Numerical optimization algorithm for rotationally invariant multi-orbital slave-boson method

Ya-Min Quan a,∗, Qing-wei Wang a, Da-Yong Liu a, Xiang-Long Yu a, Liang-Jian Zou a,b,∗

a Key Laboratory of Materials Physics, Institute of Solid State Physics, Chinese Academy of Sciences, Hefei 230031, China
b University of Science and Technology of China, Hefei 230026, China

A R T I C L E I N F O

Article history:
Received 8 December 2014
Received in revised form 3 February 2015
Accepted 4 February 2015
Available online 14 February 2015

Keywords:
Slave boson
Numerical optimization algorithm
Hubbard model
Metal–insulator transition

A B S T R A C T

We develop a generalized numerical optimization algorithm for the rotationally invariant multi-orbital slave boson approach, which is applicable for arbitrary boundary constraints of high-dimensional objective function by combining several classical optimization techniques. After constructing the calculation architecture of rotationally invariant multi-orbital slave boson model, we apply this optimization algorithm to find the stable ground state and magnetic configuration of two-orbital Hubbard models. The numerical results are consistent with available solutions, confirming the correctness and accuracy of our present algorithm. Furthermore, we utilize it to explore the effects of the transverse Hund’s coupling terms on metal–insulator transition, orbital selective Mott phase and magnetism. These results show the quick convergency and robust stable character of our algorithm in searching the optimized solution of strongly correlated electron systems.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Introducing auxiliary bosons to strongly correlated fermionic systems is an important method for many-body physics, such as the slave boson, the slave spin and slave rotor techniques [1–3]. Amongst these approaches, the Kotliar–Ruckenstein slave boson approach [1] is a very useful tool in dealing with correlated electron system. In this approach, the Hilbert space is enlarged to include fermionic and auxiliary slave bosonic degrees of freedom. The fermionic degree of freedom describes Landau quasiparticles, and the bosonic degrees describe the local states [4]. By setting a series of constraints, which select the physical states out of enlarged Hilbert space, the local physical configurations can be coherently described by the fermionic and bosonic degrees of freedom.

Slave-boson mean-field theory is mainly developed to calculate the quasiparticle (QP) weight of correlated electron systems, and a part of spin and orbital fluctuations is taken into account. The Fermi surfaces of the correlated systems are determined by the QP weight and Lagrange multipliers, which globally ensure the physical constraints. The slave-boson mean-field theory is able to eliminate local repulsive interaction through introducing local constraints, and projects a highly correlated system into an uncorrelated state [1]. Thus it is convenient to deal with complicated spin and orbital-ordered states [5–7], superconducting state [4] and spin liquid state [8]. The high-energy fluctuation processes could also be described by the auxiliary boson fields [9].

The single-orbital slave-boson functional-integral method introduced by Kotliar and Ruckenstein [1] was extended to two-orbital degenerate Hubbard model by Hasegawa et al. [5,10,11], hence could be used to investigate multi-orbital metal–insulator transitions. However, the multi-orbital Kotliar–Ruckenstein slave boson (KRSB) method is only suitable to handle with the model Hamiltonian with density–density interactions [1,12]. The rotationally invariant slave boson (RISB) method was proposed by Wölfle et al. [13,14] and generalized to multi-orbital case by Lechermann et al. [12,15–17]. In the complicated magnetic configurations or multiorbital systems, the number of slave bosons increases exponentially with the increase of the number of freedom degrees, the slave-boson mean-field approximation still costs much computation resource when searching for the ground state of multi-orbital correlated systems, even in the KRSB framework [18,19]. Meanwhile, when we further consider the contributions of the spin flip and pair hopping terms to reveal the roles of spin and orbital fluctuations, the RISB method should be applied, and we need to optimize a total-energy problem with many slave boson variables. In this case, the classical single optimization technique hardly finishes such a task. So it is desirable to develop a

∗ Corresponding authors.
E-mail addresses: ymquan@theory.issp.ac.cn (Y.-M. Quan), zou@theory.issp.ac.cn (L.-J. Zou).
http://dx.doi.org/10.1016/j.cpc.2015.02.003
0010-4655/© 2015 Elsevier B.V. All rights reserved.
more efficient and stable numerical algorithm for both the KRSB and RISB methods.

In this paper, we develop a numerical optimization algorithm by synthesizing the pattern search method [20], the gradient method [21] and the Rosenbrock technique [22]. This algorithm is applicable for arbitrary boundary constraint of the objective function. Then we employ it to search for the ground state of the two-orbital Hubbard model in the RISB framework, and discuss the effects of spin–flip and pair-hopping Hund's rule coupling terms on Mott metal–insulator transition and magnetic moments. The comparisons between our solutions and available results confirm that the present algorithm is accurate, efficient and stable for numerically optimizing the total energy of multi-orbital slave boson approach. The rest of this paper is organized as follows: firstly we describe the model Hamiltonian and theoretical approach in Section 2; then the numerical optimization algorithm and numerical method for two-orbital slave-boson approach are described in Section 3; in Section 4, the numerical accuracy and stability of our algorithm applied for two-orbital RISB model are analyzed; finally the concluding remarks are given in Section 5.

