Helicity modulus in the two-dimensional Hubbard model

Denteneer, P.J.H.; An, G.; Leeuwen, J.M.J. van

Helicity modulus in the two-dimensional Hubbard model

P. J. H. Denteneer, Guozhong An,* and J. M. J. van Leeuwen

Instituut-Lorentz, University of Leiden, P.O. Box 9506, 2300 RA Leiden, The Netherlands

(Received 11 June 1992; revised manuscript received 7 October 1992)

The helicity modulus, which is the stiffness associated with a twisted order parameter, for the two-dimensional Hubbard model is calculated for the equivalent cases of (i) attractive on-site interaction (negative U) with arbitrary strength, arbitrary electron density, and zero magnetic field and (ii) repulsive on-site interaction (positive U) with arbitrary strength, at half-filling and in an arbitrary magnetic field. An explicit formula for the helicity modulus is derived using the Bogoliubov-Hartree-Fock approximation. An improved value for the helicity modulus is obtained by performing variational Monte Carlo calculations using a Gutzwiller projected trial wave function. To within a small correction term the helicity modulus is found to be given by $-\frac{1}{8}$ of the average kinetic energy. The variational Monte Carlo calculation is found to increase the value of the helicity modulus by a small amount (about 5% for intermediate values of the interaction strength) compared to the results from the Bogoliubov-Hartree-Fock approximation. In the case of attractive interaction, from a comparison with the Kosterlitz-Thouless relation between critical temperature and helicity modulus, the critical temperature for a Kosterlitz-Thouless transition is calculated and a phase diagram is obtained. An optimal critical temperature is found for an intermediate value of U. We discuss connections of our results with results in the literature on the Hubbard model using the random-phase approximation and quantum Monte Carlo calculations.

I. INTRODUCTION

The Hubbard model[1] is the simplest model to describe correlated electron behavior in a solid, incorporating both the effects due to localized and itinerant (band) electrons. In the past few years, the Hubbard model has attracted much attention because it might play a role in understanding high-temperature superconductivity.[2] Because the materials exhibiting this type of superconductivity are built up out of layers which can be thought of as independent, the relevant physics may be found in the Hubbard model in two spatial dimensions.

The concept of the helicity modulus was introduced by Fisher, Barber, and Jasnow[3] as a measure of the response of a system in an ordered phase to a "twist" of the order parameter. For spin systems, the helicity modulus is also called spin-stiffness constant, whereas for a Bose fluid, the helicity modulus is essentially equivalent to the superfluid density. It is common usage to denote all of these response functions by ρ_s, a practice we will follow in this paper.[4,5] Since the symmetry of the order parameter is important for the character of possible phase transitions, knowledge of the response function corresponding to distorting the order parameter may prove beneficial for understanding the phase diagram of the Hubbard model. For instance, in two spatial dimensions and having a two-component order parameter a Kosterlitz-Thouless phase transition is known to occur.[6] In such a topological phase transition, which entails the binding/unbinding transition of vortex-antivortex pairs, a definite relation exists between ρ_s and the critical temperature T_c^{KT} at which the transition occurs. The purpose of this paper is to calculate ρ_s for the two-dimensional Hubbard model us-

ing both mean-field and variational Monte Carlo methods and study the consequences for the phase diagram.

The Hamiltonian for the Hubbard model is given by

$$\mathcal{H} = -\sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} - \mu \sum_{i\sigma} n_{i\sigma}$$
$$- \frac{h}{2} \sum_{i\sigma} \sigma n_{i\sigma}, \qquad (1)$$

where $c_{i\sigma}^\dagger$ is the operator which creates an electron at site i with spin σ ($\sigma = +1$ for spin-up and $\sigma = -1$ for spin-down electrons), $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$, t_{ij} is the one-electron transfer integral between sites j and i (t_{ij} equals t if i and j are nearest neighbors and 0 otherwise), U is the on-site interaction strength, μ is the chemical potential which controls the total electron density n, and h is a Zeeman magnetic field, which only couples to the difference in densities of the two spin species; $\tilde{\mu} \equiv \mu - U/2 = 0$ corresponds to a half-filled lattice, having (on average) one electron per lattice site ($n = 1$). We will study the model on a square lattice, which is bipartite, meaning that the lattice can be split up in two sublattices such that for all lattice sites all neighboring lattice sites are on the other sublattice.

A well-known canonical transformation exists that for bipartite lattices maps the repulsive (positive-U) Hubbard model onto the attractive (negative-U) model.[7] This "spin-down particle-hole" transformation is given by

$$\bar{c}_{i\uparrow} = c_{i\uparrow}, \quad \bar{c}_{i\downarrow} = (-1)^i c_{i\downarrow}^\dagger. \qquad (2)$$

The effect of this transformation is that the roles of magnetic field and chemical potential are interchanged, because the operator $m_{i\sigma} = n_{i\sigma} - 1/2$ is unaffected for spin

up and changes sign for spin down. More precisely, the parameter combinations $h/2$ and $\mu - U/2$ map onto each other. In other words, the negative-U Hubbard model at half-filling in a field is equivalent to the positive-U Hubbard model in zero field off half-filling. Also, the positive-U Hubbard model at half-filling in a field is equivalent to the negative-U Hubbard model of half-filling in zero field. Although a lot of work has been devoted to doped (i.e., off half-filling) repulsive Hubbard models and doping of simplified versions like the t-J model and $S = \frac{1}{2}$ Heisenberg antiferromagnet, not much is known for sure about this part of the phase diagram. In this paper, we will restrict ourselves mostly to an undoped repulsive Hubbard model in a magnetic field. As a consequence of the above, all results can be applied immediately to the doped attractive Hubbard model (in zero field). In the following, we will discuss the approximations and methods in terms of the repulsive Hubbard model. However, the results we obtain for the helicity modulus have the most interesting consequences when discussed in terms of the attractive Hubbard model.

The remainder of this paper is organized as follows. In Sec. II, we derive the Bogoliubov-Hartree-Fock approximation for the repulsive Hubbard model and calculate ρ_s in this approximation at half-filling for arbitrary magnetic field h and arbitrary values of the interaction strength U. In this approximation and for the case just mentioned, an expression in closed form for ρ_s is obtained containing an integral over the first Brillouin zone of the square lattice. In Sec. III, we give a detailed description of variational Monte Carlo calculations of ρ_s using a Gutzwiller-type trial wave function, which contains the Hartree-Fock wave function as a special case. In Sec. IV, the results for ρ_s from the Hartree-Fock calculation are combined with the Kosterlitz-Thouless theory of XY magnetism to find the critical temperature for superconductivity T_c for arbitrary electron density and interaction strength in the attractive Hubbard model. In Sec. V, we compare our results for ρ_s and T_c to previous work in the literature using other methods, including the random-phase approximation and quantum Monte Carlo calculations.

II. THE HARTREE-FOCK APPROXIMATION

In dealing with a complicated Hamiltonian like the Hubbard Hamiltonian (1), one of the first approximate approaches that already gives a great deal of insight is a mean-field approach. In quantum mechanics such an approach is called the self-consistent-field method or the Hartree-Fock approximation (HFA). This approximation can be formulated as the search for the many-particle wave function that can be written as a product of one-particle wave functions and minimizes the expectation value of $\mathcal{H}$. The method is exact in the case of noninteracting particles. The minimization condition together with the constraint that the one-particle states be orthonormal leads to equations for the one-particle wave functions. A Hartree-Fock approximated Hamiltonian $\mathcal{H}_{\text{HF}}$ is obtained by the condition that in the many-particle state just found it has the same expectation value

as the original Hamiltonian. The approach is also known as the Bogoliubov-Hartree-Fock approximation, expressing the generalization due to Bogoliubov of the original Hartree-Fock approximation, in which only the particle density is taken as self-consistent field. We carry out the above program for the repulsive (positive-U) Hubbard model using the averages of spin densities $n_{i\sigma}$ and spin-raising and spin-lowering operators, $S_i^+ = c_{i\uparrow}^\dagger c_{i\downarrow}$ and $S_i^- = c_{i\downarrow}^\dagger c_{i\uparrow}$, as fields. There is some arbitrariness in this choice, e.g., one could also consider fields associated with operators that do not conserve particle number like the pair-annihilation operator $P_i^- = c_{i\downarrow} c_{i\uparrow}$. Our choice anticipates the relevant ordering, which is antiferromagnetic. In case of attractive interactions (negative U), one would consider averages of P_i^- and its Hermitian conjugate instead of averages of S_i^+ and S_i^-. In the latter case the HFA becomes equivalent to the Bardeen-Cooper-Schrieffer (BCS) approximation to the pairing Hamiltonian in the theory of conventional superconductivity. The attractive feature of the HFA is that its formulation as a variational approach can be extended to finite temperatures. In the following, we refer to all generalized procedures also as HFA. First, we present the zero-temperature HFA in some detail, also for better comparison with the Gutzwiller wave function to be employed in Sec. III. Next, we use the extension to finite temperature (details of which are given in Appendix A) to evaluate the free energy. Finally, the helicity modulus is obtained from the difference in free energy between situations with a twisted order parameter and a constant order parameter.

A. Zero temperature

The Hartree-Fock wave function is

$$|\Psi_{\text{HF}}\rangle = \prod_{k \in \{\text{occ}\}} \gamma^\dagger(k)|0\rangle, \tag{3}$$

where $\gamma^\dagger(k)$ creates a one-electron state with label k, $\{\text{occ}\}$ represents a set of labels corresponding to occupied states, and $|0\rangle$ represents the vacuum state. Since our fields are averages of particle-number conserving operators, the electron creation and annihilation operators can be written as linear combinations of $\gamma^\dagger(k)$ and $\gamma(k)$, respectively:

$$c_{i\sigma} = \sum_k \phi_{i\sigma}(k)\gamma(k). \tag{4}$$

Note that in the case of fields which do not conserve particle number, one has to allow for linear combinations of both $\gamma^\dagger(k)$ and $\gamma(k)$. In order for the transformation (4) to be canonical, orthonormality of the $\phi_{i\sigma}(k)$ is required,

$$\sum_{i\sigma} \phi_{i\sigma}^*(k)\phi_{i\sigma}(k') = \delta_{k,k'}. \tag{5}$$

If we denote the expectation value of an operator A in the Hartree-Fock ground state by $\langle A \rangle$, we have

$$\langle \gamma^\dagger(k)\gamma(k') \rangle = \delta_{k,k'} n(k), \tag{6}$$

where $n(k)$ equals 1 for an occupied state k and 0 for an unoccupied state k.

The ground-state energy in the HFA is

$$\langle \mathcal{H} \rangle = -\sum_{i,j,\sigma,k} t_{ij}\phi_{i\sigma}^*(k)\phi_{j\sigma}(k)n(k) + U\sum_{i,k,k'} n(k)n(k')[|\phi_{i\uparrow}(k)|^2|\phi_{i\downarrow}(k')|^2 - \phi_{i\uparrow}^*(k)\phi_{i\downarrow}(k)\phi_{i\uparrow}(k')\phi_{i\downarrow}^*(k')]$$
$$-\mu\sum_{i,\sigma,k} |\phi_{i\sigma}(k)|^2 n(k) - \frac{h}{2}\sum_{i,\sigma,k} \sigma|\phi_{i\sigma}(k)|^2 n(k). \tag{7}$$

Requiring this energy to be minimal under the constraint of normalized $\phi_{i\sigma}(k)$ [for which we introduce the Lagrange multiplier $\epsilon(k)$] for all k independently, we arrive at the following Hartree-Fock equations for $\phi_{i\sigma}(k)$:

$$-\sum_j t_{ij}\phi_{j\sigma}(k) + \left[U\langle n_{i,-\sigma}\rangle - \frac{h\sigma}{2} - \mu\right]\phi_{i\sigma}(k)$$
$$-U\langle S_i^{-\sigma}\rangle\phi_{i,-\sigma}(k) = \epsilon(k)\phi_{i\sigma}(k). \tag{8}$$

We have used the following equations for the fields in deriving (8):

$$\langle n_{i\sigma}\rangle = \sum_k |\phi_{i\sigma}(k)|^2 n(k), \tag{9}$$

$$\langle S_i^\sigma\rangle = \sum_k \phi_{i\sigma}^*(k)\phi_{i,-\sigma}(k)n(k), \tag{10}$$

which justify the notation of the fields as averages. Using the equations for the fields (9) and (10) in the interaction term in $\langle \mathcal{H} \rangle$ and the HF equations (8) in the three remaining terms, we arrive at

$$\langle \mathcal{H} \rangle = \sum_k \epsilon(k)n(k) - U\sum_i (\langle n_{i\uparrow}\rangle\langle n_{i\downarrow}\rangle - \langle S_i^+\rangle\langle S_i^-\rangle). \tag{11}$$

