

**Conservation in two-particle self-consistent extensions of dynamical mean-field theory**Friedrich Krien,<sup>1</sup> Erik G. C. P. van Loon,<sup>2</sup> Hartmut Hafermann,<sup>3</sup> Junya Otsuki,<sup>4</sup> Mikhail I. Katsnelson,<sup>2</sup> and Alexander I. Lichtenstein<sup>1</sup><sup>1</sup>*Institute of Theoretical Physics, University of Hamburg, 20355 Hamburg, Germany*<sup>2</sup>*Radboud University, Institute for Molecules and Materials, NL-6525 AJ Nijmegen, The Netherlands*<sup>3</sup>*Mathematical and Algorithmic Sciences Lab, Paris Research Center, Huawei Technologies France SASU, 92100 Boulogne Billancourt, France*<sup>4</sup>*Department of Physics, Tohoku University, Sendai 980-8578, Japan*

(Received 21 June 2017; published 24 August 2017)

Extensions of dynamical mean-field theory (DMFT) make use of quantum impurity models as nonperturbative and exactly solvable reference systems which are essential to treat the strong electronic correlations. Through the introduction of retarded interactions on the impurity, these approximations can be made two-particle self-consistent. This is of interest for the Hubbard model because it allows to suppress the antiferromagnetic phase transition in two dimensions in accordance with the Mermin-Wagner theorem, and to include the effects of bosonic fluctuations. For a physically sound description of the latter, the approximation should be conserving. In this paper, we show that the mutual requirements of two-particle self-consistency and conservation lead to fundamental problems. For an approximation that is two-particle self-consistent in the charge and longitudinal spin channels, the double occupancy of the lattice and the impurity is no longer consistent when computed from single-particle properties. For the case of self-consistency in the charge and longitudinal as well as transversal spin channels, these requirements are even mutually exclusive so that no conserving approximation can exist. We illustrate these findings for a two-particle self-consistent and conserving DMFT approximation.

DOI: [10.1103/PhysRevB.96.075155](https://doi.org/10.1103/PhysRevB.96.075155)

Models for correlated electron systems, such as the Hubbard model, count among the hardest problems of contemporary condensed matter physics. At the same time, they are believed to capture the physics of fascinating phenomena such as high-temperature superconductivity [1] and the Mott transition [2]. To understand the underlying physics, it is necessary to develop methods which can capture these phenomena. Because of unavoidable approximations, however, it is not always possible to separate the physics from artifacts of the method. It is therefore desirable to design methods which satisfy basic requirements such as translational invariance, thermodynamic consistency [3,4], local conservation laws of charge and spin [5–7], and, in view of an application to high-temperature superconductivity of layered cuprates [1], the Mermin-Wagner theorem [8].

Dynamical mean-field theory (DMFT) [9] and its cluster extensions [10] have been an important step towards the understanding of correlated electron behavior, in particular the Mott transition. From a theoretical perspective, DMFT can be thought of as an approximation to the exact Luttinger-Ward functional, where all propagators are replaced by the corresponding local ones. An auxiliary problem subject to a self-consistency condition, often an Anderson impurity model (AIM), is used as a tool to sum the diagrams of this local functional exactly. As a consequence, DMFT is conserving in the Baym-Kadanoff sense [6,7,9,11].

The more recently introduced diagrammatic extensions of DMFT such as D $\Gamma$ A [12,13], the dual fermion (DF) [14], one-particle irreducible (1PI) [15], TRILEX [16,17], DMF<sup>2</sup>RG [18], and dual boson (DB) approaches [19] are an active field of research. The AIM plays a central role in these approaches. From a suitable dynamical vertex function of the AIM, nonlocal approximations to the self-energy are constructed by summing certain classes of diagrams. The

lattice self-energy is hence approximate, but incorporates long-range correlations. The use of dynamical vertices allows one to deal with strong correlations as opposed to approaches based on the bare Hubbard interaction, such as the fluctuation exchange approximation (FLEX) [20] or the two-particle self-consistent approach (TPSC) [21]. Despite significant progress in this field [22], a number of open questions remain. For example, it is not always clear how to choose the diagrams [23,24] or self-consistency conditions. More generally, the question is how to optimally exploit the AIM, whose solution we know (numerically) exactly, to construct approximations that meet the above-mentioned basic requirements.

Extended dynamical mean-field theory (EDMFT) [25–31] includes the effect of two-particle bosonic fluctuations through a local retarded interaction in the impurity model that is fixed by a corresponding self-consistency condition. As a result, EDMFT, as well as its extensions, like (E)DMFT+GW [32,33], are two-particle self-consistent. A consequence of the EDMFT self-consistency condition is that the lattice double occupancy equals that of the impurity model and hence is bounded. As Vilk and Tremblay have shown [21], any approximation which produces a bounded lattice double occupancy will suppress magnetic phase transitions in two dimensions, as required by the Mermin-Wagner theorem. This property is indeed respected by the spin-DMFT [26] (which is akin to EDMFT), while it is violated in DMFT.

A disadvantage of EDMFT is, however, that it breaks Ward identities [7,26] and therefore violates local conservation laws. This can lead to a qualitatively wrong description of the physics of the collective excitations. For example, in presence of a long-range interaction, the energy of the plasmon diverges in the long-wavelength limit. In case of a local interaction, one obtains a plasmonlike feature instead of a zero-sound mode [7].

In the dual boson (DB) approach [19,34], which is a diagrammatic extension of EDMFT, conservation in the charge channel can be restored by including certain ladder diagrams into the bosonic propagator. This provides a physically sound description of plasmons even in the correlated state [35]. Remarkably, global charge conservation is maintained in a two-particle self-consistent version of the approach [36]. The two-particle self-consistent theory also resolves an ambiguity in the computation of the double occupancy present in DMFT and yields results closer to benchmarks than either of the two DMFT values [37].

In particular for an application to superconductivity, it is desirable to include spin fluctuations as well, while maintaining the conserving character of the theory and a sound description of the collective modes. Self-consistent approaches based on an impurity model including retarded spin-spin interactions have been considered previously [25–27]. It seems appealing to include diagrammatic corrections in order to make such a theory conserving.

These considerations lead us to the following questions: Is it possible to construct a two-particle self-consistent version of DMFT which would be conserving and satisfy the Mermin-Wagner theorem? Similarly, under which conditions can we extend EDMFT and spin-DMFT to satisfy the conservation laws? Quantum impurity models are at the heart of these approaches and serve as exactly solvable reference systems [38]. They allow to treat at least part of the strong electronic correlations in a nonperturbative manner. Hence, more generally, the question is whether it is possible to construct approximations that exploit the nonperturbative starting point provided by the impurity model while maintaining the desirable properties mentioned above.

The aim of this paper is to show that one faces fundamental difficulties in an attempt to construct such approximations. In particular, we demonstrate that for a conserving approximation, imposing two-particle self-consistency in the charge and one of the spin channels leads to an inconsistency in the calculation of the potential energy due to the retarded interactions. More importantly, we prove that if one attempts to impose self-consistency in the charge and all three spin channels, no conserving approximation can exist. In essence, we find that the retarded spin interactions, introduced to make the theory two-particle self-consistent, undermine the desired feature of local conservation. We show that this limitation is rooted in the fact that the Ward identities of the lattice and of the impurity model are incompatible. As a concrete example, we construct a two-particle self-consistent DMFT which is conserving in the charge and one of the spin channels.

The paper is organized as follows: We recollect the DMFT approximation to the Hubbard model in Sec. I and examine the thermodynamic consistency of the total energy in this approximation. We introduce two-particle self-consistency in Sec. II and perform a similar analysis. A conflict between two-particle self-consistency and local conservation is related to the Ward identities of the impurity model in Sec. III. We present an application of a two-particle self-consistent DMFT in Sec. IV. We interpret our main results in Sec. V and end with conclusions. Derivations for several analytical results are provided in the Appendices A–E.

## I. DYNAMICAL MEAN-FIELD THEORY

To set the stage, we first discuss the familiar case of DMFT. For concreteness, we focus on the two-dimensional (2D) paramagnetic Hubbard model on the square lattice with nearest-neighbor hopping given by the Hamiltonian

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

Here,  $i, j$  label lattice sites. The local Hubbard interaction has strength  $U$ . We use the hopping  $t = 1$  as the unit of energy and denote Green's function as  $G_{ij}$  in real space and  $G_{\mathbf{k}}$  in momentum space, respectively (when it is not ambiguous, we omit the frequency dependence for brevity).

DMFT is a local approximation to the exact Luttinger-Ward functional  $\Phi[G_{ij}] \approx \sum_i \phi[G_{\text{loc}}]$ , which is therefore conserving in the Baym-Kadanoff sense [6,9]. As a result of the local approximation, the self-energy is local:  $\Sigma_{ij} = \delta\Phi[G_{i'j}]/\delta G_{ji} = \delta\phi[G_{\text{loc}}]/\delta G_{\text{loc}} \delta_{ji}$  and we note that the same holds for the irreducible vertex:  $-\Gamma_{ijkl} = \delta^2\Phi[G_{i'j}]/\delta G_{ji} \delta G_{lk} = \delta^2\phi[G_{\text{loc}}]/\delta G_{\text{loc}}^2 \delta_{ii} \delta_{jj} \delta_{lk}$ .<sup>1</sup> If we know the local Green's function, the problem is solved: in this case, we can evaluate the local functional and its derivatives at the local Green's function and hence compute the self-energy. Since we do not know the local Green's function *a priori* and the self-energy is a functional of the latter, we have to solve this problem self-consistently: we vary  $G_{\text{loc}}$  until the local Green's function computed from the self-energy equals  $G_{\text{loc}}$ .

We can employ an auxiliary local model as a *tool* to accomplish this and to sum the diagrams of this local functional exactly. This means letting  $\phi[G_{\text{loc}}] \equiv \phi_{\text{imp}}[g_{\text{imp}}]$  and  $\Sigma[G_{\text{loc}}] \equiv \Sigma_{\text{imp}}[g_{\text{imp}}]$ . The desired solution is evidently obtained when the DMFT self-consistency condition is satisfied,

$$g_{\text{imp},\nu} = G_{\text{loc},\nu}. \quad (2)$$

(Where unambiguous, we drop labels “imp” and “lat” in what follows.)

In practice, an Anderson impurity model (AIM) is often employed for this purpose, whose action reads as

$$S_{\text{AIM}} = - \sum_{\nu\sigma} c_{\nu\sigma}^* (i\nu + \mu - \Delta_\nu) c_{\nu\sigma} + U \sum_{\omega} n_{-\omega\uparrow} n_{\omega\downarrow}. \quad (3)$$

Here,  $\Delta_\nu$  denotes the electronic hybridization,  $\mu$  is the chemical potential, and  $\nu$  ( $\omega$ ) denote the discrete fermionic (bosonic) Matsubara frequencies  $\nu_n = (2n + 1)\pi/\beta$  and  $\omega_m = 2m\pi/\beta$ , respectively.  $\beta = 1/T$  is the inverse temperature. The AIM has the same local interaction  $U$  as the lattice model.

Let us now take a *practical* viewpoint. Assume we have a nontrivial model that we can solve exactly, such as the AIM described by the action (3). From this model we can obtain the local impurity self-energy and irreducible vertex function. We can now ask the question of how to construct a conserving approximation given these quantities.

<sup>1</sup>A minus sign arises in the functional relation of  $\Gamma$  due to the definitions chosen in this publication.

We recall that *local* conservation of charge and spin means that the following continuity equations for the charge ( $\rho^0$ ) and spin densities ( $\rho^{x,y,z}$ ) hold:

$$\partial_\tau \rho^\alpha = -[\rho^\alpha, H]. \quad (4)$$

We have introduced the index  $\alpha = 0, x, y, z$  to label the charge and spin channels. The corresponding charge and spin density operators are defined as  $\rho^\alpha = \sum_{\sigma\sigma'} c_{\sigma'}^\dagger s_{\sigma\sigma'}^\alpha c_\sigma$  with the Pauli matrices  $s^\alpha$ , such that  $\rho^0 = n = n_\uparrow + n_\downarrow$  and  $\rho^{x,y,z} = 2S^{x,y,z}$ .

On the lattice we can formulate the following Ward identities (cf. Appendix A), which are the Green's function analogs of the continuity equations (4):

$$\Sigma_{k+q} - \Sigma_k = - \sum_{k'} \Gamma_{kk'q}^\alpha [G_{k'+q} - G_{k'}]. \quad (5)$$

Here, we have introduced four-vector notation  $k \equiv (\mathbf{k}, \nu)$  and  $q \equiv (\mathbf{q}, \omega)$ . Summations over frequencies and momenta imply factors  $\beta^{-1}$  and  $N^{-1}$ , respectively, with  $N$  being the number of sites.  $\Sigma$  and  $G$  are the exact lattice self-energy and Green's function, respectively, and  $\Gamma^\alpha$  denotes the *irreducible* (horizontal) *particle-hole* vertex. The irreducible vertices in the charge and spin channels are explicitly defined as  $\Gamma^0 = \Gamma^{\uparrow\uparrow\uparrow\uparrow} + \Gamma^{\uparrow\uparrow\downarrow\downarrow}$ ,  $\Gamma^z = \Gamma^{\uparrow\uparrow\uparrow\uparrow} - \Gamma^{\uparrow\uparrow\downarrow\downarrow}$ , and  $\Gamma^x = \Gamma^y = \frac{1}{2}(\Gamma^{\uparrow\downarrow\downarrow\uparrow} + \Gamma^{\downarrow\uparrow\uparrow\downarrow}) = \Gamma^{\uparrow\downarrow\downarrow\uparrow}$ .

