Phase diagram of the Hubbard model: A variational wave-function approach

S. N. Coppersmith
AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Clare C. Yu*
T-11, MS B262, Los Alamos National Laboratory, Los Alamos, New Mexico 87545
(Received 5 January 1989)

We explore the phase diagram of the Hubbard model by generalizing the Gutzwiller variational wave function to take into account nearest-neighbor spin correlations. The energies of various variational states are calculated and compared. At half filling the lowest energy state has long-range antiferromagnetic correlations, in agreement with previous results. Away from half filling, when the antiferromagnet no longer has the lowest energy, the Gutzwiller approximation yields a lower energy than that found for small spin polarons.

I. INTRODUCTION

The recent discovery of high-temperature superconductivity has sparked renewed interest in the Hubbard model. The key question is whether or not a model involving only on-site repulsion between electrons can have a superconducting ground state. In this paper we shall investigate the properties of the Hubbard model in an attempt to determine whether various scenarios that have been proposed for its behavior are borne out by the behavior of variational wave functions. This calculation is one step in the determination of whether the Hubbard model contains the essential physics of the high-temperature superconductors. If it does not, then that would be strong evidence that a different or more complicated model is necessary to understand the superconducting oxides. Indeed, a myriad of other mechanisms have been proposed that involve more complicated starting Hamiltonians.[1]

Within the Hubbard model, the question of whether or not the on-site repulsion U is large in high-T_c materials is not trivial. Anderson and co-workers have proposed that the high-T_c materials can be described by a Hubbard model where U is much greater than the hopping integral t.[2] Their resonating-valence-bond (RVB) theory assumes that the only importance of charge fluctuations is that they induce antiferromagnetic interactions between moments on neighboring copper sites. Experiments favoring large U include neutron scattering measurements that indicate that the ordered moment on the copper in the insulator is very close to the value predicted for the spin-$\frac{1}{2}$ Heisenberg antiferromagnet.[3] On the other hand, band-structure calculations indicate that the hopping integral t is on the order of a volt, and the Hubbard U is estimated to be about 5 V.[4] Thus, one might expect a moderate U scenario to apply. Indeed, Schrieffer *et al.* have proposed that in this regime of intermediate U, the state in the absence of holes can be thought of as a spin-density wave (SDW) arising primarily from Fermi-surface nesting.[5] In this picture, the addition of holes induces the formation of "spin bags" where the SDW amplitude is suppressed. The spin bags confine the holes, yielding an attractive effective potential between them. This in turn leads to BCS pairing.

In this paper we explore the phase diagram of the Hubbard model as a function of U and doping. Within this context we focus on several questions. The first concerns the crossover between weak and strong coupling. In particular, how large must U be so that the picture of a Heisenberg spin system with itinerant holes is accurate? The second question is whether there is a regime of parameter space where spin-bag formation appears favorable. The third is how the magnetic long-range order is destroyed as mobile holes are added to a half-filled band. This question is clearly nontrivial, for Nagaoka has shown[6] that in the limit $U \rightarrow \infty$ one whole in a half-filled lattice causes the entire system to become ferromagnetic.

Our method of investigation involves the examination of variational wave functions that are generalizations of the one introduced by Gutzwiller.[7] These wave functions have variational parameters that allow for spin correlations other than those already contained in the Gutzwiller wave function. In particular, the spin correlations near the holes can differ from those far away, allowing one to mimic the behavior of a spin bag or polaron. Since experimentally the coherence length of the high-T_c materials is quite short,[8] one would expect these polarons, if they exist, to be rather small and hence, amenable to numerical analysis on finite lattices.

We find that in the half-filled case, when U is large, the minimum energy of the wave function corresponding to an antiferromagnetic state is substantially lower than that for the Gutzwiller wave function, though it is still higher than that obtained for the ground state of a two-dimensional nearest-neighbor Heisenberg model on a square lattice.[9,10] Away from half filling, for very large U a ferromagnetically correlated state appears to be favorable. As U is decreased, very close to half filling the ferromagnetic state gives way to a uniformly antiferromagnetic state, while for dopings greater than 5% or so the system appears to favor a Gutzwiller state with no additional spin correlations for moderate to small values of U. We fail to find evidence that small spin polarons (either ferromagnetic or Fermi liquidlike) appear in the two-dimensional Hubbard model for any value of U.

The paper is organized as follows. In Sec. II the Hamiltonian is presented and the wave functions we examine

are introduced and motivated. Section III describes the methods used in the calculation. A few details of the numerical analysis that are not completely standard are also included. The results are presented in Sec. IV. First, we discuss the half-filled band (one electron per site), and then systems away from half filling. Results are presented for systems with 48, 56, 62, and 64 electrons on 64 sites. Appendix A consists of a proof of a result that is used in the numerical analysis, and Appendix B describes the limitations of the method that prevents straightforward investigation of certain types of wave functions.

II. MODEL

A. Hamiltonian

We study the standard two-dimensional Hubbard model on a square lattice with the Hamiltonian

$$H=-t\sum_{\langle i,j\rangle,\sigma}c_{i\sigma}^{\dagger}c_{j\sigma}+U\sum_{i}n_{i\uparrow}n_{i\downarrow}\,, \tag{2.1}$$

where we take the hopping integral t to be unity. As usual, U describes the strength of the on-site repulsion, $c_{i\sigma}^{\dagger}$ ($c_{i\sigma}$) creates (destroys) an electron of spin σ at site i, and $n_{i\uparrow}$ ($n_{i\downarrow}$) is the number up (down) spins at site i. The bracketed sum signifies a summation over nearest neighbors.

B. Choice of variational wave functions

The wave functions examined in this paper are generalizations of the Gutzwiller wave function

$$|\psi_g\rangle=P_g|\phi\rangle=\prod_i[1-(1-g)n_{i\uparrow}n_{i\downarrow}]\,|\phi\rangle\,, \tag{2.2}$$

where the product over i covers the points in the lattice, and $|\phi\rangle$ is the product of two Slater determinants describing equal numbers of noninteracting up and down electrons. The variational parameter g controls the amount of double occupation. We wish to study the possibility of additional spin correlations between nearest-neighbor sites, which is plausible for large values of U/t. To allow for this, we examine the generalized wave function

$$|\psi_{gh}\rangle=P_g\prod_{\langle i,j\rangle}\exp[-h_{ij}(n_{i\uparrow}-n_{i\downarrow})(n_{j\uparrow}-n_{j\downarrow})]\,|\phi\rangle\,. \tag{2.3}$$

The h_{ij} parametrize the additional spin correlations in the system. If the h_{ij} are independent of position so that $h_{ij}=h$, then an interpretation of this wave function in terms of magnetically ordered states is straightforward. When h is positive, the correlations are antiferromagnetic because if a given spin is up, its nearest neighbors are more likely to be down than in the Gutzwiller wave function. Conversely, a ferromagnetic correlation corresponds to h being negative.

We will also consider the case where the h_{ij} are functions of position. In particular, we wish to consider the case where the magnetic correlations in the vicinity of the empty sites differ from those far from the holes. The motivation for this choice is the spin-bag proposal which says that far from the holes the state is antiferromagnetic, whereas near the holes this antiferromagnetic correlation is reduced. This behavior is mimicked by taking h_{ij} to have a positive value h far from the holes and a smaller (or negative) value h' near a hole. The size of the "spin bag" can be taken to be a variational parameter, but time limitations led us to fix the bag size and examine the energies of wave functions with different values of g, h, and h'. Since the experimentally measured coherence length in high-temperature superconductors is about four Cu-Cu spacings, we expect the polaron size, if it is related to superconductivity, to be about this magnitude. Examination of small polarons is reasonable for U/t on the order of ten because simple estimates lead one to expect a polaron radius that grows only as $(U/t)^{1/4}$ for large U.[11] Thus, for intermediate U we expect that any polaron should have a radius on the order of the lattice constant, which we take to be unity. Also, one expects that if a tendency to form spin polarons were present, then polarons of a fixed size would lower the energy, even if they did not optimize it completely. Therefore, we expect that our results reflect the qualitative behavior of a system with magnetic polarons.

