I. INTRODUCTION

The Hubbard Hamiltonian\(^1\) is the simplest model hope-
fully able to capture most of the essential physics of strongly
correlated electron systems, leading to collective effects such
as itinerant magnetism,\(^2\) the metal-insulator transition,\(^3\) or
the spectral-weight transfer between different energy scales.\(^4\)

With the advent of high-\(T_c\) superconductors, the study of this
model has received a considerable impetus as one of the
most qualified candidates for explaining this phe-

Despite its apparent simplicity, the Hubbard model is not
exactly solvable except in one dimension, where the Bethe
ansatz solution of Lieb and Wu\(^7\) for the ground state stands
as a firm reference point. For more than one dimension, one
must resort to either exact diagonalization,\(^8,9\) or quantum
Monte Carlo (QMC) simulations.\(^10,11\) The first one goes
without saying, but it applies only to very small clusters (up
to 16 sites with the Lanczos technique). Monte Carlo simu-
lations provide an accurate description of the electron prop-
erties at not too low temperatures (due to the sign problem)
for somewhat larger clusters. Both approaches are thus re-
stricted to finite small clusters with energy spacings between
the levels too large to allow one to draw firm conclusions
about the properties of the model in the thermodynamic
limit. A variety of approximation techniques, which are quite
able to address this limit, have been developed. Among these
we find mean-field theories like the Hartree-Fock approxima-
tion\(^12–14\) or the slave-boson approach,\(^15,16\) Green-
function decoupling schemes,\(^1\) large-\(N\) expansions,\(^17,18\)
functional-integral formulations,\(^19\) and variational
approaches.\(^20\) These approximations often lead to conflicting
results or violate well-established sum rules.

The view advocated in this paper is that a fair amount of
exact, or virtually exact, results is now available for some
cases. These results may serve as a criterion to test approxi-
mation schemes which, once tested in those cases, may then
be confidently applied to the general model for relevant
choices of the parameters. As a rule of thumb, the better a
theory is, the more exactly solvable the limits reproduced by
an approximation are. We now present a generalization of
the Hubbard I decoupling approximation\(^1\) which amounts to
allowing a correlation hole of opposite spin to propagate
along with the electron in its path around the lattice. By so
doing, this approximation (1) becomes exact in the strong-
coupling limit; (2) reproduces with great accuracy the
Hubbard Coulomb repulsion \(U\); (3) leads to a one-electron Green’s function
showing the spectral-wave transfer from the high- to the low-
energy degrees of freedom as soon as the electron density
deviates from half-filling; and (4) generates reasonably well-
behaved dynamic correlation functions. On the other hand,
lifetime effects and, therefore, the question of the Luttinger-
versus Fermi-like behavior of the model near the Fermi level
are completely outside the scope of this approximation.

The paper is organized as follows: In Sec. II, a heuristic,
rather intuitive approach to our dynamic correlation-hole ap-
proximation (CHA) is presented. Section III gives a general
discussion of the single-particle Green function, stressing
its properties in several important limiting cases. A more de-
tailed study of the half-filled system follows in Sec. IV, with
an emphasis on the ground-state energy and its first deriva-
tives (on-site correlation functions and kinetic energy),
which are compared with the Bethe ansatz results of Lieb
and Wu.\(^7\) Section V deals with the properties of the Hubbard
model doped away from half-filling, describing the doping
dependence of the band structure and the spectral-weight
transfer phenomenon so characteristic of strongly correlated
electron systems, followed by a comparison of our calculated
momentum distribution with those emerging from QMC
simulations and conserving approximations. Section VI describes our calculation of charge and spin correlations, based on the functional derivative formulation of response functions, the results being compared with QMC and slave-boson calculations. We conclude with some remarks in Sec. VII.

II. HEURISTIC DERIVATION OF THE DYNAMIC CORRELATION-HOLE APPROXIMATION

In order to fix the notation, let us recall the one-band Hubbard Hamiltonian

$$H = -t \sum_{\langle ij \rangle} c_{i \sigma}^\dagger c_{j \sigma} - \mu \sum_i n_{i \uparrow} + U \sum_i n_{i \uparrow} n_{i \downarrow}.$$ (1)

This describes a collection of electrons with chemical potential $\mu$ sitting on a lattice. They can hop $\langle ij \rangle$ between nearest-neighbor sites, but a Coulomb repulsion $U$ acts whenever two electrons of opposite spin sit on the same site. As usual, $c_{i \sigma}$ creates, while $c_{i \sigma}^\dagger$ destroys, an electron of spin $\sigma$ on the $i$th site, $n_{i \sigma} = c_{i \sigma}^\dagger c_{i \sigma}$ being the corresponding occupation number operator. As one more item on notation, the first term of $H$ gives just the noninteracting band structure. To emphasize this, sometimes $H$ is also written in the form

$$H = \sum_{\mathbf{k}s} (\epsilon_{\mathbf{k}} - \mu) n_{\mathbf{k}s} + U \sum_i n_{i \uparrow} n_{i \downarrow}.$$ (2)

In most of this paper, we shall consider a dispersion relation appropriate for a square lattice with nearest-neighbor hopping, i.e., $\epsilon_{\mathbf{k}} = -2t(\cos k_x a + \cos k_y a)$, or just the first term when dealing with a linear chain. As is well known, $U$ can be disposed of by simply changing $\omega \rightarrow \omega + \mu$ in the calculation of dynamical quantities. In what follows we shall write $\omega = \omega_0$, i.e., $\mu = 0$, which amounts to referring all the energies to the chemical potential. Only in the discussion of Figs. 5 and 6 will the Fermi level be explicitly considered, for reasons which will become evident then.

The so-called Hubbard I approximation\(^1\) for the single-particle Green function $G_{\mathbf{k}s}^H(\omega)$ reads as either

$$\frac{1}{G_{\mathbf{k}s}^H(\omega)} = \frac{1}{G_{\mathbf{a}1}(\omega)} - \epsilon_{\mathbf{k}},$$ (3)

or

$$\frac{1}{G_{\mathbf{k}s}^H(\omega)} = \omega - \epsilon_{\mathbf{k}} - U \langle n_{-\uparrow} \rangle - \sigma(\omega),$$ (3a)

where $G_{\mathbf{a}1}$ is the Green function in the separated atoms limit ($t=0$), and $\langle n_{-\uparrow} \rangle$ is the (site-independent) ground-state occupation of an electron with opposite spin which, for brevity, will be denoted simply by $\bar{n}$ ($n$ denoting the same quantity for the like spin). Finally, $\sigma(\omega)$ is the correlation part of the self-energy given by the $k$-independent expression

$$\sigma(\omega) = U^2 \frac{\bar{n}(1 - \bar{n})}{\omega - U(1 - \bar{n}) + \epsilon_{\mathbf{k}}}.$$ (3b)

The approximation can thus be considered as either an expansion about the atomic limit (3a), as usually done, or else, an expansion about the Hartree-Fock limit (3). Then the correlation self-energy is just that of the atomic limit, describing a localized immobile CHA which is left behind on the $i$th site when the electron hops from $i$ to $j$. It is, therefore, not surprising that the correlation energy is underestimated by this approximation. Table I gives the ground-state energy of a four-site chain, at half-filling, for different values of $U/t$. Compare the exact values, obtained by exact diagonalization (first column), against those of the Hubbard I approximation (HI, second column). For very small $U/t$, the latter performs quite nicely, even better that the unrestricted Hartree-Fock approximation (UHF, last column) but, as $U$ increases, both deteriorate consistently underestimating the correlation energy. For $U \geq 2t$ (half the “bandwidth”), the HI is always slightly worse than UHF. The source of this deterioration is, of course, their wrong behavior at large $U$, as Table I shows. Thus for $U = 10t$, HI and UHF give, respectively, a ground-state energy of $-0.07$ and $-0.08$ eV, against the exact value $(-0.12$ eV).