The Hartree-Fock approximation $\mathcal{H}_{\mathrm{HF}}$ to the original Hamiltonian (1) is now given by

$$\mathcal{H}_{\mathrm{HF}} = \sum_k \epsilon(k)\gamma^\dagger(k)\gamma(k) - U\sum_i(\langle n_{i\uparrow}\rangle\langle n_{i\downarrow}\rangle - \langle S_i^+\rangle\langle S_i^-\rangle). \tag{12}$$

Because of (6) obviously $\mathcal{H}_{\mathrm{HF}}$ and $\mathcal{H}$ have the same expectation value in the Hartree-Fock ground state. By substituting the inverse of (4) in (12) one obtains $\mathcal{H}_{\mathrm{HF}}$ expressed in the creation and annihilation operators for electrons:

$$\mathcal{H}_{\mathrm{HF}} = -\sum_{ij\sigma} t_{ij}c_{i\sigma}^\dagger c_{j\sigma} - \sum_i (\Gamma_i S_i^+ + \Gamma_i^* S_i^-)$$
$$-\sum_{i\sigma}(\mu - U\langle n_{i,-\sigma}\rangle)n_{i\sigma} - \frac{h}{2}\sum_{i\sigma}\sigma n_{i\sigma} + E^0, \tag{13}$$

where we have defined the parameter Γ_i:

$$\Gamma_i = U\langle S_i^-\rangle \tag{14}$$

and E^0 is a constant equal to

$$E^0 = \sum_i (|\Gamma_i|^2/U - U\langle n_{i\uparrow}\rangle\langle n_{i\downarrow}\rangle). \tag{15}$$

The Hamiltonian (13) is simplified considerably by the Hartree-Fock approximation in that the product of four creation and annihilation operators in the interaction term is replaced by a product of only two such operators. Yet the approximated Hamiltonian is sufficiently general to still allow for spin and charge configurations which vary over the lattice.[8]

In the following, we will restrict ourselves to the case of *homogeneous* average spin densities:

$$\langle n_{i\sigma}\rangle = \tfrac{1}{2}(n + \sigma m), \tag{16}$$

where n is the average density, $n = \langle n_{i\uparrow} + n_{i\downarrow}\rangle$, and m is the average magnetization density, $m = \langle n_{i\uparrow} - n_{i\downarrow}\rangle$, neither of which depend on the site i anymore. The averages of spin-raising and spin-lowering operators remain site dependent. In this case, the Hamiltonian further reduces to

$$\mathcal{H}_{\mathrm{HF}} = \sum_{ij\sigma} H_{ij\sigma}c_{i\sigma}^\dagger c_{j\sigma} - \sum_i(\Gamma_i S_i^+ + \Gamma_i^* S_i^-)$$
$$+\sum_i |\Gamma_i|^2/U - \frac{N_s U}{4}(n^2 - m^2), \tag{17}$$

where we have defined

$$H_{ij\sigma} = -t_{ij} - (\bar{\mu} + \sigma\bar{h}/2)\delta_{ij}, \tag{18}$$
$$\bar{h}/2 = h/2 + mU/2, \tag{19}$$
$$\bar{\mu} = \mu - nU/2, \tag{20}$$

and N_s is the number of lattice sites. In this notation, the Hartree-Fock equations (8) are written as

$$\epsilon(k)\phi_{i\uparrow}(k) = \sum_j H_{ij\uparrow}\phi_{j\uparrow}(k) - \Gamma_i\phi_{i\downarrow}(k), \tag{21}$$

$$\epsilon(k)\phi_{i\downarrow}(k) = -\Gamma_i^*\phi_{i\uparrow}(k) + \sum_j H_{ij\downarrow}\phi_{j\downarrow}(k). \tag{22}$$

Note that we have written the equations in a form

that clearly shows the similarity with the Bogoliubov-deGennes (BdG) equations in the BCS theory of superconductivity.[9] If we restrict ourselves further to the case of half-filling, i.e., $\mu = U/2$ and $n = 1$, and perform the transformation $\phi_{i\uparrow}(k) = u_i(k), \phi_{i\downarrow}(k) = (-1)^i v_i(k)$, and $\Gamma_i = (-1)^i \Delta_i$, we exactly recover the BdG equations:

$$\epsilon(k)u_i(k) = \sum_j H_{ij}u_j(k) - \Delta_i v_i(k), \qquad (23)$$

$$\epsilon(k)v_i(k) = -\Delta_i^* u_i(k) - \sum_j H_{ij}v_j(k), \qquad (24)$$

where $H_{ij} = -t_{ij} - (\bar{h}/2)\delta_{ij}$. Equations (23) and (24) have the property that if (u, v) is a solution with energy ϵ, then $(-v^*, u^*)$ is a solution with energy $-\epsilon$. So the $2N_e$ solutions come in pairs of opposite energy. In the ground state only the negative-energy states are occupied, i.e., $n(k) = 1$ for negative-energy states and $n(k) = 0$ for positive-energy states. The ground-state energy at half-filling is given by

$$E_g = \sum_k \epsilon(k)n(k) + \sum_i |\Delta_i|^2/U - \frac{N_sU}{4}(1 - m^2). \quad (25)$$

The ground-state energy can be found by solving Eqs. (23) and (24) under the consistency conditions (9) and (10), which may be rewritten for the case of homogeneous spin densities as equations relating m and U to $\bar{h}/2$ and Δ_i:

$$m = \sum_k [|u_i(k)|^2 - |v_i(k)|^2]n(k), \qquad (26)$$

$$\Delta_i = U\sum_k v_i^*(k)u_i(k)n(k). \qquad (27)$$

B. Finite temperatures

The generalization of the HFA as presented above to finite temperature can be made. Details are given in Appendix A. Here we use the principal results: the role of the energy is taken over by the free energy, which in the HFA is given by

$$F_{\text{HF}} = -\frac{1}{\beta}\ln\left(\text{Tr}e^{-\beta\mathcal{H}_{\text{HF}}}\right), \qquad (28)$$

and the occupation of levels, $n(k)$, is for finite T given by the Fermi-Dirac distribution:

$$n(k) = \frac{1}{1 + e^{\beta\epsilon(k)}}, \qquad (29)$$

where $1/\beta = k_BT$. For the free energy we have using (12)

$$F_{\text{HF}} = E^0 - \frac{1}{\beta}\sum_k \ln\left[1 + e^{-\beta\epsilon(k)}\right]. \qquad (30)$$

Using the property that the complete set of states k can be split into two equally large sets having opposite energies and substituting (15) for E^0, we arrive for half-filling at

$$F_{\text{HF}} = -\frac{2}{\beta}\sum_k{}'\ln\{2\cosh[\beta\epsilon(k)/2]\}$$
$$+ \sum_i |\Delta_i|^2/U - \frac{N_sU}{4}(1 - m^2). \qquad (31)$$

The prime on the sum over k denotes that it is only over negative-energy states (we have taken together the two terms with opposite energies). Substituting $n(k)$ into the consistency conditions (26) and (27) and again rewriting to a sum involving only negative-energy states gives

$$m = \sum_k{}'[|v_i(k)|^2 - |u_i(k)|^2]\tanh[\beta\epsilon(k)/2], \qquad (32)$$

$$\Delta_i = -U\sum_k{}'v_i^*(k)u_i(k)\tanh[\beta\epsilon(k)/2]. \qquad (33)$$

C. Calculation of the helicity modulus

In order to calculate the helicity modulus ρ_s we have to obtain, by definition,[3] the incremental free energy resulting from twisting the order parameter Δ_i. First, we present the solution of the BdG equations for the case that $\Delta_i = \Delta$ is constant over the lattice. Then it is seen that if Δ_i fluctuates over the lattice only through its phase the solutions are very similar to the case of constant Δ_i. Finally, we give the solutions for the specific phase fluctuation corresponding to a twist of the order parameter:

$$\Delta_j = |\Delta|e^{2i\mathbf{q}\cdot\mathbf{r}_j}. \qquad (34)$$

For constant Δ_i, if periodic boundary conditions are imposed, the solutions are of the form

$$u_j(k) = \frac{u_{\mathbf{k}}}{\sqrt{N_s}}e^{i\mathbf{k}\cdot\mathbf{r}_j}, \quad v_j(k) = \frac{v_{\mathbf{k}}}{\sqrt{N_s}}e^{i\mathbf{k}\cdot\mathbf{r}_j}, \qquad (35)$$

where the labels k have been replaced by the set of vectors $\mathbf{k}$ in reciprocal space [$\mathbf{k} = (\pi/L)(n_x, n_y)$ with $n_x, n_y = -L, \cdots, L-1$ and L the number of lattice sites along one side of the square lattice]. In the case of half-filling that we are considering, we have one electron per lattice site (on average; $n = 1$), $N_e = N_s$, and all of the $\mathbf{k}$ vectors specified above correspond to exactly one term in the (primed) sums over k in (31)–(33). Therefore, in the limit of infinitely large lattices, sums over $\mathbf{k}$ become integrals over the full first Brillouin zone (BZ) of the square lattice ($k_x, k_y \in [-\pi, \pi]$). The BdG equations reduce to a set of two equations for $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ for each $\mathbf{k}$ separately. The resulting energy levels are

$$\epsilon(\mathbf{k}) = \pm\sqrt{H^2(\mathbf{k}) + |\Delta|^2}, \qquad (36)$$

$$H(\mathbf{k}) = \sum_j H_{ij}e^{i\mathbf{k}\cdot(\mathbf{r}_j - \mathbf{r}_i)}$$
$$= -2t[\cos(k_x) + \cos(k_y)] - \bar{h}/2. \qquad (37)$$

Note that $H(\mathbf{k})$ is not site dependent. From the equations for $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ one easily derives [using $|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1$, which follows from (5)]

$$|v_{\mathbf{k}}|^2 - |u_{\mathbf{k}}|^2 = -\frac{H(\mathbf{k})}{\epsilon(\mathbf{k})}, \tag{38}$$

$$v_{\mathbf{k}}^* u_{\mathbf{k}} = -\frac{\Delta}{2\epsilon(\mathbf{k})}. \tag{39}$$

The consistency conditions are

$$m = -\frac{1}{N_s} \sum_{\mathbf{k}} \frac{H(\mathbf{k})}{\epsilon_0(\mathbf{k})} \tanh[\beta\epsilon_0(\mathbf{k})/2], \tag{40}$$

$$\frac{1}{U} = \frac{1}{2N_s} \sum_{\mathbf{k}} \frac{1}{\epsilon_0(\mathbf{k})} \tanh[\beta\epsilon_0(\mathbf{k})/2]. \tag{41}$$

By $\epsilon_0(\mathbf{k})$ we (arbitrarily) denote the positive square root in (36). The consistency equations allow for direct transformation between the parameters U/t and h [via m and (19)] occurring in the original Hamiltonian and the parameters Δ and $h/2$ occurring in the Hartree-Fock solution.