The form of the variational wave function (2.3) is not rotationally invariant. At half filling our results are consistent with the ground state of the model being an antiferromagnet so that this form should be reasonable, but away from half filling the antiferromagnetism disappears. A preferable wave function in this regime might be of the form

$$|\psi_{gh}\rangle=P_g\prod_{\langle i,j\rangle}e^{-(h_{ij}\mathbf{S}_i\cdot\mathbf{S}_j)}\,|\phi\rangle\,, \tag{2.4}$$

which is rotational invariant and very similar to the RVB wave function proposed by Anderson.[2] However, the nature of the numerical evaluation of the expectation values is such that using Eq. (2.4) would require a number of operations that grows exponentially with the system size. This difficulty is discussed in Appendix B. If the magnetic correlations are slowly decaying, then one would expect that locally there would be a well-defined spin direction, so that the error in energy made by picking out a preferred direction would be small.

III. METHODS

The energy and correlation functions of the variational wave functions were evaluated using well-known techniques first discussed by Ceperley *et al.*[12] The integrations corresponding to the evaluation of the relevant correlation functions were performed via a Monte Carlo evaluation with the weighting function $|\psi|^2$. The method is outlined in some detail by Yokoyama and Shiba,[13] and it is straightforward to generalize the method to apply it to wave functions with additional variational parameters. Following Yokoyama and Shiba, we use a square lattice with periodic boundary conditions along x and antiperiodic boundary conditions along y. This choice of boundary conditions breaks the large degeneracy at the Fermi surface, and allows us to check our results for the Gutzwiller wave function with theirs. The Monte Carlo procedure involves randomly choosing single-particle hops to

nearest-neighbor sites.

We examine energies of wave functions which describe Gutzwiller, antiferromagnetic, ferromagnetic, and polaronic states for several fillings. The Gutzwiller state is obtained from Eq. (2.2), or equivalently Eq. (2.3) with h and h' both set to zero. The ferromagnetically correlated states were obtained by letting h_{ij} be negative and independent of position, and the antiferromagnetic states had h_{ij} independent of i,j and positive. In order to search for small magnetic polarons, the wave function Eq. (2.3) was examined for systems where h_{ij} was chosen in the following manner: If the bond $\langle i,j\rangle$ is far from every hole, then h_{ij} is chosen to be the value h. However, near each hole, h_{ij} is chosen to be h' where $h'\neq h$. Figure 1 shows the bonds surrounding each hole for which $h_{ij}=h'$. Thus, the spin polarons are chosen to have a fixed diameter of four lattice constants. When the Monte Carlo sampling is performed, if an electron is moved, then the values of the h_{ij}'s are also adjusted to reflect the new configuration. Thus, the calculation is translationally invariant and the holes are free to move. The motion is fast in the sense that the surrounding spins do not lag behind the hole, i.e., there are no relaxation effects.

The procedure we used consists of choosing values for the variational parameters g, h, and h' and calculating the kinetic energy and double occupation for that wave function.

We searched both for ferromagnetic polarons and those which could be considered to be Fermi liquidlike. The antiferromagnetic state far from the hole for both cases was obtained by letting $h=0.1$. The nature of the polaron was determined by adjusting h'. Ferromagnetic polarons were investigated by letting $h'=-0.1$, while a reduced antiferromagnetic correlation (spin bag) was modeled by letting $h'=0$. As was discussed above, the polarons had a fixed size. However, since we expect the radius of the polaron to depend weakly on U, if the polarons are favorable, there should be a value of U for which the radius we have chosen is optimum. In any case, we would expect polarons of nonoptimal size to still show an energy reduction if their formation were favorable. A few runs were conducted with different values of h and h'. They are described below.

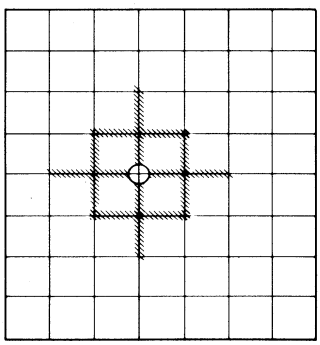

FIG. 1. Drawing of bonds around a hole that are modified in a polaronic wave function. The variational parameter h_{ij} in Eq. (2.3) is set to a value h on the bonds far from the hole, while it is taken to be a different value h' on the hatched lines.

After exploring various values of the variational parameters, the optimal set was chosen by selecting the one with the lowest-total energy. It is assumed that the correlations present in this optimized wave function reflect those that exist in the actual ground state.

A. Numerical analysis

We describe the analysis of the numerical data in some detail because it has some novel features.

For fixed values of g, h, and h', the Monte Carlo sampling yields expectation values (with some numerical uncertainty) for the kinetic energy $T=-\sum_{\langle i,j\rangle,\sigma}\langle c_{i\sigma}^{\dagger}c_{j\sigma}\rangle$ and the number of doubly occupied sites $D=\sum_i\langle n_{i\uparrow}n_{i\downarrow}\rangle$. For given values of h, h', and U, the value of g is determined by the minimization of the energy

$$0=\frac{\partial}{\partial g}(T+UD)=\frac{\partial T}{\partial g}+U\frac{\partial D}{\partial g}. \tag{3.1}$$

In Appendix A it is shown that the double occupation D is a monotonic function of the variational parameter g. This result is intuitively reasonable because, as g increases, more and more states with doubly occupied sites survive the projection operator. Therefore, Eq. (3.1) is equivalent to the statement that

$$0=\frac{\partial T}{\partial D}+U. \tag{3.2}$$

Thus, if one plots the kinetic energy as a function of the double occupation D, for a fixed h, the optimum g corresponds to the point on the graph where the slope of the curve is $-U$.

This plot is particularly useful if one considers changing the values of the other variational parameters h and h'. Suppose one calculates T and D with two sets of values of h and h' for a range of values of g. For a given D the lower-energy state corresponds to the curve which has a more negative T, i.e., a larger $|T|$. This trivial observation means that when several values of h and h' are examined, the lowest-energy state for a given value of D can be read off the graph straightforwardly.

For a fixed U we search for optimum values of the parameters in the following way: We pick values of h and h', say h_1 and h_1', plot T vs D, and find the point (D_1,T_1) where the slope equals $-U$. One then repeats this procedure for new values of h and h'. On each curve one locates the point (D_i,T_i) where the slope equals $-U$. The optimum value of D (and hence, g) corresponds to the point with the lowest value of the energy, i.e., $(-T+UD)$. For two curves that are close together, there is a simple graphical procedure one can use to do this. First, we compare the kinetic energies T_1 and T_2 corresponding to D_1. Suppose T_2 is lower (more negative). This implies the second curve is better than the first near the given value of U. D_2, the "correct" value of D on this curve, corresponds to the point where the slope equals $-U$. For instance, in Fig. 2 the $h=0$ curve corresponds to the minimal energy state for $D\gtrsim 6$.

