Gutzwiller approach to the Anderson lattice model with no orbital degeneracy

Vladimir Ž. Vulović and Elihu Abrahams
Serin Physics Laboratory, Rutgers University, Piscataway, New Jersey 08854
(Received 2 February 1987)

A new technique is used to obtain the Gutzwiller ground-state energy functional for the Anderson lattice model with no orbital degeneracy (ALM). For the Hubbard model, known expressions are derived with ease and simplicity. For the ALM, we derive the ground-state energy functional of Varma, Weber, and Randall. As a check on our Gutzwiller functional, we find an independent analytical upper bound for the ground-state energy of ALM with a dispersionless f band. For the case of a dispersionless f band and momentum-independent hybridization, in the Kondo regime, we derive analytical expressions for the ground-state energy, charge, and magnetic susceptibilities. For the special case of infinite Coulomb repulsion, we recover results of Rice and Ueda and of Fazekas and Brandow, notably the negative value of the magnetic susceptibility. The negative magnetic susceptibility persists in the entire Kondo region, i.e., finite-U effects do not stabilize the nonmagnetic Kondo state. This suggests that nonzero orbital degeneracy in the Anderson lattice model must be retained to describe heavy-fermion materials with a normal Fermi liquid ground state.

I. INTRODUCTION

The Gutzwiller trial wave function[1] (WF) has been widely used to describe ground-state properties of electron systems with a strong but screened Coulomb interaction. These are encountered in narrow-band metals and are described approximately with the Hubbard Hamiltonian:

$$H = H_0 + H_1 = H_{0\uparrow} + H_{0\downarrow} + H_1 ,$$
$$H_{0\sigma} = \sum_k \varepsilon_{a\sigma}(k) a^\dagger_{k\sigma} a_{k\sigma} , \tag{1.1}$$
$$H_1 = U\hat{D} = U \sum_{i\in\Omega} a^\dagger_{i\uparrow} a_{i\uparrow} a^\dagger_{i\downarrow} a_{i\downarrow} ,$$

where the a band is a narrow nondegenerate band with single-particle energy $\varepsilon_{a\sigma}(k)$, U is the on-site Coulomb repulsion, and Ω is the set of all lattice sites. The Gutzwiller trial WF

$$\Psi = g^{\hat{D}}\Phi \tag{1.2}$$

was originally used to investigate magnetic transition in this system. $\hat{D}$ is the total double-occupancy operator defined in (1.1) and g is a variational parameter allowed to vary between 0 and 1. Both the parameter g that describes the decrease in the number of doubly occupied sites due to H_1, and the parent state Φ are to be determined variationally by minimizing the expectation value of the energy, $E = \langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle$.

For a nonmagnetic case, one takes Φ to be the ground state of H_0. For a ferromagnetic case, with $N_\uparrow - N_\downarrow = Lm$, one takes Φ to be the lowest-energy eigenstate of H_0 for a given m. L is the number of sites in a lattice, and N_σ is the number of spin-σ electrons. This approach has been extended by Ogawa *et al.*,[2] who investigated the antiferromagnetic transition. For an antiferromagnetic parent state Φ, they take the ground state of a Hartree-Fock approximation for H. However, there are open questions regarding the application of the Gutzwiller method to antiferromagnetic states. In particular, at half-filling, Gutzwiller-type calculations[2,3] and second-order perturbation theory[4] predict that the antiferromagnetic transition occurs at $U = U_c > 0$, while mean-field[5] and path-integral calculations[6] predict $U_c = 0$.

Even the most transparent published derivations of the Gutzwiller ground-state energy functional for the Hubbard Hamiltonian[2,7] are not simple and cannot be extended in a straightforward manner to more complicated Hamiltonians. In this paper, we introduce a projection operator method to derive easily the energy functional to be minimized. The only approximation we use is that matrix elements involving real-space configurations with N_σ electrons are replaced by the same matrix element averaged over all configurations with N_σ electrons. This allows us to extend the Gutzwiller approach to a more complicated case of the Anderson lattice Hamiltonian with no orbital degeneracy.

Recently, a number of rare-earth compounds were found to have very heavy conduction electrons at low temperatures.[8] A central problem for the understanding of heavy-fermion materials is the nature of the quasiparticle-energy band structure. A theoretical model that presumably describes some properties of these materials is the Anderson lattice Hamiltonian. It describes a lattice of deep f levels of small dispersion, which experience strong Coulomb repulsion U, and hybridize with a featureless band of *sd* electrons. If one neglects orbital degeneracy of the f level one obtains the Anderson lattice model with no orbital degeneracy (ALM). Because of its relative simplicity, the ALM is of interest in its own right and has been a subject of several theoretical investigations. It is important though, to find out, to what extent the ALM accounts for heavy-fermion materials.

The ALM Hamiltonian is of the form (1.1) but with $H_{0\sigma}$ given by

$$H_{0\sigma}=\sum_k [\varepsilon_{a\sigma}(k)a^\dagger_{k\sigma}a_{k\sigma}+\varepsilon_{b\sigma}(k)b^\dagger_{k\sigma}b_{k\sigma} + V_\sigma(k)(a^\dagger_{k\sigma}b_{k\sigma}+\text{H.c.})] \ . \tag{1.3}$$

The a bands stands for a narrow nondegenerate f band and the b band represents a wide nondegenerate sd band in heavy-fermion compounds. The Coulomb repulsion for the wide band is neglected.

One approach to extract information about the ground state of this system is to use an adequately modified Gutzwiller trial function, as done by Rice and Ueda,[9,10] Varma *et al.*,[11] and Fazekas and Brandow.[12]

The general philosophy is as follows: The f band lies below the Fermi surface and in the absence of correlation U it plays no role. At nonzero U, however, the presence of doubly f-occupied sites in the uncorrelated configurations will raise the energy and it may be worthwhile to promote some electrons from the f band to the Fermi level and then take advantage of hybridization to delocalize the f electrons, thereby lowering the energy. In the Gutzwiller technique, one projects out the doubly f-occupied sites to an extent determined variationally (all of them at $U\rightarrow\infty$).

The work of Rice and Ueda[9] (RU) and Fazekas and Brandow[12] (FB) is limited to the infinite-U case; while it suggests the correct physics, the extension to finite U is required when comparison with real materials is wanted. Finite-U effects have been investigated by Varma, Weber, and Randall[11] (VWR) and Rice and Ueda.[10] However, in neither paper is the density of heavy electrons treated in a consistent manner. Rice and Ueda work with a fixed (average) number of heavy electrons; this leads to conflicting results for the hybridization matrix elements. VWR do not discuss the decrease in the density of heavy electrons caused by suppression of configurations that have many sites with two heavy electrons. To see this, let us denote by g^* and Φ^* the best choice at the variational parameter g and the parent state Φ. VWR assume that the density of heavy electrons in the state Φ^* is equal to the density of heavy electrons in the state $(g^*)^{\hat{D}}\Phi^*$. However, in general the best values are such that $g^*<1$ and that the state Φ^* has a nonzero hybridization, i.e., Φ^* is not an eigenstate of the total number of heavy electrons

$$\hat{N}_a=\hat{N}_{a\uparrow}+N_{a\downarrow} \ , \qquad \hat{N}_{a\sigma}=\sum_k a^\dagger_{k\sigma}a_{k\sigma} \ . \tag{1.4}$$

Therefore the densities

$$n^0_a=\langle\Phi^*\mid\hat{N}_a\mid\Phi^*\rangle/L\langle\Phi^*\mid\Phi^*\rangle$$

and

$$n_a=\langle\Phi^*\mid(g^*)^{2\hat{D}}\hat{N}_a\mid\Phi^*\rangle/L\langle\Phi^*\mid(g^*)^{2\hat{D}}\mid\Phi^*\rangle$$

will be different. The fact that for nonzero hybridization, $n^0_a>n_a$, will be referred to as the density depletion effect.

The structure of the paper is as follows. In Sec. II we introduce a new technique for finding a Gutzwiller ground-state energy functional for the Hubbard Hamiltonian. This approach involves no density matrices, very little combinatorics, and can be extended in a straightforward manner to the ALM Hamiltonian.

In Sec. III we obtain the Gutzwiller ground-state energy functional for the ALM. To accomplish this, we replace matrix elements in a given configuration by their average values, which are these matrix elements averaged over all configurations with the *same number* of heavy electrons. These latter quantities are evaluated in a Gaussian approximation, modified so that some important constraints are satisfied (MGA). The energy functional then has the same form as that of VWR. However, although all the observable quantities are the same in two approaches, the meaning of variational functions that enter the respective functionals is different.

We derive equations for the extrema of the energy functional and use them to show that a two-parameter variational ansatz for the (effective) hybridization angle used by VWR reproduces the exact extremum. Extremal equations reveal how the effective hybridization and the effective position of the a-level change with the discontinuity in the total density of particles at the Fermi level. We also give an alternative form of the energy functional, which contains the starting point of Rice and Ueda[9] as a special case when U is infinite and the heavy band is without dispersion.

In Sec. IV we specialize to the heavy a band with no dispersion and the light b band with a flat density of states. First we introduce a simple analytical upper bound on the ground-state energy, which completely disregards the hybridization, but is independent of the Gutzwiller approach. This approximation is at its best in the large-U limit, where hybridization effects are reduced, and it is in this region that it predicts that the magnetic susceptibility is infinite and that the change susceptibility has the noninteracting light band value. Then we return to the Gutzwiller energy functional and in the Kondo regime solve equations for the nonmagnetic minimum. This is the regime where the deviation of the a-particle density (per one spin projection) from $\frac{1}{2}$ and the density of sites occupied by two a electrons, as a function of hybridization, are exponentially small. We derive analytical expressions for the ground-state energy, and charge and magnetic susceptibilities. In the special case of infinite U our results are in agreement with RU and FB. Moreover, for any finite U such that we are still in the Kondo regime, we find that the magnetic susceptibility is negative. Thus, within the Gutzwiller mean-field theory, the ALM does not have a (nonmagnetic) Kondo ground state.

II. ENERGY FUNCTIONAL FOR THE HUBBARD MODEL

The Gutzwiller ground-state energy functional for the Hubbard model is

$$E=\frac{\langle\Psi\mid H\mid\Psi\rangle}{\langle\Psi\mid\Psi\rangle} \ , \tag{2.1}$$

where H and Ψ are given by (1.1) and (1.2). In order to evaluate matrix elements appearing in (2.1) it proves convenient to work in a basis of L sites:

$$a_i=\frac{1}{\sqrt{L}}\sum_k a_k e^{ik\cdot R_i}, \quad i=1,\ldots,L \ . \tag{2.2}$$

Then the energy E is a function of matrix elements in the state Ψ of operators 1, $a_i^\dagger a_i$, $a_i^\dagger a_j$, and $a_{i\uparrow}^\dagger a_{i\uparrow} a_{i\downarrow}^\dagger a_{i\downarrow}$, where the first three are spin diagonal. In order to evaluate these averages one needs to express the parent state Φ in a form suitable for application of the operator $g^{\hat{D}}$. We introduce an "identity insertion" technique which accomplishes this by inserting various forms of the identity operator in appropriate places. For example:

$$1_{a\uparrow} = \prod_{i\in\Omega} (a_{i\uparrow}^\dagger a_{i\uparrow} + a_{i\uparrow} a_{i\uparrow}^\dagger) = \sum_G \prod_{i\in G} a_{i\uparrow}^\dagger a_{i\uparrow} \prod_{i\in\bar{G}} a_{i\uparrow} a_{i\uparrow}^\dagger , \tag{2.3}$$

where Ω is the set of all lattice sites, $\sum_G$ denotes the summation over all possible subsets G of Ω, and $\bar{G} = \Omega \setminus G$ is the complement of G in Ω. Symbolically, $\prod_{i\in G} a_{i\uparrow}^\dagger a_{i\uparrow} \to G$ and $\prod_{i\in\Gamma} a_{i\downarrow}^\dagger a_{i\downarrow} \to \Gamma$, so that

$$1_{a\uparrow} = \sum_G G\bar{G}, \quad 1_{a\downarrow} = \sum_\Gamma \Gamma\bar{\Gamma} .$$

Then

$$|\Psi\rangle = g^{\hat{D}} |\Phi\rangle = g^{\hat{D}} 1_{a\uparrow} 1_{a\downarrow} |\Phi\rangle = \sum_G \sum_\Gamma g^{\hat{D}} G\bar{G}\Gamma\bar{\Gamma} |\Phi\rangle = \sum_G \sum_\Gamma g^{\nu(G\Gamma)} G\bar{G}\Gamma\bar{\Gamma} |\Phi\rangle , \tag{2.4}$$

where $\nu(A)$ is the number of sites in a set A. [In Eq. (2.5), $A = G\Gamma$ is the intersection of the sets G and Γ.]