In a local approximation,  $\Sigma_k \equiv \Sigma_\nu$  and  $\Gamma_{kk'q}^\alpha \equiv \gamma_{\nu\nu'\omega}^\alpha$ , such as DMFT, all momentum dependence drops out of the Ward identities (5) and we obtain<sup>2</sup>

$$\Sigma_{\nu+\omega} - \Sigma_\nu = - \sum_{\nu'} \gamma_{\nu\nu'\omega}^\alpha [G_{10c, \nu'+\omega} - G_{10c, \nu'}]. \quad (6)$$

An analogous Ward identity holds for the AIM (see Appendix E):

$$\Sigma_{\nu+\omega} - \Sigma_\nu = - \sum_{\nu'} \gamma_{\nu\nu'\omega}^\alpha [g_{\nu'+\omega} - g_{\nu'}], \quad (7)$$

where  $\Sigma_\nu, g_\nu$  and  $\gamma_{\nu\nu'\omega}^\alpha$  are the self-energy, Green's function, and the irreducible vertex of the AIM, respectively. Hence, the DMFT approximation is apparently conserving when the self-consistency condition (2) holds. Remarkably, DMFT arises when we attempt to construct a locally conserving approximation based on the AIM (3).

Let us consider further properties of the DMFT approximation. To this end, we introduce the (connected) susceptibilities

$$X_q^\alpha = -\langle \bar{\rho}_{-q}^\alpha \bar{\rho}_q^\alpha \rangle = 2 \sum_{kk'} X_{kk'q}^\alpha, \quad (8)$$

which are defined in terms of density fluctuations  $\bar{\rho}^\alpha(\tau) = \rho^\alpha(\tau) - \langle \rho^\alpha \rangle$ . Their local parts are given by  $X_{\text{loc}}^\alpha = \sum_{\mathbf{q}} X_{\mathbf{q}}^\alpha$ . The generalized susceptibility  $X_{kk'q}^\alpha$  is related to the irreducible vertex function via the integral equation

$$X_{kk'q}^\alpha = G_k G_{k+q} \left[ \beta \delta_{kk'} - \sum_{k''} \Gamma_{kk''q}^\alpha X_{k''k'q}^\alpha \right]. \quad (9)$$

<sup>2</sup>In an earlier publication by Hettler *et al.*, the right-hand side of Eq. (6) was expressed in terms of the *reducible* vertex function [62]. In that form, one does not straightforwardly realize the momentum independence of Eq. (6) which misled the authors to believe that DMFT does not satisfy the Ward identities.

Now, consider the kinetic energy of the lattice. It is expressed through single-particle quantities as  $E_{\text{kin}}^{\text{lat}} = \sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}} \langle n_{\mathbf{k}\sigma} \rangle$ . In Appendix B 1 we establish a relation that expresses the kinetic energy in terms of a two-particle quantity, more precisely the high-frequency behavior of the local susceptibility. The relation follows directly from the Ward identities (5):

$$\lim_{\omega \rightarrow \infty} (i\omega)^2 X_{\text{loc}, \omega}^\alpha = -2E_{\text{kin}}^{\text{lat}}. \quad (10)$$

As the Ward identities themselves, this relation connects single- and two-particle quantities. The local impurity Ward identities (7) imply an analogous relation (see Appendix E 2)

$$\lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_\omega^\alpha = -2E_{\text{kin}}^{\text{imp}}, \quad (11)$$

where  $\chi_\omega^\alpha = -\langle \bar{\rho}_{-\omega}^\alpha \bar{\rho}_\omega^\alpha \rangle_{\text{imp}}$  is the impurity susceptibility and the kinetic energy of the impurity model is given by [39]

$$E_{\text{kin}}^{\text{imp}} = 2 \sum_{\nu} \Delta_\nu g_\nu. \quad (12)$$

DMFT is not two-particle self-consistent. As a consequence, the impurity and local lattice susceptibility differ in general. Remarkably, however, their asymptotes are the same. Decomposing the susceptibility into a contribution from the impurity susceptibility and a momentum-dependent correction [7],  $X_{\text{loc}}^{\text{DMFT}} = \chi + X'_{\text{loc}}$ , one can show that  $X'_{\text{loc}}$  decays at least with  $\omega^{-4}$ . Therefore,  $\lim_{\omega \rightarrow \infty} (i\omega)^2 X_{\text{loc}, \omega} = \lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_\omega$ . We demonstrate this numerically in the left panel of Fig. 4 in the section on numerical results. As a consequence,  $E_{\text{kin}}^{\text{lat}} = E_{\text{kin}}^{\text{imp}}$  and the kinetic energy can be determined from the impurity model in DMFT.

Next, we consider the potential energy  $E_{\text{pot}} = U d_{\text{lat}}$  where  $d_{\text{lat}} = \langle n_\uparrow n_\downarrow \rangle$  is the double occupancy of the lattice. As a two-particle correlation function,  $d$  is naturally computed from two-particle quantities. We denote this by a superscript ‘‘2P’’ We have the following relations:

$$X_{\text{loc}, \tau=0}^0 = -\langle \bar{\rho}^0 \bar{\rho}^0 \rangle = -(\langle n \rangle + 2d_{\text{lat}}^{2P} - \langle n \rangle^2), \quad (13)$$

$$X_{\text{loc}, \tau=0}^z = -\langle \bar{\rho}^z \bar{\rho}^z \rangle = -(\langle n \rangle - 2d_{\text{lat}}^{2P}), \quad (14)$$

where  $\rho^0 = n = n_\uparrow + n_\downarrow$ ,  $\rho^z = m = n_\uparrow - n_\downarrow$ , and  $\langle m \rangle = 0$ . Hence, the double occupancy can be expressed in terms of the susceptibilities as

$$d_{\text{lat}}^{2P} = -\frac{1}{4} [X_{\text{loc}, \tau=0}^0 - X_{\text{loc}, \tau=0}^z - \langle n \rangle_{\text{lat}}^2]. \quad (15)$$

Similarly,  $d$  may be obtained from the impurity as

$$d_{\text{imp}}^{2P} = -\frac{1}{4} [X_{\tau=0}^0 - X_{\tau=0}^z - \langle n \rangle_{\text{imp}}^2]. \quad (16)$$

By virtue of the single-particle self-consistency condition (2) we have  $\langle n \rangle_{\text{lat}} = \langle n \rangle_{\text{imp}}$ . Due to the missing two-particle self-consistency in DMFT, however, the susceptibilities differ and we have in general  $d_{\text{lat}}^{2P} \neq d_{\text{imp}}^{2P}$  [37]. On the other hand, we can compute the double occupancies from single-particle quantities via the Migdal-Galitskii formula of the Hubbard model [40]

$$d_{\text{lat}}^{1P} = \frac{1}{U} \sum_{\mathbf{k}\nu} G_{\mathbf{k}\nu} \Sigma_{\mathbf{k}\nu}, \quad (17)$$

and its counterpart of the Anderson impurity model

$$d_{\text{imp}}^{\text{1P}} = \frac{1}{U} \sum_{\nu} g_{\nu} \Sigma_{\nu}. \quad (18)$$

Making use of the single-particle self-consistency condition (2) and of the locality of the self-energy,  $\Sigma_{\mathbf{k}\nu} = \Sigma_{\nu}$ , we see that lattice and impurity double occupancies computed in this way are the same.

In summary, DMFT arises when one attempts to construct a conserving, single-particle self-consistent approximation based on the AIM. The kinetic energy of the lattice model is equal to the kinetic energy of the impurity model. It can be obtained from the asymptote of the local lattice susceptibility, a general feature of conserving approximations, while in DMFT, it may also be obtained from the impurity susceptibility. An ambiguity arises in the calculation of the double occupancy from single- and two-particle quantities:  $d_{\text{imp}} = d_{\text{lat}}^{\text{1P}} \neq d_{\text{lat}}^{\text{2P}}$  in DMFT as a consequence of the lack of two-particle self-consistency.

We speak of several thermodynamically consistent [4,41] ways to obtain a quantity if these yield one and the same result. That different ways of calculating a quantity yield the same result is in general only true for an exact solution. The kinetic energy and the  $f$ -sum rule (see, e.g., [21], see Appendix B 2) are examples where thermodynamic consistency between one- and two-particle levels is ensured through the Ward identities (5). Obviously, the Ward identities are insufficient for consistency in other cases, as we have seen for the double occupancy, whose value is ambiguous in DMFT. Another important example is the inconsistency of the Schwinger-Dyson equation with the Ward identities when the reducible vertex is computed from the irreducible one through the Bethe-Salpeter equation [41]. The recently proposed QUADRILEX approach has been reported to be free of this inconsistency [42].

We will examine in the following section to what extent the deficiencies of DMFT can be cured by two-particle self-consistency.

## II. TWO-PARTICLE SELF-CONSISTENCY

Two-particle self-consistent approximations based on an impurity model go back to extended dynamical mean-field theory (EDMFT) and its precursors [25–31]. In these approximations, a frequency-dependent interaction is introduced in the impurity model, and its values are fixed through a self-consistency condition on a two-particle (bosonic) correlation function such as the susceptibility. In general, we can augment the AIM of (3) by a dynamical interaction in all four (one charge and three spin) channels as follows:

$$S_{\text{BFK}} = S_{\text{AIM}} + \frac{1}{2} \sum_{\alpha\omega} \bar{\rho}_{-\omega}^{\alpha} \Lambda_{\omega}^{\alpha} \bar{\rho}_{\omega}^{\alpha}. \quad (19)$$

We refer to this model as the Bose-Fermi-Kondo impurity (BFK) model.  $\Lambda_{\omega}^{\alpha}$  is a dynamical interaction which can be viewed as a bosonic bath or hybridization. We consider approximations to Green's function  $G$  and to the susceptibility  $X$  that are locally conserving and two-particle self-consistent. In analogy to the single-particle self-consistency condition (2), the retarded interactions in (19) are determined through the

following condition [36,43]:

$$\chi_{\omega}^{\alpha} = X_{\text{loc},\omega}^{\alpha}. \quad (20)$$

This self-consistency condition provides a bounded double occupancy by construction [cf. Eqs. (15) and (16)], which is sufficient to suppress magnetic phase transitions in two dimensions, as shown by Vilk and Tremblay [21].

The Ward identities (5) relate single-particle quantities (Green's function and self-energies) to two-particle quantities (vertex functions and susceptibilities). We study the interplay between the requirement of local conservation and the self-consistency conditions (2) and (20).

We differentiate between two kinds of approaches: (i) the case of Ising-type ( $S_z$ ) coupling is characterized by a finite  $\Lambda_{\omega}^z$  and  $\Lambda_{\omega}^0$ , while we set  $\Lambda_{\omega}^{x,y} = 0$ . That is, we require self-consistency (20) only in the charge and one of the spin channels, i.e., for  $\alpha = 0, z$ . (ii) For rotationally invariant Heisenberg-type coupling, all retarded interactions  $\Lambda_{\omega}^{\alpha}$  for  $\alpha = 0, x, y, z$  are finite and determined by (20). The retarded interactions cause shifts in the Hubbard interaction and the chemical potential which are discussed in Appendix C 1.

### A. Ising-type coupling

In the discussion of the kinetic energy in the context of DMFT we have shown that, assuming local conservation, it can be expressed in terms of the asymptotic behavior of the susceptibility. For the lattice we have by virtue of local conservation  $\lim_{\omega \rightarrow \infty} (i\omega)^2 X_{\text{loc},\omega}^{0,z} = -2E_{\text{kin}}^{\text{lat}}$ , while on the impurity,  $\lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_{\omega}^{0,z} = -2E_{\text{kin}}^{\text{imp}}$  holds (cf. Appendix C 2).

In this case, the kinetic energies computed in the two ways are equal by means of the two-particle self-consistency (20). The kinetic energy can therefore be obtained from the impurity model, as in DMFT. Note that  $E_{\text{kin}}^{\text{lat}} = E_{\text{kin}}^{\text{imp}}$  may be determined from the charge or spin susceptibility alike.

We saw previously that the double occupancy computed from two-particle quantities  $d_{\text{lat}}^{\text{2P}}$  and  $d_{\text{imp}}^{\text{2P}}$  can be expressed in terms of the local susceptibilities  $X_{\text{loc}}^{0,z}$  and  $\chi_{\omega}^{0,z}$ , respectively [Eqs. (15) and (16)]. While these differ in DMFT in general, the two-particle self-consistency ensures that  $d_{\text{lat}}^{\text{2P}} = d_{\text{imp}}^{\text{2P}}$ . Using single-particle quantities, we can still compute it on the lattice using the Migdal-Galitskii formula (17). In a local approximation to the self-energy and with the single-particle self-consistency condition (2), we can express the double occupancy in terms of the impurity self-energy and Green's function  $d_{\text{lat}}^{\text{1P}} = (1/U) \sum_{\nu} g_{\nu} \Sigma_{\nu}$ . In contrast to DMFT, however, this expression is *not* equal to the impurity double occupancy. Because the Migdal-Galitskii formula involves the potential energy, it is comprehensible that the retarded interactions will affect it. In Appendix D we derive the double occupancy for the Bose-Fermi-Kondo model, with the result

$$d_{\text{imp}}^{\text{1P}} = \frac{1}{2\tilde{U}} \left[ 2 \sum_{\nu} g_{\nu} \Sigma_{\nu} + \sum_{\omega,\alpha} \tilde{\Lambda}_{\omega}^{\alpha} (\chi_{\omega}^{\alpha} - \beta \langle \rho^{\alpha} \rangle^2 \delta_{\omega}) \right]. \quad (21)$$

Here,  $\tilde{U} = U + \Lambda_{\infty}^0 - \Lambda_{\infty}^z$  contains the asymptotic part of the retarded interaction  $\Lambda_{\omega}^{\alpha} = \Lambda_{\infty}^{\alpha} + \tilde{\Lambda}_{\omega}^{\alpha}$  (cf. Appendix C 1). The



summation in the second term in brackets in Eq. (21) in general runs over all channels,  $\alpha = 0, x, y, z$ . In the case of Ising-type coupling, only the retarded interaction in the two channels  $\alpha = 0, z$  is nonzero.