In order to estimate the energy, the T vs D curves were fit to functions of the form

$$T=a_0+a_1D^{1/2}+a_2D+a_3D^2+a_4D^3. \tag{3.3}$$

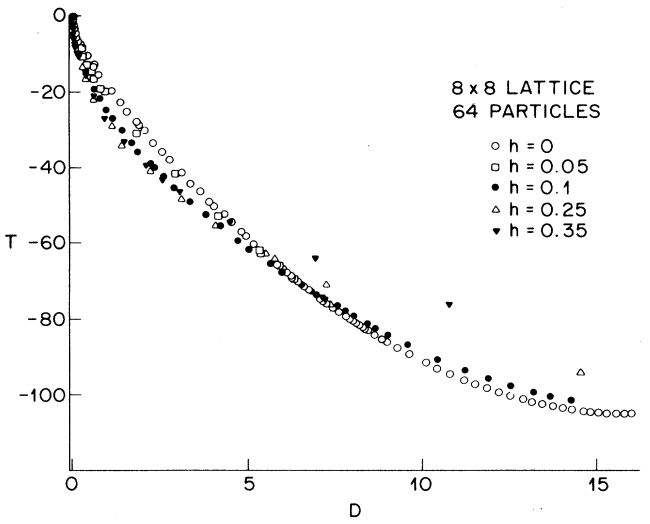

FIG. 2. Plot of the kinetic energy T as a function of the double occupation D for an 8×8 lattice with 64 particles for different values of h. The uncertainty in the data is roughly the size of the data points. In the text it is discussed how this plot can be used to read off the value of h which leads to the lowest energy.

The values of the a_i were determined using the standard method of least squares.[14] The derivative of the fitting function can be taken straightforwardly, and then it is a simple matter to add the kinetic and potential energy contributions for a given U. The uncertainties in the parameters were estimated by adjusting each one with the others held fixed at their optimum values until χ^2 varied by 100%. The uncertainties in the data were also estimated by doing several runs for fixed parameter values, each involving on the order of 10^4–10^5 Monte Carlo trials per site. The nature of the fitting procedure means that the uncertainty in the energy for a fixed value of U (which requires taking a derivative of the fitting function) is much larger than the uncertainty in the kinetic energy differences for a fixed value of the double occupation. Therefore, direct comparison between two trial wave functions which have the same U but different kinetic energies is substantially more accurate than the energy of either one separately.

IV. RESULTS

A. Half-filled case at large U

In order to compare these wave functions to those used by other authors, the half-filled case was examined for large values of U. We obtained an antiferromagnetically correlated state by letting h be positive and independent of position in Eq. (2.3). Figure 2 is a plot of the kinetic energy as a function of the double occupation for an 8×8 lattice. As described above, the lowest-energy state has the smallest value of kinetic energy T for a fixed number of doubly occupied sites. When U is small the Gutzwiller wave function has lower energy than the antiferromagnetic states. This result is reasonable because one expects the moment to be exponentially small in U in the small U limit, so a state with $h=0$ is closer to the true ground state than a state with h fixed to be 0.1. As U is increased, the moment increases and the larger values of h have lower energy than the $h=0$ state. In our calculations the antiferromagnetic state has significantly lower energy than the Gutzwiller wave function when U is large.

However, our numerical uncertainties are large enough that comparison with the best estimates of the ground-state energy of a two-dimensional Heisenberg antiferromagnet is difficult.[9] Compared to calculations on the Heisenberg model, our method is quite inefficient in the $U\rightarrow\infty$ limit since it utilizes single-electron hops and in order to exchange two particles a doubly occupied site must be created. The slowdown is mitigated somewhat because updating the Slater determinant (needed when an electron motion is accepted) requires $O(n^2)$ operations[15] while each trial involves changing one column of the determinant matrix and requires only $O(n)$ steps,[16] where n is the number of electrons. Since the acceptance ratio becomes very small in the half-filled case when U is large, the number of trials possible in a given time goes up. This improves the statistics and allows rough comparison to other methods where $U\rightarrow\infty$.

In the limit $U\rightarrow\infty$, the half-filled Hubbard model with $t=1$ is equivalent to a Heisenberg antiferromagnet with exchange constant $J=4/U$.[17] The ground-state (GS) energy of this system is

$$E_{\mathrm{GS}}=-2NJ\alpha\,, \tag{4.1}$$

where N is the number of sites and α is estimated to be 0.585 ± 0.001.[9] Thus, $E_{\mathrm{GS}}=-8N\alpha/U$.

We can determine how the energy depends on the number of doubly occupied sites D as $D\rightarrow0$ by assuming that the kinetic energy obeys $T=-AD^{\nu}$, where A and ν are constants. To find A and ν, note that the energy $T+UD$ must be a minimum, so that $\partial T/\partial D+U=0$, i.e., $A\nu D^{\nu-1}=U$. Thus, one has

$$-8N\alpha/U=-A\left(\frac{U}{A\nu}\right)^{\nu/(\nu-1)}+U\left(\frac{U}{A\nu}\right)^{1/(\nu-1)}. \tag{4.2}$$

This equation is consistent only if $\nu=\frac{1}{2}$; thus, one finds $A=(32N\alpha)^{1/2}$.

In this analysis we have assumed that the energy has the correct dependence on the parameter t^2/U for the Gutzwiller wave function and the generalizations considered here. In one dimension it is known that at half filling the energy is actually proportional to $(t^2/U)\times\ln(U/t)$.[18] However, since it is practically impossible to see this dependence in the simulations, we have chosen to ignore this possible logarithmic complication.

Figure 3 is a plot of the absolute value of the kinetic energy T as a function of $D^{1/2}$ for different values of the parameter h. On the basis of this analysis, we expect this plot to yield a straight line with slope proportional to $\sqrt{\alpha}$. The two lines on the graph correspond to $\alpha=0.585$, which is the best estimate from Heisenberg model calculations, and $\alpha=0.5$, the energy of a Néel state with no quantum

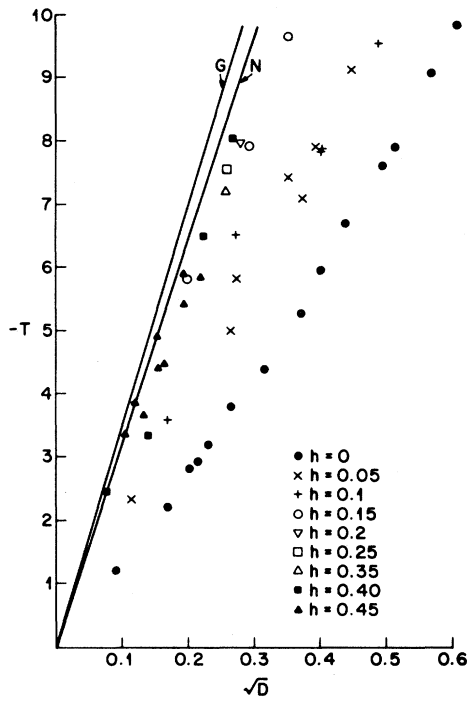

FIG. 3. Plot of the magnitude of the kinetic energy T vs the square root of the number of doubly occupied sites D for small g and several values of the magnetic correlation parameter h, calculated for an 8×8 square lattice with 64 particles. Steeper lines correspond to lower energies. The two solid lines in the figure represent the energies of (N) a Néel state with no quantum fluctuations and (G) the best estimates for the ground-state energy of the system.

fluctuations. It is clear that the Gutzwiller wave function has substantially higher energy than the wave functions with nonzero h. However, the difference is caused mostly by the poorness of the energy for the Gutzwiller wave function. The numerical errors are so large that the difference in energy between a state with no quantum fluctuations and the true ground state is smaller than our numerical uncertainty. However, examination of the spin-spin correlation function indicates that the low-energy wave functions with $h > 0.2$ have very nearly a full moment (see below), so we expect the energy of our state to be the classical Néel value. Comparison with the results of the Heisenberg model is particularly useful as a check to see if we have run long enough to reach equilibrium. When h is allowed to take on values greater than 0.2 or so, the calculation slows down markedly and the energy can have misleadingly low values for runs that are too short.