We first consider the norm of $|\Psi\rangle$:

$$\langle\Psi|\Psi\rangle = \sum_G \sum_\Gamma g^{2\nu(G\Gamma)} \langle\Phi| G\bar{G}\Gamma\bar{\Gamma} |\Phi\rangle = \sum_G \sum_\Gamma g^{2\nu(G\Gamma)} \langle S| G\bar{G} |S\rangle \langle\Sigma| \Gamma\bar{\Gamma} |\Sigma\rangle , \tag{2.5}$$

where in the second step we assume that Φ is a direct product of a spin-up state S and a spin-down state Σ (recall that the up and down spins decouple at $U=0$). Restricting ourselves to translationally invariant phases, we remark that in the Gutzwiller mean-field approach one neglects all configurational (set G) dependence of the matrix elements $\langle S| G\bar{G} |S\rangle$, except for the dependence on the number of sites in G. In the Hubbard model, this last dependence is particularly simple: $\langle S| G\bar{G} |S\rangle$ is nonzero only if $\nu(G) = N_\uparrow$, where $N_\uparrow$ is the number of electrons in state S.

The $\langle S| G\bar{G} |S\rangle$ value, averaged over all configurations G with fixed number of sites $\nu(G)$, is zero if $\nu(G) \neq N_\uparrow$. The $\nu(G) = N_\uparrow$ average value follows from

$$1 = \langle S|S\rangle = \langle S| 1_{a\uparrow} |S\rangle = \sum_G \langle S| G\bar{G} |S\rangle = \sum_G \delta_{\nu(G),N_\uparrow} \langle S| G\bar{G} |S\rangle = \binom{L}{N_\uparrow} \langle S| G\bar{G} |S\rangle_{\rm av} ,$$

since $\binom{L}{N_\uparrow}$ is the number of configurations G with $N_\uparrow$ sites. Our approximation of replacing every $\langle S| GG |S\rangle$ by its average value can be summarized by

$$\langle S| G\bar{G} |S\rangle \approx \langle S| G\bar{G} |S\rangle_{\rm av} = \frac{\delta_{\nu(G),N_\uparrow}}{\binom{L}{N_\uparrow}} . \tag{2.6}$$

Replacing all matrix elements in (2.5) by their average values, one finds

$$\langle\Psi|\Psi\rangle = \sum_{G\Gamma} g^{2\nu(G\Gamma)} \frac{\delta_{\nu(G),N_\uparrow}}{\binom{L}{N_\uparrow}} \frac{\delta_{\nu(\Gamma),N_\downarrow}}{\binom{L}{N_\downarrow}} . \tag{2.7}$$

Let us denote by $\{L/D, A_\uparrow, A_\downarrow\}$ the number of configurations $\{G,\Gamma\}$ such that $\nu(G) = A_\uparrow$, $\nu(\Gamma) = A_\downarrow$, $\nu(G\Gamma) = D$ in the space of L sites:

$$\{L/D, A_\uparrow, A_\downarrow\} = \frac{L!}{D!(A_\uparrow - D)!(A_\uparrow - D)!(L + D - A_\uparrow - A_\downarrow)!} . \tag{2.8}$$

Then, Eq. (2.7) reduces to

$$\langle\Psi|\Psi\rangle = \sum_D \{L/D, N_\uparrow, N_\downarrow\} g^{2D} \left[\binom{L}{N_\uparrow} \binom{L}{N_\downarrow} \right]^{-1} . \tag{2.9}$$

In the thermodynamic limit, the summand of Eq. (2.9) will be a sharply peaked (Gaussian) function of the intensive variable $d = D/L$. We replace the sum by its dominant term multiplied by a factor P, which comes from the Gaussian weight in D,

$$\langle\Psi|\Psi\rangle \approx P\{L/D, N_\uparrow, N_\downarrow\} g^{2D} \left[\binom{L}{N_\uparrow} \binom{L}{N_\downarrow} \right]^{-1} , \tag{2.10}$$

where the right-hand side is evaluated at its stationary point with respect to D. Actually, the prefactor P appears in all required averages and so it cancels in all contributions to $\langle\Psi|H|\Psi\rangle / \langle\Psi|\Psi\rangle$. We drop it henceforth.

The stationarity condition for Eq. (2.10) at large L is

$$g^2 = \frac{d(1 - n_{a\uparrow} - n_{a\downarrow} + d)}{(n_{a\uparrow} - d)(n_{a\downarrow} - d)} , \tag{2.11}$$

where $n_{a\sigma} = N_\sigma / L_\downarrow$. It will be referred to as the "dominant term" equation (condition) since it relates the D value of the dominant term in Eq. (2.9) to the variational parameter g.

The evaluation of other matrix elements proceeds in a similar manner. For example,

$$\langle\Psi| a_{i\uparrow}^\dagger a_{j\uparrow} |\Psi\rangle = \langle\Phi| 1_{a\uparrow} 1_{a\downarrow} g^{\hat{D}} a_{i\uparrow}^\dagger a_{j\uparrow} g^{\hat{D}} 1_{a\uparrow} 1_{a\downarrow} |\Phi\rangle = \sum_{G_1,G_2,\Gamma_1,\Gamma_2} \langle\Phi| G_1\bar{G}_1\Gamma_1\bar{\Gamma}_1 g^{\hat{D}} a_{i\uparrow}^\dagger a_{j\uparrow} g^{\hat{D}} G_2\bar{G}_2\Gamma_2\bar{\Gamma}_2 |\Phi\rangle = \sum_{G_1,G_2,\Gamma_1,\Gamma_2} g^{\nu(G_1\Gamma_1)+\nu(G_2\Gamma_2)} \langle S| G_1\bar{G}_1 a_{i\uparrow}^\dagger a_{j\uparrow} G_2\bar{G}_2 |S\rangle \langle\Sigma| \Gamma_1\bar{\Gamma}_1\Gamma_2\bar{\Gamma}_2 |\Sigma\rangle . \tag{2.12}$$

Let us introduce $G=G_1\Omega_{ij}$, where Ω_{ij} is Ω without sites i and j. Then the spin-up matrix element is zero unless $G_1=G+\{i\}$, $G_2=G+\{j\}$, and $\nu(G)=N_\uparrow-1$. The spin-down matrix element is zero unless $\Gamma_2=\Gamma_1$ ($=\Gamma$ in what follows), and $\nu(G)=N_\downarrow$. Denoting the complement of G in Ω_{ij} by $\bar{G}=\Omega_{ij}/G$, we find

$$\langle\Psi\mid a_{i\uparrow}^\dagger a_{j\uparrow}\mid\Psi\rangle=\sum_G^{(ij)}\sum_\Gamma g^{2\nu(\Gamma G)+\nu(\Gamma\cap i)+\nu(\Gamma\cap j)}\langle S\mid G\bar{G}a_{i\uparrow}^\dagger a_{j\uparrow}\mid S\rangle\delta_{\nu(G),N_\uparrow-1}\langle\Sigma\mid\Gamma\bar{\Gamma}\mid\Sigma\rangle\delta_{\nu(\Gamma),N_\downarrow}\ , \tag{2.13}$$

where $\Gamma\cap i$ is the intersection of Γ and site i, and the superscripts i and j on the summation over G remind us that G and $\bar{G}$ both belong to Ω_{ij}.

The $\langle S\mid G\bar{G}a_{i\uparrow}^\dagger a_{j\uparrow}\mid S\rangle$ value, averaged over all configurations with fixed number of sites $\nu(G)$, in the space Ω_{ij} of $L-2$ sites, is zero if $\nu(G)\neq N_\uparrow-1$. The $\nu(G)=N_\uparrow-1$ average value follows from

$$\begin{aligned}\langle\Phi\mid a_{i\uparrow}^\dagger a_{j\uparrow}\mid\Phi\rangle&=\langle S\mid a_{i\uparrow}^\dagger a_{j\uparrow}\mid S\rangle=\langle S\mid 1_{a\uparrow}a_{i\uparrow}^\dagger a_{j\uparrow}\mid S\rangle\\&=\sum_G^{(ij)}\delta_{\nu(G),N_\uparrow-1}\langle S\mid G\bar{G}a_{i\uparrow}^\dagger a_{j\downarrow}\mid S\rangle\\&=\begin{bmatrix}L-2\\N_\uparrow-1\end{bmatrix}\langle S\mid G\bar{G}a_{i\uparrow}^\dagger a_{j\uparrow}\mid S\rangle_{\rm av}\ .\end{aligned}$$

The approximation of replacing every $\langle S\mid G\bar{G}a_{a\uparrow}^\dagger a_{j\uparrow}\mid S\rangle$ by its average value can be summarized by

$$\begin{aligned}\langle S\mid G\bar{G}a_{i\uparrow}^\dagger a_{j\uparrow}\mid S\rangle&\approx\langle S\mid G\bar{G}a_{a\uparrow}^\dagger a_{j\uparrow}\mid S\rangle_{\rm av}\\&=\frac{\delta_{\nu(G),N_\uparrow-1}}{\begin{bmatrix}L-2\\N_\uparrow-1\end{bmatrix}}\langle\Phi\mid a_{i\uparrow}^\dagger a_{j\uparrow}\mid\Phi\rangle\ .\end{aligned}\tag{2.14}$$

The spin-down matrix element in (2.13) is evaluated as in (2.6). Finally, we find

$$\langle\Psi\mid a_{i\uparrow}^\dagger a_{j\uparrow}\mid\Psi\rangle=\sum_G^{(ij)}\sum_\Gamma g^{2\nu(G\Gamma)+\nu(\Gamma\cap i)+\nu(\Gamma\cap j)}\frac{\langle\Phi\mid a_{i\uparrow}^\dagger a_{j\uparrow}\mid\Phi\rangle}{\begin{bmatrix}L-2\\N_\uparrow-1\end{bmatrix}}\delta_{\nu(G),N_\uparrow-1}\frac{1}{\begin{bmatrix}L\\N_\downarrow\end{bmatrix}}\delta_{\nu(\Gamma),N_\downarrow}\ . \tag{2.15}$$

By counting the number of $\{G,\Gamma\}$ configurations in Ω_{ij} for a given configuration Γ in $\{i,j\}$, we obtain

$$\begin{aligned}\langle\Psi\mid a_{i\uparrow}^\dagger a_{j\uparrow}\mid\Psi\rangle=\sum_D&(\{L-2/D,N_\uparrow-1,N_\downarrow\}+2g\{L-2/D,N_\uparrow-1,N_\downarrow-1\}+g^2\{L-2/D,N_\uparrow-1,N_\downarrow-2\})\\&\times g^{2D}\langle\Phi\mid a_{i\uparrow}^\dagger a_{j\uparrow}\mid\Phi\rangle\left[\begin{bmatrix}L-2\\N_\uparrow-1\end{bmatrix}\begin{bmatrix}L\\N_\downarrow\end{bmatrix}\right]^{-1}\ ,\end{aligned}\tag{2.16}$$

where $D=\nu(G\Gamma)$. The dominant term in the above sum again is given by Eq. (2.11), and in the thermodynamic limit the sum on the right-hand side is approximated by its dominant term.

The other matrix elements in $\langle\Psi\mid H\mid\Psi\rangle$, $\langle\Psi\mid a_{i\uparrow}^\dagger a_{i\uparrow}\mid\Psi\rangle$, and $\langle\Psi\mid a_{i\uparrow}^\dagger a_{i\uparrow}a_{i\downarrow}^\dagger a_{i\downarrow}\mid\Psi\rangle$, are evaluated in an analogous manner. Our results may be summarized by

$$\begin{aligned}&\langle a_{i\sigma}^\dagger a_{j\sigma}\rangle=[q_\sigma(1-\delta_{ij})+\delta_{ij}]\langle\Phi\mid a_{i\sigma}^\dagger a_{j\sigma}\mid\Phi\rangle\ ,\\&\langle a_{i\uparrow}^\dagger a_{i\uparrow}a_{i\downarrow}^\dagger a_{i\downarrow}\rangle=d\ ,\end{aligned}\tag{2.17}$$

where the angular brackets mean $\langle\hat{C}\rangle=\langle\Psi\mid\hat{C}\mid\Psi\rangle/\langle\Psi\mid\Psi\rangle$, and

$$q_\sigma=\frac{n_{a\sigma}-d}{n_{a\sigma}(1-n_{a\sigma})}\left[(1-n_{a\uparrow}-n_{a\downarrow}+d)^{1/2}+\left[\frac{d(n_{a-\sigma}-d)}{n_{a\sigma}-d}\right]^{1/2}\right]\ . \tag{2.18}$$

To obtain Eq. (2.18) we have used the dominant term condition (2.11) to express g, the original variational parameter, in terms of d, the new and more convenient variational parameter.

