}.
\end{aligned}
\tag{64}
$$

Hence, equation (62a) becomes

$$
\begin{aligned}
\langle\Psi|\hat{c}_{i,\sigma}^{+}\hat{c}_{j,\sigma'}|\Psi\rangle &= \sum_{\widetilde{N}}\Omega(\widetilde{N})\prod_{\sigma''}P(L,N_{\sigma''})\eta_{\sigma''}^{2N_{\sigma''}}\\
&\times\sum_{\widetilde{M}}\prod_{I}g_I^{2M_I}C_{\widetilde{M}}(L,\widetilde{N})\,h_{i,j}^{\sigma,\sigma'}(\widetilde{N},\widetilde{M}),
\end{aligned}
\tag{65a}
$$

$$
\begin{aligned}
h_{i,j}^{\sigma,\sigma'}(\widetilde{N},\widetilde{M}) &= \omega_{i,j}^{\sigma,\sigma'}(\widetilde{N})\eta_\sigma\eta_{\sigma'}\frac{L^2}{(L-N_\sigma)\,(L-N_{\sigma'})}\\
&\times\sum_{I_i\,(\sigma\notin I_i)}\;\sum_{I_j\,(\sigma'\notin I_j)}g_{I_i\cup\sigma}g_{I_i}g_{I_j\cup\sigma'}g_{I_j}\\
&\times\left(\frac{M_\sigma}{M_\emptyset}\right)^{\delta_\sigma^{I_j}}\left(\frac{M_{\sigma'}}{M_\emptyset}\right)^{\delta_{\sigma'}^{I_i}}\frac{M_\emptyset^2}{L^2}\\
&\times\prod_{\sigma''\,(\neq\sigma,\sigma')}\left(\frac{M_{\sigma''}}{M_\emptyset}\right)^{\delta_{\sigma''}^{I_i}+\delta_{\sigma''}^{I_j}}.
\end{aligned}
\tag{65b}
$$

Apart from the factor (65b), which is of the order of unity, equation (65a) is identical to the norm (55). Thus, in the thermodynamic limit, we may replace the whole sum by its maximum term which has the same position as the norm (55). When we express the factors η_σ *via* equation (60) and use the relation

$$
\begin{aligned}
g_{I_i\cup\sigma}g_{I_i}g_{I_j\cup\sigma'}g_{I_j} &= \left(\frac{M_\emptyset}{M_\sigma}\right)^{\delta_\sigma^{I_j}}\left(\frac{M_\emptyset}{M_{\sigma'}}\right)^{\delta_{\sigma'}^{I_i}}\\
\times\prod_{\sigma''\,(\neq\sigma,\sigma')}&\left(\frac{M_\emptyset}{M_{\sigma''}}\right)^{\delta_{\sigma''}^{I_i}+\delta_{\sigma''}^{I_j}}\sqrt{\frac{M_{I_i\cup\sigma}M_{I_i}}{M_\emptyset M_\sigma}}\sqrt{\frac{M_{I_j\cup\sigma'}M_{I_j}}{M_\emptyset M_{\sigma'}}}
\end{aligned}
\tag{66}
$$

the expectation value for a hopping term finally becomes

$$
\frac{\langle\Psi|\hat{c}_{i,\sigma}^{+}\hat{c}_{j,\sigma'}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = q_{\sigma\sigma'}\langle\Phi_0|\hat{c}_{i,\sigma}^{+}\hat{c}_{j,\sigma'}|\Phi_0\rangle.
\tag{67}
$$

Here, the renormalization factors

$$
q_{\sigma\sigma} = \frac{1}{n_\sigma(1-n_\sigma)}\left[\sum_{I\,(\sigma\notin I)}\sqrt{m_{I\cup\sigma}m_I}\right]^2,
\tag{68a}
$$

$$
q_{\sigma\sigma'} = \sqrt{q_{\sigma\sigma}\,q_{\sigma'\sigma'}}
\tag{68b}
$$

are a straightforward generalization of equation (40). Note that the derivation of equations (62–68) also holds if we set $i=j$, $\sigma\neq\sigma'$, *i.e.*, our Hamiltonian (1) may eventually include a basis with more than one atom per site. Thus, in the Gutzwiller approximation the variational ground-state energy for the Hamiltonian (1) becomes

$$
\begin{aligned}
\frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} &= \sum_{i,j,\sigma,\sigma'}t_{i,j}^{\sigma,\sigma'}\sqrt{q_\sigma q_{\sigma'}}\langle\hat{c}_{i;\sigma}^{+}\hat{c}_{j;\sigma'}\rangle_0\\
&+L\sum_\sigma\epsilon_\sigma n_\sigma+L\sum_I U_I\,m_I.
\end{aligned}
\tag{69}
$$

In the first sum in (69) it is understood that either $i\neq j$ or $\sigma\neq\sigma'$. We again included the site-independent orbital on-site energies $\epsilon_\sigma\equiv t_{i,i}^{\sigma,\sigma}$ in our final expression. Note that the simple relations (58, 67) only hold for the one-particle product state $|\Phi_0\rangle$ and not for the original wave function $|\Psi_0\rangle$; compare the discussions in reference [14].