If Δ_i only differs from site to site by a phase factor,

$$\Delta_j = |\Delta| e^{i\phi_j}, \tag{42}$$

the transformation

$$\tilde{u}_j(k) = u_j(k) e^{-i\phi_j/2}, \quad \tilde{v}_j(k) = v_j(k) e^{i\phi_j/2} \tag{43}$$

leaves the BdG equations invariant except for a phase in the nondiagonal matrix elements H_{ij}:

$$\tilde{H}_{ij} = H_{ij} e^{i(\phi_j - \phi_i)/2}. \tag{44}$$

For the special choice $\phi_j = 2\mathbf{q} \cdot \mathbf{r}_j$, corresponding to a twisted order parameter Δ [see (34)], the problem becomes translationally invariant again, since the hopping matrix elements t_{ij} will only depend on the direction of hopping, not on the sites involved. The BdG equations again reduce to two equations for $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ for each $\mathbf{k}$ separately. The resulting energy levels now are

$$\epsilon_q^{\pm}(\mathbf{k}) = \chi_q(\mathbf{k}) \pm \sqrt{P_q^2(\mathbf{k}) + |\Delta|^2}, \tag{45}$$

where

$$\chi_q(\mathbf{k}) = \tfrac{1}{2}[H(\mathbf{k} + \mathbf{q}) - H(\mathbf{k} - \mathbf{q})], \tag{46}$$

$$P_q(\mathbf{k}) = \tfrac{1}{2}[H(\mathbf{k} + \mathbf{q}) + H(\mathbf{k} - \mathbf{q})].$$

Note that for $q = 0$ the energies $\epsilon_q^{\pm}(\mathbf{k})$ reduce to the energies $\epsilon(\mathbf{k})$ in (36). We require the negative-energy solutions, which for small $|\mathbf{q}|$ will be $\epsilon_q^-(\mathbf{k})$; the corresponding $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are

$$(u_{\mathbf{k}}, v_{\mathbf{k}}) = [\sqrt{\tfrac{1}{2}(1 - \alpha_{\mathbf{k}})}, \sqrt{\tfrac{1}{2}(1 + \alpha_{\mathbf{k}})}],$$

$$\alpha_{\mathbf{k}} = \frac{P_q(\mathbf{k})}{\sqrt{P_q^2(\mathbf{k}) + |\Delta|^2}}. \tag{47}$$

Substituting the solution (45) in the expression for the free energy (31) and making an expansion for small q,[10] we obtain $F(q) = F(0) + 2\rho_s N q^2 + \mathcal{O}(q^4)$ with

$$\rho_s = -\frac{t}{N} \sum_{\mathbf{k}} \left[\frac{H(\mathbf{k})}{2\epsilon_0(\mathbf{k})} \cos(k_x) \tanh[\beta\epsilon_0(\mathbf{k})/2] \right.$$
$$\left. + \frac{t\beta \sin^2(k_x)}{2\cosh^2[\beta\epsilon_0(\mathbf{k})/2]} \right]. \tag{48}$$

Using (48), ρ_s can be calculated for every combination of U/t and h as a function of temperature T. In practice, this is slightly complicated because U and h are determined by the BZ summations (40) and (41). As a result, for a chosen combination of U and h one has to adjust $\bar{h}$ and Δ for every β to obtain the same U and h. For $\bar{h} = 0$ one always has $m = 0$, because contributions from $\mathbf{k}$ and $(\pi, \pi) - \mathbf{k}$ cancel. If Δ (and therefore U) becomes very small, a very fine mesh in the BZ has to be used to converge the summations in (40) and (41). This corresponds to a very large lattice in real space and is consistent with large coherence lengths for small Δ.

In Fig. 1, we show the typical behavior of ρ_s. As a function of T [for fixed U/t and h; Fig. 1(a)], ρ_s starts from some finite value at $T = 0$ and vanishes at $T_c(\mathrm{HF})$, which is the temperature for which (40) and (41) hold with $\Delta = 0$. For small values of U/t, $T_c(\mathrm{HF})$ decreases exponentially with increasing t/U, whereas for large values of U/t (and therefore Δ) $T_c(\mathrm{HF})$ increases linearly with U. If $U = 0$, $\Delta = 0$ and therefore $\rho_s = 0$ for all T, as is seen by noting that the summand in (48) is a derivative with respect to p_x of a periodic function of p_x. As a function of U/t, for finite T (T and h fixed), ρ_s decreases (as t^2/U) with increasing U and with decreasing U drops to zero at the point where Δ goes to zero because the finite value of T equals $T_c(\mathrm{HF})$ for the value of U concerned. For $T = 0$ and $h = 0$, this value of U is zero; in the limit of U going to zero, ρ_s then goes to the finite value of $2t/\pi^2$, while being strictly zero for $U = 0$ [Fig. 1(b)]. The fact that ρ_s approaches a finite value in the limit $U \to 0$ can be understood by realizing that although Δ becomes very small in this limit, the coherence length becomes very large and a finite value for the stiffness results. Finally, ρ_s decreases monotonically as a function of h and vanishes at some critical field [Fig. 1(c)].

By the mapping between positive- and negative-U Hubbard models given in the Introduction, the above properties of ρ_s can be translated immediately into properties of ρ_s in a negative-U Hubbard model with arbitrary density and zero field. For instance, in this case ρ_s vanishes at some critical value of the chemical potential, which corresponds to zero density (without electrons there is no gap parameter and therefore no stiffness).

In Sec. IV, ρ_s as calculated in the HFA will be used to obtain the critical temperature for superconductivity for a Kosterlitz-Thouless transition in the negative-U Hubbard model for arbitrary interaction strength and density. In Sec. V, we will compare our result for ρ_s with results on the Hubbard model using different approaches. As we will show, ρ_s as given in (48) can be obtained in a specific limit of the dynamic transverse susceptibility $\chi^{+-}(\mathbf{q}, \omega)$
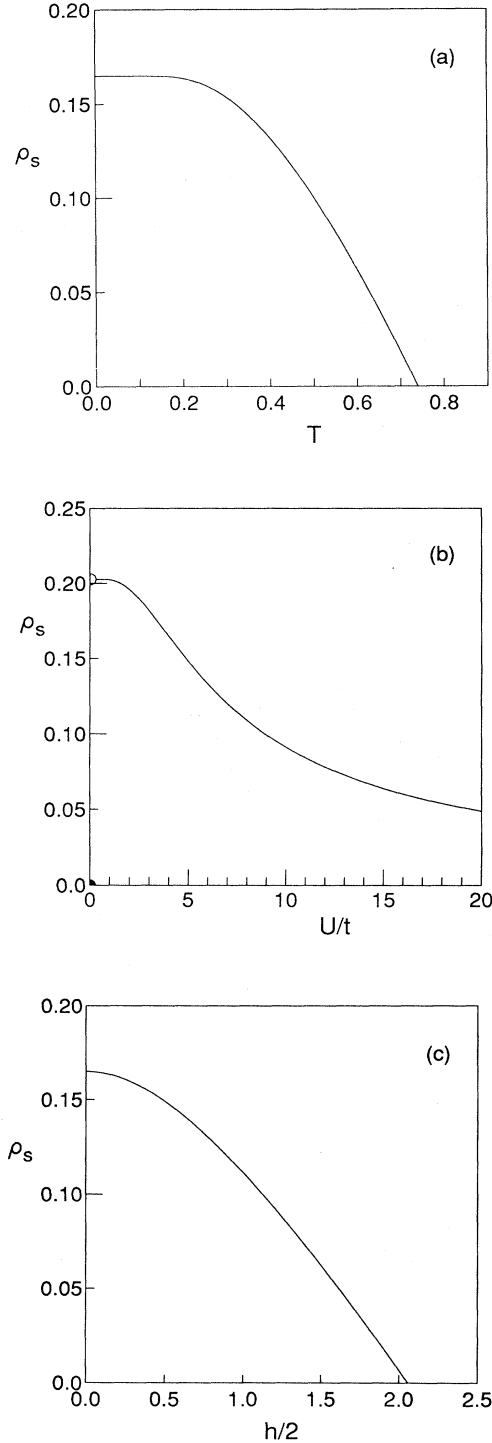
FIG. 1. The helicity modulus ρ_s (in units of t) for the repulsive half-filled Hubbard model in the Hartree-Fock approximation as a function of temperature T, interaction strength U/t, and Zeeman magnetic field h: (a) As a function of T for fixed $U/t = 4$ and fixed $h = 0$ [ρ_s vanishes at the Hartree-Fock critical temperature T_c(HF)], (b) as a function of U/t for fixed $T = 0$ and fixed $h = 0$ (for $U/t = 0$, ρ_s is zero by definition, however in the limit of vanishing U/t, ρ_s approaches the finite value of $2t/\pi^2$), (c) as a function of h for fixed $T = 0$ and fixed $U/t = 4$ (ρ_s vanishes at a critical field h_c).

as calculated in the random-phase approximation by several authors.[11,12]

III. GUTZWILLER VARIATIONAL MONTE CARLO METHOD

If the ground-state wave function for the many-electron system is approximated by an antisymmetrized product of one-electron wave functions, as is done in the Hartree-Fock approximation (HFA) of Sec. II [see (3)], the average of the double occupancy operator $n_{i\uparrow}n_{i\downarrow}$ is simply $n^2/4$ for a paramagnetic state. For large U, the potential energy of $Un^2/4$ will be large and the system is forced to build up magnetic correlations to avoid large positive energies. Therefore, the HFA favors magnetic correlations. Gutzwiller has pointed out[13] that double occupancy may also be avoided by introducing paramagnetic correlations into the wave function. In the wave function Gutzwiller proposed, this was accomplished by multiplying the *noninteracting* wave function by a factor g^D, with $0 < g < 1$ and D the number of doubly occupied sites. The Gutzwiller factor g is seen as a variational parameter in the search for the ground-state wave function for the interacting system. In this spirit and slightly generalizing the approach of Yokoyama and Shiba[14] (which already is a generalization of the Gutzwiller approach), we use the following wave function:

$$|\Psi_{\mathrm{GW}}\rangle = \prod_i [1 - (1 - g)n_{i\uparrow}n_{i\downarrow}]|\Psi_{\mathrm{HF}}\rangle. \tag{49}$$

In this way, the extra Gutzwiller correlations are built in directly into the Hartree-Fock (HF) wave function. The Gutzwiller-type wave function (49) (to be called simply Gutzwiller wave function in the following) is now a variational wave function containing three parameters g, Δ, and $\bar{h}$:

$$|\Psi_{\mathrm{GW}}\rangle = g^D|\Psi_{\mathrm{HF}}(\Delta, \bar{h})\rangle. \tag{50}$$

Ψ_{HF} is a Slater determinant made up out of the one-electron wave functions $\phi_{i\sigma}$ which, via the u_i and v_i functions, depend on Δ and $\bar{h}$. As in the preceding section, calculations are restricted to half-filling; therefore no parameter $\bar{\mu}$ appears. In the HFA, Δ and $\bar{h}$ are fixed by the choice of U and h via the consistency conditions. In the variational approach, they are parameters to be optimized so as to minimize the energy; also the relation (19) between $\bar{h}$ and h need no longer be fulfilled. The extension with respect to the approach of Yokoyama and Shiba lies in the fact that we allow for nonzero magnetic fields and consequently have one more variational parameter. Since through the mapping onto a negative-U Hubbard model nonzero field amounts to going off half-filling in that model, this freedom may well be of interest.

It is clear that setting $g = 1$ reduces our Gutzwiller wave function to the Hartree-Fock wave function of Sec. II. This fact serves as a nontrivial test for both the variational and Monte Carlo (MC) part of our calculation. In the variational part of the calculation, fixing g

at 1 should result in values for Δ and $\bar{h}$ consistent with the chosen values of U and h via the HF consistency conditions (40) and (41). In the Monte Carlo part of the calculation, using $g = 1$ and the HF values for Δ and $\bar{h}$ should result in a ground-state energy that equals the HF energy calculated through (25) to within the statistical error bar. Both tests were successfully performed.

Since the Gutzwiller wave function is an improvement over the HF wave function (the larger variational freedom will result in a lower energy), one may hope that a calculation of ρ_s using the improved wave function will result in a ρ_s more closely resembling the exact ρ_s. Below we describe variational Monte Carlo calculations of ρ_s using the Gutzwiller wave function (50).

A. Variational Monte Carlo method with determinantal wave function

Although the variational Monte Carlo (VMC) method we use is very similar to the one of Refs. 14 and 15, in this section we will give a concise description which at the same time provides the framework in which the calculation of the helicity modulus takes place.

The Monte Carlo method is an efficient method to compute expectation values given some wave function for the ground state. For instance, one is interested in the ground-state energy E_g ($T = 0$) for some Hamiltonian $\mathcal{H}$:

$$E_g = \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{1}{W} \sum_R |\Psi(R)|^2 \frac{\mathcal{H}\Psi(R)}{\Psi(R)}, \qquad (51)$$

with $W = \sum_R |\Psi(R)|^2$ and R denoting a specific configuration of the electrons in position and spin space. E_g is written as the average of $\mathcal{H}\Psi(R)/\Psi(R)$ over configurations R with probability distribution function $p(R) = |\Psi(R)|^2/W$. The average is taken over a finite number of configurations R selected from a random walk through configuration space according to the Metropolis algorithm: a MC move from R_{n-1} to R_n is always accepted if $p(R_n)/p(R_{n-1}) > 1$; if $p(R_n)/p(R_{n-1}) < 1$ such a move is only accepted if $p(R_n)/p(R_{n-1}) > r$, where r is a random number taken from a set uniformly distributed between 0 and 1. In this way, only the more important configurations for the average are sampled, whereas at the same time getting stuck in a locally favorable configuration is avoided. As a matter of course, also expectation values of other operators than the Hamiltonian may be evaluated in this way.