Our results make clear that for very large U at half filling it is much more effective to consider a canonically transformed Hamiltonian such as that used by Gros, Joynt, and Rice[17] and Yokoyama and Shiba.[19] This transformed Hamiltonian, which is the result of projecting onto the singly occupied subspace of the Hubbard Hamiltonian, accounts for the binding of holes with neighboring doubly occupied sites in a perturbative fashion. This is the important physics for large U at half filling. The Gutzwiller wave function does not include such binding when double occupation is allowed.[20] Therefore, it gives a higher energy for the original Hubbard Hamiltonian than the totally projected wave function with no double occupation gives for the transformed Hamiltonian. However, the canonical transformation is known only order by order in t/U and for intermediate values of U does not yield a variational upper bound to the energy. Thus, it is advantageous to allow double occupation when U has intermediate values. Unfortunately, it is difficult to estimate when the canonical transformation breaks down because one does not know how many of the doubly occupied sites in the untransformed Gutzwiller wave function are accounted for "virtually" using the canonical transformation.

The behavior of the Gutzwiller wave function at half filling without additional spin correlations has been investigated by Yokoyama and Shiba.[13(a)] The Mott transition does not appear in a simple manner because the double occupation is zero only as $U\rightarrow\infty$. Gros, Joynt, and Rice[17] have argued that the doubly occupied sites are actually "virtual" because they are bound to holes. However, our calculations do not shed further light on this issue, in part because the antiferromagnetic states have lower energy than the Gutzwiller wave function in this parameter region.

The correlation function $\langle S_z(0)S_z(\mathbf{r})\rangle$ for various values of h was also calculated. The form of our wave function (where only nearest-neighbor correlations are explicitly included) does not obviously imply that long-range order exists for nonzero values of the parameter h. The degree of spin ordering depends on both g and h. When g is small, fairly small values of h (on the order of 0.1) lead to spin-spin correlation functions that do not have any sign of decaying to zero at large distances. When h is made still smaller, finite-size effects prevent a definite conclusion as to whether long-range magnetic order exists. Although the amount of information contained in this correlation function is necessarily limited because of the small system sizes used in the calculations, it is easy to obtain an upper bound on the size of the ordered moment that can be compared to the results of other methods.

The minimum value of $\langle S_z(0)S_z(\mathbf{r})\rangle$ and the energy per site in units of J for an 8×8 system at half filling is shown in Fig. 4. It is clear that the moment is a very strong function of h. This result reflects the exponential dependence of the projection operator on h: Once h is greater than 0.2 or so, the ratio e^{4h} between the likelihood of having antiparallel and parallel nearest-neighbor spins is so large that the system locks into configurations with alternating up and down spins for basically all values of g. Variational estimates[10] and Monte Carlo results[9] for the magnetic moment lead us to expect that the staggered magnetization is about 0.6 of the Néel result due to quantum fluctuations. For large U the wave function we study appears to have its minimum energy when the ordered moment is larger, namely the Néel value of unity. This implies an energy per spin as $U\rightarrow\infty$ of -0.25 J.

We can compare our results to those obtained by Yokoyama and Shiba[13(b)] using a Gutzwiller-projected spin-density-wave state. The qualitative features of the phase diagram are quite similar, though the energy we ob-

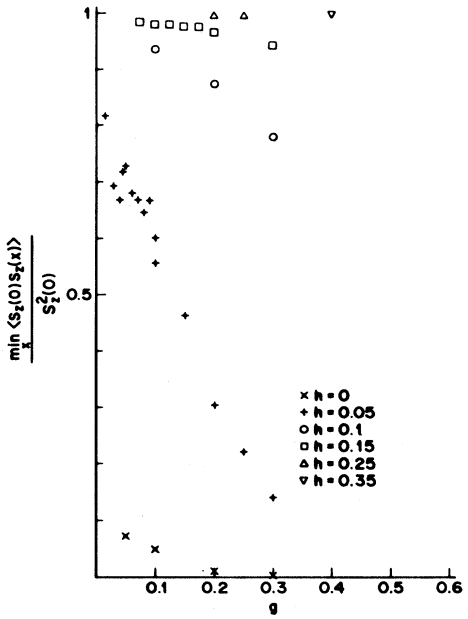

FIG. 4. Plot of the minimum of the absolute value of the spin-spin correlation function $\langle S_z(0)S_z(j)\rangle$ normalized to its value at $j=0$ for an 8×8 system with 64 particles as a function of g for different values of h. For h on the order of 0.2, there is no sign of decay in the spin order.

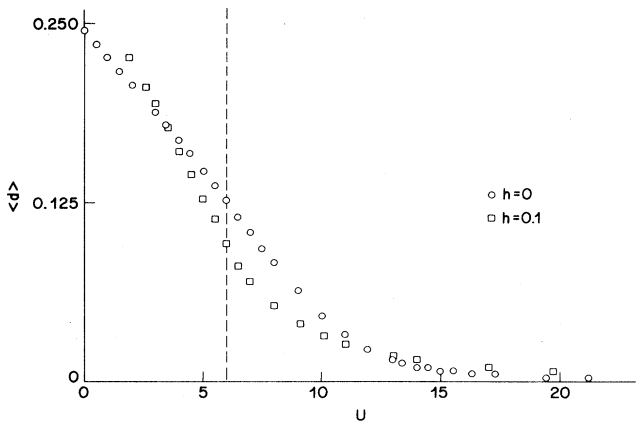

FIG. 5. Plot of the fraction of sites that are doubly occupied as a function of U for a half-filled 8×8 system with $h=0$ and $h=0.1$. The dashed line corresponds to $U\sim6$, where the $h=0$ and $h=0.1$ states have the same energy. The number of doubly occupied sites falls to half its noninteracting value at $U\sim6$ also, which can be viewed as the location of the crossover between weak and strong coupling.

tained using the real-space Jastrow operators to put in the spin-spin correlations is somewhat poorer than the energy they obtained by Gutzwiller projecting a spin-density wave state. In particular, Yokoyama and Shiba also find that for large U their variational wave function has full Néel order with no quantum fluctuations.

We now consider how the behavior of the Hubbard model changes as U is varied. There is no sharp crossover from weak coupling to strong coupling because the nearest-neighbor Hubbard model on a square lattice has antiferromagnetic long-range order for all positive values of U. Therefore, we let the number of doubly occupied sites be a measure of how small t/U is. For intermediate coupling ($U\sim4$), on an 8×8 lattice we find that the energy is minimized for the parameter values $g=0.55$ and $h=0$. The expectation value for the fraction of doubly occupied sites is roughly 0.17. In contrast, for free electrons, on average $\frac{1}{4}$ of the sites are doubly occupied. Figure 5 is a plot of the number of doubly occupied sites on a half-filled lattice as a function of U/t. In addition to showing the Gutzwiller result, the results for an antiferromagnetically correlated wave function are also presented. The number of doubly occupied sites falls to half its free-electron value for $U\approx6$, which is also where states with appreciable magnetic order begin to have substantially lower energy than the Gutzwiller state.

One can compare these results with those obtained using the Gutzwiller approximation to evaluate the expectation values of the double occupancy and the kinetic energy.[21] Using this approximation, one finds that the double occupation is zero for $U>13.137t$ (in contrast to the exact evaluation, where the double occupation is zero only when $U\rightarrow\infty$). The double occupation is half of its value for noninteracting particles when $U\cong6.6$, which is quite close to the value obtained numerically for the exact evaluation of the wave function.