Because we assume that we solve the impurity model exactly, we have  $d_{\text{imp}}^{1\text{P}} = d_{\text{imp}}^{2\text{P}} \equiv d_{\text{imp}}$  (so that we can drop the superscript indices). While  $d_{\text{lat}}^{2\text{P}} = d_{\text{imp}}$ , the second term in (21) will in general lead to  $d_{\text{lat}}^{1\text{P}} \neq d_{\text{imp}}$ . The double occupancy computed from two-particle quantities  $d^{2\text{P}}$  is consistent with that of the impurity (because of two-particle self-consistency), while that obtained from single-particle quantities is not. This situation is exactly opposite to DMFT, where  $d^{1\text{P}}$  is consistent. We see that while the retarded interactions allow us to enforce consistency of  $d^{2\text{P}}$ , they simultaneously undermine the consistency of  $d^{1\text{P}}$ . We demonstrate this numerically in Fig. 6 for the two-particle self-consistent approximation presented in Sec. IV representative of Ising-type coupling and compare to DMFT.

### B. Heisenberg-type coupling

In the case of a Heisenberg-type coupling, all retarded interactions  $\Lambda_{\omega}^{\alpha}$  are in general nonzero. We fix their values through the self-consistency condition (20), as before, and consider the SU(2)-symmetric case with  $\Lambda^x = \Lambda^y = \Lambda^z$ . We further assume that the Ward identities hold. As a consequence, the relation (cf. Appendix B 1)

$$\lim_{\omega \rightarrow \infty} (i\omega)^2 X_{\text{loc},\omega}^{\alpha} = -2E_{\text{kin}} \quad (22)$$

holds in all channels  $\alpha = 0, x, y, z$ .

On the impurity model, contrary to the case of Ising-type coupling, we now have separate relations for the charge and spin susceptibilities (Appendix C 2):

$$\lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_{\omega}^0 = -4 \sum_{\nu} \Delta_{\nu} g_{\nu}, \quad (23)$$

$$\lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_{\omega}^z = -4 \sum_{\nu} \Delta_{\nu} g_{\nu} + 4 \sum_{\omega', \alpha=x,y} \tilde{\Lambda}_{\omega'}^{\alpha} \chi_{\omega'}^{\alpha}. \quad (24)$$

The corresponding relations for  $\chi^{x,y}$  are obtained by permuting  $x, y, z$  in (24). We see that in presence of retarded spin interactions, the asymptote of the impurity spin susceptibility in Eq. (24) no longer equals the kinetic energy. In addition, the asymptotes of the charge and spin channels are different. By virtue of the self-consistency condition  $\chi_{\omega}^{\alpha} = X_{\text{loc},\omega}^{\alpha}$ , this must also hold for the asymptote of  $X_{\text{loc},\omega}^{\alpha}$ . Equation (22), on the other hand, implies that the asymptote of the local susceptibility must be equal in all channels. We therefore conclude that there is no two-particle self-consistent approximation employing the self-consistency condition (20), which at the same time is locally conserving.<sup>3</sup> We note that in the case of Heisenberg-type coupling, the conclusions regarding the potential energy remain the same as in the Ising-type coupling. In particular, Eq. (21) still holds. We provide a numerical

<sup>3</sup>This only holds if  $\sum_{\omega', \alpha=x,y} \tilde{\Lambda}_{\omega'}^{\alpha} \chi_{\omega'}^{\alpha}$  is nonzero. The term does not vanish in our calculations and can only do so for an unphysical  $\tilde{\Lambda}_{\omega}^{\alpha}$  which changes its sign for different  $\omega$  [as  $\chi$  does not, cf. Eq. (C5)].

example of Eqs. (23) and (24) in the right panel of Fig. 8 in Appendix C 2.

### III. WARD IDENTITIES AND RETARDED SPIN-SPIN INTERACTIONS

In this section, we show that the Ward identities of the Bose-Fermi-Kondo model are incompatible with the Ward identities (5) of the Hubbard model. We identify this as the root cause of our earlier finding in the previous section that no locally conserving approximation can be obtained in case of a Heisenberg-type coupling.

In the Hamiltonian formulation of the Bose-Fermi-Kondo model (cf. Appendix (C2)), the retarded interactions enter as density-boson couplings  $\propto \phi^{\alpha} \rho^{\alpha}$ . Here,  $\phi = b^{\dagger} + b$  are bosonic operators which commute with all fermions. Integrating out the bosons in the functional integral yields the effective impurity action (19).

Since the Ward identities are Green's function equivalents of the continuity equations, they describe conservation of the charge and spin currents. These currents may be caused by kinetic and interaction contributions.

Regarding the latter, we notice two properties: (i) none of the retarded interactions contribute to the charge current, that is,  $[\rho^0, \phi^{\alpha} \rho^{\alpha}] = \phi^{\alpha} [n, \rho^{\alpha}] = 0$  for  $\alpha = 0, x, y, z$ . (ii) The spin current on the other hand has contributions from the retarded spin-spin interactions  $\Lambda^{\beta}$  due to noncommutativity of the spin operators  $[\rho^{\alpha}, \phi^{\beta} \rho^{\beta}] = 2i\phi^{\beta} \sum_{\gamma} \varepsilon_{\alpha\beta\gamma} \rho^{\gamma}$  for  $\alpha, \beta = x, y, z$ . We show in Appendix E that the resulting Ward identities contain an additional term that couples the retarded spin interaction to a three-particle correlation function [45]. As a consequence, they cannot be brought into the form of the local Ward identities (7). We emphasize that this does not imply a violation of spin conservation in the Bose-Fermi-Kondo model. The issue is instead that the Ward identities accounting for spin conservation simply have a form different from Eq. (7). It therefore seems plausible that conservation on the level of the BFK model does not imply that the local Ward identities (7) are fulfilled. That they are indeed violated in general is illustrated numerically in Fig. 1 by plotting the left- and right-hand sides of Eq. (7) for finite  $\tilde{\Lambda}^{x,y,z}$ .

In order to understand the consequences for constructing conserving approximations based on an impurity model, we recall that in a local approximation to the self-energy and irreducible vertex function the local Ward identities (6) are sufficient to guarantee that the approximation is conserving. In the case of DMFT, with the self-consistency condition  $G_{\text{loc}} = g$ , they coincide with the Ward identities of the AIM, so that DMFT is conserving. In presence of retarded spin-spin interactions this is no longer the case and, as we have seen numerically, this equation in general is violated. This can be seen as follows: Eq. (7) implies that the tails of the local susceptibilities must be identical independent of the channel index  $\alpha$ . We show this in Appendices B 1 and E 2. In the previous section, we have seen, however, that for the Heisenberg-type coupling they are different because of the retarded interaction [cf. Eqs. (23) and (24)]. Equation (7) must therefore be violated and the approximation is not conserving.

In the case of Ising-type coupling, the retarded interaction  $\Lambda^z$  in the longitudinal spin channel contributes to the currents

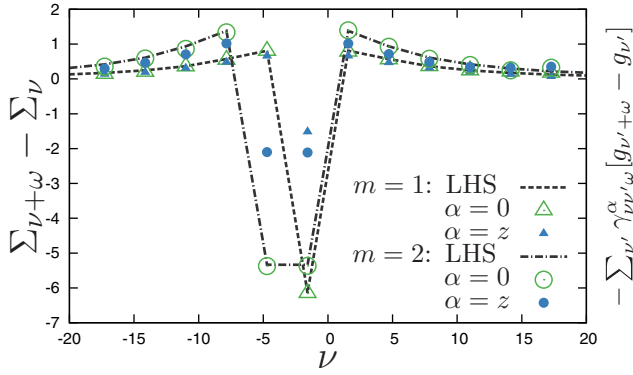


FIG. 1. Test of Eq. (7) for an isotropic retarded spin-spin interaction  $\tilde{\Lambda}^x = \tilde{\Lambda}^y = \tilde{\Lambda}^z$  and a retarded charge-charge interaction  $\tilde{\Lambda}^0$ . The imaginary part of the left-hand side (dashed black lines) and of the right-hand side (symbols) of Eq. (7) is drawn at the first two bosonic Matsubara frequencies  $\omega_{m=1,2}$ . Equation (7) holds in the charge channel (open green symbols) but is violated in the spin channels (filled blue symbols). This test was performed at  $\beta = 2$  and  $U = 6$  with a conducting bath  $\Delta$ . The violation of Eq. (7) in the spin channels depends on the magnitude of  $\tilde{\Lambda}^{x,y,z}$ , which was chosen large for demonstration purposes. The data shown in this figure were produced with the CTQMC solver presented in Ref. [44].

in the transversal spin channels of the impurity. The violation of the local Ward identities (7) thus affects only the transversal spin channels, while the longitudinal spin channel itself remains unaffected. That the Ward identity in the longitudinal spin channel indeed holds under these circumstances is demonstrated in Fig. 2.

#### IV. AN EXAMPLE: TWO-PARTICLE SELF-CONSISTENT DMFT

We have discussed the general conditions for a conserving and two-particle self-consistent approximation in Sec. II. Here, we construct a concrete example. As we have seen, an approximation that satisfies the two-particle self-consistency condition (20) can be conserving in the charge and at most one

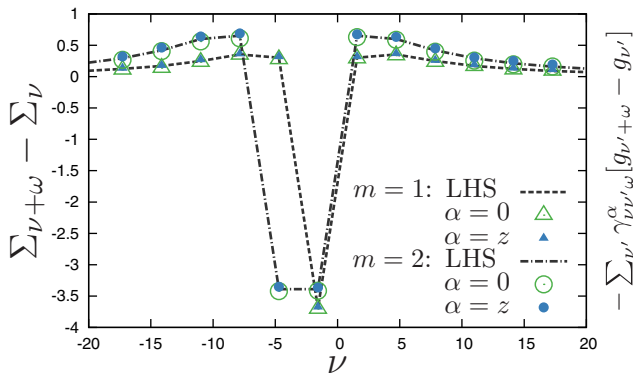


FIG. 2. Test of Eq. (7) for a retarded spin-spin interaction  $\tilde{\Lambda}^z$  in the  $z$  channel and a retarded charge-charge interaction  $\tilde{\Lambda}^0$ . Equation (7) holds in the channels  $\alpha = 0, z$ . Parameters as in Fig. 1, except  $\tilde{\Lambda}^x = \tilde{\Lambda}^y = 0$ .

of the spin channels. One may refer to this approximation as two-particle self-consistent DMFT.

We compute the lattice susceptibility in this approach according to

$$X_{q\omega}^\alpha = [(X_{q\omega}^{\text{DMFT},\alpha})^{-1} + \Lambda_\omega^\alpha]^{-1}. \quad (25)$$

The particular form (25) of the susceptibility can be motivated in the DB approach [19]. In this form, the retarded interaction is reminiscent of the Moriya  $\Lambda$  correction employed in D $\Gamma$ A. Here  $\Lambda$ , however, depends on frequency, while in D $\Gamma$ A it is instantaneous. We emphasize that the way of calculating the susceptibility and its particular form do not change the conserving character of the theory (see Sec. II A and results in Sec. IV A below), but will of course affect the results.

In the above,  $X^{\text{DMFT},\alpha}$  denotes the susceptibility computed as in DMFT in the standard way [9], including vertex corrections. This amounts to approximating the irreducible vertex function of the lattice with its local counterpart on the impurity  $\Gamma_{kk'q}^\alpha \equiv \gamma_{vv'\omega}^\alpha$  in the channels  $\alpha = 0, z$ . We compute the generalized susceptibility from the integral equation  $X_{kk'q}^\alpha = G_k G_{k+q} [\beta \delta_{kk'} - \sum_{k''} \gamma_{vv''\omega}^\alpha X_{k''k'q}^\alpha]$ . The susceptibilities are obtained from the latter by tracing out  $k, k'$ :  $X_q^{\text{DMFT},\alpha} = 2 \sum_{kk'} X_{kk'q}^\alpha$ . We emphasize that the label ‘‘DMFT’’ merely indicates that  $X^{\text{DMFT},\alpha}$  is computed as in DMFT. Its value will differ from the DMFT susceptibility because the impurity model is different.