If the wave function $\Psi(R)$ can be written as a determinant, as is the case for the Gutzwiller wave function (50), a particularly efficient method to compute the ratios $p(R')/p(R)$ exists. In our case, $\Psi(R)$ is of the form

$$\Psi(R) = g^{D(R)} d(R) = g^{D(R)} \det(D_{ij}), \qquad (52)$$

where $D(R)$ is the number of doubly occupied states in configuration R and the matrix elements D_{ij} in the Slater determinant $d(R)$ are related to the one-electron wave functions $\phi_{r\sigma}$ in (4) by

$$D_{ij} = \phi_{\xi_j}(k_i), \qquad (53)$$

where we have introduced the composite index ξ_j indicating both the position r_j and the spin σ_j of electron j. The index i denotes the one-electron energy labeled k_i. In order to compute $p(R')/p(R)$, we need to compute

$$\frac{\Psi(R')}{\Psi(R)} = g^{D(R')-D(R)} \frac{d(R')}{d(R)}. \qquad (54)$$

In performing the random walk through configuration space, we allow for two types of MC moves: (i) *flip*, a change of the spin of one electron, and (ii) *hop*, a jump of one electron to a neighboring site. For both moves it holds that the determinants $d(R')$ and $d(R)$ differ only in the one column j corresponding to the electron of which either the spin or position was changed. In that case, the ratio of determinants is simply given by

$$\frac{d(R')}{d(R)} = \sum_k \bar{D}_{kj}(R) D_{jk}(R'), \qquad (55)$$

where j denotes the column that changed by the MC move and $\bar{D}$ denotes the inverse of the transpose of D:

$$\sum_k \bar{D}_{ik}(R) D_{jk}(R) = \delta_{ij} \quad \text{for all } R. \qquad (56)$$

The efficiency of the method lies in the fact that the matrix D only has to be inverted at the beginning of the random walk in order to obtain $\bar{D}$; $\bar{D}$ can be updated by matrix multiplication (instead of inversion) after every accepted move using the formula given in Ref. 16.

In principle, the optimal variational parameters are to be found by performing a MC calculation of the energy for a large set of parameter combinations $(g, \Delta, \bar{h})$. We have used an alternative optimization method proposed by Umrigar, Wilson, and Wilkins[17] for wave functions with a large number of parameters. In the spirit of this method, we minimize the energy averaged over a fixed set of configurations. This procedure has the advantage of requiring less computing time and of using correlated samples to arrive at the optimal parameters. We typically use 500 configurations, whereas a MC run samples of the order of 10^5 configurations to arrive at energies with sufficiently small statistical errors to distinguish between different parameter combinations. To avoid getting stuck in local minima as much as possible the above optimization is repeated a number of times using different sets of fixed configurations. As another check on the result of the optimization full MC calculations are performed not only with the set of optimal parameters, but also using slight deviations from this optimal set.

B. Calculation of the helicity modulus

In the preceding section, we discussed how to evaluate the ground-state energy using a Gutzwiller trial wave function with optimized parameters. In order to do this the expectation value of the Hamilton operator is computed by MC integration over configuration space. It is not immediately obvious of which operator one should evaluate the expectation value in order to obtain the helicity modulus ρ_s. In this section, we show that ρ_s equals

$-\frac{1}{8}$ of the kinetic energy plus a correction term.

To calculate ρ_s in the ground state, i.e., for $T = 0$, by definition one must compute the energy difference $E(q) - E(0)$, where $E(q)$ and $E(0)$ are the ground-state energies for the situations where a twist of the order parameter (34) is present or absent, respectively. In the language of Monte Carlo calculations, $E(q)$ is the minimum over a set of wave functions ψ_q, which depend on q because of the constraint that the ground-state average of the operator $\mathcal{A}_j = (-1)^j c_{j\downarrow}^\dagger c_{j\uparrow}$ has a wavelike' fluctuation over the lattice:

$$\langle \psi_q | \mathcal{A}_j | \psi_q \rangle = |\Delta| e^{2i\mathbf{q}\cdot\mathbf{r}_j}, \tag{57}$$

$$E(q) = \min \frac{\langle \psi_q | \mathcal{H} | \psi_q \rangle}{\langle \psi_q | \psi_q \rangle}. \tag{58}$$

Note that $\mathcal{H}$ is q independent and the q dependence of $E(q)$ comes solely from the wave function. One way to proceed is to take for ψ_q the wave function for the "untwisted" problem (g^D times Slater determinant) and replace in the Slater determinant the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ by their q-dependent counterparts in the twisted problem, given by (47). Although this approach is feasible, it is not very convenient for two reasons: the first is that in order to compute ρ_s, results for two separate MC calculations each having a statistical error have to be subtracted, leading to large uncertainties in the resulting difference. The other reason is that the energy difference has to be found in the limit of small q. Since these calculations are performed on finite lattices and we have to impose periodic (or antiperiodic) boundary conditions, it is questionable whether q can be chosen small enough to extract a reliable value of ρ_s. We now present an alternative procedure, which is akin to a procedure by Kohn[18] and which allows one to extract the small-q limit of the energy difference directly from a single MC calculation. In this way, we arrive at the MC equivalent ($T = 0$ only) of the Hartree-Fock formula for ρ_s (48), which came directly from the small-q limit of the difference of *free* energies.

Suppose a unitary transformation $\mathcal{U}$ can be found that "gauges away" the wavelike constraint (57): i.e.,

$$|\psi_q'\rangle = \mathcal{U}|\psi_q\rangle \tag{59}$$

such that

$$\mathcal{A}_j' = \mathcal{U}\mathcal{A}_j\mathcal{U}^\dagger = \mathcal{A}_j e^{2i\mathbf{q}\cdot\mathbf{r}_j}. \tag{60}$$

Then the problem (57) and (58) transforms to

$$\langle \psi_q' | \mathcal{A}_j | \psi_q' \rangle = |\Delta|, \tag{61}$$

$$E(q) = \min \frac{\langle \psi_q' | \mathcal{H}'(q) | \psi_q' \rangle}{\langle \psi_q' | \psi_q' \rangle}, \tag{62}$$

where $\mathcal{H}'(q) = \mathcal{U}\mathcal{H}\mathcal{U}^\dagger$ is q dependent because of the q dependence of $\mathcal{U}$. Indeed, such a transformation exists:

$$\mathcal{U} = \prod_j e^{in_j\sigma\mathbf{q}\cdot\mathbf{r}_j}, \tag{63}$$

where n_j is the operator that counts the number of electrons at site j. One easily verifies that only the kinetic-energy (or hopping) term in the Hamilton operator is not

invariant under the unitary transformation (63), since $\mathcal{U}$ commutes with $n_{i\uparrow}$ and $n_{i\downarrow}$ for all i. The hopping matrix element acquires a phase which only depends on the direction of the hop but not on the sites to or from which the electron hops:

$$t_{ij}' = t_{ij} e^{i\sigma\mathbf{q}\cdot(\mathbf{r}_i - \mathbf{r}_j)}. \tag{64}$$

Because the Hamiltonian $\mathcal{H}'(q)$ is translationally invariant in the sense just described and furthermore the constraint (61) is translationally invariant, we can use for $|\psi_q'\rangle$ the Gutzwiller wave function for the "untwisted" problem (g^D times Slater determinant) and replace in the Slater determinant the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ by their q-dependent counterparts in the twisted problem [given by (47)]. The difference with the original problem is that we have traded the position-dependent constraint and q-independent Hamiltonian for a position-independent constraint and a q-dependent Hamiltonian. The q dependence of the Hamiltonian, however, is a very simple one [see (64)].

We can formally expand the wave function and Hamiltonian for small q:

$$\psi_q'(R) = \Psi_0(R) + q^2\Psi_2(R) + \mathcal{O}(q^4), \tag{65}$$

$$\psi_q'(R)\mathcal{H}'(q)\psi_q'(R') = H_0(R, R') + qH_1(R, R') \\ + q^2H_2(R, R') + \mathcal{O}(q^3), \tag{66}$$

$$E(q) = E_0 + 2\rho_s N_s q^2 + \mathcal{O}(q^4). \tag{67}$$

The odd powers of q do not appear in the wave function because the Hartree-Fock and Gutzwiller wave functions are even in $\mathbf{q}$, as can be seen from (47). The operator $\mathcal{H}_1$ is a spin-current operator; its expectation value in the HF and GW ground states is zero. Therefore, the term proportional to q in the energy vanishes. The formula for ρ_s now follows readily:

$$2N_s\rho_s = \frac{1}{W_0} \sum_{R,R'} \Psi_0^*(R)H_2(R, R')\Psi_0(R') \\ + \frac{1}{W_0} \sum_{R,R'} [\Psi_2^*(R)\{H_0(R, R') \\ - E_0\delta(R, R')\}\Psi_0(R') + \text{c.c.}], \tag{68}$$

where c.c. denotes the complex conjugate of the preceding term,

$$E_0 = \frac{1}{W_0} \sum_{R,R'} \Psi_0^*(R)H_0(R, R')\Psi_0(R') \tag{69}$$

and

$$W_0 = \sum_R |\Psi_0(R)|^2. \tag{70}$$

By choosing $\mathbf{q}$ at an angle of $\pi/4$ with an arbitrary bond of the square lattice, we find from (64) the small-q expansion of t_{ij}':

$$t'_{ij} = t_{ij}[1 + i\sigma \mathbf{q} \cdot (\mathbf{r}_i - \mathbf{r}_j) - q^2/4] + \mathcal{O}(q^3). \quad (71)$$

Therefore, we have

$$H_2(R, R') = -\tfrac{1}{4}T(R, R'). \quad (72)$$

Defining the average of F over configurations R with probability distribution $|\Psi_0(R)|^2$ as

$$\langle\langle F(R)\rangle\rangle = \frac{1}{W_0}\sum_R |\Psi_0(R)|^2 F(R), \quad (73)$$

the following formula for ρ_s emerges:

$$\rho_s = -\frac{1}{8N_s}\langle\langle T_L(R)\rangle\rangle$$
$$+\frac{1}{N_s}\left\langle\left\langle \mathrm{Re}\left\{\frac{\Psi_2(R)}{\Psi_0(R)}\right\}[E_L(R) - E_0]\right\rangle\right\rangle, \quad (74)$$

where *local* (i.e., referring to one configuration) kinetic and total energies in the untwisted state are defined:

$$T_L(R) = \sum_{R'}\frac{T(R, R')\Psi_0(R')}{\Psi_0(R)}$$

and $\quad (75)$

$$E_L(R) = \sum_{R'}\frac{H_0(R, R')\Psi_0(R')}{\Psi_0(R)}.$$

Note that $E_0 = \langle\langle E_L(R)\rangle\rangle$. Therefore, ρ_s has been written as $-\tfrac{1}{8}$ of the average of the kinetic energy per site in the untwisted state plus a correction term. In order to calculate the correction term, one needs E_0, which can be obtained from a MC calculation using the Gutzwiller wave function without q dependence, as well as the sum of ratios of determinants $\Psi_2(R)/\Psi_0(R)$. $\Psi_2(R)$ is a sum of N_e determinants (N_e is the number of electrons) in which in each determinant another single column is changed with respect to $\Psi_0(R)$. The change entails replacing the q^0 term of $u_{\mathbf{k}}$ or $v_{\mathbf{k}}$ (depending on the spin of the corresponding electron) by the q^2 term. The procedure to calculate a ratio of determinants which only differ in one column was discussed above for the case in which the change corresponded to a new configuration (hop or spin flip of an electron). The q^2 terms of $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$, $u_{\mathbf{k}}^{(2)}$ and $v_{\mathbf{k}}^{(2)}$, respectively, can be found from (47):

$$u_{\mathbf{k}}^{(2)} = \frac{t(\mathbf{k})}{16u_{\mathbf{k}}(q=0)}\frac{|\Delta|^2}{\epsilon_0^3(\mathbf{k})},$$
$$\quad (76)$$
$$v_{\mathbf{k}}^{(2)} = -\frac{t(\mathbf{k})}{16v_{\mathbf{k}}(q=0)}\frac{|\Delta|^2}{\epsilon_0^3(\mathbf{k})},$$

where

$$t(\mathbf{k}) = -2t[\cos(k_x) + \cos(k_y)], \quad (77)$$

$$\epsilon_0(\mathbf{k}) = \sqrt{[t(\mathbf{k}) - \bar{h}/2]^2 + |\Delta|^2}. \quad (78)$$

We note here that in actual calculations the correction term in the formula for ρ_s always turns out to be almost zero. Therefore, for the HF and GW wave function, ρ_s is practically given by $-\tfrac{1}{8}$ of the kinetic energy per site. For the exact wave function, ρ_s would contain a contribution from the spin-current operator as well.[28]

C. Results of VMC calculations
for the helicity modulus

All of our calculations presented here have been performed on an 8×8 lattice at half-filling, i.e., with 64 electrons. Below we argue that for the values of U/t considered the results do not change in going to larger lattices. Because of the small size of the lattice one does have to be careful in the choice of boundary conditions.[14,20] Our choice of periodic boundary condition in the x direction and antiperiodic boundary condition in the y direction conveniently removes the degeneracy for the HF wave function. After having found optimal values of g, Δ, and $\bar{h}$, as described in Sec. III A, we perform a few long MC runs of 1.2×10^6 steps (a few hundred MC steps are done for thermalization purposes, i.e., to remove dependence on the start configuration). The quantities we are interested in, like total energy, sublattice magnetization, kinetic energy, and magnetization in the z direction, are "measured" with a frequency of once every 10 steps. These values are gathered in groups of 3000 members. For each group, the averages and standard deviations of all quantities of interest are computed. The group averages (40 in total) can be considered as independent measurements of which we compute the "grand" average and corresponding standard deviation. Using this procedure and the numbers mentioned, the statistical errors become sufficiently small. We have typically performed eight independent MC runs as described above for every combination of U/t and h, with slightly differing parameter sets $(g, \Delta, \bar{h})$. We have found that values for the total energy per site E/N_s and helicity modulus ρ_s are insensitive to slightly perturbing the parameter set, but that the sublattice magnetization $\langle S^+\rangle$ and the magnetization in the z direction $S_z \equiv m/2$ are very sensitive to the parameters Δ and $\bar{h}$, respectively.