To improve this situation, while still retaining a simple theory, we allow the correlation hole to propagate with the electron throughout the lattice. Instead of the localized $k$-independent correlation self-energy (3b) we try the ansatz

$$\sigma_k(\omega) = U^2 \frac{\bar{n}(1 - \bar{n})}{\omega - U(1 - \bar{n}) + \epsilon_{\mathbf{k}}} + \frac{1}{G_{\mathbf{k}s}^{-1}(\omega)} = \omega - \epsilon_{\mathbf{k}} - U \bar{n} - \sigma_k(\omega),$$ (4)

where we have inserted the hole hopping ($-\epsilon_{\mathbf{k}}$). Our one-electron Green function is thus given by

$$G_{\mathbf{k}s}^{-1}(\omega) = \omega - \epsilon_{\mathbf{k}} - U \bar{n} - \sigma_k(\omega),$$ (5)

whose physical content is very simple: As an electron moves through the electron band $U \bar{n} + \epsilon_{\mathbf{k}}$, it “sees” its own correlation hole moving in the hole band $U(1 - \bar{n}) - \epsilon_{\mathbf{k}}$.

Since the electron is permanently accompanied by its correlation hole in this approximation, it is not surprising to find the correlation energy somewhat overestimated. In the third column of Table I, the corresponding ground-state energies are given. They are very close to, but somewhat lower than, the exact values for large $U$, say $U > 10t$, and exactly reproduce the infinite-$U$ limit for this four-site chain ($-12t^2/U$). As $U$ approaches the bandwidth, $U = 4t$, the agreement deteriorates. However, since both the $U = \infty$ and $U = 0$ limits are satisfied, the discrepancy is never larger than 10%, even in the most unfavorable cases. The reason for this
fair agreement will be explained in Sec. III where the one-electron Green function is analyzed in some detail. An attempt to derive ansatz (5) from the equation of motion method is relegated to the Appendix.

III. SINGLE-PARTICLE GREEN FUNCTION

The single-particle Green function (5) with the correlation self-energy (4) can be readily cast into the form

$$G_{k}(\omega) = \frac{\omega - \frac{U}{2} + \bar{\epsilon}_{k}}{(\omega - \frac{U}{2})^2 - \bar{\epsilon}_{k}^2 - U^2 \bar{n}(1 - \bar{n})} = \frac{z^{(+)}_{k}}{\omega - E^{(+)}_{k}} + \frac{z^{(-)}_{k}}{\omega - E^{(-)}_{k}},$$

(6)

which describes two Hubbard bands (HB) and their corresponding spectral weights (residues) given by

$$E^{(\pm)}_{k} = \frac{1}{2} U \pm \sqrt{\frac{\epsilon_{k}}{2} + U^2 \bar{n}(1 - \bar{n})},$$

(6a)

and

$$z^{(\pm)}_{k} = \frac{1}{2} \left( 1 \pm \frac{\bar{\epsilon}_{k}}{\sqrt{\frac{\epsilon_{k}}{2} + U^2 \bar{n}(1 - \bar{n})}} \right),$$

(6b)

where \(\bar{\epsilon}_{k}\) is the displaced band structure

$$\bar{\epsilon}_{k} = \epsilon_{k} - \frac{1}{2} U (1 - 2 \bar{n}).$$

(6c)

Notice that

$$E^{(+)}_{k} + E^{(-)}_{k} = U,$$

(7)

i.e., the two bands are specularly symmetric with respect to \(\omega = U/2\), and

$$z^{(+)}_{k} + z^{(-)}_{k} = 1,$$

(8)

as it should be. The spectral weight of the lower HB is, of course, the usual momentum distribution \(n_{k}\), so that one can write \(z^{(-)}_{k} = n_{k}\) and \(z^{(+)}_{k} = 1 - n_{k}\). Equations (6) imply that both the Hubbard bands and their spectral weights depend on \(\bar{n}\), i.e., on the average occupation of the opposite spin. This will be important later in the discussion of the spectral-weight transfer between the HB’s as the system departs from half-filling, \(\bar{n} \neq 0.5\) (Sec. V). Let us now see some limiting cases.

(i) \(U = 0\). The displaced band structure (6c) becomes just the noninteracting band structure, and the first line of Eq. (6) at once gives \(G_{k} = (\omega - \epsilon_{k})^{-1}\), i.e., just the noninteracting \(G_{k}\). It is, however, more instructive to look at Eqs. (6a) and (6b). They become

$$E^{(\pm)}_{k} = \pm |\epsilon_{k}|,$$

(9)

$$z^{(\pm)}_{k} = \frac{1}{2} \{ 1 \pm \text{sgn}(\epsilon_{k}) \},$$

(10)

which are just the \(\epsilon_{k}\) and \(n_{k}\) for the noninteracting system remapped into the small Brillouin zone (BZ) \(|k| \leq \pi/2\). Notice, in particular, that \(E^{(\pm)}_{k}\) just describes the empty (occupied) part of the noninteracting band which, upon turning on \(U\), evolves into the upper (lower) Hubbard band.

(ii) \(U = \infty\) (or \(\epsilon_{k} = 0\)). The square root in Eqs. (6) gives just \(0.5U\), and therefore, \(E^{(\pm)}_{k} = U/2\) and 0, while \(z^{(\pm)}_{k} = \bar{n}\) and \(1 - \bar{n}\), respectively, i.e.,

$$G_{k}(\omega) = \frac{-\bar{n}}{\omega - U} + \frac{1 - \bar{n}}{\omega} = \sigma_{ab},$$

(11)

just the atomic limit, as expected. Notice that this same limit applies whenever \(\epsilon_{k}\) happens to vanish, irrespective of the value of \(U\) and \(t\). This occurs, e.g., at \(ka = \pi/2\) in a one-dimensional (1D) chain \(\epsilon_{k} = -2t \cos ka\), at the point \((\pi, 0), (0, \pi), (\pi/2, \pi/2)\), of a square band, etc. At these special points any \(\bar{n} \neq 0\) implies a strong-coupling situation.

(iii) Finally \(\sigma_{k}\) vanishes trivially for \(\bar{n} = 0.5\) (half-filling), retrieving the Hartree-Fock approximation.

