# **Spectral density of the Hubbard model by the continued fraction method**

R. Hayn,<sup>1</sup> P. Lombardo,<sup>1</sup> and K. Matho<sup>2</sup>

1 *Laboratoire Matériaux et Microélectronique de Provence, Faculté St. Jérôme, Case 142, F-13397 Marseille Cedex 20, France*

<sup>2</sup>*Centre de Recherches sur les Très Basses Températures, Laboratoire associé à l'Université Joseph Fourier, CNRS, Boîte Postale 166,*

*F-38042 Grenoble Cedex 9, France*

(Received 27 April 2006; revised manuscript received 27 July 2006; published 30 November 2006)

We present the continued fraction method (CFM) as a microscopic approximation to the spectral density of the Hubbard model in the correlated metal phase away from half filling. The quantity expanded as a continued fraction is the single-particle Green's function. Leading spectral moments are taken into account through a set of real expansion coefficients, as known from the projection technique. Further stages are added to the continued fraction, with complex coefficients, thus defining a terminator function. This enables us to treat the entire spectral range of the Green's function on equal footing and determine the energy scale of the Fermi liquid quasiparticles by minimizing the total energy. The solution is free of phenomenological parameters and remains well defined in the strong-coupling limit, near the doping-controlled metal-insulator transition. Our results for the density of states agree reasonably with those from several variants of the dynamical mean-field theory. The CFM requires minimal numerical effort and can be generalized in several ways that are interesting for applications to real materials.

DOI: [10.1103/PhysRevB.74.205124](http://dx.doi.org/10.1103/PhysRevB.74.205124)

PACS number(s): 71.10.Fd, 75.20.Hr

## **I. INTRODUCTION**

The Hubbard Hamiltonian<sup>1</sup> is certainly the most important model in the field of strongly correlated electrons. The spectral function for the addition or removal of a single electron near half filling serves as a paradigm for the excitation spectrum of highly correlated electrons in the vicinity of a Mott transition. It was a great success of the dynamical mean-field theory<sup>2</sup> (DMFT) to connect the high- and low-energy parts of the spectral function in a nonperturbative solution for arbitrary interaction strength. In particular, it was confirmed that the coherent low-energy excitations in the metallic phase follow the same dynamics as the Kondo resonance in the Anderson impurity model, the other generic Hamiltonian for correlated electrons that is much better understood.<sup>3</sup>

The analog to the Kondo resonance in the impurity model is a quasiparticle (QP) band in the lattice model. Both straddle the chemical potential  $\mu$ , i.e., the lowest excitations are gapless. In the strong-coupling limit, the spectral weight *Z* of the QP band is small, relative to two sidebands, further removed from  $\mu$ . These sidebands are called the Hubbard bands because they are roughly reminiscent of the Hubbard I solution.<sup>1</sup> In the doping-controlled regime, $4$  one sideband always overlaps with the QPs; the other represents true high-energy excitations across the correlation gap.

Hubbard I is an approximation close to the atomic limit, but nevertheless taking exact spectral moments of the itinerant propagator up to the second order into account. It can be considered the ancestor of the projection technique<sup>5</sup> which systematically incorporates spectral moments of higher order. These approximations have severe deficiencies in the low-energy sector, unless the moment series can be effectively summed up. In particular, generating a third pole in the spectral function from the high-energy side alone leads to uncontrolled results.

When the density of states, as obtained within the DMFT, is resolved with respect to the wave number *k*, more details about the coexistence of this Kondo resonance with atomiclike features in a lattice system are revealed. The lowest excitations are true Fermi liquid QPs. (i) The finite density of states at  $\epsilon = \mu$  corresponds to long-lived excitations. (ii) These are located in *k* space on a Fermi surface that satisfies Luttinger's theorem.<sup>6</sup> (iii) As a function of the distance *k*  $-k_F$  from the Fermi surface, the excitation energy has a linear, strongly reduced dispersion. (iv) The damping is quadratic in  $k−k<sub>F</sub>$  but strongly enhanced, meaning that the linear and quadratic terms are of the same order at a very small energy scale, the coherence energy  $\Delta^*$ . The two atomiclike excitations turn out to be strongly damped, even when *k* is on the Fermi surface. Their peaks disperse with *k* but spectral tails spread over the entire bandwidth.

The DMFT thus unites atomic and itinerant features in a nonperturbative approximation. It is exact only in dimension  $d=\infty$ . As a generic scenario, it is expected to hold down to  $d=2$ , albeit with the caveat that the DMFT suppresses additional structure due to bosonic couplings. Earlier approximations at finite *d* already yielded<sup>7,[8](#page-9-7)</sup> QPs and established a connection to the Kondo effect.<sup>3</sup> The high prestige of the DMFT is due to its ability to produce a self-consistent, numerically manageable approximation to the spectral function for all energies, in particular to the parameter *Z* that governs the low-energy sector. This has opened a path to realistic modeling of correlated materials beyond the Hubbard model in such methods as the local density approximation  $(LDA)+DMFT.<sup>9</sup>$ 

It is nevertheless desirable for several reasons to pursue alternative methods in parallel. First, a *k*-independent selfenergy, which is the proper result at  $d = \infty$ , does not allow one to explain phenomena that depend on the different symmetry directions in the Brillouin zone, especially in hightemperature superconductors and other low-dimensional systems. Cluster extensions of the DMFT go in the direction of lifting this restriction, $10$  but these generalizations are numerically even more demanding than the DMFT itself. The precise solution of a many-body Kondo problem is required at each iteration step toward self-consistency. In practice, when designing the "impurity solver," a trade-off exists between improving the low-energy, low-temperature solution and exactly satisfying global sum rules. Such numerical problems are presently a bottleneck for extensions of the DMFT to larger clusters or to LDA+DMFT with charge transfer into ligand bands. A variational aspect was recently found, which may allow circumvention of some of the numerical problems.<sup>11</sup>

In this paper, we present a continued fraction method (CFM) and implement it for the doping-controlled metallic regime near the Mott transition. Similar to other recent attempts, $12-16$  we start from the projection technique, applied to the *k*-resolved single-particle Green's function. The notations are introduced in Sec. II. In Sec. III, the connections between the moment and continued fraction expansions as well as the Padé approximant are established. In the Padé approximant, qualitatively important features of the macroscopic system, such as damping, are missing. They can only be captured by resummation of the continued fraction to infinite order. The concept of a terminating function, by which an approximate resummation is achieved, is common to many methods based on the continued fraction. As a general scheme, we define our CFM by allowing only such terminating functions that preserve the structure of a truncated continued fraction, but with complex coefficients. Useful recursion relations, that are properties of Padé approximants, can thus be carried over and the solution for the Green's function can be constrained by high- as well as low-energy sum rules.

Previous solutions obtained with this ansatz $13$  were partly phenomenological, because the strong-coupling renormalization *Z* needed to be inferred from a separate Gutzwiller approximation, or else was left open for fitting to experiments.<sup>17[,18](#page-10-7)</sup> A closed solution is now achieved by minimizing the total energy in the presence of sum rules, for which the necessary self-consistency loops are introduced. The self-consistent *Z* falls below the Gutzwiller value and has a doping dependence close to that for the exact Kondo scale.<sup>19</sup> This is now a true microscopic approximation, depending only on the parameters in the Hamiltonian.