Thus, for values of U on the order of 5, the system is just crossing over from the large U to small U limits and may not be well described by a Heisenberg-hopping Hamiltonian where only the single occupied subspace is considered. This conclusion is of course dependent on the accuracy of the determination of the Hamiltonian parameters U and t.

B. Results at less than half filling

We now discuss our results for systems with less than one fermion per site. (Particle-hole symmetry ensures that systems with more than one fermion per site will behave similarly.) The calculation consists of evaluating the energy Eq. (2.1) for the wave function Eq. (2.3) with different values of the variational parameters. We then examine the state of minimum energy and assume it reflects the preferred configuration of the particles.

1. Dependence of double occupation on U

First we show how the number of doubly occupied sites D depends on the repulsion U for the Gutzwiller wave function. The purpose of this is once again to estimate whether the system away from half filling is better described by a strong-coupling or weak-coupling picture. Figure 6 shows the dependence of D on U for 62 particles on an 8×8 lattice; results for other fillings are qualitatively similar. Once again we find that D falls to half its noninteracting value when U is on the order of 5–6. Therefore, the intermediate-coupling regime appears to be

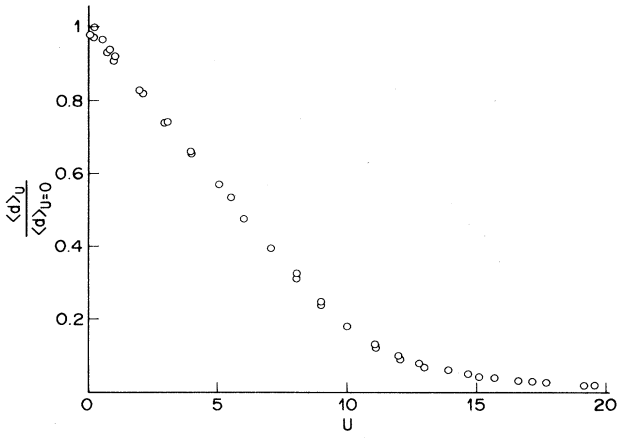

FIG. 6. Plot of the number of doubly occupied sites normalized to its value in the noninteracting system as a function of U for the Gutzwiller wave function for a system with 62 particles on an 8×8 lattice. The double occupation falls to half its value for noninteracting electrons when $U\sim 6$, similarly to the result in the half-filled case. Results for other fillings are similar.

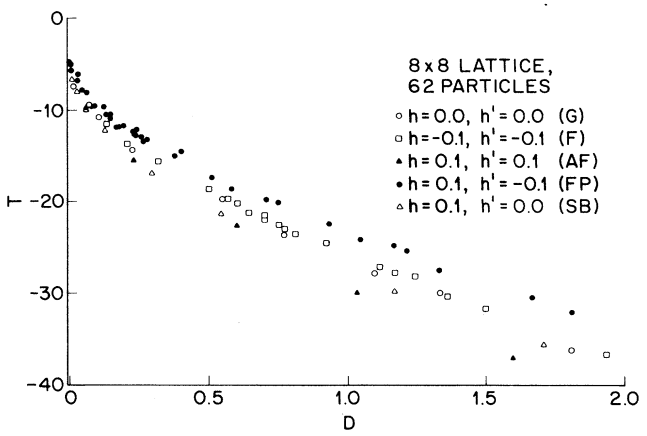

FIG. 7. Plot of the numerical data of the kinetic energy T as a function of the number of doubly occupied sites D for large U on an 8×8 lattice with 62 particles.

the one relevant to high-temperature superconductors.

These results for the double occupation as a function of U can be compared to those obtained using the Gutzwiller approximation.[21] For this filling $\rho=\frac{62}{128}\approx 0.484$, the Gutzwiller approximation yields that D falls to $N\rho^2/2$, half its noninteracting value, when $U\cong 6.2$ (N is the number of lattice sites). As the filling ρ is decreased further, the value of U where D is $N\rho^2/2$ falls slightly, reaching 4.67 for $\rho=\frac{3}{8}$. Therefore, in this intermediate-coupling region the Gutzwiller approximation provides a reasonably good description of the system's behavior.

2. *U dependence of minimum energy state*

We determined the energies of the variational wave function Eq. (2.3) away from half filling for a variety of values of the parameters h and h' and a range of g. When U was extremely large, the Gutzwiller state was higher in energy than a state where h and h' were both negative. This result is consistent with Nagaoka's theorem, which states that one hole in a half-filled band in the limit $U\rightarrow\infty$ causes the system to become ferromagnetic. As U is reduced the ferromagnetic state no longer has the lowest energy. This result can be understood in terms of the standard picture of ferromagnetic polarons that shrink as U is decreased. For U less than 8 or so, the Gutzwiller state has lower energy than any state with nonzero h or h'. For $8<U<30$ or so, the nature of the lowest-energy state depends on the filling.

(a) *62 particles in an* 8×8 *lattice.* Figure 7 is a plot of the numerical results for the expectation value of the kinetic energy T versus the double occupation D for small g (large U) and the choices $h=h'=0$ (Gutzwiller), $h=h'=0.1$ (antiferromagnetic), $h=0.1$, $h'=-0.1$ (ferromagnetic polaron), and $h=0.1$, $h'=0$ (spin bag). For most of the range of parameters shown the uniform antiferromagnetic state with $h=h'=0.1$ has lowest energy. The numerical uncertainties at very large U were so large that the phase boundary between the ferromagnetic state and antiferromagnetic state could not be resolved (though we were able to obtain sufficient statistics to conclude that the ferromagnetic state is favored for $U\rightarrow\infty$). However, we were able to obtain adequate statistics to investigate the crossover between the Gutzwiller and ferromagnetic state; the relevant fits are shown in Fig. 8. It can be seen that the ferromagnetic state has lower energy than the Gutzwiller wave function for $U\gtrsim 35$.

For a fixed hole density, if the holes remain well separated, then the range in U of the ferromagnetic region would not depend on the system size. (In terms of notation let us say that this range extends from $U_{\min}$ to ∞.) On the other hand, if the holes are all found in one polaron, the size of the ferromagnetic region in terms of U would shrink as the system size is increased. In order to investigate the dependence of $U_{\min}$ on the system size, we performed the calculation on a 16×16 system with 248

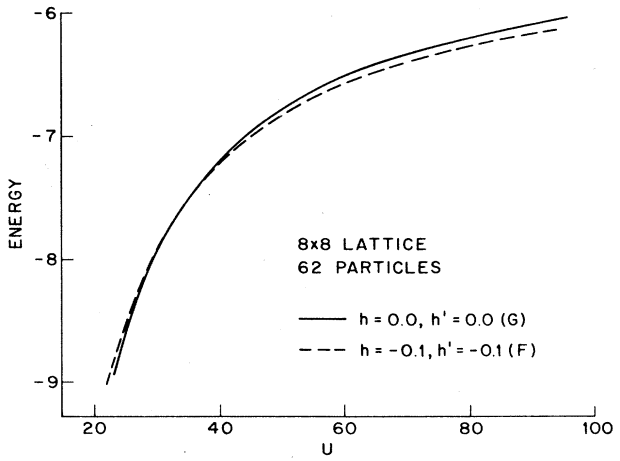

FIG. 8. Plot of the energy vs U for large U for wave functions with $h=h'=0$ and $h=h'=-0.1$. The ferromagnetic state has lower energy for $U\gtrsim 35$.

particles, once again comparing a ferromagnetically correlated state to the Gutzwiller wave function. The numerical uncertainties were so large that a definitive statement is impossible, but the data are consistent with the results for the 8×8 system, meaning the width of the ferromagnetic region in terms of U does not shrink as the system is made larger.