The BFK model can be solved accurately using a suitably generalized continuous-time quantum Monte Carlo (CTQMC) algorithm. In weak-coupling CTQMC, the inclusion of these terms is straightforward [46]. In strong-coupling CTQMC the impurity model can be solved in the segment representation when only  $\Lambda^0$  or  $\Lambda^z$  are included [47–49]. For the general case of a vector bosonic field (not considered in the numerical results of this section), the algorithm simultaneously performs a hybridization expansion and an interaction expansion with respect to the spin-off-diagonal interactions  $\Lambda^{x,y}$  [44]. Here, we compute the correlation functions  $g_\nu$  and  $\chi_\omega^\alpha$ , the self-energy  $\Sigma_\nu$ , and the irreducible vertex function  $\gamma_{vv'\omega}^\alpha$  for  $\alpha = 0, z$  using a strong coupling quantum Monte Carlo solver [49] based on the ALPS libraries [50] with improved estimators adapted to treat the retarded interactions [51,52].

The calculation procedure is as follows: We start from initial values for the hybridization  $\Delta_\nu$  and retarded interactions  $\Lambda_\omega^\alpha$ , which specify the BFK impurity model (19). After solving the model, we evaluate the lattice susceptibility (25). The Green’s function is computed from the impurity self-energy in the same way as in DMFT:

$$G_{\mathbf{k}\nu}^{-1} = i\nu + \mu - \epsilon_{\mathbf{k}} - \Sigma_\nu^{\text{imp}}. \quad (26)$$

The local parts of  $G_{\mathbf{k}\nu}$  and  $X_{q\omega}^\alpha$  will in general be different from the impurity quantities  $g_\nu$  and  $\chi_\omega^\alpha$ . We update the hybridization  $\Delta_\nu$  and retarded interactions  $\Lambda_\omega^0, \Lambda_\omega^z$  simultaneously and iteratively, until the conditions  $G_{\text{loc},\nu} = g_\nu$  and  $X_{\text{loc},\omega}^\alpha = \chi_\omega^\alpha$  for  $\alpha = 0, z$  are satisfied.

#### A. Numerical results

Let us now turn to the discussion of numerical results of the two-particle self-consistent DMFT. In the following, we use

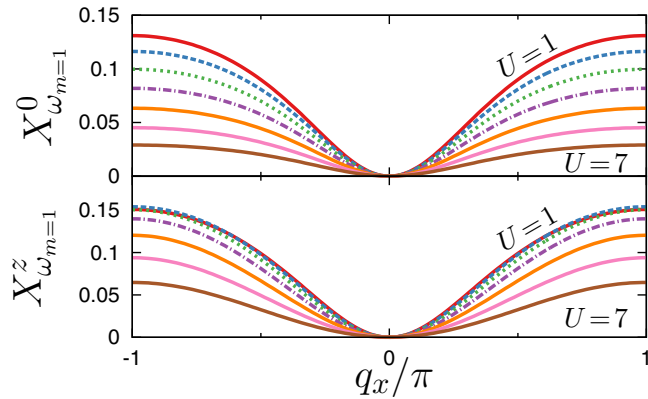


FIG. 3. Susceptibility at the first Matsubara frequency in the two-particle self-consistent DMFT. Shown is a momentum cross section at  $q_y = 0$  for different values of  $U$  ( $\beta = 2$ ). The vanishing of the susceptibility for  $\mathbf{q} \rightarrow 0$  with  $X \sim |\mathbf{q}|^2$  is a necessary condition for global conservation (see text).

parameters  $U = 6$ ,  $T = 0.5$  (in units of  $t$ ), which is somewhat above the DMFT Néel temperature  $T \approx 0.35$ .

In Fig. 2 we illustrate numerically that contrary to Heisenberg-type coupling (cf. Fig. 1), the local Ward identities (7) hold in the considered channels,  $\alpha = 0, z$ . As shown in Appendix B 1, this implies  $X_{\mathbf{q}=0, \omega \neq 0}^{\text{DMFT}, \alpha} = 0$ . Inserting this into Eq. (25) it follows that  $X_{\mathbf{q}=0, \omega \neq 0}^\alpha = 0$ , which is a necessary condition for conservation of the total density, i.e.,  $\omega \rho_{\mathbf{q}=0, \omega} = 0$ . Thus,  $X^0$  and  $X^z$  are at least globally conserving.

Figure 3 illustrates that  $X^0$  and  $X^z$  indeed vanish in the limit  $|\mathbf{q}| \rightarrow 0$  for finite frequencies ( $\omega_{m=1}$  in this case). We note that due to the retarded spin interaction  $\tilde{\Lambda}^z$  this approximation is not conserving in the  $x$  and  $y$  channels (cf. discussion in Sec. III and Appendix E).

The right panel of Fig. 4 demonstrates the equivalence of the impurity and local lattice susceptibility in two-particle self-consistent DMFT. In this approximation, the charge and the longitudinal spin susceptibility approach the same asymptote. This is required by the conservation laws and also satisfied in DMFT (see left panel, cf. Secs. I and II A).

In the top left panel of Fig. 5 we compare the static spin susceptibility  $X_{\mathbf{Q}, \omega=0}$  at  $\mathbf{Q} = (\pi, \pi)$  in DMFT, two-particle self-consistent DMFT, and the quantity  $X_{\mathbf{Q}, \omega=0}^{\text{DMFT}}$  in (25) at and near half-filling. The increase of  $X_{\mathbf{Q}, \omega=0}^{\text{DMFT}}$  compared to its value in standard DMFT is consistent with an enhanced local interaction on the impurity model  $\tilde{U} - U = \Lambda_\infty^0 - \Lambda_\infty^z > 0$  as seen in the bottom right panel of Fig. 5. Concomitantly, we also find a larger leading eigenvalue of the Bethe-Salpeter equation in the two-particle self-consistent DMFT (not shown).

In the top right panel, we see that the local susceptibility of the converged solution lies between the two values of DMFT. The effect of the two-particle self-consistency is larger close to half-filling, where antiferromagnetic fluctuations are strongest. Compared to DMFT, the increase of  $\chi_{\omega=0}$  in the two-particle self-consistent method correlates with the large enhancement  $U \rightarrow \tilde{U}$  of the onsite interaction in the impurity model.

Finally, we see in the top left panel of Fig. 5 that  $X_{\mathbf{Q}, \omega=0}^z$  is significantly reduced compared to  $X_{\mathbf{Q}, \omega=0}^{\text{DMFT}}$  due to  $\Lambda$  which acts as a cutoff. This reduction (marked by downward arrows)

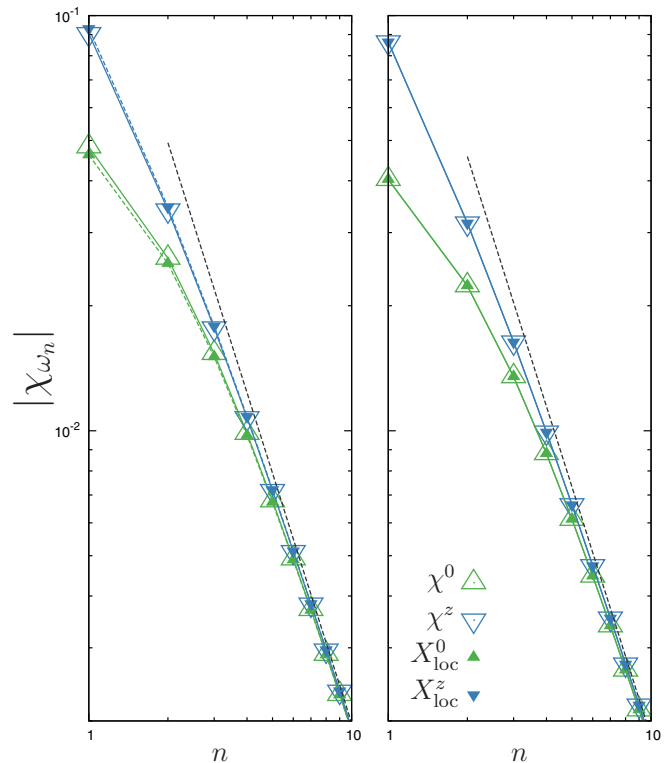


FIG. 4. High-frequency behavior of  $\chi$  (open triangles, bold lines) and  $X_{\text{loc}}$  (filled triangles, dashed lines) in DMFT (left) and in the two-particle self-consistent DMFT (right). The dashed black lines indicate the asymptotes  $-2E_{\text{kin}}/\omega^2$  computed from (11). The charge (green) and spin (blue) susceptibilities approach the same asymptote in both approximations.

becomes larger with a larger absolute value of the cutoff  $\Lambda_{\omega=0}$ , as can be seen in the bottom left panel.

While the lattice susceptibility (25) can in principle be defined without the dynamical cutoff  $\Lambda$ , this solution gives results closer to benchmarks: The double occupancy computed from the susceptibilities according to Eq. (15) or (16) gives a result that is closer to DCA benchmarks than either of the two values that are obtained in DMFT [37]. Despite two-particle self-consistency, the double occupancy is nevertheless inconsistent between one- and two-particle levels, as discussed in Sec. II A and demonstrated in Fig. 6.

We note that this approximation has several issues: The dynamic part of the retarded interaction in the spin channel  $\tilde{\Lambda}_\omega = \Lambda_\omega - \Lambda_\infty$  is positive. This corresponds to negative energies of the bosons [cf. (C5)] and is unphysical. The impurity model can nevertheless be solved in the QMC solver in the segment picture. Second, the asymptotic behavior of the self-energy is modified due to the retarded interactions [52]. Since DMFT produces the correct asymptotic behavior [13], the high-frequency tail of the self-energy in this approximation is no longer exact. Finally, even though the approximation suppresses a magnetic phase transition in two dimensions, the momentum-independent cutoff leads to an unphysical plateau of the susceptibility  $X_{\mathbf{q}, \omega=0} = \Lambda_{\omega=0}^{-1}$  for all momenta  $\mathbf{q}$  in the vicinity of  $\mathbf{Q}$  for which  $X_{\mathbf{q}, \omega=0}^{\text{DMFT}}$  diverges when approaching the DMFT Néel temperature.

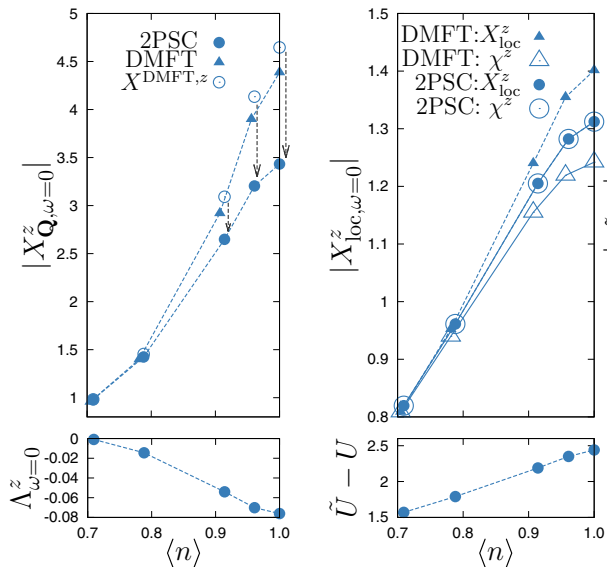


FIG. 5. Top: static susceptibilities of the Hubbard model (1) at  $U = 6$ ,  $\beta = 2$  in DMFT (triangles) and in the 2P self-consistent approximation (25) (circles) as a function of the density  $\langle n \rangle$ . Left: lattice susceptibility  $X_{\mathbf{Q}, \omega=0}^z$  at  $\mathbf{Q} = (\pi, \pi)$ . Right: local lattice ( $X_{loc}^z$ ) and impurity ( $\chi^z$ ) susceptibility. In DMFT (open and filled triangles), the local susceptibility  $X_{loc}^z$  (full triangles, dashed lines) is larger than the impurity susceptibility  $\chi^z$  (open triangles, bold lines). They coincide in the 2P self-consistent approximation (open and filled circles). Bottom, left: static component of the retarded spin-spin interaction  $\Lambda_{\omega=0}^z$ . Bottom, right: effective Hubbard repulsion  $\tilde{U}$ .

## V. DISCUSSION

We have discussed the conservation of charge and spin in two-particle self-consistent extensions of DMFT for the Hubbard model. For large interaction, the Hubbard model approximately maps to the Heisenberg model and is hence

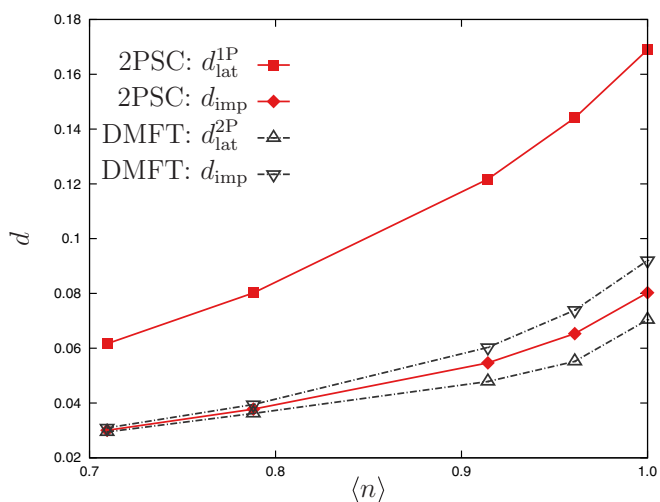


FIG. 6. Double occupancy as a function of filling. The figure illustrates the inconsistency of  $d_{imp}$  and  $d_{lat}^{2P}$  in DMFT and of  $d_{imp}$  and  $d_{lat}^{1P}$  in 2PSC. The impurity double occupancy computed from the susceptibilities (16) and the Migdal-Galitskii formula (21) yield the same result because we solve the impurity model exactly.