2. Rotationally invariant slave boson formulas of two-orbital Hubbard model

To describe the low-energy physical processes in strongly correlated electron systems, we usually adopt the multi-orbital Hubbard model Hamiltonian, which can be written in the following form:

\[ H = H_0 + H_1 \]

with

\[ H_0 = - \sum_{i,j,\alpha,\beta,\sigma} \left( \epsilon_{ij}^{\alpha,\beta} \hat{d}_{i\alpha}^{\dagger} \hat{d}_{j\beta} + h.c. \right) + \sum_{i,\alpha,\sigma} (\epsilon_{i\alpha} - \mu) n_{i\alpha\sigma} \]

and

\[ H_1 = U \sum_{\langle i,j \rangle} n_{i\alpha\downarrow} n_{j\beta\uparrow} + \sum_{i,\alpha,\sigma,\alpha'\beta'} (U' - j_s h_{\alpha\sigma\alpha'}^s) n_{i\alpha\sigma} n_{i\beta\sigma'} + j_F \sum_{i,\alpha\neq\beta} \left( d_{i\alpha\uparrow}^{\dagger} d_{i\beta\downarrow} + d_{i\beta\downarrow}^{\dagger} d_{i\beta\uparrow} + d_{i\alpha\uparrow}^{\dagger} d_{i\alpha\downarrow} + d_{i\downarrow}^{\dagger} d_{i\uparrow} \right) \]

where \( \hat{d}_{i\alpha}^{\dagger} \) creates an electron with the orbital index \( \alpha \) and spin \( \sigma \) at the lattice site \( i \), \( n_{i\alpha\sigma} \) denotes the corresponding occupation number operator, \( \epsilon_{ij}^{\alpha,\beta} \) is the energy level of the \( \alpha \) orbital, \( \mu \) is the chemical potential. The hopping integral between two orbitals \( \alpha \) and \( \beta \) is denoted by \( \epsilon_{ij}^{\alpha,\beta} \), and the intra-orbital (inter-orbital) Coulomb repulsion is \( U (U') \). The density–density interaction term, spin–flip term and pair-hopping term of the Hund’s rule coupling are denoted by \( j_s \), \( j_F \) and \( j_Z \) respectively. Throughout this paper we set \( U = U' = 2j_s \) and \( j_F = 0.25U \), and the two orbitals are degenerate, i.e. \( \epsilon_1 = \epsilon_2 \).

Next, we project the original Hamiltonian (1) into the RISB representation. In the RISB framework (see the Appendix), sixteen Fock states in two-orbital case are shown in Fig. 1, which have been used for the two-orbital KRSB model [23], the corresponding slave bosons are plotted. It is obviously that the electron occupations associated with physical space are identical to those associated with QP space in KRSB method. We chose these Fock states as the bases of the physical and QP spaces. The matrix that is made up of 256 slave bosons is shown in Fig. 2. The local atomic Fock bases in physical and QP spaces are also plotted at the heads of each row and line. The same sequences of the Fock bases in the physical and QP spaces are adopted, thus the KRSB fields only appear in the diagonal. The off-diagonal slave bosons are classified into six groups, according to their described physics as shown Fig. 2. For example, the off-diagonal slave bosons in the SR region describe the physics with regard to the spin-rotation symmetry. The off-diagonal slave bosons between different lattices are neglected since the interaction between different lattices is not considered in the present two-orbital Hubbard model. As shown in the Appendix, in the rotational-invariant saddle-point approximation, the complicated model Hamiltonian given by Eqs. (1) and (2) reduces to an optimization problem of total energy functional, see Eqs. (A.12) and (A.13).

In the saddle-point approximation, the total energy of two-orbital Hubbard model is given by:

\[ E_{\text{tot}} = E_{\text{kin}} + E_{\text{loc}} \]

with

\[ E_{\text{kin}} = \left( \begin{array}{c} R_A^2 \\ R_B^2 \end{array} \right) \epsilon(k) \begin{pmatrix} R_A & 0 \\ 0 & R_B \end{pmatrix} + \begin{pmatrix} A_A & 0 \\ 0 & A_B \end{pmatrix} \left( \begin{array}{c} 0 \\ -\mu \end{array} \right) \]

and

\[ E_{\text{loc}} = \text{tr} \left[ \Phi_A \begin{pmatrix} \Phi_B & 0 \\ 0 & \Phi_B \end{pmatrix} \begin{pmatrix} 0 \\ \Phi_B \end{pmatrix} \right] + U\lambda \sum_{\alpha \lambda m} |\phi_{\alpha \lambda m}|^2 - 2 \]

\[-\sum_{\alpha \beta m} \left( A_A \begin{pmatrix} 0 \\ \alpha \beta \end{pmatrix} \phi_{\alpha \lambda m} \begin{pmatrix} 0 \\ \alpha \beta \end{pmatrix} \begin{pmatrix} 0 \\ n_{\alpha \beta} \end{pmatrix} \right)_{\alpha \beta m} \]

Here \( \epsilon(k) \) is the energy dispersion matrix of the system. \( R_A \) and \( R_B \) in Eq. (5) are the renormalization factor matrices of lattices A and B defined by Eq. (A.5), which is written in a more compact matrix form as follows:

\[ R_A = C_{\alpha} M_{\alpha} \]

with

\[ C_{\alpha} = \sum_{\phi} \left( D_{\alpha} \circ (\Phi_{\alpha} \Phi_{\alpha}) \right)_{\phi} \]

where \( \Phi_{\alpha} \) is the slave boson matrix, and the symbol ‘\( \circ \)’ denotes Hadamard product. The index \( \alpha \) denotes sublattice A or sublattice B.