From (2.17) it follows that

$$\langle a_{k\sigma}^\dagger a_{k\sigma}\rangle=(1-q_\sigma)n_{a\sigma}+q_\sigma\langle\Phi\mid a_{k\sigma}^\dagger a_{k\sigma}\mid\Phi\rangle\ , \tag{2.19}$$

which shows that q_σ is the discontinuity in the spin-σ particle density at the Fermi surface. The ground-state energy functional for the Hubbard model follows from Eqs. (1.1), (2.1), and (2.17) through (2.19):

$$\begin{aligned}E&=\sum_{k,\sigma}\varepsilon_{a\sigma}(k)\langle a_{k\sigma}^\dagger a_{k\sigma}\rangle+U\sum_i\langle a_{i\uparrow}^\dagger a_{i\uparrow}a_{j\uparrow}^\dagger a_{j\uparrow}\rangle\\&=\sum_\sigma[(1-q_\sigma)n_{a\sigma}L\bar{\varepsilon}_{a\sigma}+q_\sigma\sum_k\varepsilon_{a\sigma}(k)n_{a\sigma}^0(k)]+ULd\ ,\end{aligned}\tag{2.20}$$

where

$$n_{a\sigma}^0(k)=\langle\Phi\mid a_{k\sigma}^\dagger a_{k\sigma}\mid\Phi\rangle\ ,$$

and

$$\bar{\varepsilon}_{a\sigma}=(1/L)\textstyle\sum_k\varepsilon_{a\sigma}(k)$$

is the mean position of the spin-σ a level. Expression (2.20) is the standard result for the Gutzwiller ground-state energy functional for the Hubbard model.[1,2,7]

III. ENERGY FUNCTIONAL FOR THE ANDERSON LATTICE MODEL

For the ground state of the Anderson lattice, a natural form of the Gutzwiller trial wave function is also given by Eq. (1.2) where, for the nonmagnetic case, the parent state

Φ is the ground state of the $U=0$ problem. However, for nonzero U, the minimum expectation value of energy can be considerably decreased if we enlarge the space of trial parent states.[11] We shall allow Φ to vary among a class of eigenstates of some trial Hamiltonian H_0', which differs from the original $U=0$ Hamiltonian (1.3) by the replacements $(\varepsilon_a, V)\rightarrow(\varepsilon_a', V')$ which are variational functions of k. This approach allows some quasiparticle renormalizations of the single-particle part of the effective Hamiltonian due to correlation effects, and is similar in spirit to the effective Hamiltonian approach of RU.[9]

The expectation value of energy is a function of matrix elements in state Ψ of the following operators: 1, $a_i^\dagger a_i$, $b_i^\dagger b_i$, $a_i^\dagger b_i$, $a_i^\dagger a_j$, $b_i^\dagger b_j$, $a_i^\dagger b_j$, and $a_{i\uparrow}^\dagger a_{i\uparrow} a_{i\downarrow}^\dagger a_{i\downarrow}$, where the first seven are spin diagonal. Below, we derive an expression for the norm of Ψ. The matrix element $\langle\Psi\,|\,a_{i\sigma}^\dagger b_{j\sigma}\,|\,\Psi\rangle$ is derived in Appendix A. Other matrix elements may be evaluated in a similar manner.

Equation (2.5) for the norm of Ψ still holds in the present case. Again, we restrict ourselves to translationally invariant phases. In the Gutzwiller mean-field approach, one neglects all configurational (set G) dependence of the matrix elements $\langle S\,|\,G\bar{G}\,|\,S\rangle$. However, the dependence on the number of sites in G should not be neglected. In particular, when hybridization V goes to zero one would like all matrix elements $\langle S\,|\,G\bar{G}\,|\,S\rangle$ involving G with $\nu(G)$ different from $N_{a\uparrow}^0=\langle S\,|\,\hat{N}_{a\uparrow}\,|\,S\rangle$ to vanish (we assume states S and Σ are normalized to unity).

We now argue that while neglecting the configuration dependence, it is possible to approximate Eq. (2.5) for $\langle\Psi\,|\,\Psi\rangle$ by its dominant term in such a way as to include the desired dependence on the number of sites. To achieve this, we introduce a state $|\,A_\uparrow\rangle$ which is the projection of $|\,S\rangle$ onto the space of the number $A_\uparrow$ of spin-up a particles:

$$|\,A_\uparrow\rangle=\sum_G \delta_{A_\uparrow,\nu(G)} G\bar{G}\,|\,S\rangle\ . \tag{3.1}$$

The norm of this state is

$$\langle A_\uparrow\,|\,A_\uparrow\rangle=\sum_G \delta_{A_\uparrow,\nu(G)}\langle S\,|\,G\bar{G}\,|\,S\rangle\ . \tag{3.2}$$

In the spirit of the Gutzwiller approach, we neglect the configuration dependence of $\langle S\,|\,G\bar{G}\,|\,S\rangle$ and replace it by its average over all configurations having $\nu(G)=A_\uparrow$ spin-up a particles. From Eq. (3.2) we obtain

$$\langle S\,|\,G\bar{G}\,|\,S\rangle\approx\langle S\,|\,G\bar{G}\,|\,S\rangle_{\rm av}=\frac{\langle A_\uparrow\,|\,A_\uparrow\rangle}{\begin{pmatrix}L\\A_\uparrow\end{pmatrix}}\ . \tag{3.3}$$

With Eqs. (2.8) and (3.3), we evaluate (2.5) as

$$\langle\Psi\,|\,\Psi\rangle=\sum_{A_\uparrow,A_\downarrow,D}\{L/D,A_\uparrow,A_\downarrow\}g^{2D}\langle A_\uparrow\,|\,A_\uparrow\rangle\langle A_\downarrow\,|\,A_\downarrow\rangle\times\left[\begin{pmatrix}L\\A_\uparrow\end{pmatrix}\begin{pmatrix}L\\A_\downarrow\end{pmatrix}\right]^{-1}\ . \tag{3.4}$$

In the thermodynamic limit, the summand of Eq. (3.4) will be a sharply peaked (Gaussian) function of the intensive variables $n_{a\sigma}=A_\sigma/L$ and $d=D/L$. We replace the sum by its dominant term

$$\langle\Psi\,|\,\Psi\rangle\approx\{L/D,A_\uparrow,A_\downarrow\}g^{2D}\langle A_\uparrow\,|\,A_\uparrow\rangle\langle A_\downarrow\,|\,A_\downarrow\rangle\times\left[\begin{pmatrix}L\\A_\uparrow\end{pmatrix}\begin{pmatrix}L\\A_\downarrow\end{pmatrix}\right]^{-1}\ , \tag{3.5}$$

where the right-hand side is evaluated at its stationary point with respect to A_σ, D. In the above equation and elsewhere we ignore a factor P which comes from the Gaussian weights in A_σ, D. As in the Hubbard case, P appears in all required averages and so cancels in all contributions to $\langle\Psi\,|\,H\,|\,\Psi\rangle/\langle\Psi\,|\,\Psi\rangle$.

The dominant term equations for the summand of Eq. (3.4), or the stationarity conditions for Eq. (3.5) at large L, are

$$\begin{aligned} g^2&=\frac{d(1-n_{a\uparrow}-n_{a\downarrow}+d)}{(n_{a\uparrow}-d)(n_{a\downarrow}-d)}\ ,\\ f'(A_\sigma)&=\ln\frac{(n_{a\sigma}-d)(1-n_{a\sigma})}{n_{a\sigma}(1-n_{a\uparrow}-n_{a\downarrow}+d)}\ , \end{aligned} \tag{3.6}$$

where $f'(A_\sigma)=f(A_\sigma+1)-f(A_\sigma)$ and $f(A_\sigma)=\ln\langle A_\sigma\,|\,A_\sigma\rangle$. The quantity $f'(A_\sigma)$ has a finite nonzero limit when $L\rightarrow\infty$ with $n_{a\sigma}$ fixed, although $f(A_\sigma)$ does not. We shall return to the determination of $f'(A_\sigma)$ below.

The treatment of the other matrix elements proceeds in a similar manner. We give a detailed evaluation of the most intriguing, $\langle\Psi\,|\,a_{i\sigma}^\dagger b_{j\sigma}\,|\,\Psi\rangle$, in Appendix A. For the others, we quote the results in Appendix A. The dominant term conditions for all the matrix elements have the same form as those for $\langle\Psi\,|\,\Psi\rangle$ in Eq. (3.6) except that the function $f(A_\sigma)$ is modified as follows:

$$\begin{aligned} f(A_\sigma)&=\ln\langle A_\sigma\,|\,A_\sigma\rangle\ \text{ for 1, } a_{i\sigma}^\dagger a_{i\sigma}\ ,\\ f_h(A_\sigma)&=\ln\langle A_\sigma+1\,|\,a_{i\sigma}^\dagger b_{j\sigma}\,|\,A_\sigma\rangle\ \text{ for } a_{i\sigma}^\dagger b_{j\sigma}\ ,\\ f_a(A_\sigma)&=\ln\langle A_\sigma\,|\,a_{i\sigma}^\dagger a_{j\sigma}\,|\,A_\sigma\rangle\ \text{ for } a_{i\sigma}^\dagger a_{j\sigma},\ i\neq j\ ,\\ f_b(A_\sigma)&=\ln\langle A_\sigma\,|\,b_{i\sigma}^\dagger b_{j\sigma}\,|\,A_\sigma\rangle\ \text{ for } b_{i\sigma}^\dagger b_{j\sigma}\ . \end{aligned} \tag{3.7}$$

Functions f are in general complex, but in the thermodynamic limit their derivatives f' are real. In order that matrix elements of the above bilinear operators correspond to the same $d=D/L$ and $n_{a\sigma}=A_\sigma/L$, it is obvious from (3.6) that it is necessary and sufficient that

$$f'(A_\sigma)=f_h'(A_\sigma)=f_a'(A_\sigma)=f_b'(A_\sigma)\ . \tag{3.8}$$

Let us assume that this is the case. Then, using (3.6), we find

$$\begin{aligned} \langle a_{i\sigma}^\dagger b_{j\sigma}\rangle&=\sqrt{q_\sigma}\langle a_{i\sigma}^\dagger b_{j\sigma}\rangle_A\ ,\\ \langle b_{i\sigma}^\dagger a_{j\sigma}\rangle&=\sqrt{q_\sigma}\langle b_{i\sigma}^\dagger a_{j\sigma}\rangle_A\ ,\\ \langle a_{i\sigma}^\dagger a_{j\sigma}\rangle&=[q_\sigma(1-\delta_{ij})+\delta_{ij}]\langle a_{i\sigma}^\dagger a_{j\sigma}\rangle_A\ ,\\ \langle b_{i\sigma}^\dagger b_{j\sigma}\rangle&=\langle b_{i\sigma}^\dagger b_{j\sigma}\rangle_A\ , \end{aligned} \tag{3.9}$$

where q_σ is defined by Eq. (2.19), and A averages are defined as follows:

$$\langle\hat{C}_\sigma\rangle_A=\frac{\langle A_\sigma+\lambda\,|\,\hat{C}_\sigma\,|\,A_\sigma\rangle}{(\langle A_\sigma+\lambda\,|\,A_\sigma+\lambda\rangle\langle A_\sigma\,|\,A_\sigma\rangle)^{1/2}}\qquad\text{for } [\hat{N}_{a\sigma},\hat{C}_\sigma]=\lambda\hat{C}_\sigma\ . \tag{3.10}$$

Note that $\lambda=0$ for $a^\dagger a$ and $b^\dagger b$, and $\lambda=1$ (-1) for $a^\dagger b$ $(b^\dagger a)$.