TABLE I. A summary of the main results for DMFT and 2PSC methods, respectively. The second column indicates which retarded interactions  $\Lambda^\alpha$  act on the impurity. The third and fourth columns show if the double occupancy  $d$  is consistent between impurity and lattice on the 1P and 2P levels, respectively (see Secs. I and II). The remaining columns list the channels in which local conservation is satisfied [see Eqs. (5)–(7) for DMFT, and Secs. II B and III for 2PSC].

	$\Lambda$	$d^{1P}$	$d^{2P}$	Charge	Spin $z$	Spin $x, y$
DMFT		✓		✓	✓	✓
Ising	$0, z$		✓	✓	✓	
Heisenberg	$0, x, y, z$		✓	✓		

dominated by spin fluctuations. The motivation for including a retarded spin-spin interaction into the impurity model is to account for these fluctuations.

As we have seen, however, introducing a retarded interaction in the longitudinal spin channel leads to a violation of conservation in the transversal spin channels of the lattice approximation. Moreover, a retarded interaction in all three spin channels violates conservation on the lattice in all spin channels (cf. Table I).

We have argued that this is related to the fact that the Ward identities of the lattice and impurity are incompatible. To make sense of this physically, we recall that an interaction only conserves the local charge or spin density if it commutes with the corresponding observable. The retarded charge-charge interaction commutes with the charge density and therefore preserves charge on the impurity. In other words, the bosons that mediate the retarded interaction do not carry a charge. On the other hand, the transversal components of the retarded spin-spin interaction do *not* commute with the longitudinal spin density operator. Consequently, the spin bosons carry spin: acting with the operator  $S^+$  or  $S^-$  on the impurity flips the spin of an electron by one quantum, which is carried by the boson. This leads to spin currents onto and off the impurity, which manifest themselves in the impurity Ward identities. These currents have no analog in the Hubbard model, where the motion of spin inevitably involves motion of charge. The latter is accounted for by the fermionic hybridization function. In essence, the introduction of the retarded spin-spin interactions in order to achieve two-particle self-consistency causes a “spin leak” in the lattice approximation. One may speculate that if the interaction of a lattice model conserves a local density, the local reference system should have the same property.

Even though the Hubbard model maps to a Heisenberg model for strong coupling, spin conservation is violated due to the exchange interaction on the impurity. This is no contradiction because the Heisenberg model is an effective low-energy model. The Ward identities, however, imply the equivalence of the charge and spin susceptibilities and excitations at high energies for any finite value of  $U$  [cf. Eq. (10) in Sec. I]. Indeed, the effective exchange coupling  $J = -4t^2/U$  involves two hopping processes and thus virtual high-energy charge excitations.

In the  $t$ - $J$  model, on the other hand, part of the spin currents is caused by the exchange interaction on the lattice. This part



is decoupled from the charge current. Contrary to the Hubbard model, a spin current on the impurity that is decoupled from the charge current *a priori* poses no problem and is even necessary. However, there remains the problem of finding a two-particle self-consistency condition which satisfies the Ward identities of the  $t$ - $J$  model.

This brings us to a last point, namely, a possible way out of this dilemma. We recall that our conclusions about the conserving character of the considered approximations follow from the two-particle self-consistency condition  $\chi_\omega^\alpha = X_{\text{loc},\omega}^\alpha$ , which seems like a natural choice. We can therefore not rule out the possibility that a different prescription exists, such that conservation is satisfied. In view of the above arguments, this seems unlikely in case of the Hubbard model, but more promising for the  $t$ - $J$  model.

### VI. CONCLUSIONS

We have investigated the interplay between the requirement of conservation of an approximation and two-particle self-consistency. Retarded interactions are required to enforce two-particle self-consistency, but their presence leads to problems. While the ambiguity in computing the DMFT double occupancy from two-particle quantities is resolved, the retarded interaction instead introduces an ambiguity in the calculation of the double occupancy from single-particle quantities.

More importantly, the Ward identities of the resulting impurity model are no longer compatible with the lattice Ward identities. As a consequence, we found that it is impossible to construct a two-particle self-consistent approximation to the Hubbard model which simultaneously fulfills the lattice Ward identities in the charge and all spin channels.

A conserving two-particle self-consistent approximation can be obtained when restricting self-consistency to the charge and one of the spin channels. We have used this to construct a two-particle self-consistent version of DMFT, which provably obeys global conservation laws and which resolves the ambiguity in the calculation of the double occupancy on the two-particle level. While this approximation suppresses a magnetic phase transition in two dimensions and yields results for the double occupancy which are closer to benchmarks than either of the two DMFT values, it however has several issues which make it impractical, in particular at low temperature. Our results imply constraints for the construction of two-particle self-consistent diagrammatic extensions.

Finally, we have seen that DMFT arises naturally when constructing a conserving approximation based on the Anderson impurity model. It may be possible to derive the cellular DMFT from the Ward identities, avoiding the cavity construction.

### ACKNOWLEDGMENTS

We thank the anonymous referees for constructive suggestions that have led to an improvement of this work. F.K. likes to thank G. Rohringer and A. Toschi for a useful discussion on Ward identities. F.K. and A.I.L. are financially supported by the DFG-SFB668 program. E.G.C.P.v.L. and

M.I.K. acknowledge support from ERC Advanced Grant No. 338957 FEMTO/NANO. Computational resources were provided by the HLRN-cluster under Project No. hhp00040.

### APPENDIX A: LATTICE WARD IDENTITIES

In this appendix, we detail the derivation of the Ward identities of a quantum lattice model with Hamiltonian  $H = H_0 + H_{\text{int}}$ , where  $H_0 = \sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma}$  is the noninteracting Hamiltonian and  $H_{\text{int}}$  is the interaction part. The Ward identities for the continuum can be found in textbooks [53,54]. Derivations for quantum lattice systems have been given in, e.g., [55–60] or [7]. The first steps of the following derivation were also done in [57], avoiding the introduction of a current operator. This makes the resulting Ward identities independent of the particular form of the dispersion  $\varepsilon_{\mathbf{k}}$ . More importantly, this allows us to derive analogous Ward identities for quantum impurity models in Appendix E.

The Ward identities can be viewed as sum rules for the four-point correlation function

$$G_{\mathbf{k}\mathbf{k}'\mathbf{q}}^{(2),\alpha}(\tau_1, \tau_2, \tau_3, \tau_4) = -\frac{1}{2} \sum_{\sigma_1\sigma_1'\sigma_2\sigma_2'} s_{\sigma_1'\sigma_1}^\alpha s_{\sigma_2'\sigma_2}^\alpha \times \langle T_\tau c_{\mathbf{k}\sigma_1}(\tau_1) c_{\mathbf{k}+\mathbf{q},\sigma_1'}^\dagger(\tau_2) c_{\mathbf{k}'\mathbf{q},\sigma_2}(\tau_3) c_{\mathbf{k}'\sigma_2'}^\dagger(\tau_4) \rangle, \quad (\text{A1})$$

which relate it to single-particle quantities. To obtain them, we examine the time derivative of this correlation function at equal times  $\tau_3 = \tau_4 = \tau$ , which allows us to express the result in terms of the time derivative of the density operators  $\rho_{\mathbf{q}}^\alpha = \sum_{\mathbf{k}\sigma\sigma'} c_{\mathbf{k}\sigma}^\dagger s_{\sigma\sigma'}^\alpha c_{\mathbf{k}+\mathbf{q},\sigma'}$ :

$$\begin{aligned} \partial_\tau \sum_{\mathbf{k}'} G_{\mathbf{k}\mathbf{k}'\mathbf{q}}^{(2),\alpha}(\tau_1, \tau_2, \tau, \tau) &= \partial_\tau \frac{1}{2} \sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha \langle T_\tau c_{\mathbf{k}\sigma}(\tau_1) c_{\mathbf{k}+\mathbf{q},\sigma'}^\dagger(\tau_2) \rho_{\mathbf{q}}^\alpha(\tau) \rangle \\ &= \frac{1}{2} \sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha \{ \langle T_\tau c_{\mathbf{k}\sigma}(\tau_1) [\rho_{\mathbf{q}}^\alpha(\tau), c_{\mathbf{k}+\mathbf{q},\sigma'}^\dagger(\tau)] \rangle \delta_{\tau,\tau_2} \\ &\quad + \langle T_\tau c_{\mathbf{k}+\mathbf{q},\sigma'}^\dagger(\tau_2) [c_{\mathbf{k}\sigma}(\tau), \rho_{\mathbf{q}}^\alpha(\tau)] \rangle \delta_{\tau,\tau_1} \\ &\quad - \langle T_\tau c_{\mathbf{k}\sigma}(\tau_1) c_{\mathbf{k}+\mathbf{q},\sigma'}^\dagger(\tau_2) [\rho_{\mathbf{q}}^\alpha(\tau), H] \rangle \}. \end{aligned} \quad (\text{A2})$$

The  $\delta$  functions arise because the time derivative does not commute with the time-ordering operator  $T_\tau$ . In the last line we have replaced the time derivative of the density operators using the continuity equation  $\partial_\tau \rho^\alpha = -[\rho^\alpha, H]$ . The commutators give  $[\rho_{\mathbf{q}}^\alpha, c_{\mathbf{k}+\mathbf{q},\sigma'}^\dagger] = \sum_{\sigma''} s_{\sigma''\sigma'}^\alpha c_{\mathbf{k},\sigma''}^\dagger$  and  $[c_{\mathbf{k}\sigma}, \rho_{\mathbf{q}}^\alpha] = \sum_{\sigma''} s_{\sigma\sigma''}^\alpha c_{\mathbf{k}+\mathbf{q},\sigma''}$ . We identify Green's function  $G_{\mathbf{k}\sigma}(\tau - \tau') \delta_{\sigma\sigma'} = -\langle T_\tau c_{\mathbf{k}\sigma}(\tau) c_{\mathbf{k}\sigma'}^\dagger(\tau') \rangle$  and bring the last term in Eq. (A2) to the left-hand side (LHS) to obtain the intermediate

result

$$\begin{aligned}
 & \partial_\tau \sum_{\mathbf{k}'} G_{\mathbf{k}\mathbf{k}'\mathbf{q}}^{(2),\alpha}(\tau_1, \tau_2, \tau, \tau) \\
 & + \frac{1}{2} \sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha \langle T_\tau c_{\mathbf{k}\sigma}(\tau_1) c_{\mathbf{k}+\mathbf{q},\sigma'}^\dagger(\tau_2) [\rho_{\mathbf{q}}^\alpha(\tau), H] \rangle \\
 & = \frac{1}{2} \sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha s_{\sigma'\sigma}^\alpha [G_{\mathbf{k}+\mathbf{q},\sigma'}(\tau-\tau_2) \delta_{\tau,\tau_1} - G_{\mathbf{k}\sigma}(\tau_1-\tau) \delta_{\tau,\tau_2}].
 \end{aligned} \tag{A3}$$

Assuming paramagnetism,  $G_\uparrow = G_\downarrow \equiv G$ , we can use  $\sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha s_{\sigma'\sigma}^\alpha = 2$ . We further use  $[\rho_{\mathbf{q}}^\alpha, H_0] = \sum_{\mathbf{k}'\sigma\sigma'} (\varepsilon_{\mathbf{k}'+\mathbf{q}} - \varepsilon_{\mathbf{k}'}) c_{\mathbf{k}'\sigma}^\dagger s_{\sigma'\sigma}^\alpha c_{\mathbf{k}'+\mathbf{q},\sigma'}$  to separate the noninteracting from the interacting Hamiltonian:

$$\begin{aligned}
 & \sum_{\mathbf{k}'} (\partial_\tau + \varepsilon_{\mathbf{k}'+\mathbf{q}} - \varepsilon_{\mathbf{k}'}) G_{\mathbf{k}\mathbf{k}'\mathbf{q}}^{(2),\alpha}(\tau_1, \tau_2, \tau, \tau) \\
 & + \frac{1}{2} \sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha \langle T_\tau c_{\mathbf{k}\sigma}(\tau_1) c_{\mathbf{k}+\mathbf{q},\sigma'}^\dagger(\tau_2) [\rho_{\mathbf{q}}^\alpha(\tau), H_{\text{int}}] \rangle \\
 & = G_{\mathbf{k}+\mathbf{q}}(\tau - \tau_2) \delta_{\tau,\tau_1} - G_{\mathbf{k}}(\tau_1 - \tau) \delta_{\tau,\tau_2}.
 \end{aligned} \tag{A4}$$

In the last step, we make use of the Fourier transform

$$\begin{aligned}
 & G^{(2)}(\tau_1, \tau_2, \tau_3, \tau_4) \\
 & = \sum_{\nu\nu'\omega} G_{\nu\nu'\omega}^{(2)} e^{-i[\nu\tau_1 - (\nu+\omega)\tau_2 + (\nu'+\omega)\tau_3 - \nu'\tau_4]},
 \end{aligned} \tag{A5}$$

introduce the short notation  $k = (\mathbf{k}, \nu)$ ,  $q = (\mathbf{q}, \omega)$ , and substitute the generalized susceptibility  $X_{kk'q}^\alpha = G_{kk'q}^{(2),\alpha} + 2\beta G_k G_{k'} \delta_q \delta_\alpha$  to obtain the Ward identities<sup>4</sup>

$$\begin{aligned}
 G_{k+q} - G_k & = \sum_{k'} X_{kk'q}^\alpha [\varepsilon_{k'+q} - \varepsilon_{k'} - i\omega] \\
 & + \frac{1}{2} \sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha \langle c_{k\sigma} c_{k+q,\sigma'}^\dagger [\rho_{\mathbf{q}}^\alpha, H_{\text{int}}] \rangle.
 \end{aligned} \tag{A6}$$

The first term on the right-hand side (RHS) can be recognized as the contribution from the noninteracting Hamiltonian, whereas the second contribution originates from the interaction.