For the rest of the discussion, it is convenient to consider separately \(\bar{n} = 0.5\) (half-filling) and \(\bar{n} \neq 0.5\) (away from half-filling).

IV. HALF-FILLED CASE

With \(\bar{n} = 0.5\), the second term in Eq. (6c) vanishes, giving \(\bar{\epsilon}_{k} = \epsilon_{k}\), and everything is simplified a great deal. We now have

$$E^{(\pm)}_{k} = \frac{1}{2} (U \pm \sqrt{U^2 + 4 \epsilon^{2}_{k}}),$$

(12)

and

$$z^{(\pm)}_{k} = \frac{1}{2} \left( 1 \pm \frac{2 \epsilon_{k}}{\sqrt{U^2 + 4 \epsilon^{2}_{k}}} \right).$$

(13)

Each band is now specularly symmetric with respect to \(\epsilon_{k} = 0\), i.e., at the special points mentioned above. They delimit the smaller BZ associated with this symmetry, i.e., \(ka < \pi/2\) for the linear chain or \(|k, a \pm k, a| \leq \pi\) for the square lattice, etc. This symmetry, sometimes called nesting for certain band structures, was not invoked at all in designing our approximate self-energy [Eq. (4)]. It just arose itself in the half-filled case. A gap is opened at the BZ boundary of magnitude

$$E_{g} = E^{(+)_{k}}_{k} - E^{(-)_{k}}_{k} = U,$$

(14)

which decreases linearly with \(U\). At both sides of the gap, \(z^{(+)_{k}}_{k} = 0.5\).

For large \(U \gg |\epsilon_{k}|\), we are in the strong-coupling regime, and Eqs. (10) and (11) give, to the lowest order in \(\epsilon_{k} / U\),

$$E^{(\pm)}_{k} = \frac{1}{2} \left( U \pm \frac{1}{2} U \left( 1 + \frac{2 \epsilon^{2}_{k}}{U^2} \right) \right),$$

(15)

and

$$z^{(\pm)}_{k} = \frac{1}{2} \left( 1 \pm \frac{2 \epsilon_{k}}{U} \right),$$

(16)

i.e., two narrow bands showing spin-wave dispersions, \(U + \epsilon^{2}_{k} / U\) and \(- \epsilon^{2}_{k} / U\), respectively. Taking, e.g., the lower band for a 1D chain, we have
\[ E_k^{(\pm)} = -\frac{4t^2}{U} \cos^2(ka) = -J \cos^2(ka), \]  
(17)

just a band of width \( J \) (the doping-induced antiferromagnetic exchange) along the small BZ. Notice that Eq. (16) implies that both bands have almost the same spectral weight

\[ \varepsilon_k^{(\pm)} \approx \frac{1}{2} |J t \cos(ka)|, \]  
(18)
i.e., 0.5 corrected by a small oscillatory tail. This is characteristic of the strong-coupling regime. The Green function (6) thus behaves correctly in this regime. This is why the ground-state energy is so accurately reproduced for large \( U \) (see below).

As \( U \) decreases, however, both spectral weights start to differ increasingly and, for any \( k \), we always have a strong band and a weak band, or satellite or shadow, band, as it is sometimes called, entirely due to correlation. The bandwidth steadily increases. For very small \( U \ll t \), finally, one has the weak-coupling results

\[ E_k^{(\pm)} = \frac{U}{2} \pm \varepsilon_k \left( 1 + \frac{U^2}{8 \varepsilon_k^2} \right), \]  
(19)

\[ \varepsilon_k^{(\pm)} \approx \frac{1}{2} \pm \frac{1}{2} \text{sgn}(\varepsilon_k) \left( 1 - \frac{U^2}{8 \varepsilon_k^2} \right), \]  
(20)

which give just the Hartree-Fock result. Notice, however, that the limit breaks down if \( \varepsilon_k=0 \), i.e., close to the BZ boundary. Then any \( U \neq 0 \) implies strong coupling, so strong in fact that a gap is opened. For a 1D chain this happens just at the Fermi level. This actually means that the series expansion is not permissible there, and one simply has to apply Eqs. (12) and (13) directly. The purpose of this apparently trivial discussion, however, has been twofold: first to show clearly that our approximation (6) for the Green function satisfies both the weak- and strong-coupling limits, and, second, the kind of pitfalls one can make when taking the weak-coupling limit.

Once the Green function is known, the ground-state (GS) energy follows at once from Galitskii and Migdal’s theorem:\footnote{5}

\[ E_{gs} = \sum_{k \epsilon} \int \frac{d\omega}{2\pi \imath} \frac{\omega + \varepsilon_k}{2} G_{ks}(\omega), \]  
(21)

where the contour of integration encloses the occupied poles of \( G \). From Eqs. (11) and (12), we therefore have

\[ E_{gs} = \sum_k \left( E_k^{(-)} + \varepsilon_k \right) \xi_k^{(-)} = \frac{1}{4} \sum_k \left( U - \sqrt{U^2 + 4\varepsilon_k^2} - \frac{4\varepsilon_k^2}{\sqrt{U^2 + 4\varepsilon_k^2}} \right). \]  
(22)

All the momentum sums run over the full Brillouin zone \( |k| \approx \pi/a \). Just for this special case of half-filling, Eq. (22) depends on \( \varepsilon_k^2 \) only and, therefore, the small zone \( \pm \pi/a \) may be used. The factor in front of Eq. (22) would then be 0.5. For \( \pi/2a \), however, this symmetry is lost and the large zone must be used.

The appropriate limits are, of course, satisfied. For \( U = 0 \) one recovers the noninteracting GS. When \( U = \infty \), \( E_{gs} = 0 \). Finally, if \( Ut \approx 1 \) we find, expanding the square roots to first order in \( (t/U)^2 \),

\[ E_{gs} = \frac{3}{2} \sum_k \frac{\varepsilon_k^2}{U}, \]  
(23)

which for, e.g., a hypercubic lattice in \( p \) dimensions,

\[ \varepsilon_k = -2t \sum_{n=1}^p \cos \varphi_{nk}, \]

with \( \varphi_{nk} = k_n a \), gives

\[ E_{gs} = -3 \sum_{n} \int \frac{T}{U} \frac{d\varphi_n}{2\pi} \cos^2 \varphi_n \]

\[ = -3 \sum_{n} \frac{4t^2}{U} p = -NpJ \langle S^2 \rangle, \]  
(24)

where \( N \) is the number of sites. This is the GS energy of the undistorted Neél state. In Eq. (24), \( \langle S^2 \rangle \) is the (site-independent) average value of the squared spin \( \langle S_j S_i \rangle \) on the \( i \)th site (just \( \frac{1}{4} \) when \( U \to \infty \)).