Fig. 1. The atomic spin and orbital configurations of two-orbital system and the corresponding slave bosons. Here \( e, p, b, d, t, q \) denote states from empty to four occupations in the corresponding graphs. The first group numbers in the indexes of slave bosons, for example \( e_0(00,00), e_0(00,00), \), describe the occupations on electron species in the sequence of (spin up of orbital 1, spin down of orbital 1, spin up of orbital 2 and spin down of orbital 2) in the physical Fock state, the second group numbers in the indexes have the similar meaning in the QP Fock state, where 0 denotes empty and 1 is occupied.
The expressions of matrices $D_a$ and $F_a$ read
\[ D_{aL} = |A| f_{aL}^f |B \]  
and
\[ F_{amn} = (m|f_{am}^f |n), \]  
respectively. Here $|A\rangle$ and $|B\rangle$ stand for the physical Fock basis states. $|m\rangle$ and $|n\rangle$ are the QP Fock basis states. $f_{am}^f$ is the auxiliary fermion creation operator. The matrix $M_a = (\sqrt{A_a})^{-1}$, where the unitary matrix $A_a = \frac{1}{2}[\hat{\Delta}^{(p)}_a \hat{\Delta}^{(h)}_a + \hat{\Delta}^{(h)}_a \hat{\Delta}^{(p)}_a]$ and $\sqrt{A_a} = \mathbb{P}_a^a(A_a)^{1/2}.\mathbb{P}_a$. The matrix $\hat{\Phi}_a$ transforms $A_a$ into diagonal eigenvalue matrix $\hat{\Phi}_a$.

The QP density matrices are expressed in the form of
\[ \hat{\Delta}^{(p)}_{amn} |\Phi_a\rangle = tr[\hat{\Phi}_a^\dagger K^{(p)}_a |\Phi_a\rangle] \]  
and
\[ \hat{\Delta}^{(h)}_{amn} |\Phi_a\rangle = tr[\hat{\Phi}_a^\dagger K^{(h)}_a |\Phi_a\rangle]. \]

where
\[ K^{(p)}_{amn} = (m|f_{am}^f |n) \]  
\[ K^{(h)}_{amn} = (m|f_{am}^h |n). \]

Here $p$ and $h$ stand for the QP and quasihole situations, respectively. Correspondingly, the renormalization factors are expressed by the slave boson matrix in Eqs. (7) to (14) so as to simplify the numerical calculation. $A_a$ and $B_a$ in Eq. (5) are the Lagrange multipliers, which are defined in Eq. (15) (see Box I).

$E_{\text{subl}}$ and $E_{\text{qoh}}$ in Eq. (5) are the matrices of the sublattice energy levels; $E^{(b)}_A$, and $E^{(b)}_B$ in Eq. (6) are the local Hamiltonian matrices on sublattices $A$ and $B$, respectively. The elements of the matrices $E^{(b)}_A$ and $E^{(b)}_B$ read
\[ E^{(b)}_{\text{subl}} = \sum_{\alpha} n_{\alpha a} \left[ U[A] \sum_{\alpha} n_{\alpha a} |B\rangle + U[A] \sum_{\alpha} n_{\alpha a} |B\rangle \right] 
+ \sum_{\alpha} n_{\alpha a} |B\rangle \]  
\[ + \sum_{\alpha} \langle\alpha > \beta \rangle \sum_{\alpha} d_{\alpha a}^\dagger d_{\alpha b}^\dagger d_{\alpha a} d_{\alpha b} |B\rangle \]  
\[ + \sum_{\alpha} \langle\alpha > \beta \rangle \sum_{\alpha} d_{\alpha a}^\dagger d_{\alpha b}^\dagger d_{\alpha a} d_{\alpha b} |B\rangle \]  
\[ (16) \]
\[ n_a |\Phi_a\rangle = \hat{\Delta}^{(p)}_a |\Phi_a\rangle. \]  
\[ (17) \]
To guarantee the QP space is identical to the physical space, 32 fermion number constraints in sublattices A and B are enforced by 32 Lagrange multipliers, such as η_{al1}, η_{al2}, etc. When the off-diagonal Lagrange multiplier vanishes, the values of corresponding off-diagonal slave bosons shown in Table 1 also become zero.

In the antiferromagnetic situation, we consider the symmetric relationship between the slave bosons of two sublattices, for example, the values of slave bosons of two sublattices with reverse spin are the same. This could greatly save computational costs. However, this symmetry relationship becomes complicated in the rotationally invariant case, since the off-diagonal slave bosons shown in Fig. 2 are involved. Thus we should derive it through gauge symmetry. Considering that an SU(N)-symmetric matrix U rotates sublattice A to sublattice B, where N is twice the number of orbital, the transformation of the auxiliary fermion creation operators is as follows

\[ f_{l\alpha}^\dagger = \sum_\rho U_{\rho\beta} f_{\beta\alpha}^\dagger. \]  