In Secs. IV,V, and VI we have investigated the terminating functions that correspond to adding one or two stages with complex coefficients to the continued fraction. We show how the Fermi surface singularity, the enclosed Luttinger volume in *k* space, and the Fermi liquid damping can be modeled rigorously. We assess the quality of our approximations by comparing the density of states with DMFT results for two variants of the impurity solver, namely, the numerical renormalization group<sup>20</sup> (NRG) and the noncrossing approximation  $(NCA).^{21}$  $(NCA).^{21}$  $(NCA).^{21}$ 

The success of the CFM with respect to the Hubbard model allows us to draw some optimistic conclusions about possible generalizations toward more realistic models, describing correlation effects in a multiband electronic environment. This will be outlined as part of the conclusions.

## **II. HAMILTONIAN, GREEN'S FUNCTION, AND GENERALITIES ABOUT CONTINUED FRACTIONS**

The Hubbard model for a grand canonical ensemble of electrons on a lattice of *N* sites  $(N \rightarrow \infty)$  is written in the usual notation

$$
\hat{H} = H - \mu \hat{N} = \sum_{k,\sigma} (E_k - \mu) c_{k\sigma}^{\dagger} c_{k\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}.
$$
 (1)

The kinetic energy consists of itinerant Bloch states with energies  $E_k$  and wave number  $k$ , running through one Brillouin zone. The bandwidth is 2*D* and, when not specified otherwise, *D* is used as unit of both energy and frequency  $(h=1)$ . We formulate the method for an arbitrary density of Bloch states. Numerical examples later on will be calculated for a semielliptic density.

The chemical potential  $\mu$  is self-consistently determined to satisfy the condition

$$
n = \langle \hat{N} \rangle / N = 2m = \sum_{\sigma} \overline{\langle c_{k\sigma}^{\dagger} c_{k\sigma} \rangle},
$$
 (2)

where the filling factor  $n (0 \le n \le 2)$  is part of the input. The chemical potential for the  $U=0$  limit is designated as  $\mu_0$ . For  $U \neq 0$ , the right hand side is calculated with our method. The overbar and the angular brackets signify the Brillouin zone average and ensemble average, respectively. The filling factor per spin direction in the spin-degenerate phase is *m*=*n*/2.

We approximate the advanced single-particle Green's function

$$
G(k,\omega) = i \int_{-\infty}^{0} e^{i\omega t} \langle [c_{k\sigma}(t), c_{k\sigma}^{\dagger}(0)] \rangle, \tag{3}
$$

from which the momentum distribution  $\langle c_{k\sigma}^{\dagger} c_{k\sigma} \rangle$  and other observables are calculated. The spin index is dropped in the unpolarized phase. The time-dependent fermionic destruction operator  $c_{k\sigma}(t)$  is in the Heisenberg representation with  $\hat{H}$ and the square brackets indicate the anticommutator. The complex frequency  $\omega$  has  $\mu$  as origin. Asymptotically, for large  $\omega$ , we have  $G(k, \omega) \approx 1/\omega$ . The coefficient 1 reflects the moment  $M_0 = 1$  or spectral norm, as required by the Pauli principle. For this relation between the leading coefficient and the norm to remain valid in an approximation, it is necessary and sufficient to conserve the Herglotz property. In the case of the advanced Green's function, it means that the relation Im  $G(k, \omega) > 0$  must be obeyed throughout the entire half plane Im  $\omega$  < 0. The physical meaning of the Herglotz property is causality and it automatically entails the existence of Kramers-Kronig relations between the real and imaginary parts. A great advantage of our method is the possibility to make straightforward evaluations along the real axis. Since this limit has to be approached from within the domain of analyticity, the notation  $\omega = \epsilon - i0^+$  with real energy  $\epsilon = E - \mu$  is introduced. The *k*-resolved spectral function is

$$
A(k,\epsilon) = \frac{1}{\pi} \operatorname{Im} G(k,\epsilon - i0^{+}).
$$
 (4)

<span id="page-2-1"></span>At  $U=0$  it has a single sharp peak at the excitation energy

$$
\eta_k = E_k - \mu_0 \propto (k - k_F),\tag{5}
$$

which also serves to measure distance in *k* space, at least in the vicinity of the Fermi surface.

The continued fraction expansion, on which our method is based in a crucial way, has already a long tradition in solid state physics, in the one-electron problem with disorder, $^{23}$  as well as in the many-electron problem. $24$  The continued fraction is generated by various procedures like tridiagonalization, recursion, or Lanczos methods. The Hubbard I Green's function is the simplest example of a continued fraction that has been truncated at low order. The exact Green's function for the Hubbard model on finite clusters is a continued fraction which naturally ends at very high order. The continued fraction for the infinite system does not end. Properties of the thermodynamic limit, such as damping due to electronelectron scattering, emerge only after resummation of the continued fraction. Approximate resummation is achieved by the terminating function, an analytic function which also has the Herglotz property.

A well-chosen terminator is thus expected to bring two improvements to the approximation for the Green's function in dimensions  $d \ge 2$ . (i) From a set of discrete, more or less intense, and more or less densely spaced Dirac peaks the final shape of the continuous spectral density emerges (see Ref. [25](#page-10-13) for tight-binding-like models and Ref. [26](#page-10-14) for strongly correlated electrons). (ii) A Fermi surface discontinuity emerges in the momentum distribution  $\langle c_{k\sigma}^{\dagger} c_{k\sigma} \rangle$  at temperatures below the strong-coupling energy scale  $\Delta^*$ . The Fermi liquid discontinuity and the correct Fermi surface volume will be incorporated in our ansatz. This means that we take the Luttinger theorem for granted and use it as a principle, even for strong coupling where there is no rigorous proof. The energy  $\Delta^*$  then comes out as part of the self-consistent solution.

#### **III. HIGH-ENERGY PART**

The first moment or center of gravity of  $G(k, \omega)$  is

$$
\omega_1 = E_k + mU - \mu. \tag{6}
$$

<span id="page-2-2"></span>It disperses like the unrenormalized Bloch energy *Ek*. In models with a more general interaction, a *k*-dependent Hartree-Fock shift is also present which, for on-site repulsion, reduces to a constant Hartree shift *mU*. The self-consistent  $\mu$  is the only unknown.

The high-energy expansion about the center of gravity is

$$
G(k,\omega) = \frac{1}{\omega - \omega_1} + \frac{M_2}{(\omega - \omega_1)^3} + \frac{M_3}{(\omega - \omega_1)^4} + \cdots
$$
 (7)

Its coefficients

$$
M_{\lambda} = \int_{-\infty}^{\infty} d\epsilon (\epsilon - \omega_1)^{\lambda} A(k, \epsilon), \quad \lambda = 2, 3 \dots,
$$
 (8)

are called the central moments  $(M_1=0, \text{ by definition})$ . They can be related to correlation functions which occur in the short time, or Liouville expansion of the operator  $c_{k\sigma}(t)$  and are evaluated in the limit *t*=0.

It is remarkable that the variance  $s_2$  of  $A(k, \epsilon)$ , defined by the second central moment

$$
M_2 = s_2^2 = m(1 - m)U^2
$$
 (9)

<span id="page-2-3"></span>is *k* independent in any dimension *d*, not only  $d = \infty$ . All the terms in the high-energy expansion are sensitive to the lowenergy sector, be it only via the self-consistent  $\mu$ .

We now turn to the continued fraction expansion which is closely related to the moment expansion. Formally, it is initiated by using  $\omega_1$  and  $s_2$  to write the Green's function as

$$
G(k,\omega)^{-1} = \omega - \omega_1 - s_2^2 G_1(\omega). \tag{10}
$$

In this identity,  $G_1(\omega)$  is again a Herglotz function with asymptotics  $G_1(\omega) \approx 1/\omega$ . Iterations, pushing the continued fraction further down step by step, require knowledge of the center of gravity  $\omega_{2l-1}$  and the variance  $s_{2l}$  of  $G_{l-1}(\omega)$ , to write

$$
G_{l-1}(\omega)^{-1} = \omega - \omega_{2l-1} - s_{2l}^2 G_l(\omega), \quad l = 2, 3 \dots \quad (11)
$$

The two new expansion coefficients depend only on the central moments  $M_{\lambda}$  up to the order  $\lambda = 2l - 1$  and  $\lambda = 2l$  of their respective indices.