In Tables I and II, we compile results of VMC calculations for both the ground-state energy per site E/N_s and the helicity modulus ρ_s for a set of values of the interaction strength U/t in zero field (Table I) and in a small field $h/t = 0.34$ (Table II). We also compare with the corresponding results from the Hartree-Fock calculation,[19] which for meaningful comparison have also been calculated on 8×8 lattices and with periodic (antiperiodic) boundary conditions in the x direction (y direction). In the Hartree-Fock calculation, it is easy to investigate the finite-size effect. We find that for the values of U in the tables the effect of going to larger lattices occurs in decimals not displayed in the tables. For $U < 3$ the size effect becomes significant. We expect the finite-size effect to be roughly the same in the VMC calculation as in the HF

TABLE I. Variational Monte Carlo (VMC) results for total energy per site E/N_s, helicity modulus ρ_s, and sublattice magnetization $\langle S^+ \rangle$ for the repulsive Hubbard model at half-filling in zero magnetic field ($h = 0$) compared to corresponding Hartree-Fock (HF) results (below VMC results between brackets). Both VMC and HF results are obtained on 8×8 lattices with periodic boundary conditions in the x direction and antiperiodic boundary conditions in the y direction. The optimal parameters g and Δ used in the Gutzwiller wave function (see text) are also given. For the HF wave function $g = 1$. All quantities except g are in units of the hopping integral t.

U	$-E/N_s$	ρ_s	$\langle S^+ \rangle$	g	Δ
3	0.993	0.188	0.24	0.69	0.32
	(0.948)	(0.183)	(0.28)		(0.84)
4	0.842	0.172	0.30	0.64	0.48
	(0.797)	(0.165)	(0.34)		(1.38)
5	0.724	0.156	0.36	0.61	0.78
	(0.682)	(0.148)	(0.39)		(1.93)
7	0.554	0.127	0.42	0.53	1.10
	(0.522)	(0.120)	(0.43)		(3.03)
8	0.493	0.117	0.43	0.49	1.16
	(0.466)	(0.109)	(0.45)		(3.57)
10	0.400	0.098	0.45	0.41	1.34
	(0.382)	(0.091)	(0.46)		(4.64)

TABLE II. As Table I, but for magnetic field $h = 0.34$. In this case, the Gutzwiller and Hartree-Fock wave functions contain an extra parameter $\bar{h}/2$ (see text). Indicated in brackets after the VMC results for Δ and $\bar{h}/2$ are the statistical errors in the last displayed digit for these quantities. All other VMC quantities have a statistical error smaller than 1 in the last displayed digit.

U	$-E/N_s$	ρ_s	$\langle S^+ \rangle$	g	Δ	$\bar{h}/2$
3	1.000	0.183	0.26	0.67	0.35(4)	0.28(3)
	(0.958)	(0.181)	(0.28)		(0.828)	(0.356)
4	0.854	0.169	0.31	0.63	0.50(7)	0.38(3)
	(0.809)	(0.163)	(0.34)		(1.357)	(0.456)
5	0.736	0.154	0.35	0.62	0.76(2)	0.38(1)
	(0.695)	(0.146)	(0.38)		(1.898)	(0.561)
7	0.569	0.122	0.41	0.51	1.02(6)	0.50(5)
	(0.538)	(0.117)	(0.42)		(2.970)	(0.822)
8	0.509	0.112	0.42	0.47	1.14(6)	0.58(5)
	(0.483)	(0.105)	(0.44)		(3.491)	(0.983)
10	0.419	0.094	0.44	0.45	1.44(8)	0.54(8)
	(0.402)	(0.087)	(0.45)		(4.500)	(1.365)

calculation. Our results in Table I ($h = 0$) are in agreement with those in Ref. 14, where it was also shown that finite-size effects are minor for $U > 3$ on 8×8 lattices (note that the sublattice magnetization M_s in Ref. 14 is twice $\langle S^+ \rangle$). From the tables one sees that for smaller values of U, g approaches 1, indicating correctly that HF becomes a better approximation. Also, for $U \to 0$, both the HF and the Gutzwiller wave function approach the exact wave function so that the energy difference becomes smaller. If U becomes large, the HF wave function is not a good approximation anymore, but the distinction between HF and Gutzwiller approximation becomes small in terms of the energy because HF already avoids double occupation and the reduction of the wave function that Gutzwiller adds becomes irrelevant. Therefore, also for large U the energy difference between HF and Gutzwiller becomes small. The same reasoning applies to the sublattice magnetization $\langle S^+ \rangle$. Although the VMC result for the energy is always lower than the HF result, the gain is not very substantial. We remark that in going off half-filling much larger gains in energy can be obtained in going from the HF to the Gutzwiller wave function. For instance, two of us calculated before[20] that for a homogeneous system with 112 sites and 104 electrons the energy per site decreases from -0.558 to -0.655 for $U/t = 7$ and

from -0.410 to -0.504 for $U/t = 10$.[21]

The new feature with respect to previous VMC calculations, apart from the generalization to nonzero magnetic field h, is the calculation of ρ_s. We find that the value of ρ_s is increased by a small amount (of the order of 5% for $4 < U/t < 8$) in the VMC calculation using the Gutzwiller wave function as compared to the Hartree-Fock result. Depending on taste one can draw two conclusions: the HF approximation already gives a fairly accurate value for ρ_s, or the Gutzwiller wave function does not improve very much on the HF wave function. In other words, it is still possible that introducing other types of correlations in the wave function will change ρ_s more drastically, but we have shown that the HF value for ρ_s is fairly insensitive to the Gutzwiller-type of correlations.

IV. OPTIMAL SUPERCONDUCTING CRITICAL TEMPERATURE IN THE ATTRACTIVE HUBBARD MODEL

The helicity modulus ρ_s, for instance, as calculated in Sec. II, may be used to obtain an estimate of the critical temperature T_c for superconductivity in the attractive Hubbard model for arbitrary values of the interaction strength U/t and arbitrary electron density n. A brief account of this procedure was given before.[22]

In this section on the attractive Hubbard model we

can use the formulas derived in Sec. II for the repulsive Hubbard model if some "translations" are made. According to the "spin-down particle-hole" transformation one must replace in (37), (40), (41), and (48) the magnetization m by the deviation from half-filling $n - 1$ and the effective Zeeman field $\bar{h}/2$ by the effective chemical potential $\bar{\mu}$ [see (19) and (20)].

For weak attraction, the Hartree-Fock approximation is a good approximation and for the Hubbard model on a square lattice gives results that are qualitatively similar to the results of the BCS theory for superconductivity (see, e.g., Ref. 23). For instance, for small $|U|/t$ the gap parameter at zero temperature $\Delta(0)$ decreases exponentially for decreasing $|U|/t$. From (40) and (41) one obtains

$$\Delta(0) = 32te^{-2\pi\sqrt{t/|U|}} \qquad \text{for } n = 1, \tag{79}$$

$$\Delta(0) = 9te^{-8t/|U|} \qquad \text{for } n = 0.5. \tag{80}$$

The critical temperature for superconductivity $T_c(\text{HF})$ follows from (40) and (41) with the additional condition $\Delta(T_c) = 0$. We find $\Delta(0)/k_B T_c = 1.80$ in the small-U limit, whereas standard BCS theory gives 1.764 for this ratio. The temperature dependence of the gap near T_c is given by

$$\frac{\Delta(T)}{\Delta(0)} = \alpha \left(1 - \frac{T}{T_c}\right)^{1/2}, \tag{81}$$

where in the small-U limit $\alpha = 1.74$, exactly as in the BCS theory. It is somewhat remarkable that this behavior near T_c is very robust for all values of $|U|/t$: in the limit of $|U|/t \to \infty$, α is minimal: $\alpha = \sqrt{3} = 1.73$, whereas α reaches a maximum for $|U|/t \simeq 5.6$ of only 1.75. Also, the ratio $\Delta(0)/k_B T_c$ never exceeds the value 2, which is the limiting value for $|U|/t \to \infty$.

It has been argued and illustrated by several researchers[24,7,25] that the BCS (or HF) approximation to the wave function remains an appropriate form for the wave function for all values of $|U|/t$. In the large-U regime, the wave function would then describe a gas of tightly bound electron pairs which will exhibit local pair (LP) or bipolaronic superconductivity. The superconducting transition then corresponds to the superfluid transition of a hard-core Bose gas of such pairs (no two pairs may occupy the same site because of the Pauli exclusion principle). Since in the limit of strong attraction the model can be mapped onto a pseudo-spin-model with effective interaction constant $J = 4t^2/|U|$, T_c (proportional to J) (Ref. 7) will decrease for increasing $|U|/t$ and can no longer correspond to $T_c(\text{HF})$ (which increases linearly with $|U|/t$). For larger $|U|/t$, low-energy excitations will destroy the order for a much lower temperature than $T_c(\text{HF})$, which is the temperature at which the now tightly bound pairs break up. The evolution from Cooper-pair superconductivity for small U to LP superconductivity for large U is thought to be smooth. Since both in the small- and large-U limit T_c vanishes, evidently there must be an optimal T_c for an intermediate value of $|U|/t$. In this section, we present a scheme from which for the Cooper-pair superconductor, the LP superconductor, and everything in between the critical

temperature can be extracted. This scheme makes use of the helicity modulus.