In the Hubbard model, the interaction conserves the densities  $\rho^\alpha$ , that is  $[\rho^\alpha, U n_\uparrow n_\downarrow] = 0$  for  $\alpha = 0, x, y, z$ . Hence, the contribution of the interaction to the current in the Ward identities, the second line of (A6), vanishes. This allows to recast Eq. (A6) into a relation between the irreducible vertex  $\Gamma$ , Green's function  $G$ , and the self-energy  $\Sigma$ : the generalized susceptibility  $X_{kk'q}^\alpha$  is related to the irreducible vertex  $\Gamma_{kk'q}^\alpha$  via the integral equation  $X_{kk'q}^\alpha = G_k G_{k+q} [\beta \delta_{kk'} - \sum_{k''} \Gamma_{kk''q}^\alpha X_{k''k'q}^\alpha]$ . We insert this relation into Eq. (A6) and divide by  $G_k G_{k+q}$ . Using  $\varepsilon_{k+q} - \varepsilon_k - i\omega = [G_k^0]^{-1} - [G_{k+q}^0]^{-1}$  and  $[G_k^0]^{-1} - G_k^{-1} = \Sigma_k$  one has

$$\Sigma_{k+q} - \Sigma_k = - \sum_{k'} \Gamma_{kk'q}^\alpha \sum_{k''} X_{k'k''q}^\alpha [\varepsilon_{k''+q} - \varepsilon_{k''} - i\omega].$$

Since  $[\rho^\alpha, H_{\text{int}}] = 0$ , one can in turn insert the Ward identities (A6) on the RHS to obtain the desired relation

$$\Sigma_{k+q} - \Sigma_k = - \sum_{k'} \Gamma_{kk'q}^\alpha [G_{k'+q} - G_{k'}]. \tag{A7}$$

## APPENDIX B: SUSCEPTIBILITY ASYMPTOTE AND $f$ -SUM RULE

Assuming that the interaction does not contribute to the currents, i.e., letting  $[\rho^\alpha, H_{\text{int}}] = 0$  in the Ward identities (A6), we derive the  $(i\omega)^{-2}$  coefficient of the lattice susceptibility of locally conserving approximations. We prove that in this case local conservation implies the  $f$ -sum rule.

### 1. Susceptibility asymptote

We recognize that the second line of (A6) vanishes due to  $[\rho^\alpha, H_{\text{int}}] = 0$ . Summing over  $k$  yields zero on the LHS and we are left with

$$0 = \sum_{kk'} X_{kk'q}^\alpha [\varepsilon_{k'+q} - \varepsilon_{k'} - i\omega]. \tag{B1}$$

We use  $X_q^\alpha = 2 \sum_{kk'} X_{kk'q}^\alpha$  to arrive at an exact expression for the susceptibility

$$i\omega X_q^\alpha = 2 \sum_{kk'} X_{kk'q}^\alpha [\varepsilon_{k'+q} - \varepsilon_{k'}]. \tag{B2}$$

This implies for the homogeneous limit  $\mathbf{q} \rightarrow 0$  for finite frequency that

$$X_{\mathbf{q}=0, \omega \neq 0}^\alpha = - \langle (\rho_{\mathbf{q}=0, \omega \neq 0}^\alpha)^2 \rangle = 0. \tag{B3}$$

$\rho_{\mathbf{q}=0}^\alpha = \sum_i \rho_i^\alpha$  is the operator of total charge (spin) which is conserved if the continuity equations hold globally, that is,  $\partial_\tau \rho_{\mathbf{q}=0, \tau}^\alpha = 0$ . Thus,  $\omega \rho_{\mathbf{q}=0, \omega}^\alpha = 0$  or (B3) reflects that the approximation to  $X_{\mathbf{q}\omega}^\alpha$  conserves the total charge or spin, respectively. Therefore, as expected, local conservation implies global conservation.

To obtain the  $(i\omega)^{-2}$  coefficient of  $X_q^\alpha$ , we expand  $X_{kk'q}^\alpha = G_k G_{k+q} [\beta \delta_{kk'} - \sum_{k''} \Gamma_{kk''q}^\alpha X_{k''k'q}^\alpha]$  on the RHS of Eq. (B2) to order  $O(\omega^{-1})$ : According to the following Eq. (B4), the interacting bubble  $G_k G_{k+q}$  decays at least as  $\omega^{-1}$ . Since we can treat  $\Gamma$  as a constant at  $\omega \rightarrow \infty$ , vertex corrections to  $X_{kk'q}^\alpha$  are negligible at order  $O(\omega^{-1})$  and one is left with

$$\begin{aligned}
 X_{kk'q}^\alpha & = \beta G_k G_{k+q} \delta_{kk'} + O(\omega^{-2}) \\
 & = \frac{\beta}{i\omega} \frac{G_k - G_{k+q}}{1 + (\varepsilon_k - \varepsilon_{k+q} + \Sigma_k - \Sigma_{k+q})/i\omega} \delta_{kk'} + O(\omega^{-2}) \\
 & = \frac{\beta}{i\omega} (G_k - G_{k+q}) \delta_{kk'} + O(\omega^{-2}).
 \end{aligned} \tag{B4}$$

The Green's function  $G_{k+q}$  in the last line contributes since the term can be of order  $O(1)$  for  $k \approx -q$ . We return to the usual frequency and momentum notation  $k = (\mathbf{k}, \nu)$ ,  $q = (\mathbf{q}, \omega)$ , and insert (B4) into the RHS of (B2). This yields the asymptotic

<sup>4</sup>The term  $\propto \delta_{q=(\mathbf{q}, \omega), (0,0)}$  does not contribute to the Ward identities.

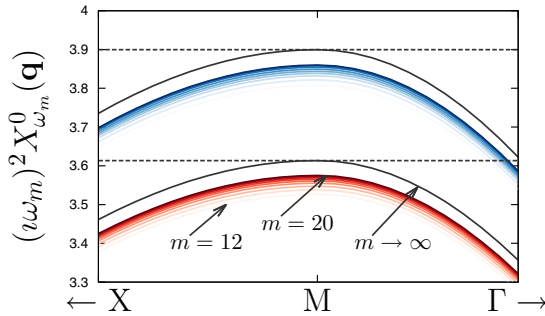


FIG. 7. Convergence of  $(t\omega_m)^2 X_{q\omega_m}^0$  to the RHS of the  $f$ -sum rule (B8). The figure shows the charge susceptibility in DMFT (blue) and two-particle self-consistent DMFT (red). Darker colors indicate a larger Matsubara index  $m \leq 20$ . Solid black curves mark the analytical expression for  $m \rightarrow \infty$ , Eq. (B7). Dashed lines indicate the value  $-4E_{\text{kin}}$  (half-filling,  $U = 6$ ,  $\beta = 2$ ).

coefficient of  $X_{q\omega}^\alpha$  at order  $O(\omega^{-2})$ :

$$\begin{aligned} \lim_{\omega \rightarrow \infty} (t\omega)^2 X_{q\omega}^\alpha &= \lim_{\omega \rightarrow \infty} 2 \sum_{\nu \mathbf{k}} (G_{\mathbf{k}\nu} - G_{\mathbf{k}+\mathbf{q}, \nu+\omega}) [\varepsilon_{\mathbf{k}+\mathbf{q}} - \varepsilon_{\mathbf{k}}] \\ &= 2 \sum_{\nu \mathbf{k}} (G_{\mathbf{k}\nu} \varepsilon_{\mathbf{k}+\mathbf{q}} + G_{\mathbf{k}+\mathbf{q}, \nu} \varepsilon_{\mathbf{k}} - 2G_{\mathbf{k}\nu} \varepsilon_{\mathbf{k}}) \\ &= \sum_{\mathbf{k}\sigma} \langle n_{\mathbf{k}\sigma} \rangle (\varepsilon_{\mathbf{k}+\mathbf{q}} + \varepsilon_{\mathbf{k}-\mathbf{q}} - 2\varepsilon_{\mathbf{k}}). \end{aligned} \quad (\text{B5})$$

Here, we have used  $\sum_{\nu} G_{\mathbf{k}\nu\sigma} e^{i\nu\omega} = \langle n_{\mathbf{k}\sigma} \rangle$ .<sup>5</sup> We show below that this equation is in fact the  $f$ -sum rule. For the local susceptibility  $X_{\text{loc},\omega}^\alpha = \sum_{\mathbf{q}} X_{q\omega}^\alpha$  it follows using  $\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} = 0$  and  $E_{\text{kin}} = \sum_{\mathbf{k}\sigma} \langle n_{\mathbf{k}\sigma} \rangle \varepsilon_{\mathbf{k}}$ , that

$$\lim_{\omega \rightarrow \infty} (t\omega)^2 X_{\text{loc},\omega}^\alpha = -2E_{\text{kin}}, \quad (\text{B6})$$

which is used several times in the main text.

Lastly, we simplify the asymptotic coefficient on the RHS of Eq. (B5) for a square lattice with nearest-neighbor hopping  $\varepsilon_{\mathbf{k}} = -2t[\cos k_x + \cos k_y]$ . In this case one can use goniometric equalities to extract the dependence on  $\mathbf{q}$  from Eq. (B5) as

$$\begin{aligned} \lim_{\omega \rightarrow \infty} (t\omega)^2 X_{q\omega}^\alpha &= 2 \sum_{\mathbf{k}\sigma} \langle n_{\mathbf{k}\sigma} \rangle (\varepsilon_{\mathbf{k}+\mathbf{q}} - \varepsilon_{\mathbf{k}}) \\ &= -4t \sum_{\mathbf{k}\sigma} \langle n_{\mathbf{k}\sigma} \rangle \sum_{i=x,y} [\cos k_i (\cos q_i - 1) - \sin k_i \sin q_i] \\ &= -4t \sum_{i=x,y} (\cos q_i - 1) \sum_{\mathbf{k}\sigma} \langle n_{\mathbf{k}\sigma} \rangle \cos k_i \\ &= \sum_{i=x,y} (\cos q_i - 1) \sum_{\mathbf{k}\sigma} \langle n_{\mathbf{k}\sigma} \rangle \varepsilon_{\mathbf{k}} \\ &= (\cos q_x + \cos q_y - 2) E_{\text{kin}}. \end{aligned} \quad (\text{B7})$$

<sup>5</sup>The factor  $e^{i\nu\omega}$  was left out for readability. It needs to be inserted in (A6) before summation over  $k$ .

It was used in the first line that  $\varepsilon_{-\mathbf{k}} = \varepsilon_{\mathbf{k}}$ , in the second line that  $\sum_{\mathbf{k}} \langle n_{\mathbf{k}} \rangle \sin k_i = 0$ , and in the third line that  $\sum_{\mathbf{k}} \langle n_{\mathbf{k}} \rangle \cos k_x = \sum_{\mathbf{k}} \langle n_{\mathbf{k}} \rangle \cos k_y$ , all valid by symmetry of the lattice.

Figure 7 numerically illustrates the validity of Eqs. (B5) and (B7) by plotting both sides of the equation (for the LHS, we take large but finite values of  $\omega$ ) in DMFT and two-particle self-consistent DMFT (cf. Sec. IV). At the  $M$  point, the RHS equals  $-4E_{\text{kin}}$ , which is marked by dashed horizontal lines. This value is larger in DMFT and given by  $-8 \sum_{\nu} \Delta_{\nu} g_{\nu}$  since the kinetic energy can be computed from the impurity (cf. Secs. I and II A).

## 2. $f$ -sum rule

The so-called  $f$ -sum rule (see, for example, [21])

$$\begin{aligned} -4 \lim_{\eta \rightarrow 0} \frac{1}{\beta} \sum_{n>0}^{\infty} \omega_n \sin(\eta\omega_n) X_{q\omega_n}^\alpha & \\ = \sum_{\mathbf{k}\sigma} \langle n_{\mathbf{k}\sigma} \rangle (\varepsilon_{\mathbf{k}+\mathbf{q}} + \varepsilon_{\mathbf{k}-\mathbf{q}} - 2\varepsilon_{\mathbf{k}}) & \end{aligned} \quad (\text{B8})$$

is a relation between the 2P response (LHS) and 1P quantities (RHS). To avoid confusion, we write the factor  $\beta^{-1}$  in front of the sum explicitly.

The Ward identities of the Hubbard model imply the  $f$ -sum rule [21]. It was also mentioned in [21] that the LHS of Eq. (B8) is entirely determined by the leading  $(t\omega)^{-2}$  coefficient of  $X_{q\omega}^\alpha$ . We can see this directly by comparing (B5) with the RHS of Eq. (B8). Here, we show that the sum on the LHS of (B8) indeed singles out the leading coefficient in the high-frequency expansion of the susceptibility.