For U less than 35 or so, the data unambiguously imply that the uniform antiferromagnetic state has lowest energy. Figure 9 is a plot of the energies of several different wave functions as a function of U for intermediate U. The antiferromagnetic state has lowest energy for $U \gtrsim 8$. The crossover at $U \sim 8$ to the Gutzwiller state does not necessarily imply that the antiferromagnetic order disappears, for a similar behavior is exhibited at half filling, where it is known that the antiferromagnetism persists to the weak-coupling limit.

(b) *56 particles in an* 8×8 *lattice.* Figure 10 shows the numerical results for the kinetic energy as a function of the double occupation for an 8×8 system with 56 particles for small g. From this plot of the raw numerical data it can already be seen that the only states with low energy for any U are the Gutzwiller state and the ferromagnetic state. Figure 11 is a plot of the energy of the ferromagnetic state and the Gutzwiller state as a function of U. When U is greater than 16 or so the ferromagnetically correlated state has the lowest energy, while for smaller U the Gutzwiller state has the lowest energy. The antiferromagnetic state and both the polaronic wave functions all have higher energy than the states shown.

(c) *48 particles in an* 8×8 *lattice.* Figure 12 shows the numerical results for the kinetic energy as a function of double occupation for an 8×8 system with 48 particles for small g. Once again it is clear from the plot that the only states which compete for the lowest energy are the ferromagnetically correlated wave function ($h = h' = -0.1$)

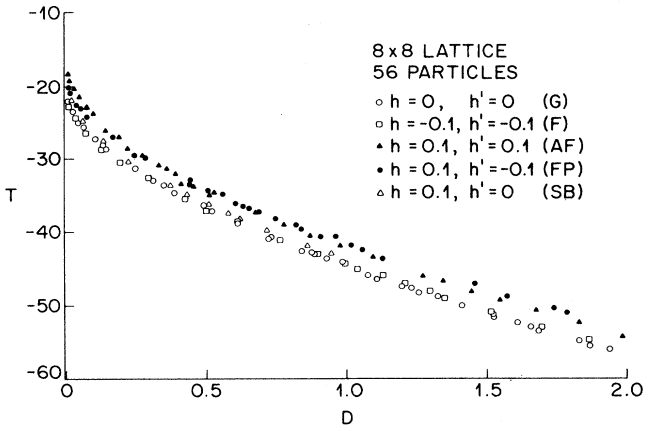

FIG. 10. Plot of the numerical data of the kinetic energy T as a function of the number of doubly occupied sites D for large U on an 8×8 lattice with 56 particles.

and the Gutzwiller wave function ($h = h' = 0$). Figure 13 is a plot of the energy of various wave functions as a function of U, where once again we find that the lowest-energy state is ferromagnetically correlated for large U (here, $U \gtrsim 11$) and has no additional spin correlations for intermediate and small U.

(d) *Summary of U dependence.* For all systems that are not half filled the ferromagnetic state has lowest energy as $U \to \infty$. Very close to half filling the ferromagnetic state gives way to a uniformly antiferromagnetic state. When U is reduced to below 10 or so, the Gutzwiller state then has lower energy.[22] For the other fillings the ferromagnetic state has lowest energy until, as U is reduced, the Gutzwiller state becomes lower in energy.

In all cases we were unable to lower the energy of the

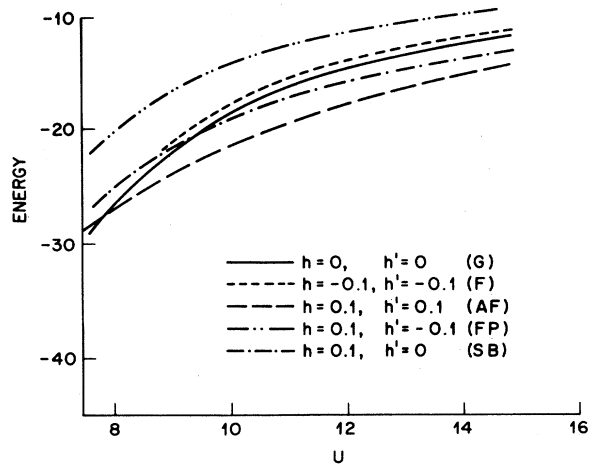

FIG. 9. Plot of the energy vs U for states with $h = h' = 0$ (Gutzwiller), $h = h' = 0.1$ (antiferromagnet), $h = h' = -0.1$ (ferromagnet), $h = 0.1$, $h' = -0.1$ (ferromagnetic polaron), and $h = 0.1$, $h' = 0$ (spin bag). The uniform antiferromagnetic state has lowest energy for $U \gtrsim 8$, and the Gutzwiller state has lowest energy for $U \lesssim 8$. In this parameter range all other states have significantly higher energy.

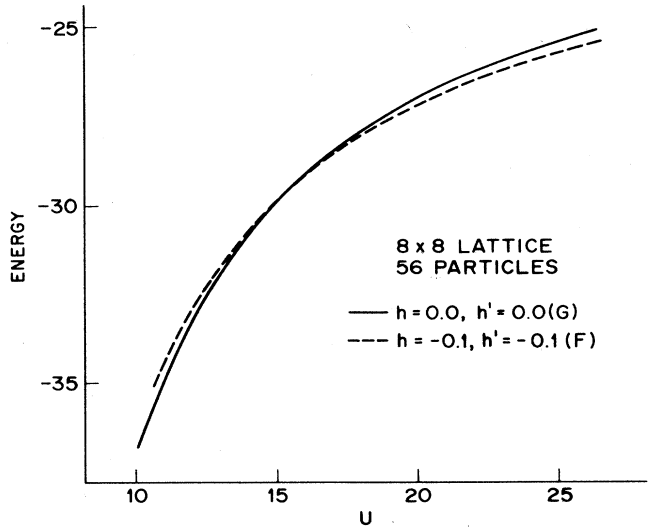

FIG. 11. Plot of the energy vs U for states with $h = h' = 0$ (Gutzwiller) and $h = h' = -0.1$ (ferromagnet) for a system of 56 particles on an 8×8 lattice. The ferromagnet has lower energy for $U \gtrsim 16$. All other wave functions investigated had substantially higher energy than the two shown in the figure.

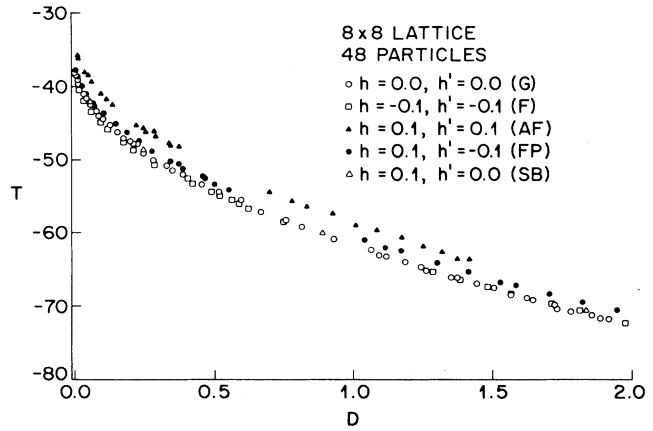

FIG. 12. Plot of the numerical data of the kinetic energy T as a function of the number of doubly occupied sites D for large U on an 8×8 lattice with 48 particles.

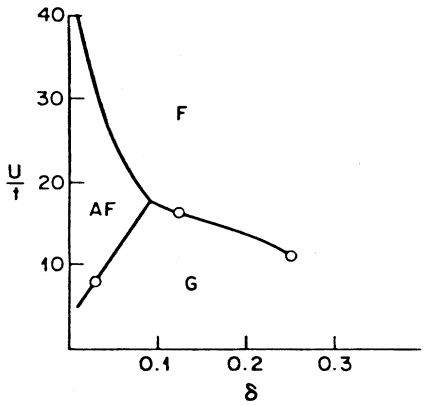

FIG. 14. Schematic phase diagram as a function of U and filling δ. The large ferromagnetic region may reflect paramagnon processes that induce ferromagnetic short-range correlations far from half filling.