Another route to derive equation (69) is the slave-boson mean-field approximation for degenerate Hubbard models, which was independently proposed in references [22,23]. Applications of the final result (69) can be found in references [3,20], where the Mott transition and itinerant ferromagnetism in degenerate band systems were addressed; see also references [4,22–24]. Furthermore, antiferromagnetism in a degenerate two-band system at half band-filling was studied in reference [22].

Recently, we showed [21] that the energy expression (69) becomes *exact* for the Gutzwiller-correlated wave functions (14) in the limit of infinite dimensions. In addition, this approach allows to go beyond systems with translational invariance, *i.e.*, all kinds of symmetry breaking are covered by this general approach; see [21] for details and a critical discussion of previous generalizations of the Gutzwiller approximation to multi-band Hubbard models.

Thus far we neglected all terms in the Hamiltonian and, consequently, in the variational wave function which involve the local transfer of two electrons between different orbitals; see equation (2). When these important terms in the local interaction are taken into account, a new class of Gutzwiller-correlated wave functions must be introduced. Their exact evaluation in the limit of infinite dimensions was accomplished in [5]. This problem seems to be infeasible within the Gutzwiller approximation scheme as outlined above.

5 Summary

In this work we presented a detailed derivation of the Gutzwiller approximation for multi-band Hubbard models. In contrast to earlier approaches [11,12] we introduced an individual variational parameter for each possible orbital occupancy on a lattice site. Our mathematically well-defined formulation of the Gutzwiller approximation allowed us to include inter-orbital electron transfer terms. The systematic treatment of these terms posed serious problems for the phenomenological counting approaches [4,19].

First applications to a two-band model [3,20,22] showed that there are significant differences between the multi-band and the one-band case. The Brinkman–Rice metal-to-insulator transition is found to be discontinuous, and itinerant ferromagnetism is dominantly triggered by the intra-atomic Hund's-rule coupling. Therefore, we expect that the application of our method to real materials will provide valuable new insight into the physics of multi-band systems.

I thank W. Weber for helpful discussions, and F. Gebhard for a critical reading of the manuscript. I am very grateful to P. Nozières for an invitation to the Institut Max-von-Laue–Paul-Langevin (ILL), Grenoble, where this project was completed.

References

1. J. Hubbard, Proc. R. Soc. A **276**, 238 (1963); *ibid.* A **277**, 237 (1964); *ibid.* A **281**, 401 (1964).
2. F. Gebhard, *The Mott Metal–Insulator Transition – Models and Methods*, Springer Tracts in Modern Physics **137** (Springer, Berlin, 1997).
3. J. Bünemann, W. Weber, Physica B **230–232**, 412 (1997).
4. T. Okabe, J. Phys. Soc. Jpn **65**, 1056 (1996).
5. J. Bünemann, W. Weber, F. Gebhard, Phys. Rev. B **57**, 6896 (1998).
6. M.C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963); Phys. Rev. **134**, A923 (1964); *ibid.* **137**, A1726 (1965).
7. W. Metzner, D. Vollhardt, Phys. Rev. Lett. **59**, 121 (1987); Phys. Rev. B **37**, 7382 (1988); F. Gebhard, D. Vollhardt, Phys. Rev. Lett. **59**, 1472 (1987); Phys. Rev. B **38**, 6911 (1988).
8. W. Metzner, D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989); F. Gebhard, Phys. Rev. B **41**, 9452 (1990).
9. T. Ogawa, K. Kanda, T. Matsubara, Prog. Theor. Phys. **53**, 614 (1975).
10. F. Takano, M. Uchinami, Prog. Theor. Phys. **53**, 1267 (1975); J. Florencio, K.A. Chao, Phys. Rev. B **14**, 3121 (1976).
11. K.A. Chao, M.C. Gutzwiller, J. Appl. Phys. **42**, 1420 (1971).
12. K.A. Chao, Phys. Rev. B **4**, 4034 (1971); *ibid.* **8**, 1088 (1973); J. Phys. C **7**, 127 (1974).
13. D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984).
14. V.Z. Vulović, E. Abrahams, Phys. Rev. B **36**, 2614 (1987).
15. C.M. Varma, W. Weber, L.J. Randall, Phys. Rev. B **33**, 1015 (1985).
16. T.M. Rice, K. Ueda, Phys. Rev. Lett. **55**, 995, 2093 (E) (1985); Phys. Rev. B **34**, 6420 (1986).
17. B.H. Brandow, Phys. Rev. B **33**, 215 (1986).
18. P. Fazekas, Solid State Commun. **60**, 431 (1986).
19. T. Okabe, J. Phys. Soc. Jpn **66**, 2129 (1997).
20. J. Bünemann, W. Weber, Phys. Rev. B **55**, 4011 (1997).
21. J. Bünemann, F. Gebhard, W. Weber, J. Phys. Cond. Matt. **8**, 7343 (1997).
22. H. Hasegawa, J. Phys. Soc. Jpn **66**, 1391 (1997); Phys. Rev. B **56**, 1196 (1997).
23. R. Frésard, G. Kotliar, Phys. Rev. B **56**, 12909 (1997).
24. J.P. Lu, Phys. Rev. B **49**, 5687 (1994); Mod. Phys. B **10**, 3717 (1996).