According to Lechermann’s RISB theory [12], the transformation of the Fock states from lattice A to lattice B reads

\[ |m\rangle = \mathcal{U}(U)_{mm'}|m'\rangle, \]  

where \(|m'\rangle\) and \(|m\rangle\) stand for the QP Fock basis states of lattice A and lattice B, respectively. \(U\) and \(\mathcal{U}\) are the representation of SU(N) group elements in the QP Hilbert space and QP Fock space, respectively. We also transform the physical basis as that in the QP Hilbert space. The transformation of physical Fock bases from lattice A to lattice B is thus read

\[ \mathcal{U}(U)_{mm'} = \langle m'|m\rangle = \langle \bar{A}|\bar{A}\rangle, \]  

where \(\bar{A}\) and \(\bar{A}\) denote the physical Fock basis states of lattice A and lattice B, respectively. To enforce the gauge invariance of the Hamiltonian, the renormalization factors and electronic operators, the slave boson matrix should be transformed as follows

\[ \Phi_{\beta} = U(U)\Phi_{\alpha}U(U)^\dagger \]  

with

\[ \mathcal{U}(U)_{mm'} = \mathcal{U}(U)_{\beta\alpha}^\dagger. \]  

For a unitary rotation operator \(U\) of the QP operators and Hamiltonian, the elements of matrix \(\mathcal{U}(U)_{mm'}\) are expressed as

\[ \mathcal{U}(U)_{mm'} = \sum_\rho \left( U_{\rho\beta_1} U_{\beta_2 \rho} \cdots U_{\beta_N \rho} \right) (-1)^{\delta_{\beta_1 \beta_2} \cdots \delta_{\beta_N}} \times \left( \prod_{j=1}^N \delta_{\beta_j m_j', \beta_j m_j} \right). \]  

where \(P\) is a permutation of \((\beta_1, \beta_2, \ldots, \beta_N)\); \(m_j\) is the number of the \(j\)th species electrons of the Fock state \(|m\rangle\) and \(m_j'\) is the number of the \(\beta_j\)th species electrons of the Fock state \(|m'\rangle\). The \(\delta\) function term \(\delta_{\beta_j m_j, \beta_j m_j'}\) in Eq. (23) mean that only the terms occupied by electron are considered. Eq. (23) is expressed in the form of a determinant of \(\mathcal{U}(U)_{mm'} = \det(G(U))\), where

\[ G(U) = \begin{pmatrix} U_{11} & \cdots & U_{1N} \\ \vdots & \ddots & \vdots \\ U_{N1} & \cdots & U_{NN} \end{pmatrix}. \]  

When \(m_j = 0\) (\(m_j' = 0\), we get \(\delta_{\beta_j m_j, \beta_j m_j'} = 0\). Thus the i row (j column) could be removed from the matrix \(G(U)\) since their elements all vanish. Notice that \(\det(G(U)) = 0\) if \(G(U)\) is not a square matrix. The transformation of the Lagrange multiplier matrices of lattices A and B is given by \(A_{\beta} = U A_{\beta} U^\dagger\). With these preparations we transform a quantum many-body problem into a numerical optimization problem of a total-energy function, and will show how to find the optimized value of the total-energy function with many physical and boundary constraints.

3. Optimization algorithm and numerical method

3.1. Optimization algorithm based on pattern search method

We aim at the minimization of the total energy defined by Eqs. (5) and (6) with the physical constraints given by Eqs. (A.10) and (A.11). Our numerical algorithm is mainly based on the pattern search method [20], which is a local direct-search optimization algorithm that does not require the analytical gradient of the objective function. To overcome the shortages of the pattern search method, we also combine it with the gradient method [21] and the Rosenbrock technique [22].

Pattern search (PS) algorithm starts with an initial point \(x_0\) to evaluate the objective function. After the initialization, the PS
algorithm searches along the coordinate axis around the optimized point \( \bar{x}_k \) with step length \( \delta \) shown in Fig. 3. If a new detection point \( \bar{y}_k \) is found with less value compared with the initial point, then we move the initial point to the detection point \( \bar{x}_{k+1} = \bar{y}_k \) where \( \bar{x}_{k+1} \) is the new optimized points for next searching step. If the point \( \bar{y}_k \) with less value of objective function is not found, the searching step length is shortened and a new search for a point towards optimality continues. The algorithm will stop if the tolerance constraints are met or a maximum number of iterations is exceeded.

From the PS algorithm, we find that the convergency becomes worse when the optimized point \( \bar{x}_k \) moves along the valley of the objective function near the vicinity of the optimized solution, as seen in Fig. 3. Therefore, we combine gradient method into the PS algorithm. The minus gradient direction is calculated by the difference method and the optimized point \( \bar{x}_k \) moves one step along this direction in one iteration. However, the detection point \( \bar{y}_k \) may not be found in the vicinity of the solution since the detected directions are only along the global coordinate, as shown in Fig. 3. To overcome this problem, we further combine Rosenbrock method into the PS algorithm, which was proposed by Rosenbrock firstly [22]. With this method, a new local coordinate system for the next step is established according to the present point \( \bar{x}_k \) and the last step point \( \bar{x}_{k-1} \) shown in Fig. 3. By combining all these algorithm methods, we list the steps of the present optimization method as follows:

**Step 1.** Start with an initial solution called Present Point and calculate the objective function \( F(\bar{x}) \).

**Step 2.** Establish the local coordinate system according to the point \( \bar{x}_{k-1} \) and point \( \bar{x}_k \) by using Schmidt orthogonalization. Here it should be pointed out that the first local coordinate axis should be in the direction from \( \bar{x}_{k-1} \) to \( \bar{x}_k \).

**Step 3.** Calculate the gradient \( \nabla F(\bar{x}_k) \) by the difference method. If the detection point \( \bar{y}_k \) satisfied \( F(\bar{y}_k) < F(\bar{x}_k) \) is not found, then the gradient is zero.