The limit  $\eta \rightarrow 0$  of every summand on the LHS of Eq. (B8) is zero, whereas the limit of the sum is not. The LHS is convergent but not necessarily absolutely convergent. We drop the labels  $\alpha$  and  $\mathbf{q}$  temporarily and expand  $X_{q\omega}^\alpha$  for large frequencies. Since  $X(\omega_n) = X(-\omega_n)$ , only even powers of  $\omega_n$  appear in the expansion  $X(\omega_n) = \sum_{k=1}^{\infty} a_{2k} \omega_n^{-2k}$ . Inserting this expression into the  $f$ -sum rule (B8), one has on the LHS

$$-4 \sum_{k=1}^{\infty} a_{2k} \lim_{\eta \rightarrow 0} \frac{1}{\beta} \sum_{n=1}^{\infty} \frac{\sin(\eta\omega_n)}{\omega_n^{2k-1}}. \quad (\text{B9})$$

We hence need to evaluate

$$\lim_{\eta \rightarrow 0} \sum_{n=1}^{\infty} \frac{\sin(\eta n)}{n^{2k-1}} = \begin{cases} \lim_{\eta \rightarrow 0} \frac{\pi - \eta}{2} = \frac{\pi}{2} & \text{if } k = 1, \\ 0 & \text{if } k > 1. \end{cases} \quad (\text{B10})$$

For  $k = 1$ , the limit  $\eta \rightarrow 0$  and the summation over  $n$  must not be interchanged since  $\sum_{n=1}^{\infty} |\sin(\eta n)|/n$  diverges for  $0 < \eta < \pi$ . For the higher-order coefficients, the sum in Eq. (B10) is absolutely convergent, the limit and the sum can be interchanged, leading to zero. Correspondingly, the LHS of (B8) becomes

$$-4a_2 \lim_{\eta \rightarrow 0} \frac{1}{\beta} \sum_{n=1}^{\infty} \frac{\sin(\eta n)}{2\pi n/\beta} = -a_2. \quad (\text{B11})$$

As expected, only the  $(i\omega)^{-2}$  coefficient  $-a_2 = \lim_{\omega \rightarrow \infty} (t\omega)^2 X(t\omega)$  determines the LHS of the  $f$ -sum rule. As we have seen in the last paragraph, the  $f$ -sum rule follows from local conservation.

The above result is useful because a straightforward numerical evaluation of the  $f$ -sum rule (B8) suffers from oscillatory behavior of the LHS with the cutoff frequency. On the other hand,  $(i\omega)^2 X_{q\omega}^\alpha$  approaches the limit  $\omega \rightarrow \infty$  smoothly, which is illustrated in Fig. 7. For an accurate extrapolation to this limit, one needs to account for the effect of finite-frequency cutoffs in the vertex corrections to  $X_{q\omega}^\alpha$ .

### APPENDIX C: ASYMPTOTE OF THE IMPURITY SUSCEPTIBILITY

We determine the  $(i\omega)^{-2}$  asymptote of the impurity susceptibility  $\chi^\alpha$ . It can be shown from the Lehmann representation of  $\chi^\alpha$  that under paramagnetism its asymptote takes on the following form:

$$\lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_\omega^\alpha = \langle \langle [\bar{\rho}^\alpha, H_{\text{imp}}], \bar{\rho}^\alpha \rangle \rangle. \quad (\text{C1})$$

#### 1. Impurity Hamiltonian

In order to evaluate (C1) in the Bose-Fermi-Kondo model (19), we need to use its Hamiltonian formulation

$$H_{\text{imp}} = H_{\text{at}} + H_\Delta^0 + H_\Delta + H_\Lambda^0 + H_\Lambda. \quad (\text{C2})$$

The first three components

$$\begin{aligned} H_{\text{at}} &= -\tilde{\mu}n + \tilde{U}n_\uparrow n_\downarrow, \\ H_\Delta^0 &= \sum_{k\sigma} \epsilon_k f_{k\sigma}^\dagger f_{k\sigma}, \\ H_\Delta &= \sum_{k\sigma} (v_k c_\sigma^\dagger f_{k\sigma} + v_k^* f_{k\sigma}^\dagger c_\sigma) \end{aligned} \quad (\text{C3})$$

are the constituents of the Anderson impurity model, where a correlated impurity  $H_{\text{at}}$  is coupled to a noninteracting bath  $H_\Delta^0$  via the hybridization  $H_\Delta$ . In the BFK we further have the bosonic contributions

$$\begin{aligned} H_\Lambda^0 &= \sum_\alpha H_\Lambda^{0,\alpha}, \quad H_\Lambda = \sum_\alpha H_\Lambda^\alpha, \\ H_\Lambda^{0,\alpha} &= \sum_q \Omega_q^\alpha (b_q^\alpha)^\dagger b_q^\alpha, \\ H_\Lambda^\alpha &= \sum_q w_q^\alpha \bar{\rho}^\alpha \phi_q^\alpha, \quad \phi_q^\alpha = [(b_q^\alpha)^\dagger + b_q^\alpha], \end{aligned}$$

these couple the correlated site to bosonic baths  $H_\Lambda^0$  via the density-boson interactions  $H_\Lambda$ . Labeling the quantum numbers of the fermionic and bosonic baths with  $\mathbf{k}$  and  $\mathbf{q}$ , respectively, the situation described by the BFK is summarized as follows: The matrix elements  $v_k$  ( $w_q^\alpha$ ) couple the correlated fermions  $c^\dagger, c$  (density  $\rho^\alpha$ ) to a bath of noninteracting fermions  $f^\dagger, f$  (bosons  $\phi^\alpha$ ) with the spectrum  $\epsilon_k$  ( $\Omega_q^\alpha$ ).

Integrating out the baths in the path-integral formalism yields the effective action of the BFK (3), with the hybridization functions

$$\Delta_{\nu\sigma} = \sum_k |v_k|^2 \mathcal{G}_{k\nu\sigma}, \quad (\text{C4})$$

$$\tilde{\Lambda}_\omega^\alpha = \sum_q (w_q^\alpha)^2 \mathcal{D}_{q\omega}^\alpha. \quad (\text{C5})$$

Here,  $\mathcal{G}_{k\nu} = 1/(i\nu - \epsilon_k)$  and  $\mathcal{D}_{q\omega}^\alpha = -2\Omega_q^\alpha / [\omega^2 + (\Omega_q^\alpha)^2]$  denote the bath Green's functions. The self-consistent interaction  $\Lambda^\alpha = \Lambda_\infty^\alpha + \tilde{\Lambda}_\omega^\alpha$  has a constant part  $\Lambda_\infty^\alpha$  and a dynamic part  $\tilde{\Lambda}_\omega^\alpha$ , which has to vanish at  $\omega \rightarrow \infty$  [cf. (C5)]. The constant parts are absorbed into the Hubbard interaction  $\tilde{U} = U + \Lambda_\infty^0 - f\Lambda_\infty^z$  where  $f = 1$  in the Ising-type coupling and  $f = 3$  in the Heisenberg-type coupling (cf. Sec. II). The interactions  $\Lambda^\alpha$  change the chemical potential  $\tilde{\mu} = \mu + \mu_{\text{shift}}$ , which is shifted such that half-filling is obtained at  $\tilde{\mu} = \tilde{U}/2$ .

#### 2. Susceptibility asymptotes

To determine the asymptotic coefficient (C1), we need to calculate

$$C^\alpha = \langle \langle [\bar{\rho}^\alpha, H_{\text{imp}}], \bar{\rho}^\alpha \rangle \rangle.$$

Only the operators  $H_\Delta, H_\Lambda$  contribute to this expression:

$$C^\alpha = C_\Delta + \tilde{C}_\Lambda^\alpha, \quad (\text{C6})$$

$$C_\Delta = \langle \langle [\bar{\rho}^\alpha, H_\Delta], \bar{\rho}^\alpha \rangle \rangle = -\langle H_\Delta \rangle, \quad (\text{C7})$$

$$\begin{aligned} \tilde{C}_\Lambda^\alpha &= \sum_\gamma \langle \langle [\bar{\rho}^\alpha, H_\Lambda^\gamma], \bar{\rho}^\alpha \rangle \rangle = \begin{cases} 0 & \text{if } \alpha = 0, \\ \sum_{\gamma \neq \alpha} C_\Lambda^\gamma & \text{if } \alpha = x, y, z, \end{cases} \\ C_\Lambda^\gamma &= -4\langle H_\Lambda^\gamma \rangle. \end{aligned} \quad (\text{C8})$$

We calculate these expectation values from the impurity action. To this end, we denote the action of the Hamiltonian  $H_{\text{imp}}^0 = H_{\text{at}} + H_\Delta^0 + H_\Lambda^0$  as  $S^0$  and add sources  $J, J^\alpha$  to the actions corresponding to the operators  $H_\Delta$  and  $H_\Lambda = \sum_\alpha H_\Lambda^\alpha$ , respectively:

$$\begin{aligned} S^\Delta(J) &= \int_0^\beta d\tau \left\{ \sum_{k\sigma} [v_k + J_{k\sigma\tau}] c_{\sigma\tau}^* f_{k\sigma\tau} + v_k^* f_{k\sigma\tau}^* c_{\sigma\tau} \right\}, \\ S^\Lambda(J^\alpha) &= \int_0^\beta d\tau \sum_q [w_q^\alpha + J_{q\tau}^\alpha] \bar{\rho}_\tau^\alpha \phi_{q\tau}^\alpha. \end{aligned}$$

The expectation values can then be obtained as functional derivatives ( $S_{\text{imp}}^{\Delta,\Lambda} = S^0 + S^\Delta + S^\Lambda$ )

$$\begin{aligned} -\langle H_\Delta \rangle &= -\sum_{k\sigma} \{v_k \langle c_\sigma^\dagger f_{k\sigma} \rangle + v_k^* \langle f_{k\sigma}^\dagger c_\sigma \rangle\}, \\ \langle c_\sigma^\dagger f_{k\sigma} \rangle &= -\frac{1}{\mathcal{Z}} \frac{\delta}{\delta J_{k\sigma\tau=0}} \Big|_{J=0} \int \mathcal{D}[c, f, b] e^{-S_{\text{imp}}^{\Delta,\Lambda}}, \end{aligned} \quad (\text{C9})$$

where  $\mathcal{Z}$  is the grand partition sum, and

$$\begin{aligned} -\langle H_\Lambda^\gamma \rangle &= -\sum_q w_q^\gamma \langle \bar{\rho}^\gamma \phi_q^\gamma \rangle, \\ \langle \bar{\rho}^\gamma \phi_q^\gamma \rangle &= -\frac{1}{\mathcal{Z}} \frac{\delta}{\delta J_{q\tau=0}^\gamma} \Big|_{J^\gamma=0} \int \mathcal{D}[c, f, b] e^{-S_{\text{imp}}^{\Delta,\Lambda}}. \end{aligned} \quad (\text{C10})$$

The next step is to integrate out the fermionic and bosonic baths  $\int \mathcal{D}[c, f, b] e^{-S_{\text{imp}}^{\Delta,\Lambda}} = \mathcal{Z}_{f,b} \int \mathcal{D}[c] e^{-S_{\text{imp}}}$ . This gives rise to retarded couplings of  $c_\tau^*, c_{\tau'}$  and  $\rho_\tau^\alpha, \rho_{\tau'}^\alpha$  via the respective



bath Green's functions  $\mathcal{G}_{\tau-\tau'}$  and  $\mathcal{D}_{\tau-\tau'}$  [see below (C5)<sup>6</sup>]:

$$\begin{aligned} S'_{\text{imp}}(J, J^\alpha) &= \int_0^\beta d\tau \{c_{\sigma\tau}^* (\partial_\tau - \tilde{\mu}) c_{\sigma\tau} + U n_{\uparrow\tau} n_{\downarrow\tau}\} \\ &+ \int_0^\beta d\tau d\tau' \sum_{k\sigma} [v_k + J_{k\sigma\tau}] c_{\sigma\tau}^* \mathcal{G}_{k\sigma, \tau-\tau'} c_{\sigma\tau'} v_k^* \\ &+ \frac{1}{2} \int_0^\beta d\tau d\tau' \sum_{q\alpha} [w_q^\alpha + J_{q\tau}^\alpha] \bar{\rho}_\tau^\alpha \mathcal{D}_{q, \tau-\tau'}^\alpha \bar{\rho}_{\tau'}^\alpha [w_q^\alpha + J_{q\tau'}^\alpha]. \end{aligned}$$

Performing the derivatives with respect to the sources  $J, J^\alpha$  relates the desired expectation values  $\langle c_\sigma^\dagger f_{k\sigma} \rangle$  and  $\langle \bar{\rho}^\gamma \phi_q^\gamma \rangle$  to the correlation functions  $g$  and  $\chi$ :

$$\begin{aligned} \langle c_\sigma^\dagger f_{k\sigma} \rangle &= v_k^* \int_0^\beta d\tau' \mathcal{G}_{k\sigma, 0-\tau'} \langle c_{\sigma\tau=0}^* c_{\sigma\tau'} \rangle \\ &= v_k^* \sum_v \mathcal{G}_{k\sigma v} g_{v\sigma}, \end{aligned} \quad (\text{C11})$$

$$\begin{aligned} \langle \bar{\rho}^\gamma \phi_q^\gamma \rangle &= w_q^\gamma \int_0^\beta d\tau' \mathcal{D}_{q, 0-\tau'}^\gamma \langle \bar{\rho}_{\tau=0}^\gamma \bar{\rho}_{\tau'}^\gamma \rangle \\ &= -w_q^\gamma \sum_\omega \mathcal{D}_{q, \omega}^\gamma \chi_\omega^\gamma. \end{aligned} \quad (\text{C12})$$