Gutzwiller wave function by adding the variational parameter h'. A state with either no magnetic correlations or uniform magnetic correlations was always lower in energy than a state with magnetic correlations that depend on the hole locations. This statement is true for all the systems we studied (8×8 samples with 48, 56, 60, and 62 particles).

One surprising result is that the ferromagnetic region, which has quite a small range in U for the system with 62 particles, grows as the filling is reduced. This result could be the manifestation of paramagnon processes[23] which promote short-range ferromagnetic correlations. These processes are expected to be more important as one moves away from half filling. One can draw the rough phase diagram based on the results we have obtained shown in Fig. 14.

It is possible to examine still more values of h and h',

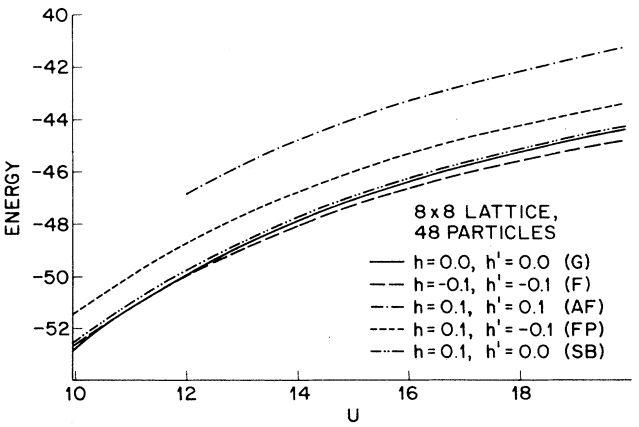

FIG. 13. Plot of the energy vs U for states with $h=h'=0$ (Gutzwiller), $h=h'=0.1$ (antiferromagnet), $h=h'=-0.1$ (ferromagnet), $h=0.1$, $h'=-0.1$ (ferromagnetic polaron), and $h=0.1$, $h'=0$ (spin bag) for a system of 48 particles on an 8×8 lattice. The uniform ferromagnetic state has lowest energy for $U\gtrsim11$, and the Gutzwiller state has lowest energy for $U\lesssim11$. All other wave functions investigated had higher energy.

but the amount of computation required makes a significantly broader search of parameter space unfeasible. In addition, we could have introduced more variational parameters, but once again this process is extremely time consuming, and there are no indications that this step would lead to an energetically favorable small polaron state.

V. INTERPRETATION AND DISCUSSION

In this paper we have investigated the behavior of the Hubbard model over a range of values for the doping and U by generalizing the Gutzwiller variational wave function. In addition to comparing the results for large U with results for the Heisenberg model, we have searched for small magnetic polarons in the intermediate U region.

Our main result is that we find no evidence in favor of the formation of small spin polarons centered around holes in the repulsive Hubbard model on a square lattice. Away from half filling, when U is very large, our results are consistent with a state in which large ferromagnetic polarons are stable. However, when U is reduced so that the ferromagnetic state no longer has the lowest energy, the lowest-energy state is either uniformly antiferromagnetic (very close to half filling) or Gutzwiller. The polaronic wave functions that we examined were not the lowest-energy state for any value of U. We interpret this result to mean that small polarons cost too much energy to be formed in the intermediate U regime.

A. Relation to other work

Given the huge effort expended on the Hubbard model in the last two years, it is natural that our calculation is related to several other investigations. First, several other calculations using Monte Carlo evaluation of trial wave functions have been reported.[24] However, with the exception of Ref. 13, most of them have tended to focus on the large U limit of the Hubbard model. They eliminate the

need for the variational parameter g by using a canonical transformation that is valid for small t/U. This transformation clearly renders the calculation much more efficient when U is large, but in the intermediate U region its validity is not obvious. Also, these investigations did not focus on the possible existence of small magnetic polarons.

Shraiman and Siggia[11] and Trugman[25] have investigated the possibility of pairs of holes binding in an antiferromagnetic matrix. In addition it has been proposed that topological excitations which would destroy the antiferromagnetism could be stabilized by the presence of holes.[26] These calculations are very interesting and yield insight into the large U limit of the Hubbard model. However, these calculations do not specifically address the intermediate U regime.

Exact diagonalization studies[27] have found evidence that the spin-bag picture may apply to one hole in a 4×4 lattice. Further work is necessary to see if this picture applies to more than one hole.

We view our calculation as complementary to these approaches. While the variational aspect of the approach means that we cannot make definitive statements, our results can give information about the behavior of the model over a large range of U.

ACKNOWLEDGMENTS

We would like to thank Ravin Bhatt for suggesting a search for small magnetic polarons and Farid Abraham for making available to us Ref. 15. S.N.C. acknowledges the hospitality of Los Alamos National Laboratory and the Aspen Center for Physics, where part of this work was performed.

APPENDIX A: RELATION BETWEEN DOUBLE OCCUPANCY AND g

This appendix outlines the proof that the number of doubly occupied sites is a monotonic function of the variational parameter g for the variational wave functions studied in this paper.

The wave functions studied here can be written in the form

$$|\psi\rangle=\exp\left(-\eta\sum_i n_{i\uparrow}n_{i\downarrow}\right)\exp\left(-\sum_{\langle j,k\rangle}h_{jk}(n_{j\uparrow}-n_{j\downarrow})(n_{k\uparrow}-n_{k\downarrow})\right)|\phi\rangle\,, \tag{A1}$$

where $|\phi\rangle$ is a product of Slater determinants describing free electrons with $\langle\phi|\phi\rangle=1$, $\langle j,k\rangle$ denotes a sum over nearest-neighbor pairs, and $e^{-\eta}$ is identified with g in Eq. (2.2). Since the term describing spin correlations is independent of η, one can write

$$|\psi\rangle=\exp\left(-\eta\sum_i n_{i\uparrow}n_{i\downarrow}\right)|Y\rangle\,, \tag{A2}$$

with $\langle Y|Y\rangle>0$.

The expectation value of the number of doubly occupied sites $\bar{d}=\sum_i\langle n_{i\uparrow}n_{i\downarrow}\rangle$ is

$$\bar{d}=\frac{\sum_i\langle\psi|n_{i\uparrow}n_{i\downarrow}|\psi\rangle}{\langle\psi|\psi\rangle}=\sum_i\frac{\langle Y|\exp\left(-\eta\sum_k n_{k\uparrow}n_{k\downarrow}\right)n_{i\uparrow}n_{i\downarrow}\exp\left(-\eta\sum_j n_{j\uparrow}n_{j\downarrow}\right)|Y\rangle}{\langle Y|\exp\left(-\eta\sum_j n_{j\uparrow}n_{j\downarrow}\right)\exp\left(-\eta\sum_k n_{k\uparrow}n_{k\downarrow}\right)|Y\rangle}\,, \tag{A3}$$

or $\bar{d}=-\frac{1}{2}\partial/\partial\eta(\ln\langle\psi|\psi\rangle)$. Therefore,

$$\frac{\partial\bar{d}}{\partial\eta}(\ln\langle\psi|\psi\rangle)=2\bar{d}^2-\tfrac{1}{2}\frac{1}{\langle\psi|\psi\rangle}\frac{\partial^2}{\partial\eta^2}\langle\psi|\psi\rangle\,. \tag{A4}$$