Differentiating \( E_{gs} \) with respect to \( U \) and \( t \), one obtains the on-site correlations and the kinetic energy, respectively. Thus, the double occupancy is given by

\[ d_i = \langle n_i, n_i \rangle = \frac{1}{N} \frac{\partial E_{gs}}{\partial U}, \]  
(25)

which determines at once the on-site charge and spin correlations:

\[ \langle q_i^2 \rangle = \langle (n_i, n_i) \rangle = \langle q_i \rangle + 2d_i, \]

\[ \langle \sigma_1^2 \rangle = \langle (n_i, n_i) \rangle = \langle q_i \rangle - 2d_i. \]  
(26)

On the other hand, differentiating with respect to \( t \) yields

\[ t \frac{\partial E_{gs}}{\partial t} = -t \sum_{(j,js)} \langle e_j^c e_j \rangle = T = IT_0, \]  
(27)

where \( T_0 = T(U=0) \), \( T \) being the kinetic energy. The ratio \( I = T/T_0 \) measures the degree of itinerancy of the system. For \( U = 0 \), \( I = 1 \), and the system is fully delocalized. As \( U \) increases, \( I \) decreases, the system becoming increasingly localized. An alternative way of phrasing this is that \( I \) gives the reduction in effective hopping since, as \( U \) increases, the electrons are finding ever more difficult to hop between sites. Equations (25) and (27) are just a statement of the Hellman-Feynman theorem for the Hubbard Hamiltonian. They are not independent since, taking the ground-state average of \( H \), Eq. (1), we must have

\[ E_{gs} = t \frac{\partial E_{gs}}{\partial t} + U \frac{\partial E_{gs}}{\partial U}, \]  
(28)
which means simply that $E_{gs}$ is a homogeneous function of the first degree in the variables $U$ and $t$. Therefore,

$$I = \frac{E_{gs} - UN_d}{T_0} = \frac{\varepsilon_{gs} - Ud}{\varepsilon_{gs}^0},$$

(29)

in terms of quantities per site. (Notice that $T_0$ is just the whole GS energy at $U=0$.)

**A. One dimension: The linear chain**

The GS energy (22) can be expressed in closed form for the infinite chain in terms of complete elliptic integrals. We shall then compare with Lieb and Wu’s exact result,

$$E_{gs}^{LW} = -4N \int_0^\infty d\omega \frac{J_0(\omega)J_1(\omega)}{\omega \left[1-\exp\left(-\frac{1}{2}\frac{U}{t}\right)\right]},$$

(30)

where $J_n$ are the Bessel function of integer order. Equation (22) can be rewritten as

$$E_{gs} = \frac{1}{4} \sum_k \left[ U - 2 \sqrt{U^2 + 4\varepsilon_k^2} + \frac{U^2}{\sqrt{U^2 + 4\varepsilon_k^2}} \right],$$

(31)

which for the infinite chain reads (per site)

$$\varepsilon_{gs} = \frac{1}{4} \int_0^\pi \frac{d\varphi}{2\pi} \left[ U - 2 \sqrt{U^2 + 16t^2 \cos^2 \varphi} + \frac{U^2}{\sqrt{U^2 + 16t^2 \cos^2 \varphi}} \right] = \frac{U}{4} - \frac{4t}{\pi m} \frac{E(m)}{8\pi t} K(m),$$

(32)

where $\varphi = ka$, $m = 4t(U^2 + 16t^2)^{-1/2}$, and $K$ and $E$ are the elliptic integrals of the first and second kinds, respectively.

For $U/t \gg 1$, our result (32) goes, of course, into Eq. (24) for $p=1$, i.e., $\varepsilon_{gs} = 0.75J$ (per site) while Lieb and Wu’s (LW) expression (30) goes into $\varepsilon_{gs} = -J \ln 2$. It is apparent that the contribution to the GS energy of the zero-point quantum spin fluctuations is absent in Eq. (32). We can try to remedy this omission by rescaling Eq. (32) so that its large-$U$ behavior be that of LW’s solution. This is most simply done by rescaling $U$ to a new value $U_{eff} = (3/4 \ln 2)U$. Table II gives the GS energy and double occupancy as $U$ increases from 0 to 100. Three calculations are compared: (a) the LW result [Eq. (30)], (b) our result [Eq. (32)], and (c) Eq. (32), with the rescaled $U_{eff}$. Notice, first of all, that the three sets of values are very close to each other. For small $U$, $E^b$ differs only slightly from the LW value, but it starts to deteriorate for $U > 2t$. Just the opposite behavior is found in $E^c$: it becomes increasingly better for increasing $U > 2t$. With rescaling to fit the large-$U$ behavior of the LW solution, the agreement of $E^b$ at low $U$ is spoiled. Even so, since $E^c$ must tend to the exact $E_{gs}(U=0) = -1.27$, it never deviates much from the LW value $E^a$ ($\sim 1\%$ for $U/t = 1$). Concerning the double occupancy, the deviations range between 5% and 10% until fairly large $U$. One must go up to $U/t > 20$ to find $d^c$ values sufficiently close to the LW values $d^a$. The performance of $d^c$ is thus worse than that of $E^c$. This is only to be expected. It simply tells us that the wave function implicit in

**B. Two dimensions: The square lattice**

Since no closed expressions are now available, apart from Eq. (24) for the large-$U/t$ limit, one must resort to numerical evaluation of Eq. (22) and its derivatives. Figures 1–4, which refer to a $4 \times 4$ lattice with periodic boundary conditions, give the GS energy [Eq. (22)], the double occupancy $d$, the squared local moment $m^2 = \langle \sigma_i^z \rangle$, and the ratio of kinetic energies, $I = I^0/T_0$, for increasing values of $U/t$. ($t = 1$ from now onwards.)

The full curve in Fig. 1 represents the numerical evaluation of Eq. (22), the white circles corresponding to the exact diagonalization (Lanczos) results of Ref. 23, while the dotted curve gives the UHF energies of Louis et al. Notice the

![FIG. 1. Ground-state energy $E_{gs}$ (in eV) vs $U/t$ ($t=1$) for a $4 \times 4$ cluster with periodic boundary conditions. This work, Eq. (22) (full line). Exact diagonalization from Ref. 23 (open circles). Unrestricted Hartree-Fock from Ref. 24 (dotted line).](image-url)
(somewhat surprising) extraordinary good fit of our calculation to the exact Lanczos calculation along the whole range of \( U \) values. The UHF curve deviates up slightly, but is also very close to the exact curve. The full curves in Figs. 2–4 represent our evaluation of Eqs. (25)–(27), the white circles corresponding this time to the QMC simulations of White et al.\(^\text{11}\) As \( U \) increases, \( d \) must decrease since the electrons are ever less able to double occupy the lattice sites. Consequently, the local (on-site) squared moment [Eq. (26)], increases from its \( U=0 \) values, 0.5 up to unity at infinite \( U \). The agreement between the full line and the white circles is very good for both small and large \( U \), but it deteriorates in the intermediate region, \( U/t \) between 2 and 10. The agreement for \( I \), the relative kinetic energy or effective hopping, in Fig. 4 is less good along the whole range of \( U/t \) values represented (except at \( U=0 \), of course). We have checked, however, that relationship (29) is satisfied very accurately. We are talking, in any case, of small differences which, in the less favorable case (\( U/t \gtrsim 6 \)), amount to a deviation of \( \sim 8\% \). Our curve, therefore, also interpolates smoothly between the two dotted curves labeled \( a \) and \( b \), which are a second-order perturbation result and a strong-coupling result, respectively [Eqs. (36) and (38), and Fig. 7 of Ref. 11].