We next have to find a suitable approximation for the A averages which appear in Eqs. (3.6) and (3.9). This is carried out in Appendix B. In Appendix B, we returned to the Bloch basis and considered matrix elements of $a^\dagger_{k\sigma}b_{k\sigma}$, $a^\dagger_{k\sigma}a_{k\sigma}$ and $b^\dagger_{k\sigma}b_{k\sigma}$. The results of Appendix B, the modified Gaussian approximation (MGA), can be combined with Eqs. (3.9) to give the relevant expectation values in $\langle H\rangle$:

$$\begin{aligned}
&\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle=(1-q_\sigma)n_{a\sigma}+q_\sigma n^0_{ak\sigma}e^{\mu_\sigma/2}/z_{ak\sigma}\ ,\\
&\langle a^\dagger_{k\sigma}b_{k\sigma}\rangle=(q_\sigma)^{1/2}\rho^0_{k\sigma}/z_{ak\sigma}\ ,\\
&\langle b^\dagger_{k\sigma}b_{k\sigma}\rangle=n^0_{bk\sigma}e^{-\mu_\sigma/2}/z_{bk\sigma}\ ,\\
&\langle a^\dagger_{i\uparrow}a_{i\uparrow}a^\dagger_{i\downarrow}a_{i\downarrow}\rangle=d\ ,
\end{aligned}\tag{3.11}$$

where

$$\begin{aligned}
&z_{ak\sigma}=n^0_{ak\sigma}e^{\mu_\sigma/2}+(1-n^0_{ak\sigma})e^{-\mu_\sigma/2}\ ,\\
&z_{bk\sigma}=n^0_{bk\sigma}e^{-\mu_\sigma/2}+(1-n^0_{bk\sigma})e^{\mu_\sigma/2}\ ,\\
&n^0_{ak\sigma}=\langle\Phi\,|\,a^\dagger_{k\sigma}a_{k\sigma}\,|\,\Phi\rangle\ ,\\
&n^0_{bk\sigma}=\langle\Phi\,|\,b^\dagger_{k\sigma}b_{k\sigma}\,|\,\Phi\rangle\ ,\\
&\rho^0_{k\sigma}=\langle\Phi\,|\,a^\dagger_{k\sigma}b_{k\sigma}\,|\,\Phi\rangle\ .
\end{aligned}$$

Thus, n_a counts the a electrons and n_b the b electrons. The MGA dominant term conditions are

$$\begin{aligned}
&g^2=\frac{d(1-n_{a\uparrow}-n_{a\downarrow}+d)}{(n_{a\uparrow}-d)(n_{a\downarrow}-d)}\ ,\\
&e^{\mu_\sigma}=\frac{n_{a\sigma}(1-n_{a\uparrow}-n_{a\downarrow}+d)}{(n_{a\sigma}-d)(1-n_{a\sigma})}\ .
\end{aligned}\tag{3.12}$$

The parameters μ_σ are obtained in terms of d from the second of Eqs. (3.12) and the condition

$$A_\sigma=\sum_k\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle_A=\sum_k n^0_{ak\sigma}e^{\mu_\sigma/2}/z_{ak\sigma}\ .\tag{3.13}$$

Comparing our results with those of VWR, we notice that VWR in expressions for $\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle$, $\langle a^\dagger_{k\sigma}b_{k\sigma}\rangle$, and $\langle b^\dagger_{k\sigma}b_{k\sigma}\rangle$ have $n^0_{ak\sigma}$, $\rho^0_{k\sigma}$, and $n^0_{bk\sigma}$ where in Eq. (3.11) we have

$$\begin{aligned}
&\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle_A=n^0_{ak\sigma}e^{\mu_\sigma/2}/z_{ak\sigma}\ ,\\
&\langle a^\dagger_{k\sigma}b_{k\sigma}\rangle_A=\rho^0_{k\sigma}/z_{ak\sigma}\ ,\\
&\langle b^\dagger_{k\sigma}b_{k\sigma}\rangle_A=n^0_{bk\sigma}e^{-\mu_\sigma/2}/z_{bk\sigma}\ .
\end{aligned}\tag{3.14}$$

We recall that $n^0_{ak\sigma}$, $\rho^0_{k\sigma}$, and n^0_{bk} stand for averages of $a^\dagger_{k\sigma}a_{k\sigma}$, $a^\dagger_{k\sigma}b_{k\sigma}$, and $b^\dagger_{k\sigma}b_{k\sigma}$ in the state Φ, where Φ is the ground state of the trial noninteracting Hamiltonian

$$H'_0=H_0(\varepsilon_a\to\varepsilon'_a,\ V\to V')\ .$$

In order to draw a comparison with the VWR method, we need to check if A averages in Eq. (3.14) can be written as matrix elements in some state $\tilde\Phi$, which is the ground state of some effective noninteracting Hamiltonian

$$\tilde H_0=H_0(\varepsilon_a\to\tilde\varepsilon_a,\ V\to\tilde V)\ .$$

The answer is affirmative, as we shall demonstrate in what follows.

The trial noninteracting Hamiltonian, $H'_0=H_0(\varepsilon_a\to\varepsilon'_a,\ V\to V')$, is diagonalized by a unitary transformation

$$\begin{bmatrix}\alpha_{k\sigma}\\ \beta_{k\sigma}\end{bmatrix}=\begin{bmatrix}\cos\theta_{k\sigma} & \sin\theta_{k\sigma}\\ -\sin\theta_{k\sigma} & \cos\theta_{k\sigma}\end{bmatrix}\begin{bmatrix}a_{k\sigma}\\ b_{k\sigma}\end{bmatrix}\ ,\tag{3.15}$$

where α and β denote the upper and the lower band of H'_0, and the hybridization angle $\theta_{k\sigma}$ is given by $\tan(2\theta_{k\sigma})=2V'_\sigma(k)/[\varepsilon'_{a\sigma}(k)-\varepsilon_{b\sigma}(k)]$. Since Φ is an eigenstate of H'_0, then $n^0_{\alpha k\sigma}=\langle\Phi\,|\,\alpha^\dagger_{k\sigma}\alpha_{k\sigma}\,|\,\Phi\rangle$ and $n^0_{\beta k\sigma}=\langle\Phi\,|\,\beta^\dagger_{k\sigma}\beta_{k\sigma}\,|\,\Phi\rangle$ take values 0 and 1 only, while $\langle\Phi\,|\,\alpha^\dagger_{k\sigma}\beta_{k\sigma}\,|\,\Phi\rangle$ is zero. The corresponding expressions for $n^0_{\alpha k\sigma}$, $\rho^0_{k\sigma}$, and $n^0_{\beta k\sigma}$ are

$$\begin{aligned}
&n^0_{ak\sigma}=n^0_{\beta k\sigma}\sin^2\theta_{k\sigma}+n^0_{\alpha k\sigma}\cos^2\theta_{k\sigma}\ ,\\
&\rho^0_{k\sigma}=(n^0_{\alpha k\sigma}-n^0_{\beta k\sigma})\sin\theta_{k\sigma}\cos\theta_{k\sigma}\ ,\\
&n^0_{bk\sigma}=n^0_{\beta k\sigma}\cos^2\theta_{k\sigma}+n^0_{\alpha k\sigma}\sin^2\theta_{k\sigma}\ .
\end{aligned}\tag{3.16}$$

Equation (3.16) shows that the two variational functions $\varepsilon'_{a\sigma}(k)$ and $V'_\sigma(k)$, that define H'_0 and the trial parent state Φ, in the ground-state energy functional E appear only in the combination $\theta_{k\sigma}$. This means that the density of doubly occupied sites d and the hybridization angle function $\theta_{k\sigma}$ comprise the complete set of independent variational parameters with respect to which E is to be minimized.

When $n^0_{k\sigma}=n^0_{\alpha k\sigma}+n^0_{\beta k\sigma}$ is 0 or 2, from (3.14) and (3.16) one easily finds that $\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle_A=\langle b^\dagger_{k\sigma}b_{k\sigma}\rangle_A=\frac{1}{2}n^0_{k\sigma}$ and $\langle a^\dagger_{k\sigma}b_{k\sigma}\rangle_A=0$. When $n^0_{k\sigma}$ is 1, using (3.14) and (3.16) one can check that $z_{ak\sigma}=z_{bk\sigma}$, $0<\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle_A<1$, $0<\langle b^\dagger_{k\sigma}b_{k\sigma}\rangle_A<1$, $\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle_A+\langle b^\dagger_{k\sigma}b_{k\sigma}\rangle_A=1$, and $\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle_A\langle b^\dagger_{k\sigma}b_{k\sigma}\rangle_A=\langle a^\dagger_{k\sigma}b_{k\sigma}\rangle^2_A$. These relations show that for arbitrary momentum k the following representation for the A averages is possible:

$$\begin{aligned}
&\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle_A=n^0_{\beta k\sigma}\sin^2\tilde\theta_{k\sigma}+n^0_{\alpha k\sigma}\cos^2\tilde\theta_{k\sigma}\ ,\\
&\langle a^\dagger_{k\sigma}b_{k\sigma}\rangle_A=(n^0_{\alpha k\sigma}-n^0_{\beta k\sigma})\sin\tilde\theta_{k\sigma}\cos\tilde\theta_{k\sigma}\ ,\\
&\langle b^\dagger_{k\sigma}b_{k\sigma}\rangle_A=n^0_{\beta k\sigma}\cos^2\tilde\theta_{k\sigma}+n^0_{\alpha k\sigma}\sin^2\tilde\theta_{k\sigma}\ .
\end{aligned}\tag{3.17}$$

Comparison of Eq. (3.16) with Eq. (3.17) shows that there exists some effective noninteracting Hamiltonian $\tilde H_0=H_0(\varepsilon_a\to\tilde\varepsilon_a,\ V\to\tilde V)$, with $\tilde\Phi$ as its ground state, such that $\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle_A$, $\langle a^\dagger_{k\sigma}b_{k\sigma}\rangle_A$, and $\langle b^\dagger_{k\sigma}b_{k\sigma}\rangle_A$ are matrix elements of $a^\dagger_{k\sigma}a_{k\sigma}$, $a^\dagger_{k\sigma}b_{k\sigma}$, and $b^\dagger_{k\sigma}b_{k\sigma}$ in the state $\tilde\Phi$. The effective hybridization function $\tilde\theta_{k\sigma}$ that appears in Eq. (3.17) is related to functions $\tilde\varepsilon_{a\sigma}(k)$ and $\tilde V_\sigma(k)$ that define the effective Hamiltonian $\tilde H_0$ through

$$\tan(2\tilde\theta_{k\sigma})=\frac{2\tilde V_\sigma(k)}{\tilde\varepsilon_{a\sigma}(k)-\varepsilon_{b\sigma}(k)}\ .$$

From Eqs. (3.14), (3.16), and (3.17), for momenta k such that $n^0_{k\sigma}=1$, one can relate $\tilde\theta_{k\sigma}$ and the hybridization angle $\theta_{k\sigma}$ characterizing the parent state:

$$\tan\tilde\theta_{k\sigma}=e^{\mu_\sigma(n^0_{\beta k\sigma}-n^0_{\alpha k\sigma})/2}\tan\theta_{k\sigma}\ .\tag{3.18}$$

We notice that it is possible to use d and $\tilde\theta_{k\sigma}$ instead of d and $\theta_{k\sigma}$ as independent variational parameters. If we

do that then our expression for the energy functional becomes identical to that of VWR. However, there is an important difference in the interpretation of the variational function that enters the ground-state energy functional. In the approach of VWR the variational function is the hybridization angle $\theta_{k\sigma}$ characterizing the parent state Φ of the trial wave function $\Psi=g^{\hat{D}}\Phi$. In our approach it is the function $\tilde{\theta}_{k\sigma}$ that depends both on the parent state Φ (through angle $\theta_{k\sigma}$) and the amount of density depletion (through μ_σ).

Once we choose d and $\tilde{\theta}_{k\sigma}$ as independent variational parameters, the density depletion variables μ_σ completely disappear from the energy functional. In particular, the global constraint, Eq. (3.13), then serves to define $n_{a\sigma}=n_{a\sigma}(\tilde{\theta}_{k\sigma})$ while before it served to determine μ_σ in terms of d and $\theta_{k\sigma}$.

The energy functional in $(d,\tilde{\theta}_{k\sigma})$ variables is given by

$$E=\sum_\sigma \Big[(1-q_\sigma)n_{a\sigma}L\bar{\varepsilon}_{a\sigma}+q_\sigma\sum_k \varepsilon_{a\sigma}(k)\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle_A + \sum_k \varepsilon_{b\sigma}(k)\langle b^\dagger_{k\sigma}b_{k\sigma}\rangle_A +2(q_\sigma)^{1/2}\sum_k V_\sigma(k)\langle a^\dagger_{k\sigma}b_{k\sigma}\rangle_A\Big]+ULd\ , \tag{3.19}$$

where $\bar{\varepsilon}_{a\sigma}=(1/L)\sum_k \varepsilon_{a\sigma}(k)$ and the A averages are given in Eq. (3.17).