By truncating the continued fraction, i.e., by setting  $G_l(\omega) \equiv 0$ , an approximation to the Green's function is obtained that is the quotient of two polynomials, one with *l*− 1 real zeros in the numerator and another with *l* real zeros in the denominator. The Green's function thus has *l* poles on the real axis, alternating with *l*− 1 zeros. This object is defined as the Padé approximant  $\langle l-1|l\rangle$ <sup>[27](#page-10-15)</sup> It represents the optimal use one can make of a set of known spectral moments up to  $M_{2l-1}$ . Constructing the Padé approximant to a Green's function is essentially more difficult, because the moments themselves are not yet known. Numerical values, found for them by self-consistency conditions within a fixed order, turn out to be quite inexact. This fact is often ignored when it is claimed that a certain high-energy approximation obeys a set of "exact" sum rules.

<span id="page-2-0"></span>We now discuss some well-known results concerning approximations at the second stage of the continued fraction. As a still exact representation of the Green's function we have

$$
G(k,\omega) = \frac{1}{\omega - \omega_1 - \frac{s_2^2}{\omega - \omega_3 - s_4^2 G_2(\omega)}}.
$$
 (12)

The relations between the first few terms are

$$
s_2^2 = M_2,
$$

$$
\omega_3 = \omega_1 + M_3/M_2,
$$

$$
s_4^2 = M_4/M_2 - M_2 - (M_3/M_2)^2.
$$
 (13)

Besides the variance, quantities used to further characterize the internal shape of a spectrum are the skewness  $\gamma = M_3 / M_2^{3/2}$  and the kurtosis  $\kappa = M_4 / M_2^2 - 3$ . In terms of these, we have  $\omega_3 = \omega_1 + \gamma s_2$  and  $s_4^2 = s_2^2(\kappa + 2 - \gamma^2)$ . From the third moment one finds the coefficient

$$
\omega_3 = (1 - m)U + B_3 - \mu. \tag{14}
$$

This continued fraction coefficient is the first quantity in the expansion with a nontrivial *k* dependence. The full correlation function appearing in the third moment was first derived in Ref. [28](#page-10-16) and determined self-consistently for a short linear chain in Ref. [29.](#page-10-17) The shift in the spectral skewness, caused by  $B_3$ , regulates the dynamical weight transfer between the Hubbard peaks at finite *U*. [30](#page-10-18) One can decompose

$$
B_3 = W_0 + W_3(k) \tag{15}
$$

<span id="page-3-7"></span>in such a way that the term  $W_3(k)$  vanishes in high dimensions. For making contact with the DMFT we will presently neglect it and adopt the expression $^{28}$ 

$$
B_3 = \frac{(2m-1)}{2m(1-m)} \langle \hat{T} \rangle, \tag{16}
$$

<span id="page-3-4"></span>by which it is linked self-consistently to the expectation value of the kinetic energy  $\langle \hat{T} \rangle$ .

Concerning the behavior of the fourth moment, not even the correlation functions involved in its self-consistency loop have as yet been evaluated. Again, the actual numerical value of  $s_4^2$  is also expected to be sensitive to the low-energy sector and, in low-dimensional systems, *k* dependent.

Given this situation, approximations on the level of Eq. ([12](#page-2-0)) are at present inevitable. Straightforward truncation,  $G_2(\omega)$ =0, leads to the Padé approximant  $\langle 1|2 \rangle$ . This solution with two Dirac peaks goes beyond Hubbard I, because the dynamical weight transfer is taken into account. The first example of an approximate resummation of the continued fraction is the alloy analogy, developed in the paper called Hubbard III.<sup>31</sup> Following Hubbard's notation, we approximate  $s_4^2 G_2(\omega)$  by a *k*-independent terminating function  $\Omega_H(\omega)$ , which has to be a Herglotz function.

The alloy analogy satisfies at least the task (i) of a terminating function, namely, to generate finite damping. Far away from  $\mu$ , where the excitations are incoherent, it actually represents a physically correct picture. We therefore keep the result  $\Omega_H(\omega) \to iD$  for large  $\omega$  from Hubbard III. The physical reason why the damping is of the order of the bare bandwidth *D* is that the mean free path is as short as one lattice constant. In practice, we incorporate the high-energy damping in an effective  $\omega_3$ ,

$$
\bar{\omega}_3 = \omega_3 + iD,\tag{17}
$$

and henceforth deal with a terminator that decays as  $1/\omega$ . This way, we conserve the sum rules, encapsuled in the central moments  $M_0$  to  $M_3$ . Since Hubbard III is unrealistic at low energies, we do not pursue it any further. Nevertheless, it should be noted that Hubbard III generates a branch cut in  $\Omega_H(\omega)$ , causing the imaginary part to drop back to zero and

a correlation gap with sharp edges to appear, at least in the zero-temperature limit. This property of Hubbard III is also not expected to survive in improved approximations for the metallic phase. We will address the consequences that the absence of a branch cut has for the shape of the density of states, both in the CFM and in the DMFT.

<span id="page-3-0"></span>To sum up, our approximation to the Green's function is formally similar to that of Hubbard III,

$$
G(k,\omega) = \frac{1}{\omega - \omega_1 - \frac{s_2^2}{\omega - \bar{\omega}_3 - \Omega(\omega)}},
$$
(18)

but with a terminating function,  $\Omega_H(\omega) = \Omega(\omega) + iD$  that retains the strong damping of the alloy analogy only at high energy. Two successive implementations of the terminating function with appropriate Fermi liquid properties at low energy are the subject of the following sections.

### **IV. LOW-ENERGY PART**

A Fermi surface discontinuity is strictly realized only in the zero-temperature limit and in a system with no residual disorder. Since  $T=0$  solutions are hardest to obtain with DMFT and, on the contrary, easily implemented with our method, we concentrate in the following on this limit. We write the standard microscopic definition of a self-energy as a complex correction to the bare excitation  $\eta_k$ :

$$
G(k,\omega)^{-1} = \omega - \eta_k - \Sigma(k,\omega) \tag{19}
$$

<span id="page-3-6"></span>and compare with the inverse of Eq.  $(18)$  $(18)$  $(18)$ . The high-energy limit  $\Sigma(k, \infty)$  is the difference between two dispersive quantities. In the present case, Eqs.  $(5)$  $(5)$  $(5)$  and  $(6)$  $(6)$  $(6)$  have identical dispersion and

$$
p_1 = \eta_k - \omega_1 = \mu - \mu_0 - mU \tag{20}
$$

<span id="page-3-5"></span>is, in fact, constant. Within the other approximations, discussed in the preceding section, we then obtain the *k*-independent self-energy

$$
\Sigma(\omega) = -p_1 + \frac{s_2^2}{\omega - \bar{\omega}_3 - \Omega(\omega)}.
$$
 (21)

<span id="page-3-3"></span>In this case, as in the DMFT, the Fermi surface has the exact shape of the uncorrelated system. It is given by all *k* points where  $\eta_k = 0$  in Eq. ([5](#page-2-1)). The QP peak of weight *Z* at the Fermi level and the step of amplitude *Z* in the momentum distribution are fixed by the conditions

$$
\Sigma(0) = 0 \tag{22}
$$

<span id="page-3-2"></span><span id="page-3-1"></span>and

$$
\frac{d\Sigma}{d\omega}(0) = \alpha = 1 - 1/Z < 0. \tag{23}
$$

At finite *T* or in the presence of a residual diffusive mean free path, Im  $\Sigma(0)$  remains finite.