Our conclusion is thus very much like in the 1D case. Our dynamic CHA gives very accurate GS energies but less accurate (although still fair) GS derivatives, i.e., correlation functions. This is perhaps a signal that the implicit GS wave function associated with Eq. (5) is not altogether satisfactory (see Sec. VII below).

V. AWAY FROM HALF-FILLING

When \( \bar{n} \neq 0.5 \), the symmetry about \( \epsilon_k=0 \) is lost, and the general expressions (6) apply and we have to work in the full BZ. Now we face a somewhat surprising situation at first sight, namely, that the number of states per spin in each band, \( N'(\pm) = \sum_k \xi_k^{(\pm)} \), depends on the occupation, although, of course, Eq. (7) ensures us that the total number of states in both bands together equals \( N \), the number of lattice sites. Since the same situation applies to the Hubbard I approximation [Eq. (3)], Hubbard himself\(^\text{10}\) thought this was a weakness of the approximation, no simple interpretation of the weights \( \xi_k^{(\pm)} \) being available at the time. Thus it remained largely ignored. However, recent experiments in the wake arisen by the study of the high-\( T_c \) copper oxides found spectral-weight transfer from the upper HB to the lower HB near the Fermi edge. Such was the case of O 1s x-ray-
absorption study\textsuperscript{25,26} of La\textsubscript{2-x}Sr\textsubscript{x}CuO\textsubscript{4} upon hole doping in the O 2s band. A similar behavior has been found in other correlated systems.\textsuperscript{27,28} A general discussion of the physical origin of this redistribution of intersites and its doping dependence was made in Ref. 4. This occurrence of spectral-weight transfer in now commonly accepted as a fingerprint for correlations effects, and has been observed in several numerical calculations of correlated electron systems.\textsuperscript{29}

In order to study this transfer of spectral weight in our dynamic CHA, let us look more closely at Eqs. (6). First of

FIG. 5. Doping dependence of the Hubbard bands $E_k$ (in eV) for a linear chain with ten sites. From left to right, the panels are for $x = -0.5$ (a), $x = -0.1$ (b), $x = 0$ (c), $x = 0.1$ (d), and $x = 0.5$ (e). The horizontal dotted lines mark the Fermi level.
all, we examine the band structure.

A. Doping dependence of the band structure

The energy gap is just the min \( (E_k^+ - E_k^-) \) and occurs at \( \tilde{E}_k = 0 \), i.e., at

\[
e_k = \frac{1}{2} U(1 - 2\tilde{n}) = -\frac{1}{2} Ux,
\]

taking the value

\[
E_g = 2U\sqrt{\tilde{n}(1-\tilde{n})} = U\sqrt{1-x^2},
\]

where we have assumed \( \tilde{n} = n = q/2 \) and defined \( q = 1 + x \), \( x \) being the charge deviation from half-filling, or degree of doping. The spectral weights at both sides of the gap are then 0.5, i.e., both magnitude of the gap and the locus of points \( k_0 \) in the BZ with spectral weight 0.5 depend on the degree of doping. For the 1D chain, e.g., Eq. (33) gives

\[
\cos k_0 a = \frac{U}{4t} x
\]

i.e., \( k_0 a \approx \pi/2 \) for \( x \approx 0 \). The points \( k_0 \) lie inside and outside the small BZ for electron and hole doping. For \( t > 0 \) the situation is clearly the reverse one. But this set of points might not exist at all if \( U \) and/or \( x \) are large enough so that Eq. (35) ceases to hold. The gap will then be located at \( k = 0 \) for electron doping, or \( k = \pi/a \) for hole doping. For a square lattice, this would happen at the center or at the corners of the 2D BZ, respectively. The gap is then (\( |x| < 1 \))

\[
E_g = \sqrt{U^2 + 16t^2 - 8Ut|x|}
\]

for the linear chain, and

\[
E_g = \sqrt{U^2 + 64t^2 - 16Ut|x|}
\]
for the square lattice. For $|x|=1$, $n$ takes the values 0 or 1, $\alpha_k$ vanishes, and one of the bands disappears (the corresponding $z_k$ vanishes).

These features are clearly apparent in Fig. 5, which displays the two HB's of a ten-site linear chain for $U=10$ and $t=1$ and several occupations. The central panel shows the half-filled chain with ten electrons ($n=0.1$ and $x=0$). Notice the specular symmetry of both bands with respect to $ka=\pm \pi/2$, already discussed. At both sides, several cases of electron (right panels) and hole (left panels) doping are shown, namely, two situations immediately away from half-filling ($n=0.55$ and $x=0$ and 1, and $n=0.45$ and $x=-0.1$) as well as two more of fairly large doping ($n=0.75$ and $x=0.5$, and $n=0.25$ and $x=-0.5$). Immediately away from half-filling, Eq. (35) is clearly satisfied by two pairs of points, $k_0\delta=\pm 0.42\pi$ for electron doping and $k_\delta=\pm 0.58\pi$ for hole doping. At these points the lower and upper HB's go through their maximum and minimum values. At $x=\pm 0.5$, Eq. (35) in no longer satisfied, and $z_k^{\pm(-)}$ never go through 0.5. As stated above, and is fairly apparent in Fig. 5, the minimum value of the gap occurs now at $ka=0$ or $\pi$ for electron or hole doping [Eq. (36)]. Notice the reversal in band curvature when going from electron to hole doping.

Some general comments are now in order. For low doping rates, the bands are very flat in the neighborhood of the points $k_0$, which is only natural as they are going through an extremum value. Although these portions of the bands lie very close to the Fermi level, we have not been able to establish any connection, in the analogous case of the square lattice, with the famous flatbands of the copper oxides.30 Both facts seem quite unrelated, there being no simple relationship between the points $k_0$ and the Fermi level. This has to be found self-consistently, as usual, so as to accommodate the right number of electrons in the bonds. It is marked by the dotted horizontal lines in the five panels of Fig. 5. In drawing this figure, we have chosen the midgap as origin of energies so that we can easily follow the Fermi-level change with doping. Notice that the bands move only slightly, while varying in shape and width. It is the Fermi level which moves. This takes us to the following subject, namely, the transfer of spectral weight when the half-filled system is doped.

\section*{B. Transfer of spectral weight in the doped chain}

Figure 6 shows the density of states $N(E)$ for the case discussed in Fig. 5. As usual, $N(E)$ is given by

\begin{equation}
N(E) = -\frac{1}{\pi N} \sum_k \text{Im} G_k(E + i \eta) = \frac{1}{N} \sum_k \{z_k^{+(+)} \delta(E-E_k^{(+)} + z_k^{(-)} \delta(E-E_k^{(-)})\},
\end{equation}

where the $k$ sums run over the full BZ, and $\eta \to 0$. We have found it convenient to arrange the figure vertically with five panels, in correspondence with Fig. 5. Thus the central panel corresponds to half-filling, the upper panels correspond to hole doping, and the lower panels correspond to electron doping. The transfer of spectral weight is apparent. For $x=0$, the two HB's are identical, and symmetrically disposed with respect to the midgap. For $x>0$ (electron doping) the upper structure increases both in width and in area, to the detriment of the lower one, the opposite being true for hole doping. At the same time, the Fermi level (vertical dotted line) moves to the upper band for electron doping while moving to the lower band for hole doping. It is clear that some states have been transferred from one of the bands to the other.