**Step 4.** Calculate the new detection point \( \bar{y}_{k+1} = \bar{x}_k - \delta \nabla F(\bar{x}_k) \).

**Step 5.** Continue the optimization unless some termination criterion is met.

**Step 6.** If \( F(\bar{y}_{k+1}) < F(\bar{x}_k) \), set \( \bar{x}_{k+1} = \bar{y}_{k+1} \). Otherwise, shorten the step length \( \delta \). Go to step 2.

**Step 7.** Set \( \bar{x}_{k+1} = \bar{x}_k + \xi (\bar{x}_{k+1} - \bar{x}_k) \), where \( \xi \) is the accelerating factor. Go to step 2.

**End.**

The algorithms in Step 3 and Step 4 show that the detection point and optimized point should move along the minus gradient direction, though the accelerating algorithm in Step 7 deviates the optimized point from this direction. We find that the path of the optimized point may not be seriously changed and the convergence can still complete with the accelerating factor \( \xi < 1 \) [18].

The PS algorithm is a traditional optimization method for unconstrained optimization problem. However, the realistic optimization problem usually suffers from strong boundary constrains. Thus the present modified PS algorithm should be generalized to include the boundary constraint condition. A lot of theories have been proposed for derivative-free bound constrained optimization algorithm [24,25]. In this paper, we suggest a more adaptive technique, as addressed in what follows. Considering a detection point \( \bar{y}_k \) surrounded by equipotential surface and boundary shown in Fig. 4, the detection point \( \bar{y}_k \) on the boundary moves one step length inward the boundary on the equipotential surface. Then the optimization procedure within the boundary continues as unconstrained case. Since the function value does not increase when the detection point moves on the equipotential surface, the objected function can be optimized on the boundary surface. The steps of the boundary constrained optimization algorithm are listed as follows:

**Step 1.** Optimize the objective function in the feasible region.

**Step 2.** When the detection point \( \bar{y}_{k+1} \) shifts on the boundary, calculate the directions \( \bar{e}_i \) inward the restrained boundary along each local coordinate axis that moves inward the equipotential surface.

**Step 3.** Calculate the unit vector \( \bar{e} = \sum_{i \in j=1} \bar{e}_i \), where \( \bar{x}_{k+1} - \bar{x}_k \neq 0 \). Move the detection point to \( \bar{y}_{k+1} = \bar{y}_{k+1} + \delta \bar{e} \), where \( \delta \) is the present step length.

**Step 4.** Set \( \bar{x}_{k+1} = \bar{y}_{k+1} \). Go to Step 1

**End.**

In the present algorithm, the optimized point \( \bar{x}_k \) shifts on the boundary surface, as shown in Fig. 4.

The brief flowchart of the PS algorithm is shown in Fig. 5. Firstly from Step 1 to Step 3 of the local coordinate systems and the gradient of the objective function are calculated in turn with the initial values or the last cycle values. In Step 4, we calculate a new detection point, and evaluate a lot of convergent conditions, such as the minima step length and the maximum number of iteration, in the Step 5 and Step 6. If the detection point with less value is found on the boundary, the boundary constrained optimization algorithm is applied in Step 7. It is obviously that the accelerator algorithm in Step 7 only takes effect in the feasible region. When the constraint


\[ \frac{\partial E_{gs}}{\partial \phi_{\alpha m}} = \sum_{a \neq \alpha m} \frac{\partial \epsilon_a(k)}{\partial \phi_{\alpha m}} + \frac{\partial E_{loc}}{\partial \phi_{\alpha m}} = 0, \]  

(27)

where

\[ \frac{\partial E_{loc}}{\partial \phi_{\alpha m}} = \text{tr}\left[ \left( \Phi_\alpha^0 \Phi_\beta^0 \right) \left( E_\beta^{(b)} \right) \left( \Phi_\alpha^0 \right) \right] + U \lambda \sum_{a \in \epsilon} \frac{\partial \phi_{\epsilon a}}{\partial \phi_{\alpha m}} \]  

(28)

\[ \frac{\partial E_{gs}}{\partial \eta_{\alpha \beta}} = \sum_{a \neq \alpha m} \frac{\partial \epsilon_a(k)}{\partial \eta_{\alpha \beta}} + \frac{\partial E_{gs}}{\partial \eta_{\alpha \beta}} = 0. \]  

(29)

and

\[ \frac{\partial C_a}{\partial \phi_{\alpha m}} = \sum_{mn} |D_a \circ (\Psi_\alpha^0 \phi_{\epsilon a} + \Phi_\alpha^0 \phi_{\epsilon a})|_{mn}, \]  

(30)

where \( \phi_{\epsilon a} \) is the element of the slave boson matrix \( \Phi_\epsilon \). In other words, it is the slave boson associated to the pair of atomic multiplet \( |\alpha \rangle \) and QP Fock state \( |m\rangle \). The partial derivatives \( \partial \Phi_\alpha^0 / \partial \phi_{\epsilon a} \) and \( \partial \phi_{\epsilon a} / \partial \Phi_\alpha^0 \) have been given by Lechermann et al. in Ref. [12].