Guided by the insight that the strong-coupling peak is distinct from the Hubbard peaks, we can formulate a

<span id="page-4-0"></span>minimal ansatz for the terminating function $13$  as

$$
\Omega(\omega) = \frac{(\overline{s}_4)^2}{\omega - \overline{\omega}_5}.
$$
\n(24)

Adding a new stage to the continued fraction is the proper way to "add" a pole to the Green's function. When this terminator is inserted in Eq.  $(18)$  $(18)$  $(18)$ , it generates a Green's function with three complex zeros in the denominator and two zeros in the numerator, i.e., the same structure as the Padé approximant  $\langle 2|3 \rangle$ . The connection of the parameters  $\overline{s}_4^2$  and  $\bar{\omega}_5$  to central moments  $M_4$  and  $M_5$  is lost. In fact, the very existence of moments beyond  $M_3$  has been sacrificed by admitting  $\bar{\omega}_3$ ,  $\bar{s}_4$ , and  $\bar{\omega}_5$  as complex quantities. They now have to be determined from conditions  $(22)$  $(22)$  $(22)$  and  $(23)$  $(23)$  $(23)$ .

<span id="page-4-1"></span>For the Herglotz property one finds

$$
(\operatorname{Im} \overline{s}_4)^2 / \operatorname{Im} \overline{\omega}_5 \le \operatorname{Im} \overline{\omega}_3 = D \tag{25}
$$

as a necessary and sufficient condition. This causes all three poles to lie in the upper half plane. Further, it guarantees a normalized, positive semidefinite  $A(k, \epsilon)$ , which also implies quite intricate relations between the complex residues.

Now, the important point is the following: This simple ansatz is so heavily constrained by sum rules that it offers a self-consistent solution of the problem, without any free parameters. It remains to substantiate this claim and then to discuss the quality of the solution.

After inserting Eq.  $(24)$  $(24)$  $(24)$  in Eq.  $(21)$  $(21)$  $(21)$ , the conditions  $(22)$  $(22)$  $(22)$ and  $(23)$  $(23)$  $(23)$  can be brought into a system of two linear equations for the unknowns  $\overline{s}_4^2$  and  $\overline{\omega}_5$ . The determinant of this system is

$$
\det_2 = -p_1^2 - \alpha s_2^2,\tag{26}
$$

<span id="page-4-4"></span>and the Herglotz property requires  $det_2$  > 0. This is a constraint on the QP weight *Z*: Instead of  $Z \le 1$  (Pauli principle) we have  $Z \leq s_2^2/(s_2^2+p_1^2)$ . Closer inspection reveals that it means the QP cannot take more spectral weight than the peak in the Padé approximant  $\langle 1|2 \rangle$  that is nearest to  $\mu$ . Since around half filling this weight stays above  $1/2$ , it is indeed only a weak constraint.

The solution

$$
\bar{s}_4^2 = \frac{p_2^2}{\det_2} \tag{27}
$$

and

$$
\overline{\omega}_5 = -\frac{p_1 p_2}{\det_2} \tag{28}
$$

<span id="page-4-5"></span>is expressed in terms of the complex quantity

$$
p_2 = -(p_1\bar{\omega}_3 + s_2^2). \tag{29}
$$

It satisfies the Herglotz condition  $(25)$  $(25)$  $(25)$  with the equality sign. This is a consequence of our strong  $T=0$  constraint  $\Sigma(0)$  $= 0$ , concerning both the real and imaginary parts. The selfenergy is now parametrized, up to *Z*, which remains free within a restrained interval and will be determined by minimizing the total energy.

We note, before closing this section, that Ref. [13](#page-10-5) allows us to define one-pole terminations for the more general case of a truncated Green's function that is expressed as a higher-order Padé approximant. The general algorithm is given by which Eqs.  $(22)$  $(22)$  $(22)$  and  $(23)$  $(23)$  $(23)$  can be satisfied.

### **V. NUMERICAL PROCEDURE**

The uncorrelated chemical potential as function of the filling,  $\mu_0(n)$ , depends only on the kinetic energy part and is determined once for all. The density of states per lattice site in the  $U=0$  limit

$$
\rho_0(\varepsilon) = \frac{2}{\pi} \operatorname{Im} F_0(\varepsilon - i0^+) \tag{30}
$$

is obtained from the on-site Green's function

$$
F_0(\omega) = \frac{1}{N} \sum_k \frac{1}{\omega + \mu_0 - E_k}.
$$
 (31)

The factor of 2 comes from summing over spin directions. The density of states of the correlated system

$$
\rho(\varepsilon) = \frac{2}{\pi} \operatorname{Im} F(\varepsilon - i0^+) \tag{32}
$$

is obtained from  $F(\omega) = G(k, \omega)$ , the on-site Green's function in real space, which is independent of the site index. For a  $k$ -independent self-energy such as Eq.  $(21)$  $(21)$  $(21)$ , the on-site Green's functions  $F(\omega)$  and  $F_0(\omega)$  are related to each other by

$$
F(\omega) = F_0(\omega - \Sigma(\omega)).
$$
 (33)

The *k* summations can then be carried out by using the analytic function that represents the solution for  $F_0(\omega)$  in the limit  $N \rightarrow \infty$ .

<span id="page-4-2"></span>We now turn to the discussion of the self-consistency loops. The condition for  $\mu$  is implemented at  $T=0$  by the integral

$$
\int_{-\infty}^{0} d \epsilon \rho(\epsilon) = n.
$$
 (34)

According to Eq.  $(16)$  $(16)$  $(16)$ , the term  $B_3$  from the third moment requires the self-consistent determination of the kinetic energy,  $\langle \hat{T} \rangle = 2 \int_{-\infty}^{0} d\epsilon \, \overline{A(k, \epsilon) E_k}$ . One finds

$$
\langle \hat{T} \rangle = \frac{2}{\pi} \int_{-\infty}^{0} d\epsilon \operatorname{Im} \{ \tilde{\omega} F_0[\omega - \Sigma(\omega)] - 1 \},
$$

$$
\widetilde{\omega} = \epsilon - i0^+ + \mu_0 - \Sigma(\epsilon - i0^+). \tag{35}
$$

<span id="page-4-3"></span>Finally, the total energy is

$$
E_{tot} = \frac{1}{2} \left( \langle \hat{T} \rangle + \int_{-\infty}^{0} d\epsilon \, \epsilon \rho(\epsilon) \right). \tag{36}
$$

The integrals in Eqs.  $(34)$  $(34)$  $(34)$ – $(36)$  $(36)$  $(36)$  are carried out numerically. For the calculations in this paper we took the on-site Green's function

<span id="page-5-0"></span>

FIG. 1. Energy minimum  $E_{tot}(Z)$  for  $U=4$ ,  $D=1$ , and  $\mu_0$ =−0.166 corresponding to *n*=0.79. The choice of *D*=1 means that all energies are measured in units of the half-width *D* of the uncorrelated band.

$$
F_0(\omega) = \frac{2}{(\omega + \mu_0) \left(1 + \sqrt{1 - \frac{1}{(\omega + \mu_0)^2}}\right)}.
$$
 (37)

In the context of  $d=\infty$ , it is the Green's function for a Bethe lattice. A half-width  $D=1$  is now used as energy unit. For the Herglotz property, it is important to choose the square root with a positive real part. The model density of states belonging to this Green's function,

$$
\rho_0(\epsilon) = \frac{4}{\pi} \sqrt{1 - (\epsilon + \mu_0)^2},\tag{38}
$$

is the semielliptic function which was also used by Hubbard.