Once the best (extremal) values of d and $\tilde{\theta}_{k\sigma}$ for the above energy functional are known one can find parameters μ_σ from the second dominant term equation (3.12). This and Eq. (3.18) then allows one to find the hybridization function $\theta_{k\sigma}$ characterizing the parent state Φ, use Eq. (3.16) to calculate the density of a particles in this state, and compare it with n_a which is the density of a particles in the state Ψ.

Since the parameters μ_σ drop out of the energy functional (3.19) one may feel uncomfortable with the above prescription for evaluating $(1/L)(\langle\Phi|\hat{N}_{a\sigma}|\Phi\rangle-\langle\Psi|\hat{N}_{a\sigma}|\Psi\rangle)/\langle\Psi|\Psi\rangle$ and doubt the very existence of the density depletion effect. After all, we find the same energy functional as VWR, who do not consider the A_σ stationarity equation and make no distinction between $\tilde{\theta}_{k\sigma}$ and $\theta_{k\sigma}$.

The density depletion effect may be examined by looking at the g dependence of the quantity

$$n_{a\sigma}(g)=\frac{\langle\Phi|g^{\hat{D}}\hat{N}_{a\sigma}g^{\hat{D}}|\Phi\rangle}{L\langle\Phi|g^{2\hat{D}}|\Phi\rangle}\ . \tag{3.20}$$

For the fixed best parent state Φ, this quantity evaluated at the best g value, and at $g=1$, should have the same value if the density depletion effect is absent.

For sufficiently small U, the best value of g is close to 1. Then for a fixed finite L, the right-hand side of Eq. (3.20) can be expanded in a Taylor series which gives without approximation

$$n_{a\sigma}(g)=n^0_{a\sigma}+2(g-1)n^0_{a-\sigma}\frac{1}{L}\sum_k n^0_{ak\sigma}(1-n^0_{ak\sigma}) +O[(g-1)^2]\ .$$

We recall that $n^0_{ak\sigma}=\langle\Phi|a^\dagger_{k\sigma}a_{k\sigma}|\Phi\rangle$ and that $n^0_{a\sigma}=(1/L)\langle\Phi|\hat{N}_{a\sigma}|\Phi\rangle$ is the $a\sigma$ density in the best parent state Φ. The coefficient of the $g-1$ term has a finite limit when $L\to\infty$ for $n_\sigma=A_\sigma/L$ fixed. Not having checked the L dependence of higher-order coefficients of the $g-1$ expansion, this does constitute a rigorous proof that $n_{a\sigma}(g)=(1/L)\langle\hat{N}_{a\sigma}\rangle_g$ is g dependent in the thermodynamic limit. However, precisely the same small $g-1$ behavior follows from the MGA dominant term equations (3.12).

In order to obtain further evidence for the density depletion effect we have examined the $n_{a\sigma}=n_{a\sigma}(g)$ dependence for $L=2$ and $N_\uparrow=N_\downarrow=4$ numerically, using the Gutzwiller expression for the trial wave function and assuming that $\varepsilon'_{a\sigma}(k)$ and $V'_\sigma(k)$ are spin and k independent while $\varepsilon_{b\sigma}(k)$ is spin independent. We find that the g dependence of n_a for this small sample qualitatively agrees with a corresponding MGA result for an infinite sample at 50% electron filling (two electrons per spin projection per site).

The energy functional to be minimized, in $(d,\tilde{\theta}_{k\sigma})$ variables, is given by Eqs. (3.19) and (3.17). It depends on the function $\tilde{\theta}_{k\sigma}$, but only for those pairs (k,σ) such that the total particle number in the parent state, $n^0_{k\sigma}=\langle\Phi|a^\dagger_{k\sigma}a_{k\sigma}+b^\dagger_{k\sigma}b_{k\sigma}|\Phi\rangle$, is equal to 1. For such pairs (k,σ) the stationarity equation with respect to $\tilde{\theta}_{k\sigma}$ can be written as

$$\tan(2\tilde{\theta}_{k\sigma})=\frac{2(q_\sigma)^{1/2}V_\sigma(k)}{q_\sigma\varepsilon_{a\sigma}(k)-\varepsilon_{b\sigma}(k)+C_{a\sigma}+C_{V\sigma}}\ , \tag{3.21}$$

where

$$C_{a\sigma}=\sum_\rho\frac{\partial\ln q_\rho}{\partial n_{a\sigma}}(h_{a\rho}-n_{a\rho}\bar{\varepsilon}_{a\rho})+(1-q_\sigma)\bar{\varepsilon}_{a\sigma}\ ,$$
$$C_{V\sigma}=\tfrac{1}{2}\sum_\rho\frac{\partial\ln q_\rho}{\partial n_{a\sigma}}h_{V\rho}\ . \tag{3.22}$$

Above,

$$h_{a\rho}=(1/L)\sum_k\varepsilon_{a\rho}(k)\langle a^\dagger_{k\rho}a_{k\rho}\rangle$$

and

$$h_{V\rho}=(2/L)\sum_k V_\rho(k)\langle a^\dagger_{k\rho}b_{k\rho}\rangle$$

are spin-ρ average a-particle energy and spin-ρ average hybridization energy, while $\bar{\varepsilon}_{a\rho}=(1/L)\sum_k\varepsilon_{a\rho}(k)$ is the mean position of the spin-ρ a level.

We remark that in general, regions in k space where $n^0_{k\sigma}$ takes respective values 0, 1, and 2, do not have to be invariant under changes in U, since the best parent state Φ itself changes with U. Moreover, the entire derivation of the energy functional is valid even when we enlarge the class of parent states Φ to include *all* eigenstates of the trial Hamiltonian H'_0. (Then the state $\tilde{\Phi}$ is a corresponding eigenstate of $\tilde{H}_0$.) A motivation for this more general class of trial parent states comes from the possibility that at large U it may be advantageous to transfer some of the particles from the flat portion of the β band, where they experience a large Coulomb repulsion, to the steep portion of the α band, where they have larger single-particle ener-

gy but the Coulomb repulsion is absent. If this were to happen one would have, for example, a possibility of interaction-induced conductivity: an insulator at low values of U might become a conductor at higher values of U.

We remark that in the present case of several bands, the Luttinger theorem[13] says only that the total volume under the Fermi surface is unchanged by interactions; it does not forbid the above exotic scenario. However, in what follows we shall discuss a less unusual case and assume that regions in k space, where $n^0_{k\sigma}=0$, 1, and 2 do not change with U.

We notice that Eq. (3.21) indicates that the effective hybridization $\tilde{V}_\sigma(k)$ and the effective dispersion of a particles $\tilde{\varepsilon}_{a\sigma}(k)$ can be chosen as

$$\begin{aligned}\tilde{V}_\sigma(k)&=(q_\sigma)^{1/2}V_\sigma(k)\ ,\\ \tilde{\varepsilon}_{a\sigma}(k)&=q_\sigma\varepsilon_{a\sigma}(k)+C_{a\sigma}+C_{V\sigma}\ .\end{aligned}\tag{3.23}$$

When $n^0_{k\sigma}=n^0_{ak\sigma}+n^0_{bk\sigma}$ is two (zero) then states $ak\sigma$ and $bk\sigma$ are both occupied (empty) and the energy is independent of $\tilde{\theta}_{k\sigma}$. For such values of $k\sigma$, $\tilde{\theta}_{k\sigma}$, $\tilde{V}_\sigma(k)$, and $\tilde{\varepsilon}_{a\sigma}(k)$ are not defined, and may be chosen at one's convenience. Our choice of functions $\tilde{\theta}_{k\sigma}$, $\tilde{V}_\sigma(k)$, and $\tilde{\varepsilon}_{a\sigma}(k)$ is such that relations (3.21) and (3.23) hold regardless of whether $n^0_{k\sigma}$ is 0, 1, or 2. Equations (3.23) show that the effective noninteracting Hamiltonian $\tilde{H}_0$, in comparison with H_0, has a bandwidth reduced by q, hybridization reduced by $\sqrt{q}$, and the mean position of spin-σ a level raised by

$$C_{V\sigma}+\sum_\rho \frac{\partial \ln q_\rho}{\partial n_{a\sigma}}(h_{a\rho}-n_{a\rho}\bar{\varepsilon}_{a\rho})\ .$$

Using Eqs. (3.21)–(3.23), one can transform the ground-state energy functional (3.19) into

$$E=\left\langle\tilde{\Phi}\left|\left[H_0[V_\sigma\to(q_\sigma)^{1/2}V_\sigma]+\sum_\sigma(1-q_\sigma)\sum_k[\bar{\varepsilon}_{a\sigma}-\varepsilon_{a\sigma}(k)]a^\dagger_{k\sigma}a_{k\sigma}\right]\right|\tilde{\Phi}\right\rangle+ULd\ .\tag{3.24}$$

When $\varepsilon_{a\sigma}(k)$ is independent of k and $U\to\infty$ this form recues to the energy functional of RU.[9] Explicitly, $H_{\rm eff}(n_f)$ and $K_{\rm eff}$ of RU in our notation are $H_0[V_\sigma\to(q_\sigma)^{1/2}V_\sigma]$ and $H_0[\varepsilon_{a\sigma}\to\tilde{\varepsilon}_{a\sigma},\ V_\sigma\to(q_\sigma)^{1/2}V_\sigma]$. This latter Hamiltonian is just $\tilde{H}_0$ with $\tilde{V}$ eliminated using the first stationarity equation (3.23), and its ground state is $\tilde{\Phi}$ as well. The last two terms in Eq. (3.24) are absent in RU (Ref. 9), because they assumed a dispersionless a band and infinite U.

The energy functional (3.24) has $\tilde{V}_\sigma(k)$ eliminated through the stationarity condition $\tilde{V}_\sigma(k)=(q_\sigma)^{1/2}V_\sigma(k)$. The ground-state energy can still be obtained by minimizing it with respect to $\tilde{\theta}_\sigma(k)$ and d. But now, for the independent set of variational parameters, one may also choose $\tilde{\varepsilon}_{a\sigma}(k)$ and d, or, in the spirit of Rice and Ueda, one may choose $\langle a^\dagger_{k\sigma}a_{k\sigma}\rangle_A$ and d.

For a dispersionless a band, $h_{a\sigma}$ [defined below, Eq. (3.22)] becomes $n_{a\sigma}\varepsilon_{a\sigma}$ and Eq. (3.21) simplifies to

$$\tan(2\tilde{\theta}_{k\sigma})=\frac{2(q_\sigma)^{1/2}V_\sigma(k)}{\varepsilon_{a\sigma}-\varepsilon_{b\sigma}(k)+C_{V\sigma}}\ .\tag{3.25}$$

We recall that VWR have minimized their $\varepsilon_{a\sigma}(k)=\varepsilon_a$, $V_\sigma(k)=V$, and $\varepsilon_{b\sigma}(k)=\varepsilon_b(k)$ Anderson lattice energy functional in the space of functions θ_k that can be parametrized by

$$\tan(2\theta_k)=\frac{v}{a-\varepsilon_b(k)}\ .\tag{3.26}$$

They found the best values of parameters a and v by minimizing the ground-state energy with respect to a and v. With the understanding that their angle θ_k does not describe the parent state Φ, but corresponds to our angle $\tilde{\theta}_k$, our Eq. (3.25) shows that their ansatz for the hybridization angle has a form appropriate for finding the nonmagnetic minimum of the energy functional.

In the next section we shall also need the stationarity condition for the functional (3.19) with respect to d. It is of the form

$$0=U+\sum_\sigma\frac{\partial\ln q_\sigma}{\partial d}(h_{a\sigma}-n_{a\sigma}\bar{\varepsilon}_{a\sigma})+\sum_\sigma\frac{1}{2}\frac{\partial\ln q_\sigma}{\partial d}h_{V\sigma}\ .\tag{3.27}$$

For a dispersionless a band the middle term in the above equation vanishes.