To evaluate this spectral-weight transfer, we simply have to keep in mind that the area of each peak equals 0.5 for $x=0$. Therefore, the relative spectral-wave transfer (SWT) is simply given by (summing over spin)

\begin{equation}
S^{(\pm)} = \frac{2}{N} \sum_k \bar{e} \delta \left\langle \left\langle \pm \right\rangle \right\rangle - 1
\end{equation}

for electron and hole doping, respectively. Equation (6b) then yields

\begin{equation}
S^{(\pm)} = \pm \frac{1}{N} \sum_k \frac{\bar{e} \delta \left\langle \left\langle \pm \right\rangle \right\rangle}{\sqrt{\bar{e}^2 + \frac{1}{4} U^2(1-x^2)}},
\end{equation}

which can be rewritten as

\begin{equation}
S^{(\pm)} = \pm \int_{-\infty}^{\infty} d\epsilon N_0(\epsilon) \frac{\partial}{\partial \epsilon} \sqrt{\bar{e}^2 + \frac{1}{4} U^2(1-x^2)} = \pm \int_{-\infty}^{\infty} d\epsilon \frac{\partial N_0(\epsilon)}{\partial \epsilon} \sqrt{\bar{e}^2 + \frac{1}{4} U^2(1-x^2)},
\end{equation}

where $N_0(\epsilon)$ is the density of states of the unperturbed ($U=0$) band structure and $\bar{e} = e + U x/2$. Equation (41) shows that the SWT is specially tied to rapid changes in $N_0(\epsilon)$. In a 1D chain, thus, the main contribution comes from the band edges while the van Hove logarithmic singularity at the band center would play the dominant role in a square lattice.

For the ten-site chain with $U=10$ and $t=1$, we can evaluate directly the sum in Eq. (40) and find the results shown in Table III. Only hole-doping rates have been considered, since Eq. (39) is even in $x$. The SWT is very close to, but somewhat smaller than, the doping degree. Some authors define the low-energy SWT as the empty part of the lower HB, for hole doping, or the filled part of the upper HB for electron doping. This quantity is given below the heading LESWT in the last column of Table III. It is somewhat smaller than $2x$, in contrast to the exact diagonalization result which yields values somewhat above $2x$. An underestimation of the SWT may be a clear signal that some additional structure in the neighborhood of the Fermi level is missing in our mobile CHA.
TABLE III. Spectral-weight transfer (SWT) and low-energy spectral-weight transfer (LESWT) for different hole-doping rates. The corresponding average electron density (n) and the area of the lower Hubbard band (area) are given as a guide.

<table>
<thead>
<tr>
<th>n</th>
<th>x</th>
<th>Area SWT</th>
<th>LESWT</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.0</td>
<td>0.500</td>
<td>0.00</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2</td>
<td>0.590</td>
<td>0.180</td>
</tr>
<tr>
<td>0.3</td>
<td>0.4</td>
<td>0.682</td>
<td>0.364</td>
</tr>
<tr>
<td>0.2</td>
<td>0.6</td>
<td>0.778</td>
<td>0.556</td>
</tr>
<tr>
<td>0.1</td>
<td>0.8</td>
<td>0.882</td>
<td>0.764</td>
</tr>
<tr>
<td>0.0</td>
<td>1.0</td>
<td>1.000</td>
<td>1.00</td>
</tr>
</tbody>
</table>

C. Doping dependence of the momentum distribution

To end our discussion of the one-electron Green function, in Fig. 7 we give the momentum distribution, or occupancy factor n_k for an 8x8 square lattice, U/t = 4, and two occupations: (a) n = 0.5 (x = 0) and (b) n = 0.4375 (x = 0.125). Our results (full diamonds) are compared with those of QMC simulations, and parquet (triangles). In this and the following figures, all the quantities are shown along the Brillouin-zone contour Γ(0,0) → X(π,0) → M(π,π) → Γ(0,0).

The agreement is fairly good along the ΓM direction, for both occupations, with a clear tendency to develop a jump around the middle, i.e., near the Fermi line. Along the other directions, the agreement remains good for x = 0.125 [Fig. 7(b)] but it deteriorates just at half-filling, particularly along the ΓX direction [Fig. 7(a)], our curve being smoother around the X point. Notice that both X and (0.5π, 0.5π), about which a jump tends to develop in n_k, are special points in the sense described above (ε_k = 0).

VI. CORRELATION FUNCTIONS

Once a specific self-energy approximation has been made, namely, Eq. (4), one can easily develop conserving approximations for all the correlation functions by, e.g., using the functional derivative technique and following Baym-Kadanoff procedure. Simply stated, we can explicitly find the functional derivative of σ_{is,js}(t) with respect to an external field coupled to the electron density ⟨n_is(t)⟩. By so doing, we can go beyond the random phase approximation (RPA). Just as the RPA is the conserving approximation associated with the Hartree approximation for the single-particle G, so our approximate correlation functions will be the corresponding conserving approximations following from the dynamical CHA, for G [Eq. (5)].

We are specifically interested in the dynamic charge and spin correlation functions, which are the k,ω Fourier transform of the time-ordered linear response functions

\[ \chi_{ij}^{\alpha\beta}(t) = \frac{1}{i} \langle T q_i(t) q_j(0) \rangle \]

and

\[ \chi_{ij}^s(t) = \frac{1}{i} \langle T S_i^s(t) S_j^s(0) \rangle, \]

where, as in Eq. (26), q_i = n_{iσ} + n_{i\bar{σ}} and S_i^s = (n_{iσ} − n_{i\bar{σ}})/2, and T is Wick’s chronological operator. The contribution of the transverse spin components associated with the deviation operators S_i^s can be dealt with in a similar manner, but, for brevity’s sake, will be omitted here. We thus need to find

\[ \chi_{is,js}'(t) = (1/i) \langle T n_is(t) n_js'(0) \rangle \]

which, according to a well-known trick, is given by

\[ \chi_{is,js}'(t) = \frac{1}{i} \langle n_is(t) n_js'(0) \rangle - i \left( \frac{\delta G_{is,js}(t, t); \xi \rangle \langle \delta G_{is,js}(t, t); \xi \rangle}{\delta \xi} \right|_{\xi = 0}, \]

where G here is the nonequilibrium Green function under the action of an external potential ξ(t) = Σ_{is} ξ_is(t) n_is(t), which is let to vanish after computing the functional derivative. It should be kept in mind that Eq. (44) holds only in the imaginary time domain, 0 < t < −iβ (β = 1/kT), not to be confused with the real time of Eqs. (42) and (43).