While searching for the self-consistent  $\mu$  and  $\langle \hat{T} \rangle$  at a given input  $n$  and  $U$ , the renormalization  $Z$  is still kept as a parameter, only limited by the condition  $det_2$  > 0. With these constrained solutions for the Green's function, we calculate the total energy  $E_{tot}$ . As shown in the example of Fig. [1,](#page-5-0)  $E_{tot}$ has a well-defined minimum as a function of *Z*. Taking the value which minimizes  $E_{tot}$  fixes the last parameter  $Z$  and defines our solution for the Green's function.

The density of states obtained for the same input as in Fig. [1](#page-5-0) is shown in Fig. [2,](#page-5-1) together with the  $U=0$  limit. The QP

<span id="page-5-1"></span>

FIG. 2. Spectral density; comparison of  $\rho_0(\epsilon)$  (dashed line) and  $\rho(\epsilon)$  (full line) for the same parameters as in Fig. [1.](#page-5-0)

band has the same intensity at the Fermi level as the uncorrelated band,  $\rho(0) = \rho_0(0)$ . This invariance signals the unitary limit for the Kondo resonance in the limit *T*= 0. Thus, the reduction of the QP weight does not show up in  $\rho(0)$  but in the bandwidth, which is scaled down by *Z*. In a lattice system, this one-to-one relationship between QP weight and bandwidth only holds when the self-energy is local *k* independent).

On the *k*-resolved level, near the Fermi surface, the QP pole in the complex plane has a parabolic trajectory parametrized by  $\eta_k$ , Eq. ([5](#page-2-1)):

$$
\omega^*(k) = Z\eta_k + i\Gamma_k,\tag{39}
$$

<span id="page-5-2"></span>with a scattering rate

$$
\Gamma_k = (Z \eta_k)^2 / \Delta^*.
$$
 (40)

<span id="page-5-3"></span>The half-width for coherent states within the QP band is

$$
\Delta^* = \frac{Zs_2^2|p_2|^2}{D[(1-Z)s_2^2 - Zp_1^2]^2}.
$$
 (41)

This formula is well behaved also in the weak-coupling limit, in fact for all possible metallic, unpolarized regimes of the Hubbard model. In the strongly correlated regime, the energy scale  $\Delta^*$  is always small, due to the explicit factor Z in the numerator. Since we have modeled the ballistic limit (residual diffusive scattering rate  $\Gamma_d = 0$ ) the QP resonance in  $A(k, \epsilon)$  is a Dirac peak for  $k = k_F$  and, for  $k \neq k_F$ , it has the so called Breit-Wigner line shape (see Hedin and Lundquist<sup>32</sup> for a generic plot). This shape is due to an interference between the QP residue and the other residues. The line shape becomes approximately Lorentzian whenever a  $\Gamma_d > \Gamma_k$  is present.

Returning to the density of states, we note that the global shape of the valence spectrum for a hole-doped Mott insulator, i.e., QP band and lower Hubbard band, is well rendered by our present approximation. The sum rules up to  $M_3$  are exactly satisfied and their interplay regulates the overall skewness and the relative weight of all three features.

The one-pole terminating function has the drawback of being unable to reproduce a sharp gap formation. The high level of intensity between the QP band and the upper Hubbard band shows that the dynamical spectral transfer<sup>30</sup> is not realized completely, at least for *U*/*D*= 4. The intensity at the minimum decays like  $(U/D)^{-2}$ , so that this spurious effect disappears for larger *U*. We shall discuss the presence of residual intensity in the gap region in more detail when we compare with our second ansatz and with the DMFT.

Two remarks conclude this section. (i) The limit  $U = \infty$ describes spin and charge excitations in the subspace of singly occupied sites. It is equivalent to the *t*-*J* model with *J* = 0. The one-pole terminating function thus allows one to project out a quantitatively valid Green's function for the valence sector near this limit, up to terms of order  $(U/D)^{-2}$ . (ii) Ratios  $U/D \approx 4$  are relevant for the doping-controlled Mott transition in real materials close to criticality. Our main motivation to pursue the CFM was to investigate whether by simply adding a second complex stage to the terminator we could handle this regime in a semiquantitative way. The

derivation of the two-pole terminating function and its application to  $U/D=4$  are presented in the next section.

### **VI. IMPROVING THE DYNAMICAL WEIGHT TRANSFER**

To generalize our ansatz, we introduce algebraic expressions for  $\Omega(\omega)$ , such that Eq. ([18](#page-3-0)) can be cast into the form of a truncated continued fraction with complex coefficients. This defines the general framework of the CFM, provided the Herglotz condition is satisfied. The *k*-resolved Green's function has then the structure of a generalized higher order Padé approximant. We label the approximations CFM1, CFM2,… according to the number of complex stages or number of "added" poles. By terminating the Padé approximant  $\langle 1|2 \rangle$ , we still retain the important sum rules that govern the dynamical weight transfer. Spectral moments beyond  $M_3$ are then not exploitable, but this may not be a great sacrifice, given the difficulties known from the projection method to obtain correct values for higher moments.

What can be gained by using complex coefficients is the possibility to model constructive and destructive interference phenomena in the Green's function at intermediate energies. A single feature in the spectral function can be built up by the contributions of several poles, resulting in uncommon line shapes. An ansatz frequently employed in the phenomenological interpretation of spectra is the superposition of complex poles with real residues (superposition of Lorentzians in the spectrum). Although this allows several peaks to coalesce, it still eliminates interference. One striking example is the Fano-like interference of QP amplitude and background, described in the preceding section. This is already included in CFM1.

Likewise, the dynamical weight transfer and the formation of the Mott gap can be interpreted as a destructive interference in the intermediate energy range between the Hubbard bands. When the interference is complete the function  $G_2(\omega)$  in Eq. ([12](#page-2-0)) should acquire a branchcut and a gap interval with zero density of states and sharp edges should result. This may be possible only on the insulating side of the Mott transition and strictly at  $T=0$ . When the system is metallic and the chemical potential falls in a region of high density of states, it is satisfactory to model the correlation gap by one deep minimum at some point  $x_0$  on the energy axis. This corresponds to the doping-controlled regime, where the parameter  $|p_1|$ , Eq. ([20](#page-3-5)), is of order *U*. Studying the Mott gap within the Hubbard III model shows that the minimum always occurs near

$$
x_0 = \text{Re } \bar{\omega}_3. \tag{42}
$$

<span id="page-6-4"></span>Consequently, the position  $x_0$  is related to  $M_3$  and to the skewness. We adopt this approximation in order to reduce the number of free parameters.