It should be pointed out that the eigenvectors used for calculating the derivatives of the degenerate eigenvalues must be carefully selected since the eigenvectors are not uniquely determined. In this case, the eigenvalue derivatives should be performed in accordance with Van der AA’s theory [26]. The objective function for optimization is defined as

\[ F(\Phi_\alpha, \Phi_\beta, \Lambda_\alpha, \Lambda_\beta) = \sum_{a \neq m} \left( \frac{\partial E_{gs}}{\partial \phi_{\alpha m}} \right)^2 + \sum_{a \neq \beta} \left( \frac{\partial E_{gs}}{\partial \eta_{\alpha \beta}} \right)^2 + \left( \frac{\partial E_{gs}}{\partial \delta} \right)^2. \]  

(31)

It is obviously that the minimum of this objective function should be zero once the self-consistent conditions are satisfied.

3.3. Finding the initial value

A proper choice of the initial value could greatly accelerate the optimization to self-consistent solutions in complicated charge, magnetic and orbital configurations of strongly correlated systems. In this paper, the initial values are searched through optimizing the total energy directly and the fermion number constraints are handled by employing generalized Lagrange multiplier method. In this case, the objective function of the system is defined as

\[ F(\Phi_\alpha, \Phi_\beta, \Lambda_\alpha, \Lambda_\beta) = E_{gs} + \sigma G. \]  

(32)

where \( \sigma \) is the penalty factor and \( G \) is the penalty function, which is defined as follows

\[ G = \sum_{a \neq \alpha m} \left( \sum_k (f_{\alpha a k}^* f_{\alpha k}^0) - n_{\alpha \beta} \right)^2 + \sum_{a \neq \alpha m} \left| \phi_{\alpha m}^0 \right|^2 - 2. \]  

(33)

It is obviously that the fermion number constraints are satisfied gradually with increasing penalty factor in the process of optimization. Meanwhile, the normalization constraint given by Eq. (A.10) should be held at all time. In this case, the empty occupation slave

Fig. 5. Brief flowchart of the optimization procedure. \( N_{\text{inner}} \) is the difference between the present and last optimized points of an outer cycle. \( N_{\text{inner}} \) and \( N_{\text{outer}} \) are the corresponding maximum numbers of the inner and outer iterations.

\[ \delta < \delta_{\text{min}} \text{ is satisfied in an inner cycle, the optimization process is} \]  

repeated once again until the constraint \( \Delta_{\text{inner}} < \Delta_{\text{min}} \) is met or the maximum number of iteration \( N_{\text{outer}} \) is exceeded. In this paper, we set the minimum step length \( \delta \) and the maximum number of iteration \( N_{\text{inner}} \) as \( 1.0 \times 10^{-8} \) and 20000, respectively; in obtaining a typical ground state calculation of the two-orbital Hubbard model with the rectangle DOS, it will cost less than 10 min.

\[ 3.2. \text{Solving the saddle-point equations} \]

In the next we apply the modified pattern search method algorithm to find the minimum energy of the system defined by Eqs. (5)–(17), which is a function of the slave boson fields and order parameters. For convenience, the real and imaginary parts of the parameters are set as mutual independent variables, as shown in Table 1. In the saddle-point approximation, the self-consistent equations can be written as follows:

\[ \frac{\partial E_{gs}}{\partial \phi_{\alpha m}} = \sum_{a \neq \alpha m} \frac{\partial \epsilon_a(k)}{\partial \phi_{\alpha m}} + \frac{\partial E_{loc}}{\partial \phi_{\alpha m}} = 0, \]  

(25)

\[ \frac{\partial E_{gs}}{\partial \eta_{\alpha \beta}} = \sum_{k} \left( f_{\alpha a k}^* f_{\alpha k}^0 - n_{\alpha \beta} \right) U = 0 \]  

(26)
Hund’s coupling term can be divided into longitudinal (the Ising-type \( z \)-component) and transverse (the spin–flip and orbital pair-hopping) terms, and most of previous literatures only considered the contribution of Ising-type term. The magnetic phase diagram and metal-insulating transition of multi-orbital systems strongly depend on the Hund’s rule coupling [27–29]. To explore the effects of spin–flip and pair-hopping terms on the Mott transition and antiferromagnetic ground state, we discuss the dependence of the QP weight and magnetic moments on different Hund’s coupling terms. Throughout this paper, we assume a rectangular density of states (DOS) for the two-orbital Hubbard model, and set the half bandwidth of orbitals 1 and 2 be \( D_1 \) and \( D_2 \), respectively.

Fig. 7 displays the influences of spin–flip and pair-hopping Hund’s coupling terms on Mott metal-insulating transition. For the KRSB case \( J_x = J_y = 0 \), two orbitals become insulating at about \( U/D_1 = 3.5 \), which is in consistent with the numerical results given by Ruegg et al. [23]. Meanwhile, the OSMP region seriously narrows when the spin–flip term becomes zero, which is also in agreement with Refs. [23,30]. This could be explained as follows: the Ising-type Hund’s term of \( J_z \) favors the high-spin state and drives electrons to distribute in different orbitals with the same spin orientation. Thus the spin and orbital fluctuations are suppressed [31], and the two orbitals tend to become insulating at the same critical \( U_z \) [23], suppressing the OSMP region.