We conclude this section with a brief comparison of our work and the most recent paper of Rice and Ueda.[10] They work with a (fixed) average number of heavy electrons which does not allow for a proper derivation of the hybridization matrix elements. In particular, their expressions for $\langle b^\dagger_{i\sigma}a_{i\sigma}\rangle$ and $\langle a^\dagger_{i\sigma}b_{i\sigma}\rangle$ [Eqs. (2.10) through (2.13) of Ref. 10] are such that the identity $\langle b^\dagger_{i\sigma}a_{i\sigma}\rangle=\langle a^\dagger_{i\sigma}b_{i\sigma}\rangle^*$ is satisfied only when $d=n_{a\uparrow}n_{a\downarrow}$. The dominant term D equation then implies $g=1$, regardless of the value of U. Rice and Ueda avoid this contradiction by replacing both $\langle b^\dagger_{i\sigma}a_{i\sigma}\rangle$ and $\langle a^\dagger_{i\sigma}b_{i\sigma}\rangle$ by their geometric mean. However, in the expression for the ground-state energy functional the above matrix elements appear as a sum, not as a product. Thus the work of Rice and Ueda is a heuristic way to make plausible the ground-state energy functionals given by our Eqs. (3.19) and (3.24).

IV. MAGNETIC INSTABILITY OF HEAVY FERMIONS

In this section we specialize to $\varepsilon_{a\sigma}(k)=\varepsilon_a$, $V_\sigma(k)=V$, and $\varepsilon_{b\sigma}(k)=\varepsilon(k)$ restricted to the interval $[-W/2,W/2]$ and having a constant density of states. We first describe a simple upper bound on the ground-state energy, which completely disregards hybridization but has a virtue of be-

ing free of any Gutzwiller-type approximation.

For large U and/or small V one can expect very small effective hybridization. Then we may choose the trial ground state Ψ to be a direct product of Ψ_a and Φ_b, where Ψ_a is an arbitrary a-particle state with a sharp number of a-electrons and Φ_b is a Fermi sea of b electrons. The expectation value E_{ab} of H in this state is given by

$$\frac{1}{L}E_{ab}=2\varepsilon_a n_a+\bar{\varepsilon}_{b\uparrow}+\bar{\varepsilon}_{b\downarrow}+dU\ , \tag{4.1}$$

where $n_a=\frac{1}{2}(n_{a\uparrow}+n_{a\downarrow})$ $\bar{\varepsilon}_{b\sigma}=\frac{1}{2}Wn_{b\sigma}(1-n_{b\sigma})$, $n_{a\sigma}$ and $n_{b\sigma}$ are spin-σ densities of a and b particles, and d is the density of sites occupied by two a electrons. We recall that $n_{a\sigma}+n_{b\sigma}=n_\sigma+\frac{1}{2}\sigma m$ where $\sigma=\pm1$ and n is the total density per one spin projection and m is the total spin density.

We choose $n_{a\sigma}$ and d as variational parameters, subject to constraints $0<n_{a\sigma}<\min(1,n_\sigma)$ and $\max(0,n_{a\uparrow}+n_{a\downarrow}-1)<d<\min(n_{a\uparrow},n_{a\downarrow})$. It is an easy task to find the nonmagnetic minimum of (4.1) and evaluate the response functions

$$\chi_c=\left[\frac{1}{2L}\frac{d^2E_{ab}}{dn^2}\right]^{-1}$$

and

$$\chi_s=\left[\frac{1}{L}\frac{d^2E_{ab}}{dm^2}\right]^{-1}$$

in this state.

For very low density, such that for $U=0$ the a band is empty, there is no U dependence and the charge and spin susceptibilities have noninteracting b-band values $\chi_c=1/W$ and $\chi_s=2/W$.

For low density, such that for $U=0$ we have $n_{a\uparrow}+n_{a\downarrow}<1$, again there is no U dependence and the response functions (in the $m=0$ state) are $\chi_c=\chi_s=\infty$.

For intermediate density, such that for $U=0$ we have $1<n_{a\uparrow}+n_{a\downarrow}<2$, again $\chi_c=\chi_s=\infty$ as long as $U<U_c=\varepsilon_a+W(n-1)$. In this region $n_a=n-\frac{1}{2}-(\varepsilon_a+U)/W$ and $d=2n_a-1$ drop from their $U=0$ values to $\frac{1}{2}$ and 0, respectively. For $U>U_c$ we have $n_a=\frac{1}{2}$, $d=0$, $\chi_s=\infty$, and $\chi_c=1/W$. The system is stable against compression, but it can be polarized with no effort.

For high density, such that at $U=0$ the b band is filled above the ε_a level, for $U<U_1=n-\frac{3}{2}-\varepsilon_a/W$ we have $\chi_s=2\chi_c=2/W$ and $d=n_a=1$. In region $U_1<U<U_c$ we have $\chi_s=\chi_c=\infty$ with n_a and d decreasing linearly toward $\frac{1}{2}$ and 0. And for $U>U_c$ we have $\chi_s=\infty$, $\chi_c=2/W$ with $n_a=\frac{1}{2}$, and $d=0$.

Apart from establishing the $V\to0$ checkpoint for the Gutzwiller energy functional, this simple model is of little value for small U where hybridization is essential. However, an important point to be learned from the above analysis is that for large enough density, there is a critical U above which compressibility takes the free b-band value and the nonmagnetic state is marginally unstable against magnetism. This region corresponds to the heavy-fermion region in the Gutzwiller approach. We shall see below how small but nonzero effective hybridization in this region translates this marginal stability into a true magnetic instability.

When $n_\sigma<1$, the parent state Φ has the α band empty and β band populated for momenta such that $\varepsilon=\varepsilon(k)$ lies between $\varepsilon_0=-W/2$ and $\varepsilon_{F\sigma}=W(n_\sigma-\frac{1}{2})$. Introducing $R_{0\sigma}=[(\tilde{\varepsilon}_{a\sigma}-\varepsilon_0)^2+4\tilde{V}_\sigma^2]^{1/2}$ and $R_{F\sigma}=[(\tilde{\varepsilon}_{a\sigma}-\varepsilon_{F\sigma})^2+4\tilde{V}_\sigma^2]^{1/2}$, from Eq. (3.19) we derive the energy functional per site:

$$\begin{aligned}
h&=dU+\sum_\sigma h_\sigma\ ,\\
h_\sigma&=h_{a\sigma}+h_{b\sigma}+h_{V\sigma}\ ,\\
h_{a\sigma}&=\varepsilon_a n_{a\sigma}=\varepsilon_a\left[\frac{1}{2}n_\sigma+\frac{1}{2W}(R_{F\sigma}-R_{0\sigma})\right]\ ,\\
h_{b\sigma}&=\frac{1}{4W}(\varepsilon_{F\sigma}^2-\varepsilon_0^2)\\
&\quad-\frac{1}{4W}[(\tilde{\varepsilon}_{a\sigma}+\varepsilon_{F\sigma})R_{F\sigma}-(\tilde{\varepsilon}_{a\sigma}+\varepsilon_0)R_{0\sigma}]\\
&\quad+\frac{\tilde{V}_\sigma^2}{W}\ln\frac{R_{F\sigma}+\varepsilon_{F\sigma}-\tilde{\varepsilon}_{a\sigma}}{R_{0\sigma}+\varepsilon_0-\tilde{\varepsilon}a\sigma}\ ,\\
h_{V\sigma}&=-2(q_\sigma)^{1/2}\frac{\tilde{V}_\sigma V}{W}\ln\frac{R_{F\sigma}+\varepsilon_{F\sigma}-\tilde{\varepsilon}_{a\sigma}}{R_{0\sigma}+\varepsilon_0-\tilde{\varepsilon}_{a\sigma}}\ .
\end{aligned} \tag{4.2}$$

The stationarity conditions (3.23) and (3.27) now become

$$\begin{aligned}
\tilde{V}_\sigma&=(q_\sigma)^{1/2}V_\sigma\ ,\\
\tilde{\varepsilon}_{a\sigma}&=\varepsilon_{a\sigma}-\sum_\alpha\frac{V\tilde{V}_\alpha}{W(q_\alpha)^{1/2}}\frac{\partial q_\alpha}{\partial n_{a\sigma}}\ln\frac{R_{F\alpha}+\varepsilon_{F\alpha}-\tilde{\varepsilon}_{a\alpha}}{R_{0\alpha}+\varepsilon_0-\tilde{\varepsilon}_{a\alpha}}\ ,\\
U&=\sum_\alpha\frac{V\tilde{V}_\alpha}{W(q_\alpha)^{1/2}}\frac{\partial q_\alpha}{\partial d}\ln\frac{R_{F\alpha}+\varepsilon_{F\alpha}-\tilde{\varepsilon}_{a\alpha}}{R_{0\alpha}+\varepsilon_0-\tilde{\varepsilon}_{a\alpha}}\ .
\end{aligned} \tag{4.3}$$

In the spin symmetric ($m=0$) case, these equations imply

$$\frac{\tilde{\varepsilon}_a-\varepsilon_a}{U}=-\frac{1}{2}\frac{\partial q/\partial n_a}{\partial q/\partial d}\ . \tag{4.4}$$

Next we assume that we are in the Kondo regime: $\delta=\frac{1}{2}-n_a$ and d are non-negative and small, with *no* assumptions about their relative magnitude. To leading order, Eq. (4.4) gives

$$d=\frac{2t^2}{1-2t}\delta,\quad\text{where } t=\frac{\tilde{\varepsilon}_a^*-\varepsilon_a}{U}\ . \tag{4.5}$$

We recall that the expression for n_a in Eq. (4.2) implies that to leading order $\tilde{\varepsilon}_a=\tilde{\varepsilon}_a^*+WC\delta$. Here,

$$\tilde{\varepsilon}_a^*=W(n-1)\ , \tag{4.6a}$$

$$C=1+\frac{8V^2}{W^2}\frac{n-1}{n-\frac{1}{2}}\ . \tag{4.6b}$$

Equations (4.3) also imply

$$\frac{\varepsilon_a-\tilde{\varepsilon}_a}{\partial q/\partial n_a}=V^2\ln\frac{R_F+\varepsilon_F-\tilde{\varepsilon}_a}{R_0+\varepsilon_0-\tilde{\varepsilon}_a}\ . \tag{4.7}$$

To leading order in δ, straightforward calculation gives

$$\delta=\frac{W^2(n-\frac{1}{2})}{8V^2}(1-2t)\exp\left[-\frac{(\tilde{\varepsilon}_a^*-\varepsilon_a)W}{4V^2}(1-t)\right] . \tag{4.8}$$

The corresponding energy per site and charge susceptibility χ_c are given by

$$h=\varepsilon_a+W(n-\tfrac{3}{2})(n-\tfrac{1}{2})-\frac{8V^2}{W}\frac{\delta}{1-2t} , \tag{4.9a}$$

$$\chi_c^{-1}=W-\left[\frac{W^2}{4V^2}-\frac{2}{n-\frac{1}{2}}+\frac{2W}{(1-2t)U}\right]W\delta . \tag{4.9b}$$

The most interesting thing about finite-U corrections is the continuously varying exponent in Eq. (4.8). We expect a similar result to hold for the Anderson lattice model *with* orbital degeneracy. There, the U-dependent exponent may be observable through the magnetic susceptibility. Equation (4.9) also shows that, compared with infinite U, finite U increases both the condensation energy $[h(V\to 0)-h]$ and the charge susceptibility.

The density of a particles in the parent state Φ can be calculated as described below Eq. (3.19). In the Kondo regime, the leading-order result is

$$n_a^0=\tfrac{1}{2}+\tfrac{1}{2}r\left[\frac{\pi}{2}-\theta_0\right] , \tag{4.10}$$

where $r=(2V/W)[\sqrt{1-2t}\,/(1-t)]$ and $\theta_0=\tan^{-1}[r/(2n-1)]$. It is interesting that $n_a=\frac{1}{2}-\delta$ lies below $\frac{1}{2}$ by an amount exponentially small in V^{-2}, while n_a^0 lies above $\frac{1}{2}$ by a term of order V.

One can also evaluate discontinuities at the Fermi energy in the particle occupation numbers of a, b, α, β, $\tilde{\alpha}$, and $\tilde{\beta}$ particles. Here $\tilde{\alpha}$ and $\tilde{\beta}$ denotes operators defined as in Eq. (3.15) but with $\tilde{\theta}$ substituted for θ. At the Fermi energy $\tilde{\theta}\approx\pi/2-(2V/W)\sqrt{q}$, $\theta\approx\frac{1}{2}-(2V/W)(1-t/\sqrt{1-2t}\,)q$, which implies $\tilde{\alpha}\approx\alpha\approx b$ and $\tilde{\beta}\approx\beta\approx -a$. In obvious notation, to leading order, discontinuities are $\Delta n(\tilde{\alpha})=\Delta n(\alpha)=\Delta n(b)=(4V^2/W^2)q$ and $\Delta n(\tilde{\beta})=\Delta n(\beta)=\Delta n(a)=q$.