We demonstrate here that the strong-coupling regime, including the Mott gap and a quantitative treatment of dynamical weight transfer, is captured by the CFM2. The two-pole terminating function is of the form

$$
\Omega(\omega) = \frac{\overline{s}_4^2}{\omega - \overline{\omega}_5 - \frac{\overline{s}_6^2}{\omega - \overline{\omega}_7}}.
$$
\n(43)

<span id="page-6-0"></span>The new degrees of freedom are given by  $\bar{s}_4^2$  and  $\bar{\omega}_5$ . These will all be fixed due to some further qualitative arguments, restricting the ansatz from the start. Then,  $\bar{s}_6^2$  and  $\bar{\omega}_7$  can again be eliminated by the Fermi liquid conditions of Eqs.  $(22)$  $(22)$  $(22)$  and  $(23)$  $(23)$  $(23)$ , using the next iteration of the algorithm in Ref. [13.](#page-10-5)

The Green's function now has four poles and the Herglotz condition becomes a crucially important issue. To formulate it, for arbitrary complex values of  $\bar{\omega}_3$  to  $\bar{\omega}_7$ , seems at first sight rather difficult. The Green's function on the Fermi surface [Eq. ([19](#page-3-6)) with  $\eta_k = 0$ ] has additive coherent and incoherent contributions,

$$
\frac{1}{\omega - \Sigma(\omega)} = \frac{Z}{\omega} + G_b(\omega). \tag{44}
$$

<span id="page-6-1"></span>This decomposition, which is only possible because one pole lies infinitesimally close to the real axis, enables us to manage the Herglotz condition for  $\Sigma(\omega)$  more easily: from Eq. ([43](#page-6-0)) we obtain a background function  $G_b(\omega)$  with three poles that can be written

<span id="page-6-2"></span>
$$
G_b(\omega) = \frac{1 - Z}{\omega - \Omega_1 - \frac{\Sigma_2^2}{\omega - \Omega_3 - \frac{\Sigma_4^2}{\omega - \Omega_5}}}.
$$
(45)

The new coefficients are designated by capital greek letters. Systematically, they depend on  $Z$  and, at order  $\lambda$ , on all coefficients in Eqs. ([18](#page-3-0)) and ([43](#page-6-0)) with index  $\lambda' \leq \lambda$ . Explicitly, the first three are

$$
\Omega_1 = -\frac{p_1}{1 - Z},
$$
  

$$
\Sigma_2^2 = \frac{Z \det_2}{(1 - Z)^2},
$$
  

$$
\Omega_3 = \frac{1}{p_1} \left( \frac{\alpha p_2 s_2^2}{\det_2} + \frac{\det_2}{\alpha} \right),
$$
 (46)

in terms of the previously defined quantities Eqs.  $(9)$  $(9)$  $(9)$ ,  $(20)$  $(20)$  $(20)$ ,  $(26)$  $(26)$  $(26)$ , and  $(29)$  $(29)$  $(29)$ .

<span id="page-6-3"></span>The high-energy damping in the background function is

Im 
$$
\Omega_3 = \left(1 + \frac{p_1^2}{\det_2}\right) \text{Im } \overline{\omega}_3 > \text{Im } \overline{\omega}_3 = D.
$$
 (47)

Between the coherence energy  $\Delta^*$  in Eq. ([40](#page-5-2)) and the background function at the Fermi edge there is the relation

$$
\Delta^* \operatorname{Im} G_b(0) = Z. \tag{48}
$$

For  $\Sigma_4 = \Omega_5 = 0$ , we recover the one-pole terminating function and Eq. ([41](#page-5-3)) for  $\Delta^*$ . Since  $\Omega_1$  and  $\tilde{\Sigma}_2^2$  are real, there are now only three complex quantities and the Herglotz condition can <span id="page-7-0"></span>be specified exhaustively, analogous to Eq.  $(25)$  $(25)$  $(25)$ :

$$
(\operatorname{Im} \Sigma_4)^2 / \operatorname{Im} \Omega_5 \le \operatorname{Im} \Omega_3. \tag{49}
$$

The foregoing analysis suggests that  $\Sigma_4$  and  $\Omega_5$  are more useful than  $\bar{s}_4$  and  $\bar{\omega}_5$  as control parameters. To obtain the self-energy, one can then use Eqs. ([44](#page-6-1)) and ([45](#page-6-2)). For further discussion we parametrize

$$
\Omega_3 = X_3 + iY_3,
$$
  
\n
$$
\Sigma_4 = X_4 + iY_4,
$$
  
\n
$$
\Omega_5 = X_5 + iY_5.
$$
\n(50)

Our requirement for properly defining the dynamical weight transfer is vanishing  $Im \Sigma(\epsilon - i0^+)$  in the point  $x_0$ . This is satisfied if (and only if) the equality sign applies in ([49](#page-7-0)), leading to the joint conditions  $Y_4^2 = Y_3 Y_5$  and

$$
X_5 - \frac{Y_5}{Y_4} X_4 = x_0. \tag{51}
$$

The Herglotz property now guarantees that it is in fact a minimum. The influence of this interference on the shape of the valence spectrum is weakest for  $X_5 = 0$ . For simplicity, we also need to set  $Y_5 = Y_4 = Y_3$ , where  $Y_3$  is already defined in Eq. ([47](#page-6-3)). The last parameter  $X_4 = -x_0$  is then fixed by the point with lowest intensity inside the correlation gap.

Before continuing with this ansatz, it is important to realize that it cannot apply exactly at half filling. There, the metallic phase is obtained by driving *U*/*D* below the critical ratio (so-called bandwidth-controlled transition).<sup>[4](#page-9-3)</sup> The particle-hole symmetric density of states has a quite different morphology than what is shown in Fig. [2:](#page-5-1) the QP's are in the center and the correlation gap is split in two symmetric gaps of order *U*/2.[2](#page-9-1) Within the CFM, it can be envisaged to use a CFM3 ansatz to model two symmetric destructive interferences. A big challenge, left for later studies, is to describe the rapid movement of the QP band as function of doping, when one is near the critical *U*.

Returning to the doping controlled regime at strong coupling, the remaining free parameter *Z* is determined again by minimizing the total energy. The numerical procedure is as described before. In Fig. [3,](#page-7-1) results with the one- and twopole terminating functions (CFM1 and CFM2) are compared to the Gutzwiller approximation  $(GA)$  at constant  $U$ , for a wide range of fillings. The upper curve is the well known lower bound for the GA,  $Z=(1-n)/(1-m)$ , obtained by excluding double occupancy. By projecting out the background, the GA is known to systematically overestimate the coherent weight. The behavior that results from the CFM, i.e., lowering of *Z* and upward curvature at the approach of zero doping  $(1-n\rightarrow 0)$ , is close to that of the exact Kondo scale in the Bethe ansatz solution for the Anderson impurity.<sup>19</sup>

Figure [4](#page-7-2) shows the results for *Z* as function of intermediate to strong coupling. The CFM1 shows good qualitative behavior even for zero doping, although *Z* does not

<span id="page-7-1"></span>

FIG. 3. Spectral weight *Z* of QP pole as function of electron density *n* for *U*= 4. Comparison of the Gutzwiller approximation (GA) to our result with the one- and two-pole terminating functions (CFM1 and CFM2).

rigorously vanish at a critical *U*. The CFM2, applicable only for  $U > U_c$  and finite doping, yields reasonable values for *Z*, even at small *U*.

#### **VII. COMPARISON WITH OTHER METHODS**

With the two-pole terminating function, realistic results for the density of states in the doping controlled regime can be obtained, even close to the critical *U*. To illustrate this, we compare our CFM with the DMFT for two different impurity solvers. The impurity solvers perform the crucial step in mapping the Hubbard lattice model onto an Anderson impurity model. The effective medium surrounding a given site is determined self-consistently, still a formidable many body problem. The NRG, $^{20}$  used to solve it at the lowest temperatures and energies, requires a heavy amount of computer time. The NCA is an alternative,  $2^{1,22}$  $2^{1,22}$  $2^{1,22}$  more analytic method, less reliable for  $|\omega| \ll \Delta^*$ , but obeying high-energy sum rules well. Therefore, the NRG and NCA are expected to be complementary.

A comparison for the same parameters as before, i.e. *U*  $=$  4 and  $n = 0.79$ , is shown in Fig. [5.](#page-8-0) The NRG data are taken from Ref. [18,](#page-10-7) the NCA is our own unpublished calculation,

<span id="page-7-2"></span>

FIG. 4. Spectral weight *Z* of QP pole as function of electron correlation *U*. Comparison of the one- and two-pole terminating function's (CFM1 and CFM2) in the doped  $(n=0.898)$  and undoped  $(n=1, \text{ only CFM1})$  cases.