Switching on finite spin–flip term of the Hund’s rule coupling causes considerable spin and orbital fluctuations, and the itinerancy of wide-orbital electrons is enhanced. Thus the OSMP region enlarges in the presence of finite spin–flip term [32]. This mechanism is also supported by the analysis of the slave boson occupations shown in Fig. 8(a). We find that the off-diagonal slave boson \( dd_{x(1001,0110)} \), which is in relationship with the spin–flip term, is considerable at finite \( J_{0z} \). Since this term contributes to the spin fluctuations, the significant \( dd_{x(1001,0101)} \), shows that the spin–flip term enhances spin fluctuation. The value of triplet occupation \( tt_{1011,1011} \) is also finite in the OSMP region, as seen in Fig. 8, which mainly contributes charge fluctuations of the wide orbital. The finite value of \( tt_{1011,1011} \) illustrates that the wide orbital maintains metallic in the OSMP. In the Mott insulating region, the double occupation boson \( dd_{x(1001,0101)} \) is also considerable, as seen in Fig. 8(a). This implies that the ground state of the system in the Mott insulating phase is a superposition of high-spin and low-spin states. Therefore the magnitude of local spin significantly decreases by the spin–flip term in the Mott insulating region.

In order to reveal the effect of pair-hopping term on the OSMP, we plot the QP weights with \( J_x = J_y = 0 \) and \( J_0 = 0 \) for comparison. From Fig. 7(a), we find that the QP weight of the narrow orbital slightly increases with finite pair-hopping term in the paramagnetic metal region. In the OSMP and Mott insulating phase, the effect of pair-hopping term can be neglected. Note that the double occupations \( bb_{x(1100,0011)} \) and \( bb_{x(1100,1100)} \) only show considerable values in the paramagnetic metal region, as shown in Fig. 8(a), thus the pair-hopping between the two orbitals is suppressed absolutely by electron correlation in the OSMP and Mott insulating phases. That is the reason why there is no considerable effect of pair-hopping term on the QP weight in the OSMP and Mott insulating phase. All above results demonstrate that the OSMP region of two-orbital model strongly depends on the spin–flip Hund’s rule coupling term, since it enhances the itinerancy of the electrons in wide orbital. However, the situation of three-orbital model is very different, where the Ising-type Hund’s rule coupling term and crystal field splitting become the key factors for OSMP at integer filling [27,28,30,31].

Away from half filling, the critical value \( U_z \) of the Mott transition of the narrow orbital is not seriously affected by the spin–flip and pair-hopping terms. The off-diagonal slave boson \( bb_{x(1100,0011)} \) becomes zero when the narrow orbital becomes insulating, as shown.
Fig. 7. Correlation dependence of the QP weight for different $J_X$ and $J_P$ with $n = 2$ (a) and $n = 1.8$ (b). Other theoretical parameter $D_2 = 2D_1$.

Fig. 8. The slave boson occupations for (a) $n = 2$ and (b) $n = 1.8$ with $J_P = J_X = J_Z$. The half bandwidth $D_2 = 2D_1$ is adopted.

in Figs. 7(b) and 8(b). This is the same as the results at half filling. Fig. 7(b) also shows that the QP weights with $J_X = J_Z$ and $J_X = 0$ crossed with each other at $U/D_1 = 3.5$. These results could be addressed as follows: the spin–flip term displays different behaviors in different correlation regions; in the weak and intermediate correlation regions, the spin–flip term mainly suppresses the orbital fluctuation [32], which narrows the bandwidth; however in the strongly correlated region, the spin–flip term enhances the spin fluctuations. Thus the bandwidth of the wider orbital increases as seen in Fig. 7(b).

Furthermore, we also present the influences of spin–flip and pair-hopping terms on the sublattice magnetic moments in the antiferromagnetic phase. The magnetic moments as a function of electron correlation and particle number fillings are plotted in Fig. 9, where $\Delta m_1$ is the difference of sublattice moments between the cases with $J_X = J_P = 0$ and $J_X = J_P = J_Z$ and $\Delta m_2$ is the difference between the cases with $J_X = J_P = 0$ and $J_X = J_Z$, $J_P = 0$. From Fig. 9(a), we find that the difference reach the maximum at $U/D_1 = 1$. This arises from the fact that the orbital and spin fluctuations are considerable in the intermediate correlation region. By comparing $\Delta m_1$ with $\Delta m_2$, we find that the effect of the pair-hopping term on the magnetic moment is comparable to that induced by spin–flip term, which originates from that the orbital distributions and sublattice magnetic moments closely interact with each other.

On the other hand, in the strong correlation region, the electron is almost localized and sublattice spin is almost polarized fully, thus the influences of the spin–flip and pair-hopping terms on magnetic moment are negligible at half filling. Away from half-filling, the difference between sublattice magnetic moments for $J_X = J_Z$ and $J_X = 0$ increases with the deviation of filling number, as shown in Fig. 9(b). The reason is that the formation of local magnetic moments crucially depends on the Hund’s rule coupling when the electron filling factor is away from half filling. The sublattice magnetic moments for finite pair-hopping term with $J_X = J_P = J_Z$ shown in Fig. 9(b) demonstrate that the pair-hopping term reduces the magnetic moment. In strongly electron correlation regime such a reduction in magnetic moment alleviates.

5. Conclusion

In this paper, we notice that in the case of many orbitals at lower symmetry, most of the off-diagonal elements of the slave boson fields should be considered, the variational parameters increase so
fast that it costs too much computational resources, and the realistic computation becomes impossible. To cope with the largely growing number of the slave boson variables, we could use the projection method to excluding the high-energy processes, as done in many other numerical methods for strongly correlated electron systems [1]. Thus a numerous of slave boson variables are set to zero due to its vanishing probabilities, only those slave boson variables relevant to low-energy processes are kept; thus, further studies for the realistic systems with many orbitals at lower symmetry will become plausible.