To proceed, it is convenient to simplify the notation in order to avoid lengthy expressions. Let α = (iṣt) be a collective index denoting site i, spin s, and imaginary time t. Then Eq. (44) can be rewritten as

\[ \langle \delta n_α, \delta n_β \rangle = -i \frac{\delta G_{αα}(ξ)}{\delta ξ_β} = -i \sum_γ G_{αγ} \Lambda_{γσ β} G_{δσ}, \]

(45)
where $\delta n_a = n_a - \langle n_a \rangle$, and, therefore, Eq. (45) is just the density-fluctuation propagator. The triangular vertex $\Lambda_{\gamma\delta\beta}$ is just the functional derivative of the inverse $G$ which, from the equation of motion for $G$, reads

$$
\Lambda_{\gamma\delta\beta} = -\frac{\delta G^{-1}_{\gamma\delta}}{\delta \xi_{\beta}} = \left( \delta_{\gamma\beta} + U \frac{\delta (\delta n_{\gamma})}{\delta \xi_{\beta}} + \frac{\delta \sigma_{\gamma\delta}}{\delta \xi_{\beta}} \right) \delta_{\gamma\delta} + \frac{\delta \sigma_{\gamma\delta}}{\delta \xi_{\beta}}
$$

(46)

where the bar over $\gamma$ means an opposite spin site. As is well known, the terms within brackets on the right-hand side lead to the RPA. The self-energy derivative is readily found from Eq. (4). Fourier transforming back to space and imaginary time

$$
\sigma_{\gamma\delta} = U^2 g_{\gamma\delta}(\langle n_{\delta} \rangle) (1 - \langle n_{\delta} \rangle)
$$

(47)

where $g_{\gamma\delta}$ is the modified propagator given by

$$
\left\{ \frac{\partial}{\partial t} - \xi_{\gamma} - U(1 - \langle n_{\gamma} \rangle) \right\} \delta_{\gamma\delta} + t_{\gamma\delta} = (g^{-1})_{\gamma\delta},
$$

(48)

t_{\gamma\delta} = t_{ij} \delta_{\alpha\beta} \delta(t - t')$ being the hopping matrix. Then one has

$$
\frac{\delta \sigma_{\gamma\delta}}{\delta \xi_{\beta}} = U^2 \frac{\delta g_{\gamma\delta}(\langle n_{\delta} \rangle)}{\delta \xi_{\beta}} (1 - \langle n_{\delta} \rangle) + U^2 g_{\gamma\delta}(1 - 2 \langle n_{\delta} \rangle) \frac{\delta \langle n_{\delta} \rangle}{\delta \xi_{\beta}}
$$

(49)

and

$$
\frac{\delta g_{\gamma\delta}}{\delta \xi_{\beta}} = \sum_{\mu} g_{\gamma\mu} \left( \delta_{\mu\beta} - U \frac{\delta \langle n_{\mu} \rangle}{\delta \xi_{\beta}} \right) g_{\mu\delta}.
$$

(50)

Notice that, first, the second term on the right-hand side of Eq. (49) acts only away from half-filling and, second, the reversal in sign of the $\delta\langle n \rangle$ term in Eq. (50), in contrast to the usual plus sign, cf. Eq. (46). After Eqs. (45), (46), (49), and (50), we finally obtain the RPA-like expression

$$
\langle \delta n_a, \delta n_b \rangle = \chi_{a\beta}(1) \delta_{ab} + U \sum_{\gamma} \chi_{a\gamma}(\delta n_{\gamma}, \delta n_{b\gamma}),
$$

(51)

where, instead of the free susceptibility $\chi_{a\beta}^{(0)} = -iG_{a\beta}G_{\beta a}$, we find

$$
\chi_{a\beta}^{(1)} = -i[G_{a\beta}G_{\beta a} + \bar{n}(1 - \bar{n})F_{a\beta}F_{\beta a}^+]
$$

(52)

and

$$
\chi_{a\beta}^{(2)} = -i[G_{a\beta}G_{\beta a} - \bar{n}(1 - \bar{n})F_{a\beta}F_{\beta a}^+] - i(\bar{z} - \bar{n})(F_{a\beta}G_{\beta a} + G_{a\beta}F_{\beta a}^+),
$$

(53)

with

$$
F_{a\beta} = U(Gg)_{a\beta}.
$$

(54)

Going now through the usual trick of symmetrizing (and antisymmetrizing) Eq. (51) with respect to the spin indices and, lastly, Fourier transforming into $k$ and $\omega$ (following the well-known prescriptions for imaginary time), one finally ends up with the following expressions for the dynamic charge and spin susceptibilities

$$
\chi_{\delta}^{(1)}(\omega) = (q)^2 \delta_{\omega \omega} + 2 \chi_{\delta}^{(1)}(\omega) - \frac{\chi_{\delta}^{(2)}(\omega)}{1 - \chi_{\delta}^{(1)}(\omega)},
$$

(55)

FIG. 8. Magnetic structure factor $S_k$ for an $8 \times 8$ cluster at half-filling, and $U/t = 4$. Our result is shown by full diamonds whereas the open circles correspond to the QMC simulation (Ref. 31).

$$
\chi_k^{(1)}(\omega) = \frac{\chi_k^{(1)}(\omega)}{2(1 + U\chi_k^{(2)}(\omega))}.
$$

(56)

Figures 8–10 show our results in an $8 \times 8$ lattice for the magnetic and charge structure factors defined by

$$
S_k = 4 \int_0^\pi \frac{d\omega \pi}{2} \text{Im} \chi_k^{(1)}(\omega),
$$

and

$$
Q_k = \int_0^\pi \frac{d\omega \pi}{2} \text{Im} \chi_k^{(2)}(\omega).
$$

Figure 8 compares our results at half-filling for $S_k$ (full diamonds), with the QMC simulation of Moreo et al. Although our peak at the $M$ point is not so sharp, both plots are quite similar. Hence antiferromagnetic spin correlations of fairly long range are also reproduced by our dynamic CHA. Figure 9 displays $S_k$ for $x = 0.275$ $(2n = 0.725)$. Our results (full diamonds) are shown along with those of QMC (open circles) and the slave-boson calculation of Zimmermann et al. (open squares). The overall shape of the three plots is in fairly good agreement, although some differences at the quantitative level are clearly apparent. We find a peak and a shoulder somewhat away from $M$ and $X$, respectively, in almost quantitative agreement with the QMC calculation. Figure 10 shows the result of the same three calculations for the charge structure factor $Q_k$. The agreement between them is now even better.
away from half-filling is correctly reproduced, although somewhat underestimated. Likewise, the momentum distribution behaves nicely although with somewhat smoother behavior than QMC. The same comment applies to the spin structure factor, especially at half-filling, where the peak of $S(\pi, \pi)$ does not acquire the extreme sharpness (long-range antiferromagnetism) of QMC simulations.