<span id="page-8-0"></span>

FIG. 5. Spectral density. Comparison of the continued fraction method using the one- and two-pole terminating functions CFM1 and CFM2) with NCA and NRG (data taken from Ref. [18](#page-10-7)).

CFM1 is again the density of states from Fig. [2,](#page-5-1) and CFM2 the result with Eq.  $(43)$  $(43)$  $(43)$ . All four solutions obey the  $\rho(0) = \rho_0(0)$  condition. This confirms that temperatures in the DMFT solutions are sufficiently low to warrant a comparison with our  $T=0$  results. As manifest in the width of the OP band, the self-consistent *Z* obtained for CFM2 coincides with both versions of the DMFT. Since the NRG is expected to determine essentially the exact low energy scale, this is a good point for both the NCA and the CFM2 results.

The solutions start to differ somewhat in the gap region. Neither DMFT version shows a gap with sharp edges that would correspond to a branch cut in the self-energy. A real benchmark for low-*T* impurity solvers in the dopingcontrolled regime does not yet exist. From the NCA, we can confirm that some very low residual density inside the gap seems to be the generic situation.

In the ansatz for CFM2, the existence of a point with zero density of states is postulated. Determining its position according to Eq.  $(42)$  $(42)$  $(42)$  involves the self-consistency conditions for  $\mu$  and  $B_3$ . The quantitative agreement with the DMFT in the QP band and good overall agreement in the entire valence sector are due to this built in interference. In comparing CFM1 and CFM2, one notices a feedback of the improved gap region on the QP band: The sum rules up to  $M_3$ are satisfied for both approximations, but the dynamical weight transfer is more complete within CFM2. Removing the spurious intensity inside the gap slightly raises the QP weight (compare Fig. [3](#page-7-1)), bringing it into agreement with the NRG result.

The rather large variation among the different solutions in the region of the upper Hubbard band is remarkable and still deserves more detailed investigations. At higher temperatures  $T \geq \Delta^*$ , where quantum Monte Carlo results are available as benchmark, the NCA was found to be satisfactory.<sup>21</sup> In the present comparison, the NCA comes closer to obeying the sum rules than the NRG. As far as numerical effort is concerned, the NRG is the most demanding, followed by the NCA. The CFM2 stands up quite honorably in this comparison, especially when considering that the sum rules are rigorously incorporated, no "technical" broadening needs to be introduced, and the required computer time to achieve self-consistency is in fact negligible.

### **VIII. DISCUSSION AND OUTLOOK**

We have introduced the CFM as a nonperturbative method to calculate the self-energy of the Hubbard model in metallic regions of the phase diagram. The two variants that are implemented in this paper, CFM1 and CFM2, are labeled by the number of complex poles that appear in the ansatz for the terminating function. The ansatz itself is essentially an educated guess that incorporates Fermi liquid behavior in the low-energy sector of the self-energy. The simple algebraic structure, together with the use of the Luttinger sum rule as an input, is essential for circumventing heavy many-body calculations, such as, e.g., required for the impurity solvers within the DMFT. On the other hand, this simple structure prevents one from capturing critical phenomena very close to the metal-insulator transition, e.g., power law behavior in the position  $\mu(n)$  of the chemical potential, right at the critical  $U=U_c$ , as the filling  $n \rightarrow 1$ .

The variant CFM1 can be used everywhere outside the line  $n=1$ ,  $U > U_c$ . It covers also the weak-coupling regime, for arbitrary *n*, yielding limiting behaviors such as 1−*Z*  $\propto (U/D)^2$  for the QP weight and  $\Gamma_k \propto \eta_k^2 U^2 / D^3$  for the QP scattering rate. The real advantage of the CFM appears for intermediate and strong coupling, where perturbation theory breaks down. A typical numerical result for the density of states in CFM1 is discussed in Sec. IV, around Fig. [2.](#page-5-1) Results for *Z* over wide ranges of doping and coupling strength are discussed in Sec. VI, around Figs. [3](#page-7-1) and [4.](#page-7-2)

In the variant CFM2, a second pole is used to improve the dynamical weight transfer that manifests itself by a wellpronounced Mott gap in the strong-coupling spectra. By the nature of this ansatz, it is restricted to  $U > U_c$  and  $n \neq 1$ . In the weak-coupling regime, CFM2 still yields a reasonable result for  $Z$  (see Fig. [4](#page-7-2)), but to impose a Mott gap in the density of states is unphysical in this case. The CFM2 result for the strong-coupling density of states compares favorably with the best DMFT results at very low temperature, as shown in Fig. [5.](#page-8-0)

The high- and intermediate-energy behavior of the selfenergy can be given by a sequence of stages with real coefficients in the continued fraction expansion, going as far as moment sum rules are exploitable. The extension to moments of high order becomes possible due to the close relationship of the CFM with the numerical Lanczos procedure. Approximate coefficients can be generated by applying the Lanczos procedure to a finite cluster. A proper Fermi liquid termination of the continued fraction is then obtained from the general algorithm in Ref. [13.](#page-10-5)

The algebraic terminating function was already introduced earlier, in a phenomenological context, leaving *Z* as a free parameter. The possibilities for fitting photoemission spectra have been demonstrated. $13,17,18$  $13,17,18$  $13,17,18$  The cornerstone of the CFM as a microscopic method is now the minimization of the total energy to obtain *Z*. Given  $G(k, \omega)$  for variable *Z*, we calculate the total energy from the exact many-body expression, actually another sum rule, first established by Galitski. Notwithstanding the convincing numerical results, the deeper reason for the success of this procedure should be discussed some more. Without an explicit wave function, we are lacking a rigorous variational principle. The situation is similar to the Gutzwiller method, which starts from the Rayleigh-Ritz variational principle for a projected wave function, but then needs to abandon this principle, because the kinetic part of the energy is calculated within a further, quite rough approximation.

For the case of hole doping, which we have studied, the lower Hubbard peak corresponds to a lowering of kinetic energy by incoherent hopping over short distances. Including these processes in the energy balance, the QP weight *Z* is also lowered. These processes are projected out in a Gutzwiller approximation, Thus, by including this background and constraining it via the sum rules, we arrive at a more realistic balance between the competing kinetic and Coulomb energies.

That a lower *Z*, relative to the Gutzwiller result, is an improvement in the right direction, and that including the incoherent part is the clue to this effect, is fully confirmed by the DMFT. Here, the impurity solvers mimic the impurity Kondo effect and *ZD* is a measure for the Kondo temperature. A simple, quantitative picture arises in the limit of large degeneracy  $N_f$ <sup>[3,](#page-9-2)[19](#page-10-8)</sup> The Kondo scale in the impurity Anderson model can be obtained exactly, as function of *n* and *m*  $=n/N_f$ , by the Bethe ansatz. Concerning the total spectrum, coherent weight is of order zero in  $1/N_f$ , the leading background contribution starts at first order. Neglecting background, as for instance in the slave boson method at meanfield level, yields  $Z = (1 - n)/(1 - m)$  for the renormalization. This is also the equivalent of the Gutzwiller approximation, as plotted for  $N_f = 2$  in Fig. [3.](#page-7-1) The influence of the background is strikingly illustrated by solving for the Kondo temperature only to the first order in  $1/N_f$ .<sup>[19](#page-10-8)</sup> It causes indeed a substantial decrease, bringing the result for *Z* close to the exact Bethe ansatz value. The doping dependence with the upward curvature, as seen in our approximations CFM1 and CFM2, is also a feature of the exact result. Finally, the small correction from CFM1 to CFM2 shows a delicate interplay between the result for the low-energy scale and the chosen approximation for the dynamical weight transfer, related to the double occupancy.