We proceed by calculating the magnetic susceptibility in the nonmagnetic state. The $(d,\tilde{\theta}_{k\sigma})$ form (3.19) of the energy functional implies that the magnetic field for a given m quite generally is

$$B(m)=\frac{d(E/L)}{dm}=\tfrac{1}{2}(\varepsilon_{F\beta\uparrow}-\varepsilon_{F\beta\downarrow})$$
$$=\tfrac{1}{4}(Wm+\tilde{\varepsilon}_{a\uparrow}-\tilde{\varepsilon}_{a\downarrow}-R_{F\uparrow}+R_{F\downarrow}) , \tag{4.11}$$

where $\varepsilon_{F\beta\sigma}=(\tilde{\varepsilon}_{a\sigma}+\varepsilon_{F\sigma}-R_{F\sigma})/2$. We recall that stationarity of E with respect to $(d,\tilde{\theta}_{k\sigma})$ allowed us to replace the total m derivative in the above equation with the partial m derivative.

Taking the total m derivative of Eqs. (4.3) and evaluating them at $m=0$, we find that the derivatives of $\tilde{\varepsilon}_{a\sigma}$ and $\tilde{V}_\sigma$ are odd in σ while the derivative of d vanishes. This allows us to solve for $d\tilde{\varepsilon}_{a\uparrow}/dm$ and $d\tilde{V}_\uparrow/dm$ in a straightforward way. In the Kondo region, using the results for the $m=0$ state, we find to leading order

$$\chi_s^{-1}=8\frac{V^2}{W}\frac{\delta}{1-2t}+\frac{d\tilde{\varepsilon}_{a\uparrow}}{dm}-16\frac{V^2}{W}\frac{dn_{a\uparrow}}{dm}\delta ,$$
$$\frac{d\tilde{\varepsilon}_{a\uparrow}}{dm}=-4(1-t)\frac{1-t+t^2}{1-2t}(\tilde{\varepsilon}_a^*-\varepsilon_a)\delta-16\frac{V^2}{W}t\delta , \tag{4.12}$$
$$\frac{dn_{a\uparrow}}{dm}=\tfrac{1}{2} .$$

These equations show that finite U increases χ_s^{-1} compared with its $U=\infty$ value, but that χ_s stays negative in the entire Kondo region, where δ and d are small due to the exponential term in (4.8) (not because of the prefactor). We remark that this treatment recovers formulas of Rice and Ueda[9] and Fazekas and Brandow,[12] as a special case when $U=\infty$ $(t=0)$.

When $1<n_\sigma<\frac{3}{2}$, the parent state Φ has the β band completely filled and α band populated for momenta such that $\varepsilon=\varepsilon(k)$ lies between ε_0 and $\varepsilon_{F\sigma}=W(n_\sigma-\frac{3}{2})$. (For $n_\sigma>\frac{3}{2}$ the energy diverges as U goes to infinity.) Rather than writing down many lengthy expressions, we shall describe what changes occur in the previous formulas.

We introduce $\varepsilon_1=W/2$ and $R_{1\sigma}=[(\tilde{\varepsilon}_{a\sigma}-\varepsilon_1)^2+4\tilde{V}_\sigma^2]^{1/2}$. Then in Eqs. (4.2), (4.3), and (4.7) one should replace $R_{0\sigma}$ by $R_{F\sigma}$, ε_0 by $\varepsilon_{F\sigma}$, $R_{F\sigma}$ by $R_{1\sigma}$, and $\varepsilon_{F\sigma}$ by ε_3. The only exception is the first term of $h_{b\sigma}$, which stays unchanged (recall $\varepsilon_0=-\frac{1}{2}W$).

Equations (4.4), (4.5), (4.6a), and (4.9a) remain unchanged, while $n-\frac{1}{2}$ in Eq. (4.6b) becomes $n-\frac{3}{2}$. The density n_a^0 is given as in Eq. (4.10) but with angle θ_0 replaced with $\theta_F=\tan^{-1}(r)$. At the Fermi energy, angles describing $(\tilde{\alpha},\tilde{\beta})$ and (α,β) particles are approximately given by $\tilde{\theta}=(2V/W)\sqrt{q}$ and $\theta=\theta_F$, indicating $\tilde{\alpha}\approx a$ and $\tilde{\beta}\approx b$. To leading order, discontinuities are $\Delta n(\tilde{\alpha})=\Delta n(a)=q$, $\Delta n(\tilde{\beta})=\Delta n(b)=(4V^2/W^2)q$, $\Delta n(\alpha)=[q/(1+r^2)][1+(2V/W)r]^2$, and $\Delta n(\beta)=[q/(1+r^2)](r-2V/W)^2$. For infinite U, $r=2V/W$, and the discontinuity in the β occupation number vanishes. Presumably, this indicates that for infinite-U particles in the lower band of the parent state are localized.

Because of their importance, we write down explicitly expressions for δ, χ_c, B, and χ_s:

$$\delta=\frac{W^2(\frac{3}{2}-n)}{8V^2}(1-2t)\exp\left[-\frac{(\tilde{\varepsilon}_a^*-\varepsilon_a)W}{4V^2}(1-t)\right] ,$$
$$\chi_c^{-1}=W\left[\frac{W^2}{4V^2}-\frac{2}{n-\frac{3}{2}}+\frac{2W}{(1-2t)U}\right]W\delta ,$$
$$B(m)=\tfrac{1}{2}(\varepsilon_{F\alpha\uparrow}-\varepsilon_{F\alpha\downarrow})=\tfrac{1}{4}(Wm+\tilde{\varepsilon}_{a\uparrow}-\tilde{\varepsilon}_{a\downarrow}+R_{F\uparrow}-R_{F\downarrow}) , \tag{4.13}$$

$$\chi_s^{-1}=\frac{8V^2}{W}\frac{\delta}{1-2t}+\frac{d\tilde{\varepsilon}_{a\uparrow}}{dm}+16\frac{V^2}{W}\frac{dn_{a\uparrow}}{dm}\delta ,$$
$$\frac{d\tilde{\varepsilon}_{a\uparrow}}{dm}=-4(1-t)\frac{1-t+t^2}{1-2t}(\tilde{\varepsilon}_a^*-\varepsilon_a)\delta+16\frac{V^2}{W}(1-t)\delta ,$$
$$\frac{dn_{a\uparrow}}{dm}=\frac{1}{2} .$$

Here $\varepsilon_{F\alpha\sigma}=(\tilde{\varepsilon}_{a\sigma}+\varepsilon_{F\sigma}+R_{F\sigma})/2$. When U is infinite, $\chi_s^{-1}=32V^2/W\delta-4(\tilde{\varepsilon}_a^*-\varepsilon_a)\delta<0$. Finite U leads to larger χ_s^{-1}, but as long as we are in the Kondo regime, χ_s stays negative. A negative zero-temperature value of χ_s^{-1} means that the high-temperature nonmagnetic state becomes unstable at a nonzero critical temperature.

We have shown that within the Gutzwiller mean-field theory, the nonmagnetic heavy-fermion solution of the ALM is unstable against ferromagnetism. A natural question to ask is whether fluctuations stabilize the heavy-fermion nonmagnetic state. Recently, Shiba[14] employed a variational Monte Carlo method to evaluate ground-state properties of the $U=\infty$ nonmagnetic solution of the ALM. Because of different band-structure and dimensionality used by Shiba, one cannot obtain a useful comparison between his and our values of $\tilde{\varepsilon}_a$, $\tilde{V}$, and q in the $U=\infty$ limit. Also, at zero temperature, spin-spin correlation functions in the optimal nonmagnetic state, evaluated by Shiba, do not give a clue as to whether this state is unstable against ferromagnetism at a nonzero temperature. The variational Monte Carlo method should be extended to include spin-flip processes in order to investigate magnetic instabilities of the ALM.

It has been pointed out[9] that within the Gutzwiller approach, a large enough degeneracy in the f band (a band band in this paper) can stabilize the nonmagnetic state, even in the Kondo limit. While this is true, it is important to note that in some heavy-fermion compounds, notably those that are cerium based,[15] the lowest f state has only the Kramers degeneracy. Therefore, these would appear to remain magnetic within the Gutzwiller approximation.

APPENDIX A

Here we give the derivation of the hybridization matrix element, $\langle\Psi\,|\,a_{i\sigma}^\dagger b_{j\sigma}\,|\,\Psi\rangle$. Following the steps leading to Eq. (2.12), we find

$$\langle\Psi\,|\,a_{i\uparrow}^\dagger b_{j\uparrow}\,|\,\Psi\rangle=\sum_{G_1,G_2,\Gamma_1,\Gamma_2} g^{\nu(G_1\Gamma_1)+\nu(G_2\Gamma_2)} \times\langle S\,|\,G_1\bar{G}_1 a_{i\uparrow}^\dagger b_{j\sigma} G_2\bar{G}_2\,|\,S\rangle \times\langle\Sigma\,|\,\Gamma_1\bar{\Gamma}_1\Gamma_2\bar{\Gamma}_2\,|\,\Sigma\rangle\ . \quad \text{(A1)}$$

Let us introduce $\Gamma=\Gamma_1$ and $G=G_1\Omega_i$, where Ω_i is Ω without site i. Then the spin-up matrix element is zero unless $G_1=G+\{i\}$ and $G_2=G$. The spin-down matrix element is zero unless $\Gamma_2=\Gamma$. Denoting the complement of G in Ω_i by $\bar{G}=\Omega_i/G$ it follows that

$$\langle\Psi\,|\,a_{i\uparrow}^\dagger b_{j\uparrow}\,|\,\Psi\rangle=\sum_G^{(i)}\sum_\Gamma g^{2\nu(G\Gamma)+\nu(\Gamma\cap i)} \times\langle S\,|\,G\bar{G}a_{i\uparrow}^\dagger b_{j\uparrow}\,|\,S\rangle \times\langle\Sigma\,|\,\Gamma\bar{\Gamma}\,|\,\Sigma\rangle\ , \quad \text{(A2)}$$

where $\Gamma\cap i$ is the intersection of Γ and site i and the superscript i on the summation over G reminds that G and $\bar{G}$ both belong to Ω_i. States S and Σ in general involve a nonzero hybridization between a and b particles, therefore Cronecker deltas appearing in a related Eq. (2.13) for the Hubbard model are here absent. Using the projective state $|\,A_\downarrow\rangle$ defined in Eq. (3.1) we find

$$\langle A_\uparrow'\,|\,a_{i\uparrow}^\dagger b_{i\uparrow}\,|\,A_\uparrow\rangle=\delta_{A_\uparrow',A_\uparrow+1}\sum_G^{(i)}\delta_{A_\uparrow,\nu(G)} \times\langle S\,|\,G\bar{G}a_{i\uparrow}^\dagger b_{i\uparrow}\,|\,S\rangle\ .$$

Thus, the average of $\langle S\,|\,G\bar{G}a_{i\uparrow}^\dagger b_{i\uparrow}\,|\,S\rangle$ over states $\nu(G)=A_\uparrow$ is given by

$$\langle A_\uparrow+1\,|\,a_{i\uparrow}^\dagger b_{i\uparrow}\,|\,A_\uparrow\rangle=\begin{bmatrix}L-1\\A_\uparrow\end{bmatrix}\langle S\,|\,G\bar{G}a_{i\uparrow}^\dagger b_{j\uparrow}\,|\,S\rangle_{\rm av}\ . \quad \text{(A3)}$$

Approximating all $\langle S\,|\,G\bar{G}a_{i\uparrow}^\dagger b_{j\uparrow}\,|\,S\rangle$ and $\langle\Sigma\,|\,\Gamma\bar{\Gamma}\,|\,\Sigma\rangle$ matrix elements with their average values, with Eqs. (3.1), (3.3), and (A3), we transform Eq. (A2) into

$$\langle\Psi\,|\,a_{i\uparrow}^\dagger b_{j\uparrow}\,|\,\Psi\rangle=\sum_{A_\uparrow A_\downarrow D}(\{L-1/D,A_\uparrow,A_\downarrow\}+g(L-1/D,A_\uparrow,A_\downarrow-1\}) \times g^{2D}\langle A_\uparrow+1\,|\,a_{i\uparrow}^\dagger b_{j\uparrow}\,|\,A_\uparrow\rangle\langle A_\downarrow\,|\,A_\downarrow\rangle\left[\begin{bmatrix}L-1\\A_\uparrow\end{bmatrix}\begin{bmatrix}L\\A_\downarrow\end{bmatrix}\right]^{-1}\ . \quad \text{(A4)}$$