Given its extreme simplicity, the good performance of this approximation for global and short-range properties is remarkable, being in qualitative, and sometimes even quantitative, agreement, with QMC simulations, slave-boson calculations, or conserving approximations like FLEX, parquet, etc., with less effort. It is far better than both the UHR and the Hubbard I approximation with practically the same effort. However, the long-range lattice properties, and hence the small-$k$ behavior of the spectral functions, do not seem quite correct. This is signaled by a number of dysfunctions here and there: (i) The energy gap is always linear in $U$, even for very small $U$ much smaller than the bandwidth. (ii) The abrupt jump of the chemical potential from the midgap to one of the HB’s as soon as the system is doped, in contrast to the much smoother variation which seems to emerge from, e.g., QMC simulations. (iii) The spectral-weight transfer is underestimated. These last two observations point toward the existence of some additional structure near the Fermi level for the doped system. This would tend to both pin the Fermi level near half-filling, and also give some additional contribution to the spectral transfer. Finally, (iv) our self-energy [Eq. (4)], is real, with no allowance whatsoever for lifetime effects.

This situation is not much surprising if one looks closely at the decoupling made (see the Appendix). To start with, the whole term with the time-derivative of $n_{i,s}$ in Eq. (A3) has been dropped. This omission throws out lifetime effects, and strongly constrains charge and spin fluctuations. In fact, a cursory evaluation of this term in momentum space leads at once to the charge and spin susceptibilities. However, the approximation then ceases to be simple, and leads to some variant of the random-phase approximation. This omission is partially counterbalanced by our specific assumption [Eq. (A4)], where the strict decoupling of the Hubbard I approximation $n_{i,s} = (n_{i,s})$ is relaxed by allowing $n_{i,s}$ to fluctuate within the “resonating bonds” surrounding the $i$th lattice site. But this is a short-range assumption, and, consequently, this approximation works for short-range, high-energy properties.

It seems, therefore, that the whole realm of low-energy, near-Fermi-level, properties is outside the scope of our dynamical CHA. While this is partially so, one should keep in mind that a simple theory where both high- and low-energy properties are considered on the same footing is still lacking and most likely is an impossible task. In the process of projecting out the high-energy degrees of freedom from the initial Hamiltonian (by whatever method, be bosonization, the renormalization group, etc.) one ends up with an effective Hamiltonian which correctly describes the low-energy physics, but not the high-energy properties. The procedure can be used for any energy range leading in each case to a different effective Hamiltonian. It would thus appear that no single effective Hamiltonian can describe the properties of a highly correlated electron system in the full energy range.

Some improvements can be tried. An observation, initially due to Hubbard himself (the absence of a two-electron repulsive bound state for small $U$ in narrow bands should be reflected in the single-particle Green function and is not), ties in with Anderson’s unrenormalizable Fermi-surface phase-shift idea. This can be incorporated into our formalism most likely via the $n_{i,s}$ derivative term. A widely different kind of improvement is letting $\bar{n}$ and $\bar{n}$ to be calculated self-consistently as in the unrestricted Hartree-Fock approach. We can then look for inhomogeneous phases. Notice that, throughout this paper, we have always assumed $n = \bar{n}$, i.e., $\langle S_z \rangle = 0$, and have thus gone a long way further than with a paramagnetic HF solution. Both improvements are presently being the object of active research in our group.

ACKNOWLEDGMENT

This work was partially supported by the Spanish Ministry of Education and Science through DGICYT Project No. PB95-0069.
APPENDIX

Here we derive Eqs. (4) and (5) from the equation-of-motion (EM) method for the single-particle Green function

$$G_{is,js}(t) = \frac{1}{\mathcal{I}} \langle Tc_i(t)c_j^+(0) \rangle = \langle c_i, c_j^+ \rangle$$  \hspace{1cm} (A1)

where $T$ is Wick’s chronological operator, and the second expression on the right-hand side is a shorthand notation. The first two steps of the EM for $G$ read

$$i \frac{\partial}{\partial t} G_{is,js}(t) = \delta_{ij} \delta(t) - i \sum_G G_{is,js}(t) + U \Gamma_{is,js}(t),$$ \hspace{1cm} (A2)

where $\Gamma_{is,js}(t) = \langle n_{i,s} c_{is}, c_{js}^+ \rangle$, and

$$\left( i \frac{\partial}{\partial t} - U \right) \Gamma_{is,js} = \langle n_{i,s} \rangle \delta_{ij} \delta(t) - t \sum \langle n_{i,s} c_l, c_j^+ \rangle$$

$$+ \left( i \frac{\partial}{\partial t} c_{is}, c_{js}^+ \right).$$  \hspace{1cm} (A3)

In the Hubbard I approximation, the last term on the right-hand side, which involves the time derivative of $n_{i,s}$, is factorized out to zero, and the second term is decoupled as $-t \langle n_{i,s} \rangle \sum G_{is,js}$. Our generalization lies in the decoupling of this term. Instead of just writing $n_{i,s} = \langle n_{i,s} \rangle$, we let somewhat more flexibility to the electron density by assuming that $n_{i,s} + n_{i,\bar{s}}$ is conserved in each hop from site $i$ to site $l$, i.e., in each $il$ “bond”:

$$n_{i,s} + n_{i,\bar{s}} = \langle n_{i,s} \rangle + \langle n_{i,\bar{s}} \rangle.$$ \hspace{1cm} (A4)

Then the expression within brackets in the second term on the right-hand side of Eq. (A3), becomes

$$\langle n_{i,s} c_{is}, c_{js}^+ \rangle = \langle n_{i,s} \rangle G_{is,js} - \langle \Gamma_{is,js} - \langle n_{i,s} \rangle G_{is,js} \rangle,$$ \hspace{1cm} (A5)

the second term on the right-hand side describing the density fluctuation at the other ($l$) end of the bond. Putting Eq. (A5) into Eq. (A3), and making use of Eq. (A2), it is a simple matter to show that

$$\left( i \frac{\partial}{\partial t} - U \langle n_{i,s} \rangle \right) \Gamma_{is,js} = \langle n_{i,s} \rangle \delta_{ij} \delta(t) - t \sum \langle n_{i,s} G_{is,js} \rangle$$

$$+ i \sum X_{is,js},$$ \hspace{1cm} (A6)

and

$$\left( i \frac{\partial}{\partial t} - U (1 - \langle n_{i,s} \rangle) \right) X_{is,js} = \langle n_{i,s} \rangle \delta_{ij} \delta(t) - \sum \langle n_{i,s} G_{is,js} \rangle$$

$$+ i \sum X_{is,js},$$ \hspace{1cm} (A7)

where $X_{is,js} = \Gamma_{is,js} - \langle n_{i,s} \rangle G_{is,js}$. Fourier transforming Eqs. (A6) and (A7) with respect to both time and lattice sites, and recalling that $-t \sum_l e_k l$, one obtains

$$(\omega - e_k - U\bar{n}) G_{kl} = 1 + UX_{kl},$$ \hspace{1cm} (A8)

$$(\omega + e_k - U (1 - \bar{n})) X_{kl} = U \bar{n} (1 - \bar{n}) G_{kl}.$$ \hspace{1cm} (A9)

Equations (4) and (5) of the text now follow immediately.