The low-energy excitations that destroy the order for a lower temperature than $T_c(\text{HF})$ can be thought of as fluctuations of the order parameter which have been neglected in the HFA. Here, we consider fluctuations in the phase of the order parameter. Then we have a model similar to that of XY ferromagnetism, where the complex gap field Δ_j is the equivalent of an XY spin. For low enough T, the system is in a ferromagnetically ordered state (i.e., constant Δ_j), whereas for higher temperature in two dimensions a Kosterlitz-Thouless (KT) phase transition arises.[6] The KT transition corresponds to the binding/unbinding of certain spin configurations called vortex-antivortex pairs, which correspond in the negative-U model to certain configurations of the Δ_j involving only phase fluctuations. Thus, in this view, the superconducting state of the attractive Hubbard model is a KT phase, with corresponding algebraically decaying correlations. Exactly at the transiton temperature T_c^{KT} the following relation with the spin-stiffness or helicity modulus holds:

$$k_B T_c^{\text{KT}} = \frac{\pi}{2} \rho_s^-. \tag{82}$$

Here, ρ_s^- is the value just below the critical temperature. Since ρ_s vanishes above T_c and (82) is independent of other system parameters, this relation describes a universal jump in the superfluid density at T_c in the theory of thin superfluid helium films.[26] Although the helicity modulus in (82) is a renormalized quantity and the helicity modulus ρ_s discussed in the preceding sections is unrenormalized, we consider ρ_s as calculated by us in the HFA as a good approximation to the renormalized ρ_s. Thus we can extract an approximate T_c^{KT}. The critical temperature T_c^{KT} for a KT transition in the negative-U Hubbard model is then defined as the temperature where the relation (82) for ρ_s and the computed $\rho_s(T)$ coincide (see Fig. 2). We can carry out this procedure relatively
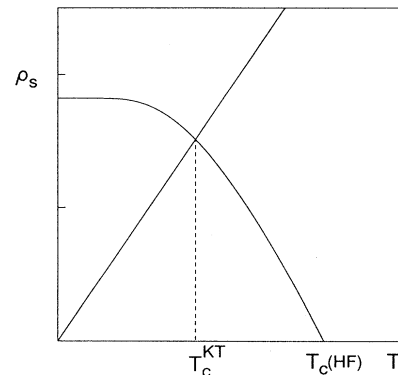

FIG. 2. Schematic representation of the procedure to extract the Kosterlitz-Thouless critical temperature T_c^{KT} from the computed helicity modulus $\rho_s(T)$ for a wavelike distortion of the homogeneous ground state. $\rho_s(T)$ vanishes at the critical temperature $T_c(\text{HF})$ of the Hartree-Fock theory. T_c^{KT} follows from the intersection of $\rho_s(T)$ with the "universal-jump" relation $\rho_s(T_c) = (2/\pi)k_B T_c$ from KT theory.

easily for arbitrary combinations of the parameters $|U|/t$ and density n. As in the case of positive U (for which the formulas were given in Sec. II), there is a slight complication in that in calculating $\rho_s(T)$ for fixed $|U|/t$ and n (which are the physical parameters), for every new T the parameters Δ and $\bar{\mu}$ have to be adjusted according to (40) and (41). In this way, the critical temperature for superconductivity T_c^{KT} is indeed always smaller than $T_c(\mathrm{HF})$; the reduction is very small for small $|U|$ $[\rho_s(T = 0)$ is sizable but drops off to zero rapidly] and very large for large $|U|$ $[\rho_s(T = 0)$ is small but practically constant up to the intersection point].

In Fig. 3(a), T_c^{KT}/t as a function of $|U|/t$ is shown for three values of the filling ($n = 1$, $n = 0.5$, and $n = 0.1$, respectively). For small $|U|$, T_c is only reduced by a small amount compared to standard BCS theory,
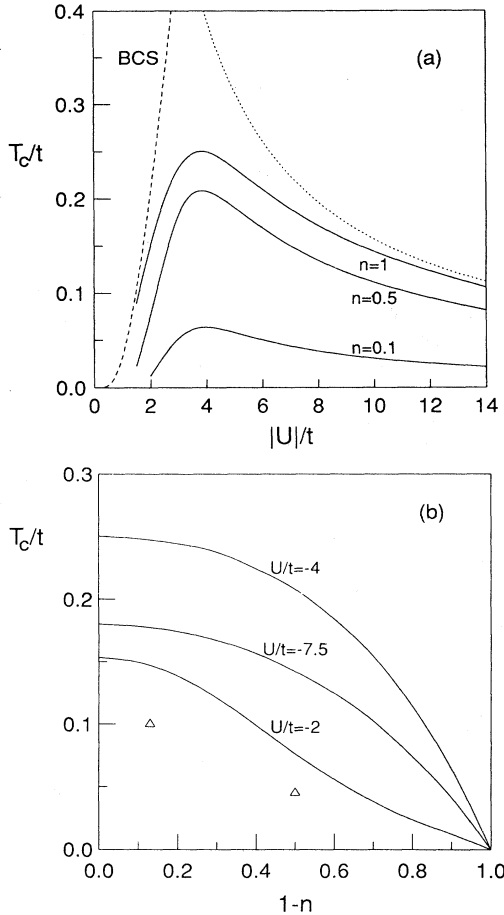

FIG. 3. Phase diagram of the 2D negative-U Hubbard model on a square lattice. (a) T_c^{KT}/t as a function of $|U|/t$ for three different fillings ($n = 1$, $n = 0.5$, and $n = 0.1$, respectively). For comparison are shown $T_c(\mathrm{HF})/t$ [or equivalently $T_c(\mathrm{BCS})/t$; dashed line] and the functional form $\pi t/2|U|$ (dotted line) to which T_c/t approaches (for half-filling) for small $|U|$ and large $|U|$, respectively. (b) T_c^{KT}/t as a function of electron density n for $U/t = -2, -4$, and -7.5. The quantum Monte Carlo results of Ref. 31 for $U/t = -4$ are denoted by triangles (see the discussion in Sec. V).

whereas for very large $|U|$ a T_c proportional to $t^2/|U|$ is indeed found, as expected because of the connection with a pseudo-spin-model with $J = 4t^2/|U|$. In Fig. 3(b), T_c/t is shown as a function of n for $|U|/t$ equals 2, 4, and 7.5. Figure 3(b) can be extended to $1 < n \leq 2$ since, because of particle-hole symmetry, $T_c(n - 1) = T_c(1 - n)$. Therefore, within the approximations that were made, we obtain the phase diagram of the 2D negative-U Hubbard model for the whole range of interaction strengths $|U|/t$ and electron densities n. The maximum T_c of $0.25t$ occurs for $|U|/t = 4$ and $n = 1$. For large $|U|$, the ratio $\Delta(0)/k_B T_c$ is of course much enhanced as compared to the BCS (or HF) result; e.g., in case of the optimal T_c ($|U|/t = 4$) $\Delta(0)/k_B T_c = 11.1$, whereas for $|U|/t = 3$ this ratio equals 7.3.

A deficiency of our result is that T_c does not vanish at half-filling, where one would expect that the Heisenberg symmetry, which the system possesses at half-filling, does not allow for the KT phase and therefore gives $T_c = 0$. The vanishing T_c is a result of the delicate symmetry between charge density and pairing correlations at half-filling. It is not surprising that we do not recover this feature since our procedure cannot easily accommodate the additional symmetry which arises at half-filling, where the XY symmetry we use for all n is extended to the higher Heisenberg symmetry. A small amount of doping already destroys the Heisenberg symmetry and results in a finite T_c; also a small deviation from ideal two dimensionality results in a finite T_c. Therefore our results are expected to be good already for a small deviation from half-filling, which in our calculation does not alter T_c very much [see Fig. 3(b)], but we miss the logarithmic drop to zero when approaching half-filling.

Another point of concern regarding our procedure is that we have borrowed the relation $k_B T_c = (\pi/2)\rho_s$ from the exact, renormalized KT theory, whereas our calculated ρ_s is unrenormalized. The fact that ρ_s is unrenormalized will decrease the estimate for T_c; the results we give are upper bounds in this respect.[6,26] We have investigated this effect by studying another, inhomogeneous, excitation of the ground state with constant Δ_i. We compute by exact diagonalization of the Bogoliubov-deGennes equations (23) and (24) on small lattices (8×8) the excitation energy ΔE of turning one Δ_i to $-\Delta_i$ as a function of $|U|/t$ (for half-filling and $T = 0$). This raise in energy we relate to a nearest-neighbor interaction constant J_{NN} between (ferromagnetic) XY spins by equating ΔE to $8J_{NN}$. Subsequently, we use the relation between the interaction constant J_{NN} and the critical temperature T_c^{XY} for the phase transition in the two-dimensional XY ferromagnet (with nearest-neighbor coupling only) as found in recent Monte Carlo calculations:[27]

$$k_B T_c^{XY} = 0.898 J_{NN}. \tag{83}$$

Again this procedure to extract T_c is not rigorous, since in our system of phase-fluctuating Δ_i the interaction is not restricted entirely to nearest neighbors. We have verified, by also computing the energy of a "Néel configuration" of Δ_i, that for $|U|/t > 3$ the next-nearest-neighbor coupling J_{NNN} is always smaller than J_{NN} by a factor of

TABLE III. Critical temperature T_c for a Koster-litz-Thouless (KT) type of phase transition for a system of phase-fluctuating gap parameters Δ as a function of U at half-filling and zero magnetic field. If the system is taken to be identical to a nearest-neighbor ferromagnetic XY model, $k_B T_c^{XY} = 0.898 J_{\rm NN}$ (taken from Ref. 27), with $J_{\rm NN}$ the coupling constant. The calculation of $J_{\rm NN}$ is explained in the text. If the system is taken to obey the universal-jump relation of KT theory, $k_B T_c^{\rm KT}$ is obtained using a Hartree-Fock calculation of the helicity modulus $\rho_s(T)$ by an intersection procedure (see Fig. 2). For comparison $\rho_s(T = 0)$ is listed in the last column; $\rho_s(T = 0)$ and $J_{\rm NN}$ become equal in the limit of large U. All results in the table are obtained on 8×8 lattices and are in units of the hopping integral t.

Δ	U	$J_{\rm NN}$	$k_B T_c^{XY}$	$k_B T_c^{\rm KT}$	$\rho_s(T = 0)$
0.8	2.87	0.0750	0.067	0.224	0.183
1.0	3.27	0.1020	0.092	0.244	0.177
1.2	3.65	0.1310	0.118	0.251	0.171
1.4	4.02	0.1376	0.124	0.251	0.164
1.6	4.40	0.1374	0.123	0.245	0.158
2.0	5.12	0.1334	0.120	0.230	0.146
2.4	5.85	0.1270	0.114	0.213	0.135
3.0	6.95	0.1160	0.104	0.190	0.121
3.3	7.50	0.1106	0.099	0.180	0.114
3.6	8.05	0.1056	0.095	0.170	0.108
4.0	8.80	0.0990	0.089	0.159	0.101
5.0	10.68	0.0852	0.077	0.136	0.086
6.0	12.59	0.0742	0.067	0.118	0.075

at least about 2 ($J_{\rm NNN}$ can have either sign). In Table III, we compare T_c^{XY} with $T_c^{\rm KT}$ obtained by the intersection procedure using ρ_s from the HFA. We also list $\rho_s(T = 0)$ as calculated in the HFA. For better comparison, both $T_c^{\rm KT}$ and $\rho_s(T = 0)$ have been calculated on 8×8 lattices. For the range of values of U considered the finite-size effect is negligible. From the table it is clear that T_c^{XY} is smaller than $T_c^{\rm KT}$ by a factor of about 2 for most of the U values considered. Note that like $T_c^{\rm KT}$, T_c^{XY} goes through a maximum, which furthermore occurs for about the same value of $|U|/t$. However, T_c^{XY} drops off much more rapidly for decreasing $|U|/t$. We also note that $\rho_s(T = 0)$ does not go through a maximum; the nonmonotonous behavior of $T_c^{\rm KT}$ is due to our intersection procedure. The fact that both critical temperatures show the same qualitative behavior can be seen as a sign of the consistency of our arguments. Finally, we note that in the large-U limit $\rho_s(T = 0)$ and $J_{\rm NN}$ coincide.

V. CONNECTION WITH PREVIOUS APPROACHES AND DISCUSSION

Several results obtained in the literature on the two-dimensional Hubbard model can be related to results for the helicity modulus ρ_s and the critical temperature T_c

for the doped negative-U model obtained in this paper. First, we discuss results to be compared with ρ_s and later on results to be compared with T_c .

We have found that ρ_s—as calculated in the HFA in Sec. II for the repulsive case for half-filling and arbitrary U/t, T, and h—for the special case of $T = 0$ and $h = 0$ coincides exactly (for all values of U/t) with a particular long-wavelength large-frequency limit of the dynamic transverse susceptibility $\chi^{+-}(\mathbf{q}, \omega)$ as calculated in the random-phase approximation (RPA):

$$\lim_{\omega \to \infty} -\frac{\omega^2}{2} \chi_{\rm RPA}^{+-}(\mathbf{q}, \omega) = \rho_s q^2 + \mathcal{O}(q^4). \tag{84}$$

This specific limit can be shown to be equal to the first moment $\langle \omega \rangle$ of the dynamic structure factor $S_\perp(\mathbf{q}, \omega)$ associated with spin-spin correlations, a relation known as the f-sum rule.[12] Equation (84) is derived in detail in Appendix B, starting from the RPA calculations as published before by several groups.[11,12] In Ref. 12, the large-U limit of the f-sum rule was found to be $\langle \omega \rangle = q^2 t^2/U$, in agreement with (84) (see Appendix B). Our results for ρ_s generalize the f-sum rule for the positive-U Hubbard model at half-filling to arbitrary U/t.