We now give all needed matrix elements in the dominant term approximation:

$$\begin{aligned}
&\langle\Psi\,|\,\Psi\rangle=\{L/D,A_\uparrow,A_\downarrow\}g^{2D}\langle A_\uparrow\,|\,A_\uparrow\rangle\langle A_\downarrow\,|\,A_\downarrow\rangle\left[\begin{bmatrix}L\\A_\uparrow\end{bmatrix}\begin{bmatrix}L\\A_\downarrow\end{bmatrix}\right]^{-1},\\
&\langle\Psi\,|\,a_{i\uparrow}^\dagger b_{j\uparrow}\,|\,\Psi\rangle=(\{L-1/D,A_\uparrow,A_\downarrow\}+g\{L-1/D,A_\uparrow A_\downarrow-1\}) \times g^{2D}\langle A_\uparrow+1\,|\,a_{i\uparrow}^\dagger b_{j\uparrow}\,|\,A_\uparrow\rangle\langle A_\downarrow\,|\,A_\downarrow\rangle\left[\begin{bmatrix}L-1\\A_\uparrow\end{bmatrix}\begin{bmatrix}L\\A_\downarrow\end{bmatrix}\right]^{-1},\\
&\langle\Psi\,|\,a_{i\uparrow}^\dagger a_{j\uparrow}\,|\,\Psi\rangle=(\{L-2/D,A_\uparrow-1,A_\downarrow\}+2g\{L-2/D,A_\uparrow-1,A_\downarrow-1\}+g^2\{L-2/D,A_\uparrow-1,A_\downarrow-2\}) \times g^{2D}\langle A_\uparrow\,|\,a_{i\uparrow}^\dagger a_{j\uparrow}\,|\,A_\uparrow\rangle\langle A_\downarrow\,|\,A_\downarrow\rangle\left[\begin{bmatrix}L-2\\A_\uparrow-1\end{bmatrix}\begin{bmatrix}L\\A_\downarrow\end{bmatrix}\right]^{-1}\ (i\neq j\text{ only}),\\
&\langle\Psi\,|\,a_{i\uparrow}^\dagger a_{i\uparrow}\,|\,\Psi\rangle=(\{L-1/D,A_\uparrow-1,A_\downarrow\}+g^2\{L-1/D,A_\uparrow-1,A_\downarrow-1\}) \times g^{2D}\langle A_\uparrow\,|\,a_{i\uparrow}^\dagger a_{i\uparrow}\,|\,A_\uparrow\rangle\langle A_\downarrow\,|\,A_\downarrow\rangle\left[\begin{bmatrix}L-1\\A_\uparrow-1\end{bmatrix}\begin{bmatrix}L\\A_\downarrow\end{bmatrix}\right]^{-1},\\
&\langle\Psi\,|\,b_{i\uparrow}^\dagger b_{j\uparrow}\,|\,\Psi\rangle=\{L/D,A_\uparrow,A_\downarrow\}g^{2D}\langle A_\uparrow A_\uparrow\rangle\langle A_\downarrow\,|\,A_\downarrow\rangle\left[\begin{bmatrix}L\\A_\uparrow\end{bmatrix}\begin{bmatrix}L\\A_\downarrow\end{bmatrix}\right]^{-1},\\
&\langle\Psi\,|\,\hat{C}\,|\,\Psi\rangle=C\{L/D,A_\uparrow,A_\downarrow\}g^{2D}\langle A_\uparrow\,|\,A_\uparrow\rangle\langle A_\downarrow\,|\,A_\downarrow\rangle\left[\begin{bmatrix}L\\A_\uparrow\end{bmatrix}\begin{bmatrix}L\\A_\downarrow\end{bmatrix}\right]^{-1},
\end{aligned} \quad \text{(A5)}$$

where $\hat{C}=\hat{D}$ or $\hat{N}_{a\sigma}$, for $C=D$ or A_σ, respectively. When $\hat{C}=\hat{N}_{b\sigma}$, we use $\hat{N}_{b\sigma}=N_\sigma-\hat{N}_{a\sigma}$, where N_σ is the total number of spin-σ particles. We note that by using the dominant-term conditions of Eq. (3.6), we verify $L\langle\Psi\,|\,a_{i\sigma}^\dagger a_{i\sigma}^\dagger\,|\,\Psi\rangle=\langle\Psi\,|\,\hat{N}_{a\sigma}\,|\,\Psi\rangle$. Also, the identity $\langle\Psi\,|\,b_{j\uparrow}^\dagger a_{i\uparrow}\,|\,\Psi\rangle=\langle\Psi\,|\,a_{i\uparrow}^\dagger b_{j\uparrow}\,|\,\Psi\rangle^*$ is satisfied at this and further stages of our calculation.

APPENDIX B

Here, we derive approximations for $f'(A_\sigma)=\ln(\langle A_\sigma+1\,|\,A_\sigma+1\rangle/\langle A_\sigma\,|\,A_\sigma\rangle)$ and A averages $\langle a_{k\sigma}^\dagger a_{k\sigma}\rangle_A$, $\langle b_{k\sigma}^\dagger b_{k\sigma}\rangle_A$, and $\langle a_{k\sigma}^\dagger b_{k\sigma}\rangle_A$ which are necessary for $\langle H\rangle$ as in Eq. (3.9). To simplify the notation, we consider spin up only and drop the σ subscript.

From Eq. (3.2), it is evident that $p(A)=\langle A\,|\,A\rangle$ is a distribution in the number A which is normalized to unity. The same is true of $p_k=\langle A\,|\,a_k^\dagger a_k\,|\,A\rangle/n_k^0$, $\bar{p}_k=\langle A\,|\,a_k a_k^\dagger\,|\,A\rangle/(1-n_{qk}^0)$, and $p_k^v=\langle A+1\,|\,a_k^\dagger b_k\,|\,A\rangle/\rho_k^0$. Here $n_{ak}^0=\langle S\,|\,a_k^\dagger a_k\,|\,S\rangle$, $\rho_k^0=\langle S\,|\,a_k^\dagger b_k\,|\,S\rangle$. We shall approximate these distributions by Gaussians with the correct first two moments. Thus, the Gaussian approximation (GA) for $p(A)$ is

$$p(A)=(2\pi\Delta)^{-1/2}\exp[-(A-\bar{A})^2/2\Delta]\ , \tag{B1}$$

where $\bar{A}=\sum_k n_{ak}^0=Ln^0$ and $\Delta=\sum_k n_{ak}^0(1-n_{ak}^0)=L\sigma^0$. Here n_a^0 is the $a_\uparrow$-particle density in the parent state $|\Phi\rangle$. The other GA distributions have the same form as Eq. (B1) but with $\bar{A}$ and Δ replaced by

$$Ln_a^0+1-n_{ak}^0,\quad L\sigma^0-n_{ak}^0(1-n_{ak}^0)\quad\text{for } p_k\ ,$$

$$Ln_a^0-n_{ak}^0,\quad L\sigma^0-n_{ak}^0(1-n_{ak}^0)\quad\text{for } \bar{p}_k \text{ and } p_k^v\ .$$

At large A,

$$f'(A)\approx p(A+1)-p(A)=-(n_a-n_a^0)/\sigma^0=-\mu\ , \tag{B2}$$

where $n_a=A/L$.

The A averages follow immediately. Thus, for example, $\langle a_k^\dagger a_k\rangle_A=n_{ak}^0 p_k(A)/p(A)$ and we have

$$\begin{aligned}
&\langle a_k^\dagger a_k\rangle_A=n_{ak}^0\exp[+\mu(1-n_{ak}^0)-\mu^2 n_{ak}^0(1-n_{ak}^0)/2]\ ,\\
&\langle a_k a_k^\dagger\rangle_A=(1-n_{ak}^0)\exp[-\mu n_{ak}^0-\mu^2 n_{ak}^0(1-n_{ak}^0)/2]\ ,\\
&\langle a_k^\dagger b_k\rangle_A=\rho_k^0\exp[\mu(1-2n_{ak}^0)/2-\mu^2 n_{ak}^0(1-n_{ak}^0)/2]\ ,\\
&\langle b_k^\dagger b_k\rangle_A=n_{bk}^0\exp[-\mu(1-n_{bk}^0)-\mu^2 n_{bk}^0(1-n_{bk}^0)/2]\ ,
\end{aligned} \tag{B3}$$

where in the last line $n_{bk}^0=\langle S\,|\,b_k^\dagger b_k\,|\,S\rangle$ is the $b_\uparrow$-particle k-state occupation in $|\Phi\rangle$.

It is evident from Eqs. (B3) that $\langle a_k^\dagger a_k+a_k a_k^\dagger\rangle_A=1$ only when $n_{ak}^0=0$ or 1. For intermediate values of n_{ak}^0, this normalization fails at $O(\mu^3)$. To insure proper normalization, we divide all GA A averages by the GA expression for $\langle a_k^\dagger a_k+a_k a_k^\dagger\rangle_A$. This defines the normalized Gaussian approximation NGA:

$$\begin{aligned}
&\langle a_k^\dagger a_k\rangle_A=n_{ak}^0 e^{\mu/2}/z_{ak}\ ,\\
&\langle a_k^\dagger b_k\rangle_A=\rho_k^0/z_{ak}\ ,\\
&\langle b_k^\dagger b_k\rangle_A=n_{bk}^0 e^{-\mu/2}/z_{bk}\ ,
\end{aligned} \tag{B4}$$

where $z_{ak}=n_{ak}^0 e^{\mu/2}+(1-n_{ak}^0)e^{-\mu/2}$ and $z_{bk}=n_{bk}^0 e^{-\mu/2}+(1-n_{bk}^0)e^{\mu/2}$.

Our final approximation, the modified Gaussian approximation (MGA), comes from the observation that $\sum_k\langle a_k^\dagger a_k\rangle_A$ evaluated in the NGA is a function of n_{ak}^0 and μ which depends on A. With the NGA value of $\mu=(n_a-n_a^0)/\sigma^0$, this quantity does not reduce identically to $A=n_aL$. The constraint

$$A=\sum_k\langle a_k^\dagger a_k\rangle_A \tag{B5}$$

is satisfied only for a particular value of A which does not depend on U. Therefore we drop the relation $\mu=(n_a-n_a^0)/\sigma^0$ and use instead Eq. (B5) as a new (implicit) definition of μ, hence the MGA.

The MGA reduces to the GA for small μ. Within the MGA, the normalization of the matrix elements of a-particle and b-particle number operators is insured. When $V=0$ and/or $U=0$, the MGA reproduces known results.

[1]M. Gutzwiller, Phys. Rev. **137**, A1726 (1965).

[2]T. Ogawa, K. Kanda, and T. Matsubara, Prog. Theor. Phys. **53**, 614 (1975).

[3]F. Takano and M. Uchinami, Prog. Theor. Phys. **53**, 1267 (1975); J. Florencio and K. A. Chao, Phys. Rev. B **14**, 3121 (1976).

[4]L. G. Caron and G. Kemeny, Phys. Rev. B **3**, 3007 (1971).

[5]D. R. Penn, Phys. Rev. **142**, 350 (1966).

[6]U. Wolff, Nucl. Phys. **B225**, 391 (1983); G. Kotliar and A. Ruckenstein, Phys. Rev. Lett. **57**, 1362 (1986).

[7]D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984).

[8]Z. Fisk, H. R. Ott, T. M. Rice, and J. L. Smith, Nature **320**, 124 (1986).

[9]T. M. Rice and K. Ueda, Phys. Rev. Lett. **55**, 995 (1985); **55**, 2093(E) (1985).

[10]T. M. Rice and K. Ueda, Phys. Rev. B **34**, 6420 (1986).

[11]C. M. Varma, W. Weber, and L. J. Randall, Phys. Rev. B **33**, 1015 (1985).

[12]P. Fazekas and B. H. Brandow (unpublished).

[13]J. M. Luttinger, Phys. Rev. **119**, 1153 (1960).

[14]H. Shiba, J. Phys. Soc. Jpn. **55**, 2765 (1986).

[15]S. Horn, E. Holland-Moritz, M. Lowenhaupt, F. Steglich, H. Scheuer, A. Benoit, and J. Flouquet, Phys. Rev. B **23**, 3171 (1981).