In summary, we develop a modified pattern search algorithm by combining several classical optimization methods. This algorithm is applicable for arbitrary boundary constraint of the objective functions, and could save great computational resources when used to optimize high-dimensional objective functions, in comparison with any one of these optimization methods. As a concrete application, we apply this algorithm on the rotationally invariant slave boson theory for optimizing the total energy and self-consistent equations of the two-orbital Hubbard models. Our results are in agreement with the literature available, verifying the validity and accuracy of this numerical algorithm. We further utilize it to reveal the influences of the spin-flipping and pair-hopping Hund’s coupling terms on orbital selective Mott phase and magnetic moments of two-orbital systems, and show that the orbital selective Mott phase region significantly enlarges in the presence of the spin–flip Hund’s coupling term, and the sublattice magnetic moments are considerably reduced by this spin–flip Hund’s coupling term in the cases of away from half filling. We expect that our numerical algorithm could also be applicable for other similar theories, such as the slave spin and the slave rotor approaches.

Acknowledgments

This work was supported by the National Science Foundation of China under Grant no. 11204311, 11274310 and 2012FXZY004. Numerical calculations were performed at the Center for Computational Science of CASHIPS.

Appendix. General rotationally invariant slave-boson formulas

In this Appendix, we describe the multi-orbital Hubbard model in rotationally invariant slave boson representation [12]. The generic multiorbital Hubbard model is written in the form of

\[ H = \sum_k \sum_{\alpha\beta} \varepsilon_{\alpha\beta}(k) d_{\alpha\beta}^\dagger d_{\alpha\beta} + \sum_i H_{\text{loc}}[f] \]  
\[ H_{\text{loc}} = \sum_{\alpha\beta} \varepsilon_{\alpha\beta} d_{\alpha\beta}^\dagger d_{\alpha\beta} + \sum_{\alpha\beta\gamma\delta} U_{\alpha\beta\gamma\delta} d_{\alpha\beta}^\dagger d_{\gamma\delta}^\dagger d_{\gamma\delta} d_{\alpha\beta}. \]  

In this expression, the indexes \( \alpha \) and \( \beta \) label the quantum states of the electrons, which include the spin, orbital and lattice indices. \( d_{\alpha\beta}^\dagger \) is the electron creation operator and \( U_{\alpha\beta\gamma\delta} \) the coefficient of the interaction Hamiltonian term. \( \varepsilon_{\alpha\beta}(k) \) is the energy dispersion of the system and \( \varepsilon_{\alpha\beta} \) the energy level.

According to Ref. [12], the wavefunction of a realistic state, which is the product of the slave bosons and the quasi-particle states in the enlarged Hilbert space, reads

\[ |A\rangle \equiv \frac{1}{\sqrt{D_A}} \sum_n \phi_n^A |\text{Vac}\rangle \otimes |\text{n}\rangle_f, \]  

where the underline in |\( A \rangle \) denotes the representative states in the enlarged Hilbert space. \( D_A \) denotes the dimension of the subspace of the Hilbert space [12]. Index \( A \) and \( n \) in \( \phi_n^A \) refer to the \( \text{A} \)th atomic Fock state in the physical space and \( \text{n} \)th atomic Fock state in theQP space, respectively. \( |\text{n}\rangle_f \) is the \( \text{n} \)th Fock state in theQP space.

In the RISB framework [12], we need to find the representation of the electron creation operator in the enlarged Hilbert space. It is obviously that the action of this operator in the enlarged Hilbert space must be the same as that of the electron creation operator in the physical space, that is, the following equation must be held:

\[ \langle A | d_{\alpha\beta}^\dagger | B \rangle = \langle A | d_{\alpha\beta}^\dagger | B \rangle, \]  

which could be also written as \( d_{\alpha\beta}^\dagger | B \rangle = \sum_{\gamma\delta} \langle A | d_{\gamma\delta}^\dagger | B \rangle | A \rangle \). It is proved that this representation in terms of bosons and quasiparticle operators reads [12]

\[ d_{\alpha\beta} = \sum_{\rho} R_{\alpha\beta}^\rho f_{\rho}. \]  

To ensure this representation is rotationally invariant, the matrix \( R_{\alpha\beta}^\rho \) is written as [12]

\[ R_{\alpha\beta}^\rho = \sum_{A\beta\mu\gamma} \phi_{\alpha\mu}^A \phi_{\beta\mu} M_{\gamma\beta}. \]
The generalized expression of the Hamiltonian in terms of the slave bosons and QP fermionic operator thus reads [12]:

\[ H_{\text{cin}} = \sum_k \sum_{\alpha'' \beta''} \langle \hat{R}_\beta | e_{\alpha'' \beta''}(k) \hat{R}_{\alpha'} | \beta' \rangle f^\dagger_{\alpha''} f_{\alpha'} \]  
\[ H_{\text{loc}} = \sum_{AB} \langle A | H_{\text{loc}} | B \rangle \sum_n \phi^\dagger_{\alpha n} \phi_{\beta n}, \]

where \( e_{\alpha'' \beta''}(k) \) is the energy dispersion of the physical electrons.

To ensure the renormalization factor matrix be diagonal in the Kotliar–Ruckenstein slave-boson situation, we adopt the same Fock basis in the physical and QP space throughout this paper.

References