Now we evaluate

$$\frac{\partial^2}{\partial\eta^2}\langle\psi|\psi\rangle=\frac{\partial^2}{\partial\eta^2}\langle Y|\exp\left(-2\eta\sum_i n_{i\uparrow}n_{i\downarrow}\right)|Y\rangle=4\sum_l\sum_m\langle Y|\exp\left(-2\eta\sum_i n_{i\uparrow}n_{i\downarrow}\right)n_{l\uparrow}n_{l\downarrow}n_{m\uparrow}n_{m\downarrow}|Y\rangle=4\langle\psi|\left(\sum_i n_{i\uparrow}n_{i\downarrow}\right)^2|\psi\rangle\,. \tag{A5}$$

Therefore, we see that

$$\frac{\partial\bar{d}}{\partial\eta}=2\left[\left\langle\sum_i n_{i\uparrow}n_{i\downarrow}\right\rangle^2-\left\langle\left(\sum_i n_{i\uparrow}n_{i\downarrow}\right)^2\right\rangle\right]=-2\left[\left\langle\left(\sum_i n_{i\uparrow}n_{i\downarrow}-\left\langle\sum_i n_{i\uparrow}n_{i\downarrow}\right\rangle\right)^2\right\rangle\right]. \tag{A6}$$

Thus, it follows that $\partial\bar{d}/\partial\eta\le0$. Equality is attained only when there are no doubly occupied sites. Therefore, $\partial\bar{d}/\partial g\ge0$, and for all $g>0$, $\partial\bar{d}/\partial\eta$ is strictly greater than zero.

APPENDIX B: LIMITATIONS OF THE MONTE CARLO METHOD

In this appendix we demonstrate why wave functions of the form $|\psi\rangle = \exp-(\sum_{\langle i,j\rangle} h_{ij}\mathbf{S}_i\cdot\mathbf{S}_j)\,|\phi\rangle$ cannot be examined using the methods of this paper.

It is simplest to consider first the method for a density projection operator of the Gutzwiller type. To decide whether to accept a hop from site a to site b of the lth electron, one first calculates the ratio of the Slater determinants with the lth electron at site a and at site b. Because one is calculating the ratio rather than the determinant itself, this step requires $O(N)$ operations, where N is the number of particles.[28] The Gutzwiller projection is accounted for by multiplying the square of the modulus of the ratio of determinants by a scalar (either g^2, g^{-2}, or 1) depending on whether the number of doubly occupied sites increases by one, decreases by one, or remains unchanged when the lth electron is moved from a to b.

Now consider the effect of the operator $\mathbf{S}_i\cdot\mathbf{S}_j = S_i^z S_j^z + \frac{1}{2}(S_i^+ S_j^- + S_i^- S_j^+)$ on the noninteracting Fermi sea. The S_z components do not cause particular complication, but the S^+ and S^- terms cause particle interchanges. When two particles are interchanged, a new Slater determinant must be formed. Therefore, instead of one Slater determinant, one has a sum over two Slater determinants to evaluate in both the numerator and denominator (one from the S_z term and the other from a S^+S^- term). When the exponential is expanded, an infinite product of terms which either do or do not interchange particles results. Thus, one is left with the ratios of infinite sums of different Slater determinants to evaluate. For the finite systems with N particles on N sites, there are configurations where the number of terms in the sums grow faster than any power of N.

Therefore, the Monte Carlo evaluation for wave functions of this form is slower by factors that are basically exponential in N.

---

*Current address: Physics Department, University of Illinois, 1110 W. Green St., Urbana, IL 61801.

[1]One choice explicitly includes the inequivalent copper and oxygen sites; see, e.g., C. M. Varma, S. Schmitt-Rink, and E. Abrahams, Solid State Commun. (to be published); V. J. Emery, Phys. Rev. Lett. **58**, 2794 (1987). Of course, many other mechanisms that are not purely electronic have also been proposed.

[2]P. W. Anderson, Science **235**, 1196 (1987); G. Baskaran, Z. Zou, and P. W. Anderson, Solid State Commun. **63**, 973 (1987).

[3]See, e.g., Y. Endoh *et al.*, Phys. Rev. B **37**, 7443 (1988).

[4]Several groups have estimated the electron bandwidth and the Coulomb repulsions on the Cu sites, the O sites, and between the Cu and O sites. One obtains t on the order of a volt and a ratio of the Cu on-site repulsion to t on the order of 8. See, e.g., A. K. McMahan, Richard M. Martin, and S. Satpathy, Phys Rev. B **38**, 6650 (1988); E. B. Stechel and D. R. Jennison, *ibid.* **38**, 4632 (1988); M. S. Hybertsen, M. Schluter, and N. E. Christensen (unpublished). The effective U in a Hubbard model (where each unit cell is treated an entity) is expected to be smaller than that on the Cu only.

[5]J. R. Schrieffer, X. G. Wen, and S. C. Zhang, Phys. Rev. Lett. **60**, 944 (1988).

[6]Y. Nagaoka, Phys. Rev. **147**, 392 (1966).

[7]M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963); M. C. Gutzwiller, Phys. Rev. **134**, A923 (1964); M. C. Gutzwiller, *ibid.* **137**, A1726 (1965).

[8]See, e.g., V. Z. Kresin and S. A. Wolf, in *Novel Superconductivity,* edited by S. A. Wolf and V. Z. Kresin (Plenum, New York, 1987), p. 287.

[9]J. D. Reger and A. P. Young, Phys. Rev. B **37**, 5978 (1988); T. Barnes and E. Swanson (unpublished); E. Gross, Sanchez-Velasco, and E. Siggia (unpublished).

[10]D. Huse and V. Elser, Phys. Rev. Lett. **60**, 2531 (1988); S. Liang, B. Doucot, and P. W. Anderson, *ibid.* **61**, 365 (1988).

[11]B. I. Shraiman and E. D. Siggia, Phys. Rev. Lett. **60**, 740 (1988).

[12]D. Ceperley, G. V. Chester, and M. Kalos, Phys. Rev. B **16**, 3081 (1977). The method was explicitly applied to the Hubbard model by P. Horsch and T. A. Kaplan, J. Phys. C **16**, L1203 (1983).

[13](a) H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **56**, 1490 (1987); (b) H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **56**, 3582 (1987).

[14]William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge Univ. Press, Cambridge, 1986), Chap. 14.

[15]N. Megiddo (unpublished).,

[16]See, e.g., Ceperley *et al.*, Ref. 12.

[17]See, e.g., C. Gros, R. Joynt, and T. M. Rice, Phys. Rev. B **36**, 381 (1987).

[18]W. Metzner and D. Vollhardt, Phys. Rev. Lett. **59**, 121 (1987); Phys. Rev. B **37**, 7382 (1988).

[19]H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **56**, 3570 (1987).

[20]T. A. Kaplan, P. Horsch, and P. Fulde, Phys. Rev. Lett. **49**, 889 (1982).

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

[22]This last result does not necessarily imply that there is no antiferromagnetism for small U, for a small spin-density-wave-type instability would cause the value of h needed to minimize the energy to be very small.

[23]N. Berk and J. R. Schrieffer, Phys. Rev. Lett. **17**, 433 (1966); S. Doniach and S. Engelsberg, *ibid.* **17**, 750 (1966).

[24]See, e.g., Refs. 13, 17, and 21.

[25]S. A. Trugman, Phys. Rev. B **37**, 1597 (1988).

[26]B. Shraiman and E. Siggia, Phys. Rev. Lett. **61**, 467 (1988).

[27]E. Dagotta, J. R. Schrieffer, A. Moreo, and T. Barnes (unpublished).

[28]Specifically, the evaluation of this ratio involves multiplying the new column of the matrix with the corresponding row in the inverse matrix. For a discussion of this point, see Ceperley *et al.*, Ref. 12.