The CFM is generalizable in many directions. The variational procedure renders it independent of the limit  $d = \infty$ . It is straightforward and, for low-dimensional systems, potentially very important to incorporate the *k* dependence in the moment  $M_3$ . The term  $W_3(k)$  in Eq. ([15](#page-3-7)) was already identified in the exact diagonalization of a short linear chain,  $29$  as causing a coupling of the QP to antiferromagnetic fluctuations. This can be generalized to fluctuations above other possible groundstates and the self-consistent determination of  $W_3(k)$  thus offers a path to describing the feedback of bosonic fluctuations on the low-energy sector. Up to now, the treatment of low-energy effects within the projection method was based more on physical intuition, or guesswork for the more critical observer, than on an objective procedure.

As an outlook, we enumerate other possibilities that are inherent in the CFM, beyond the results of this paper. They are listed roughly according to increasing effort that will be required to implement them. (i) A more detailed exploitation of spectral functions on the *k*-resolved level is immediately possible with the self-energy given in this paper. The ideal *k*-resolved spectrum  $A(k, \epsilon)$  is not directly observable. Interpretation of Raman, angle-resolved photoemission, or tunneling data requires its partial summation over selected spots in the Brillouin zone, weighted by matrix elements. (ii) All equations are formulated for an arbitrary density of Bloch states. Hubbard lattice models with a more realistic kinetic energy part, including van Hove singularities, can be implemented. (iii) The method can be applied to the generalized periodic Anderson model (PAM): These are lattice models with Hubbard repulsion among transition orbitals but, in addition, hybridization with ligand orbitals. Here, the simplest case is one transition orbital, such as  $x^2 - y^2$  in the cuprates, or a Kramers doublet for Ce compounds, coupled with any number of ligand orbitals. The exact Dyson equation for this case requires only a scalar self-energy function. Its implementation within the CFM was described in Ref. [13.](#page-10-5) Orbital degeneracy within the transition shell requires first a generalization of the terminating function to the case of matrix continued fractions. (iv) Implementation of LDA+CFM: The algebraic simplicity of the CFM allows to calculate the charge transfer effects, present in model Hamiltonians of the PAM type, on an *ab initio* level. These effects, important for many real materials, could not yet be handled successfully by LDA+DMFT. As in the PAM or in recent developments of-LDA+DMFT, the orbitally degenerate transition shell requires to deal with a matrix self-energy. (v) Not difficult to implement, but leaving the strict framework of the CFM as an algebraic method, is the inclusion of non-Fermi-liquid effects on a phenomenological level into the terminating function.<sup>13</sup>

In conclusion, we have attempted to demonstrate by means of the Hubbard model that the CFM is a powerful method. Numerically simple, due to its algebraic structure, it is still sufficiently rigorous to deal with strongly correlated electrons in mesoscopic and macroscopic samples of condensed matter.

- <span id="page-9-0"></span><sup>1</sup> J. Hubbard, Proc. R. Soc. London, Ser. A 276, 238 (1963).
- <span id="page-9-1"></span><sup>2</sup>A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. 68, 13 (1996).
- <span id="page-9-2"></span>3A. C. Hewson, *The Kondo Problem to Heavy Fermions* Cambridge University Press, Cambridge, U.K., 1997).
- <span id="page-9-3"></span>4M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039  $(1998).$
- <span id="page-9-4"></span><sup>5</sup>H. Mori, Prog. Theor. Phys. **33**, 423 (1965); see also P. Fulde, *Electron Correlations in Molecules and Solids*, 3rd ed. (Springer-Verlag, Berlin, 1995).
- <span id="page-9-5"></span><sup>6</sup> J. M. Luttinger, Phys. Rev. 119, 1153 (1960).
- <span id="page-9-6"></span><sup>7</sup>D. M. Edwards and J. A. Herz, Physica B **163**, 527 (1990).
- <span id="page-9-7"></span> ${}^{8}$ M. C. Gutzwiller, Phys. Rev. 137, A1726 (1965); for a review, see D. Vollhardt, Rev. Mod. Phys. 56, 99 (1984).
- <span id="page-10-0"></span><sup>9</sup>K. Held *et al.*, Psi-k Newsl. **56**, 65 (2003).
- <span id="page-10-1"></span>10Th. Maier, M. Jarrell, Th. Pruschke, and M. Hettler, Rev. Mod. Phys. 77, 1027 (2005).
- <span id="page-10-2"></span>11M. Potthoff, M. Aichhorn, and C. Dahnken, Phys. Rev. Lett. **91**, 206402 (2003).
- <span id="page-10-3"></span>12M. Potthoff, T. Herrmann, and W. Nolting, Eur. Phys. J. B **4**, 485  $(1998).$
- <span id="page-10-5"></span>13K. Matho, J. Electron Spectrosc. Relat. Phenom. **117-118**, 13  $(2001).$
- 14D. Villani, E. Lange, A. Avella, and G. Kotliar, Phys. Rev. Lett. 85, 804 (2000).
- 15A. Avella, F. Mancini, and R. Hayn, Eur. Phys. J. B **37**, 465  $(2004).$
- <span id="page-10-4"></span><sup>16</sup> Y. Kakehashi and P. Fulde, Phys. Rev. Lett. **94**, 156401 (2005).
- <span id="page-10-6"></span><sup>17</sup>K. Matho, Mol. Phys. Rep. **17**, 141 (1997).
- <span id="page-10-7"></span>18K. Byczuk, R. Bulla, R. Claessen, and D. Vollhardt, Int. J. Mod. Phys. B 16, 3759 (2002).
- <span id="page-10-8"></span><sup>19</sup> J. W. Rasul and A. C. Hewson, J. Phys. C 17, 3337 (1984).
- <span id="page-10-9"></span><sup>20</sup> R. Bulla, Phys. Rev. Lett. **83**, 136 (1999).
- <span id="page-10-10"></span>21Th. Pruschke, D. L. Cox, and M. Jarrell, Phys. Rev. B **47**, 3553

 $(1993).$ 

- <span id="page-10-21"></span>22Th. Pruschke and N. Grewe, Z. Phys. B: Condens. Matter **74**, 439  $(1989).$
- <span id="page-10-11"></span>23R. Haydock and V. Heine, and M. J. Kelly, J. Phys. C **5**, 2845  $(1972).$
- <span id="page-10-12"></span><sup>24</sup> E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).
- <span id="page-10-13"></span>25P. Turchi, F. Ducastelle, and G. Treglia, J. Phys. C **15**, 2891  $(1982).$
- <span id="page-10-14"></span>26R. O. Kuzian, R. Hayn, and J. Richter, Eur. Phys. J. B **35**, 21  $(2003).$
- <span id="page-10-15"></span><sup>27</sup>G. A. Baker, Jr., *Essentials of Padé Approximants* (Academic Press, New York, 1975).
- <span id="page-10-16"></span><sup>28</sup> W. Nolting and W. Borgiel, Phys. Rev. B **39**, 6962 (1989).
- <span id="page-10-17"></span> $^{29}$ B. Mehlig, H. Eskes, R. Hayn, and M. B. J. Meinders, Phys. Rev. B 52, 2463 (1995).
- <span id="page-10-18"></span>30M. B. J. Meinders, H. Eskes, and G. A. Sawatzky, Phys. Rev. B 48, 3916 (1993).
- <span id="page-10-19"></span><sup>31</sup> J. Hubbard, Proc. R. Soc. London, Ser. A **277**, 237 (1963).
- <span id="page-10-20"></span><sup>32</sup>L. Hedin and S. Lundquist, Solid State Phys. **23**, 1 (1969).