Our result for the optimal critical temperature for superconductivity in the attractive Hubbard model can also be brought in connection with results in the literature. A large number of papers has been devoted to the question of a smooth transition from a BCS-like superconducting-normal transition for small U to a Bose-Einstein-like superfluid-normal transition (or local-pair superconductor-normal transition) for large U.[7,24,25,29,30] In some of these papers even a qualitative phase diagram like Fig. 3 is drawn.[7,29,30] We have given a quantitative scheme from which T_c can be found for arbitrary $|U|/t$ and arbitrary density. This scheme is based on the helicity modulus, which we have calculated using two well-defined approximations. We also compared the approximated ρ_s with more exact calculations (see above), showing that the HFA results for ρ_s are fairly accurate. The weakness of our approach to extract T_c lies in the fact that we invoke exact theories like that of Kosterlitz and Thouless or results from accurate Monte Carlo calculations on the ferromagnetic XY model which are not exactly applicable to our model. For instance, the "universal jump" relation (82) from KT theory holds exactly only for the case that the core energy E_c associated with a vortex is infinitely large.[26] If not, T_c will be lower. For instance, the ferromagnetic XY model on a square lattice has a finite E_c. It would be interesting to see what would be the core energy for vortices in a system of gap parameters which are phase fluctuating over the lattice. That indeed the actual T_c for the Hubbard model can be much lower than extracted from $\rho_{s,\rm HFA}$ we already showed in Sec. IV, by finding T_c from an effective nearest-neighbor coupling constant for a ferromagnetic XY model. Therefore, although we cannot claim any degree of exactness of

our result for T_c, we speculate that the HFA results for ρ_s lead to an approximation for T_c , which will not have to be adjusted downward by a factor of more than about 3 (for densities $|n - 1| > 0.1$ and arbitrary U). Furthermore, our approximation to T_c can be evaluated readily for the whole parameter space $(U/t, n)$. In view of the above, the rather large difference between our result for T_c and that of quantum Monte Carlo (QMC) calculations [Ref. 31 and Fig. 3(b)] is not distressing. Furthermore, although QMC calculations are arbitrarily accurate in principle, T_c does not follow from them directly; T_c was extracted from the behavior of pairing correlations fitted to expected Kosterlitz-Thouless behavior. However, it is not obvious that the correlation function will follow KT behavior on the small lattices of maximally 8 × 8 that were used and therefore the exact T_c may differ from the QMC result. An additional drawback of QMC calculations is that it is very time consuming to compute T_c for more than a few points in the parameter space $(U/t, n)$.

For completeness we note that recently an optimal T_c for superconductivity was found in the Gorkov model of fermions with an attractive short-range interaction in two dimensions.[32] This model can be seen as an analog of the Hubbard model for the case that the fermions are not restricted to a lattice. The optimal T_c is obtained using a procedure similar to ours; first a Ginzburg-Landau theory is constructed describing the crossover from BCS superconductivity to Bose superfluidity and then the Monte Carlo result of Ref. 27 for the nearest-neighbor 2D XY model [see (83)] is used to extract T_c. An important difference with our approach is that in Ref. 32 the superfluid density ρ_s is identified directly with the interaction constant J_{NN} in the XY model. As can be seen from Table III, we find this identification to be correct only for large values of $|U|/t$.

In conclusion, we have presented detailed calculations using a variety of approximations of the helicity modulus in both repulsive (at half-filling) and attractive (in magnetic field zero) Hubbard models. We furthermore use the helicity modulus to extract a critical temperature for superconductivity in the attractive Hubbard model for arbitrary density and arbitrary interaction strength which correctly interpolates between weak- and strong-coupling limits. Using the Hartree-Fock approximation for ρ_s, which we have shown to coincide with the result in the random-phase approximation for ρ_s, and the universal-jump relation from Kosterlitz-Thouless theory, an optimal T_c of $0.25t$ is found for $U/t = -4$. We have given estimates of the deviation of this result from the (unknown) exact result. Finally, we have extensively compared our results for ρ_s and T_c with results in the literature on the Hubbard model.

Note added. After completion of our work a paper by Scalapino, White, and Zhang appeared[33] which expresses the belief that the Drude weight equals the superfluid weight in the superconducting state (see also Ref. 28). If this belief is borne out, the possibility is opened of comparing our calculations for ρ_s with the extensive exact diagonalization calculations of the Drude weight in the negative-U Hubbard model of Ref. 34. This possibility is presently being investigated by us.

APPENDIX A

We show that at finite temperatures the Hartree-Fock approximation (HFA) can be formulated as a variational search for self-consistent fields $\langle n_{i\sigma}\rangle_\beta$ and $\langle S_i^\sigma\rangle_\beta$, where the subscript denotes the temperature dependence of the fields. For finite temperatures, the quantity to be minimized turns out to be the Hartree-Fock (HF) free energy F_{HF} (to be defined below); moreover optimal fields $\langle A\rangle_\beta$ turn out to be given by HF averages:

$$\langle A\rangle_\beta = \mathrm{Tr}(\rho_{HF}A). \tag{A1}$$

In (A1), ρ_{HF} is the density matrix associated with the Hartree-Fock approximation to the Hamiltonian (to be defined below). This variational principle thus generalizes the $T = 0$ HFA of Sec. II A to finite temperatures.

Consider the functional $\mathcal{F}[\rho]$:

$$\mathcal{F}[\rho] = \mathrm{Tr}(\rho\mathcal{H} + k_B T\rho\ln\rho), \tag{A2}$$

working on the space of normalized N-body density operators ρ ($\mathrm{Tr}\rho = 1$). Using the normalization of the density operator, it follows that $\mathcal{F}[\rho]$ is minimized by the Boltzmann operator

$$\rho = \frac{1}{Z}e^{-\beta\mathcal{H}}, \tag{A3}$$

where $\beta = 1/k_B T$ and Z is the partition function:

$$Z = \mathrm{Tr}\left(e^{-\beta\mathcal{H}}\right). \tag{A4}$$

The minimum value $\mathcal{F}[\rho]$ attains is exactly the free energy:

$$F = -k_B T\ln Z. \tag{A5}$$

Now we consider the Hartree-Fock approximation to the Hubbard Hamiltonian (1) which allows for antiferromagnetic ordering, i.e., allowing for nonvanishing averages $\langle S_i^\sigma\rangle$ besides nonzero averages $\langle n_{i\sigma}\rangle$. The approximated Hamiltonian is given by [as in Sec. II; see (13), but now with temperature-dependent fields]

$$\mathcal{H}_{HF} = \sum_{i\sigma, j\sigma'} c_{i\sigma}^\dagger H_{i\sigma,j\sigma'}c_{j\sigma'}$$
$$-U\sum_i(\langle n_{i\uparrow}\rangle_\beta\langle n_{i\downarrow}\rangle_\beta - \langle S_i^+\rangle_\beta\langle S_i^-\rangle_\beta), \tag{A6}$$

with the Hermitian matrix $H_{i\sigma,j\sigma'}$ given by

$$H_{i\sigma,j\sigma'} = \left[-t_{ij} + \left(U\langle n_{i,-\sigma}\rangle_\beta - \frac{h\sigma}{2} - \mu\right)\delta_{ij}\right]\delta_{\sigma,\sigma'}$$
$$-U\langle S_i^{-\sigma}\rangle_\beta\delta_{ij}\delta_{\sigma,-\sigma'}. \tag{A7}$$

At this point the fields $\langle S_i^\sigma\rangle_\beta$ and $\langle n_{i\sigma}\rangle_\beta$ are just pa-

rameters, which will be identified with averages below. The one-electron HF energies $\epsilon(k)$ follow from the HF equations [cf. (8)]:

$$\sum_{j\sigma'} H_{i\sigma,j\sigma'} \phi_{j\sigma'}(k) = \epsilon(k)\phi_{i\sigma}(k). \qquad \text{(A8)}$$

We now proceed to show that the fields $\langle S_i^\sigma \rangle_\beta$ and $\langle n_{i\sigma} \rangle_\beta$ that optimize the HF free energy F_{HF} [defined by (A5) using $\mathcal{H}_{\text{HF}}$ in (A4) to obtain Z_{HF}] also optimize $\mathcal{F}[\rho]$. In other words, we show

$$\mathcal{F}[\rho_{\text{HF}}] = F_{\text{HF}}, \qquad \text{(A9)}$$

which is a special property of the HFA and not true for an arbitrary approximation to the Hamiltonian. Using (A6) and evaluating the trace in the basis of the $2N_s$ eigenstates of $H_{i\sigma,j\sigma'}$ (N_s is the number of sites of the lattice), the HF free energy is

$$F_{\text{HF}} = -\frac{1}{\beta} \sum_k \ln\left[1 + e^{-\beta\epsilon(k)}\right]$$
$$-U\sum_i (\langle n_{i\uparrow}\rangle_\beta \langle n_{i\downarrow}\rangle_\beta - \langle S_i^+\rangle_\beta \langle S_i^-\rangle_\beta). \qquad \text{(A10)}$$

Considering the fields $\langle S_i^\sigma \rangle_\beta$ and $\langle n_{i\sigma}\rangle_\beta$ as four independent parameters one straightforwardly finds the fields that optimize F_{HF}. As an example we take $\langle n_{i\uparrow}\rangle_\beta$. From (A10) one finds

$$0 = \frac{\partial F_{\text{HF}}}{\partial \langle n_{i\uparrow}\rangle_\beta} = \sum_k n(k)\frac{\partial\epsilon(k)}{\partial\langle n_{i\uparrow}\rangle_\beta} - U\langle n_{i\downarrow}\rangle_\beta, \qquad \text{(A11)}$$

where $n(k)$ is the Fermi-Dirac distribution:

$$n(k) = \left[1 + e^{\beta\epsilon(k)}\right]^{-1}. \qquad \text{(A12)}$$

For the derivative of $\epsilon(k)$ we use (A8), in which only the

explicit dependence of $H_{i\sigma,j\sigma'}$ on $\langle n_{i\uparrow}\rangle_\beta$ needs to be taken into account:

$$\frac{\partial\epsilon(k)}{\partial\langle n_{i\uparrow}\rangle_\beta} = \sum_{l\sigma,j\sigma'} \phi_{l\sigma}^*(k)\frac{\partial H_{l\sigma,j\sigma'}}{\partial\langle n_{i\uparrow}\rangle_\beta}\phi_{j\sigma'}(k)$$
$$= U|\phi_{i\downarrow}(k)|^2. \qquad \text{(A13)}$$

Applying similar considerations to the other fields, one arrives at the following consistency equations for the fields:

$$\langle n_{i\sigma}\rangle_\beta = \sum_k |\phi_{i\sigma}(k)|^2 n(k), \qquad \text{(A14)}$$

$$\langle S_i^\sigma\rangle_\beta = \sum_k \phi_{i\sigma}^*(k)\phi_{i,-\sigma}(k)n(k). \qquad \text{(A15)}$$

Note that Eqs. (A14) and (A15) are straightforward finite-temperature generalizations of Eqs. (9) and (10). Working again in the basis of eigenstates of $H_{i\sigma,j\sigma'}$, one easily verifies that the finite-temperature Hartree-Fock averages of $n_{i\sigma}$ and S_i^σ are given precisely by the fields in (A14) and (A15):

$$\langle n_{i\sigma}\rangle_{\text{HF}} \equiv \text{Tr}(\rho_{\text{HF}}n_{i\sigma}) = \langle n_{i\sigma}\rangle_\beta, \qquad \text{(A16)}$$
$$\langle S_i^\sigma\rangle_{\text{HF}} \equiv \text{Tr}(\rho_{\text{HF}}S_i^\sigma) = \langle S_i^\sigma\rangle_\beta. \qquad \text{(A17)}$$

Using the definition of the free-energy functional $\mathcal{F}[\rho]$ (A2), one finds that

$$\mathcal{F}[\rho_{\text{HF}}] = \text{Tr}\left[\rho_{\text{HF}}(\mathcal{H} - \mathcal{H}_{\text{HF}})\right] + F_{\text{HF}}. \qquad \text{(A18)}$$

The difference between the original and HF Hamiltonian is found just in the on-site interaction term. Evaluating the HF average of this difference, while invoking Wick's theorem to show that

$$\langle n_{i\uparrow}n_{i\downarrow}\rangle_{\text{HF}} = \langle n_{i\uparrow}\rangle_{\text{HF}}\langle n_{i\downarrow}\rangle_{\text{HF}} - \langle S_i^+\rangle_{\text{HF}}\langle S_i^-\rangle_{\text{HF}}, \qquad \text{(A19)}$$

one obtains for arbitrary $\langle S_i^\sigma\rangle_\beta$ and $\langle n_{i\sigma}\rangle_\beta$

$$\langle \mathcal{H} - \mathcal{H}_{\text{HF}}\rangle_{\text{HF}} = U\sum_i [(\langle n_{i\uparrow}\rangle_{\text{HF}} - \langle n_{i\uparrow}\rangle_\beta)(\langle n_{i\downarrow}\rangle_{\text{HF}} - \langle n_{i\downarrow}\rangle_\beta) - (\langle S_i^+\rangle_{\text{HF}} - \langle S_i^+\rangle_\beta)(\langle S_i^-\rangle_{\text{HF}} - \langle S_i^-\rangle_\beta)]. \qquad \text{(A20)}$$

Obviously the optimal fields, which obey (A16) and (A17), make the HF average of the difference zero, but, furthermore, adding (A20) to F_{HF} will not change the optimization equations (A14) and (A15) since the added term is quadratic in the deviations. Therefore, the optimized fields of the HFA optimize F_{HF} and $\mathcal{F}[\rho_{\text{HF}}]$ at the same time.

APPENDIX B

We show that the helicity modulus as obtained in the Hartree-Fock approximation in Sec. II, Eq. (48), is identical to a small-wave-vector, large-frequency limit of the dynamic transverse susceptibility as calculated in the

random-phase approximation (RPA). More specifically, we show

$$\lim_{\omega\to\infty} -\frac{\omega^2}{2}\chi^{+-}(\mathbf{q},\omega) = \rho_s q^2 + \mathcal{O}(q^4). \qquad \text{(B1)}$$

The RPA was applied to the positive-U Hubbard model at half-filling in zero magnetic field by several groups.[11,12] Here we follow Ref. 12, in which the following expression for the dynamic transverse susceptibility is given (we omit the so-called umklapp branch, since we will be interested in the small-q limit):

$$\chi^{+-}(\mathbf{q},\omega) = \frac{\chi_{11} - U(\chi_{11}\chi_{22} - \chi_{12}\chi_{21})}{(1 - U\chi_{11})(1 - U\chi_{22}) - (U\chi_{12})^2}, \qquad \text{(B2)}$$

where

$$\chi_{11}(\mathbf{q},\omega) = \frac{1}{2N}\sum_{\mathbf{p}}{}' \left(1 - \frac{\epsilon_+\epsilon_- - |\Delta|^2}{E_+E_-}\right)(f_- + f_+),$$

$$\text{(B3)}$$

$$\chi_{12}(\mathbf{q},\omega) = \frac{\Delta}{2N}\sum_{\mathbf{p}}{}'\frac{E_+ + E_-}{E_+E_-}(f_- - f_+),\qquad \text{(B4)}$$

$$\chi_{21}(\mathbf{q},\omega) = \chi_{12}(\mathbf{q},\omega),\qquad \text{(B5)}$$

$$\chi_{22}(\mathbf{q},\omega) = \chi_{11}(\mathbf{Q} - \mathbf{q},\omega).\qquad \text{(B6)}$$

The prime on the summation indicates the restriction to the magnetic Brillouin zone (MBZ): $p_x \pm p_y \in [-\pi,\pi]$ and $\mathbf{Q} = (\pi,\pi)$. The following abbreviations are introduced:

$$f_\pm \equiv (E_+ + E_- \pm \omega)^{-1},\qquad \text{(B7)}$$

$$E_\pm \equiv \sqrt{\epsilon_\pm^2 + |\Delta|^2},\qquad \text{(B8)}$$

$$\epsilon_\pm \equiv \epsilon_{\mathbf{p}\pm\mathbf{q}/2},\qquad \text{(B9)}$$

$$\epsilon_{\mathbf{p}} \equiv -2t(\cos p_x + \cos p_y).\qquad \text{(B10)}$$

Note that $\epsilon_{\mathbf{p}}$ is the $\mathbf{q} = 0$ expression for $\epsilon_\pm$. Analogously, the $\mathbf{q} = 0$ expression for $E_\pm$ is denoted by $E_{\mathbf{p}}$:

$$E_{\mathbf{p}} = \sqrt{\epsilon_{\mathbf{p}}^2 + |\Delta|^2}.\qquad \text{(B11)}$$

First, we consider the large-ω limit. The only frequency dependence is in the functions $f_\pm$. Because for large frequencies χ_{11} and χ_{22} are proportional to ω^{-2} and χ_{12} is proportional to ω^{-1}, we can neglect the product $\chi_{11}\chi_{22}$ in the numerator. Second, the leading term of the denominator in the large-ω, small-q limit is just 1. Therefore, we require only the limiting behavior of χ_{11} and χ_{12}. After a straightforward, but tedious, calculation one finds to leading order for large ω and small $\mathbf{q}$

$$\chi_{11} = -\frac{4\Delta^2}{\omega^2}\left(\frac{1}{U} + yq^2\right),\qquad \text{(B12)}$$

$$\chi_{12} = \chi_{21} = -\frac{2\Delta}{\omega U},\qquad \text{(B13)}$$

where y is given by the following sum over the MBZ:

$$y = \frac{t^2}{N}\sum_{\mathbf{p}}{}'\frac{\sin^2 p_x}{E_{\mathbf{p}}^3}.\qquad \text{(B14)}$$

We remark that the q^2/ω term in χ_{12} is proportional to a sum over the MBZ which vanishes for all values of the

parameter Δ. This is not obvious from the sum on first inspection; however, one may use the identity

$$\left[\frac{(\Delta^2 - 2\epsilon_{\mathbf{p}}^2)\sin^2 p_x}{E_{\mathbf{p}}^5} + \frac{\epsilon_{\mathbf{p}}\cos p_x}{2E_{\mathbf{p}}^3}\right] = 0,\qquad \text{(B15)}$$

where we have taken $t = 1$ and introduced the notation

$$[F(\mathbf{p})] \equiv \frac{1}{N}\sum_{\mathbf{p}}{}'F(\mathbf{p}).\qquad \text{(B16)}$$

The identity (B15) may be proved by noting that the summand is a derivative with respect to p_x of a function that is periodic in p_x. Using (B12) and (B13), we obtain the desired limit as

$$\lim_{\omega\to\infty}-\frac{\omega^2}{2}\chi^{+-}(\mathbf{q},\omega) = 2\Delta^2 yq^2 + \mathcal{O}(q^4).\qquad \text{(B17)}$$

We now proceed to show that the prefactor of the q^2 term in (B17), $2\Delta^2 y$, is exactly ρ_s as calculated in the HFA, formula (48), for $T = 0$ and $h = 0$. Using the notation introduced above, for this case ρ_s is given by

$$\rho_s(T = 0, h = 0) = \left[-\frac{\epsilon_{\mathbf{p}}\cos p_x}{E_{\mathbf{p}}}\right].\qquad \text{(B18)}$$

Note that in (48) a sum over the full Brillouin zone appears, whereas here we have a sum over the MBZ. For summands which are (i) invariant under interchange of p_x and p_y and (ii) invariant under a simultaneous change of sign of all sines and cosines the sum over the full BZ is twice the sum over the MBZ. To derive (B18) we have made use of this property. Finally, to prove (B1) we use an identity similar to (B15), which is proved analogously. With the identity

$$\left[\frac{\Delta^2\sin^2 p_x}{E_{\mathbf{p}}^3} + \frac{\epsilon_{\mathbf{p}}\cos p_x}{2E_{\mathbf{p}}}\right] = 0,\qquad \text{(B19)}$$

it follows easily that

$$\rho_s(T = 0, h = 0) = 2\Delta^2 y,\qquad \text{(B20)}$$

which together with (B17) proves (B1). Our result includes the large-U limit discussed in Ref. 12: $\langle\omega\rangle = (J/4)q^2$, where $J = 4t^2/U$ is the effective interaction constant for large U. In that limit, one finds from the above that $\rho_s = 2\Delta^2 y = t^2/U = J/4$.

---

*Present address: Shell Research BV, P.O. Box 60, 2280 AB Rijswijk, The Netherlands.

[1] M.C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963); J. Hubbard, Proc. R. Soc. (London) Ser. A **276**, 238 (1963); J. Kanamori, Prog. Theor. Phys. (Kyoto) **30**, 275 (1963); P.W. Anderson, Phys. Rev. **115**, 2 (1959).

[2] P.W. Anderson, Science **235**, 1196 (1987).

[3] M.E. Fisher, M.N. Barber, and D. Jasnow, Phys. Rev. A **8**, 1111 (1973).

[4] R.R.P. Singh and D.A. Huse, Phys. Rev. B **40**, 7247 (1989).

[5] E. Manousakis, Rev. Mod. Phys. **63**, 1 (1991).

[6] J.M. Kosterlitz and D.J. Thouless, J. Phys. C **6**, 1181 (1973).

[7] R. Micnas, J. Ranninger, and S. Robaszkiewicz, Rev. Mod. Phys. **62**, 113 (1990).

[8] J.A. Vergés, E. Louis, P.S. Lomdahl, F. Guinea, and A.R. Bishop, Phys. Rev. B **43**, 6099 (1991).

[9] P.G. deGennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, 1966).

[10] We have taken $\mathbf{q}$ to be $(q,0)$ with $q \in [0,\pi]$ but this is not essential.

[11] J.R. Schrieffer, X.G. Wen, and S.C. Zhang, Phys. Rev. B

39, 11663 (1989); A. Singh and Z. Tesanovic, *ibid.* **41**, 614 (1990).

[12] H. Monien and K.S. Bedell, Phys. Rev. B **45**, 3164 (1992).

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

[14] H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **56**, 3582 (1987).

[15] T. Giamarchi and C. Lhuillier, Phys. Rev. B **42**, 10641 (1990).

[16] D. Ceperley, G.V. Chester, and M.H. Kalos, Phys. Rev. B **16**, 3081 (1977).

[17] C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).

[18] W. Kohn, Phys. Rev. **133**, A171 (1964).

[19] Since in the VMC calculation, the average of $\mathcal{H}$ is calculated without the chemical potential term, whereas the ground-state energy in the HFA is calculated including the chemical potential, to the energy given in (25) a constant μN_e has to be added to obtain the ground-state energy at a fixed number of particles. At half-filling, this constant equals $N_s U/2$. In this way, the ground-state energy at half-filling goes to zero in the limit of infinitely large U.

[20] Guozhong An and J.M.J. van Leeuwen, Phys. Rev. B **44**, 9410 (1991).

[21] Table III in Ref. 20 contains a printing error: the ground-state energy using the Gutzwiller wave function (E_{GWF}) for the commensurate phase (GWF CM) for $U/t = 10$ should be -56.46 instead of -54.46.

[22] P.J.H. Denteneer, Guozhong An, and J.M.J. van Leeuwen,

Europhys. Lett. **16**, 5 (1991); **16**, 509(E) (1991).

[23] M. Tinkham, *Introduction to Superconductivity* (McGraw-Hill, Kogakusha, 1975).

[24] P. Nozières and S. Schmitt-Rink, J. Low-Temp. Phys. **59**, 195 (1985); A.J. Legget, *Modern Trends in the Theory of Condensed Matter* (Springer, Berlin, 1980), p. 13.

[25] M. Randeria, J.-M. Duan, and L.-Y. Shieh, Phys. Rev. B **41**, 327 (1990).

[26] D.R. Nelson, in *Fundamental Problems in Statistical Mechanics V*, edited by E.G.D. Cohen (North-Holland, Amsterdam, 1980).

[27] R. Gupta, J. DeLapp, G.G. Batrouni, G.C. Fox, C.F. Baillie, and J. Apostolakis, Phys. Rev. Lett. **61**, 1996 (1988).

[28] B.S. Shastry and B. Sutherland, Phys. Rev. Lett. **65**, 243 (1990).

[29] R.T. Scalettar, E.Y. Loh, J.E. Gubernatis, A. Moreo, S.R. White, D.J. Scalapino, R.L. Sugar, and E. Dagotto, Phys. Rev. Lett. **62**, 1407 (1989).

[30] D.I. Khomskii and A.K. Zvezdin, Solid State Commun. **66**, 651 (1988).

[31] A. Moreo and D.J. Scalapino, Phys. Rev. Lett. **66**, 946 (1991).

[32] M. Drechsler and W. Zwerger (unpublished).

[33] D.J. Scalapino, S.R. White, and S.C. Zhang, Phys. Rev. Lett. **68**, 2830 (1992).

[34] D.J. Scalapino, Physica C **185-189**, 104 (1991); E. Dagotto, A. Moreo, F. Ortolani, J. Riera, and D.J. Scalapino, Phys. Rev. B **45**, 10107 (1992).