Here, we have identified  $-\langle c_{\sigma\tau}^* c_{\sigma\tau=0} \rangle = g_{\sigma\tau}$  and  $-\langle \bar{\rho}_\tau^\gamma \bar{\rho}_{\tau=0}^\gamma \rangle = \chi_\tau^\gamma$ . Similarly,

$$\begin{aligned} \langle f_\sigma^\dagger c_{k\sigma} \rangle &= v_k \int_0^\beta d\tau \mathcal{G}_{k\sigma, \tau-0} \langle c_{\sigma\tau}^* c_{\sigma\tau=0} \rangle \\ &= v_k \sum_v \mathcal{G}_{k\sigma v} g_{v\sigma}. \end{aligned}$$

Inserting the expectation values into Eqs. (C9) and (C10), we finally determine the asymptotic coefficients

$$\begin{aligned} C_\Delta &= -\langle H_\Delta \rangle = -2 \sum_{k\sigma v} |v_k|^2 \mathcal{G}_{k\sigma v} g_{v\sigma}, \\ C_\Lambda^\gamma &= -4 \langle H_\Lambda^\gamma \rangle = 4 \sum_{q\omega} (w_q^\gamma)^2 \mathcal{D}_{q, \omega}^\gamma \chi_\omega^\gamma. \end{aligned}$$

Using the definitions of the bath Green's functions [Eqs. (C4) and (C5)], we conclude

$$C_\Delta = -2 \sum_{v\sigma} \Delta_{v\sigma} g_{v\sigma}, \quad (\text{C13})$$

$$C_\Lambda^\gamma = 4 \sum_\omega \tilde{\Lambda}_\omega^\gamma \chi_\omega^\gamma. \quad (\text{C14})$$

We examine this result for the Bose-Fermi-Kondo model (C2) in two cases: (i) Ising-type coupling: finite retarded interactions  $\tilde{\Lambda}^0, \tilde{\Lambda}^z$  in the density-type channels and  $\tilde{\Lambda}^x = \tilde{\Lambda}^y = 0$  in the transversal spin channels. Collecting Eqs. (C6), (C8), (C13), and (C14), we find that the susceptibilities of the

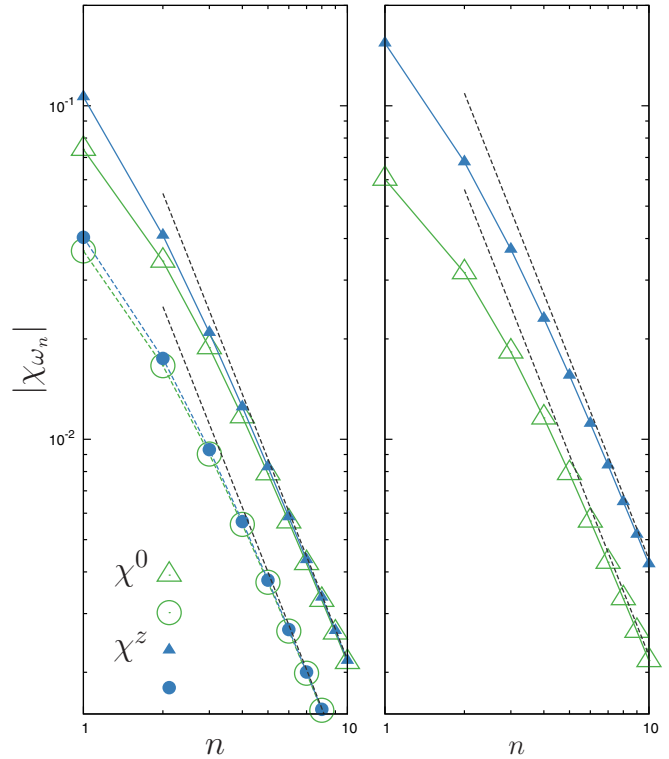


FIG. 8. Impurity susceptibility  $\chi^\alpha$  of the Bose-Fermi-Kondo model (symbols) and its exact high-frequency asymptote (dashed lines, see text). Left panel: Ising-type coupling,  $\chi^0$  (green) and  $\chi^z$  (blue) approach the same asymptote (triangles:  $\langle n \rangle = 0.54$ , circles:  $\langle n \rangle = 0.17$ ). Right panel: Heisenberg-type coupling,  $\chi^0$  and  $\chi^z$  approach different asymptotes. (Colors as in the left panel.  $\beta = 2$ ,  $U = 6$  in both panels.)

BFK assume the following asymptotic behavior for large  $\omega$ :

$$\lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_\omega^{0,z} = -2 \sum_{v\sigma} \Delta_{v\sigma} g_{v\sigma}, \quad (\text{C15})$$

$$\lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_\omega^{x,y} = -2 \sum_{v\sigma} \Delta_{v\sigma} g_{v\sigma} + 4 \sum_{\omega'} \tilde{\Lambda}_{\omega'}^z \chi_{\omega'}^z. \quad (\text{C16})$$

The Ising-type coupling is used in the two-particle self-consistent DMFT in Sec. IV. Hence, in this approximation  $\chi^0$  and  $\chi^z$  approach the asymptote given in Eq. (C15), which is demonstrated in the left panel of Fig. 8. Results for two different fillings  $\langle n \rangle$  are shown. Heisenberg-type coupling: Finite retarded interactions in all channels,  $\tilde{\Lambda}^0, \tilde{\Lambda}^x = \tilde{\Lambda}^y = \tilde{\Lambda}^z = \tilde{\Lambda}'$ :

$$\lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_\omega^0 = -2 \sum_{v\sigma} \Delta_{v\sigma} g_{v\sigma}, \quad (\text{C17})$$

$$\lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_\omega^z = -2 \sum_{v\sigma} \Delta_{v\sigma} g_{v\sigma} + 4 \sum_{\omega', \alpha=x,y} \tilde{\Lambda}_{\omega'}^\alpha \chi_{\omega'}^\alpha. \quad (\text{C18})$$

The asymptotes of  $\chi^x$  and  $\chi^y$  may be obtained by permuting the labels  $x, y, z$  in Eq. (C18).

We compare Eqs. (C17) and (C18) with numerical results in the right panel of Fig. 8. We used the CTQMC solver presented in [44] to calculate the impurity susceptibility  $\chi^\alpha$

<sup>6</sup>For readability, expectation values are still marked with the same brackets after integrating out the baths.

in the Bose-Fermi-Kondo model in Heisenberg-type coupling,  $\tilde{\Lambda}' \neq 0$ , with a conducting bath  $\Delta$  at half-filling.

#### APPENDIX D: MIGDAL FORMULA FOR THE DOUBLE OCCUPANCY

In the Hubbard model, the Migdal formula  $d = \langle n_\uparrow n_\downarrow \rangle = \text{Tr}(G\Sigma)/2U$  can be used to calculate the double occupancy from the potential energy  $Ud$  [40]. In the Bose-Fermi-Kondo model (C2), the potential energy is modified due to retarded interactions  $\Lambda_\omega = \Lambda_\infty + \tilde{\Lambda}_\omega$ . Their constant parts cause a shift  $U \rightarrow \tilde{U} = U + \Lambda_\infty^0 - f\Lambda_\infty^z$ . In the following, we determine the effect of the dynamic part  $\tilde{\Lambda}_\omega$ . We start from the equation of motion (EOM) of the impurity Green's function  $g_{-\tau} = -\langle T_\tau c_{-\tau\sigma} c_\sigma^\dagger \rangle$ :

$$\begin{aligned} -\partial_\tau g_{-\tau\sigma} &= -\delta(\tau) \langle (c_\sigma, c_\sigma^\dagger) \rangle + \langle T_\tau [c_{-\tau,\sigma}, H_{\text{imp}}] c_\sigma^\dagger \rangle \\ \Leftrightarrow \sum_\nu (-i\nu) g_{\nu\sigma} e^{i\nu\tau} &= -\delta(\tau) + \langle T_\tau [c_{-\tau,\sigma}, H_{\text{imp}}] c_\sigma^\dagger \rangle. \end{aligned} \quad (\text{D1})$$

The time derivative in the EOM is taken from the left ( $-\tau \leq 0$ ). This is done in order to approach the equal-time limit  $\tau \rightarrow 0$  without a jump in  $g(\tau)$ . To evaluate the RHS, we need to build the commutator  $[c_\sigma, H_i]$  with all the components of the impurity Hamiltonian (C2). For this, we need the commutators

$$\begin{aligned} [c_\sigma, c_{\sigma'}^\dagger] &= \delta_{\sigma,\sigma'} - 2c_{\sigma'}^\dagger c_\sigma, \\ [c_\sigma, c_{\sigma'}] &= -2c_{\sigma'} c_\sigma, \\ [c_\sigma, n_\uparrow n_\downarrow] &= (n_\uparrow \delta_{\sigma\downarrow} + n_\downarrow \delta_{\sigma\uparrow}) c_\sigma, \\ [c_\sigma, \bar{\rho}^\alpha] &= [c_\sigma, \rho^\alpha] = \sum_{\sigma'} s_{\sigma\sigma'}^\alpha c_{\sigma'}. \end{aligned}$$

Then, the commutators on the RHS of the EOM (D1) become

$$\begin{aligned} [c_\sigma, H_{\text{at}}] &= -\tilde{\mu} c_\sigma + \tilde{U} (n_\uparrow \delta_{\sigma\downarrow} + n_\downarrow \delta_{\sigma\uparrow}) c_\sigma, \\ [c_\sigma, H_\Delta] &= \sum_{k\sigma'} (v_k [c_\sigma, c_{\sigma'}^\dagger, f_{k\sigma'}] + v_k^* [c_\sigma, f_{k\sigma'}^\dagger, c_{\sigma'}]) \\ &= \sum_k v_k f_{k\sigma}, \\ [c_\sigma, H_\Lambda^\alpha] &= \sum_q w_q^\alpha \phi_q^\alpha \sum_{\sigma'} s_{\sigma\sigma'}^\alpha c_{\sigma'}. \end{aligned}$$

The other components of  $H_{\text{imp}}$  commute with  $c_\sigma$ . We insert these results into (D1) and sum over  $\sigma$  on both sides:

$$\begin{aligned} \sum_\sigma (-\partial_\tau) g_{-\tau\sigma} &= -2\delta(\tau) - \tilde{\mu} \sum_\sigma \langle T_\tau c_{-\tau,\sigma} c_\sigma^\dagger \rangle \\ &+ \tilde{U} \sum_\sigma \langle T_\tau n_{-\tau,-\sigma} c_{-\tau\sigma} c_\sigma^\dagger \rangle + \sum_{k\sigma} v_k \langle T_\tau f_{-\tau k\sigma} c_\sigma^\dagger \rangle \\ &+ \sum_{\alpha q} w_q^\alpha \left\langle T_\tau \phi_{-q}^\alpha \sum_{\sigma\sigma'} s_{\sigma\sigma'}^\alpha c_{-\tau\sigma'} c_\sigma^\dagger \right\rangle. \end{aligned} \quad (\text{D2})$$

We identify the impurity Green's function  $g_{-\tau\sigma} = -\langle T_\tau c_{-\tau,\sigma} c_\sigma^\dagger \rangle$  on the RHS and order the remaining expectation values by time ( $-\tau < 0$ ):

$$\begin{aligned} \sum_\sigma (-\partial_\tau - \tilde{\mu}) g_{-\tau\sigma} &= -2\delta(\tau) - \tilde{U} \sum_\sigma \langle T_\tau c_\sigma^\dagger n_{-\sigma} c_{-\tau\sigma} \rangle - \sum_{k\sigma} v_k \langle T_\tau c_\sigma^\dagger f_{-\tau k\sigma} \rangle \\ &- \sum_{\alpha q} w_q^\alpha \left\langle T_\tau \phi_{-q}^\alpha \sum_{\sigma\sigma'} s_{\sigma\sigma'}^\alpha c_{-\tau\sigma'} c_\sigma^\dagger \right\rangle. \end{aligned} \quad (\text{D3})$$

The aim is to take the EOM at  $\tau = 0$ , we insert  $0^+$  in places where the limit is dubious. Time-ordered products sort creation operators to the left of annihilators at equal time. Since we approach this point from  $-\tau < 0$ , the order of operators in Eq. (D3) remains unchanged in the limit, avoiding jumps in the expectation values. At equal time, we recognize the double occupancy  $\langle c_\sigma^\dagger n_{-\sigma} c_\sigma \rangle = \langle n_{-\sigma} n_\sigma \rangle = d$  and the density operators  $\sum_{\sigma\sigma'} s_{\sigma\sigma'}^\alpha c_\sigma^\dagger c_{\sigma'} = \rho^\alpha$ . The averages  $\sum_k v_k \langle c_\sigma^\dagger f_{k\sigma} \rangle = \sum_\nu \Delta_{\nu\sigma} g_{\nu\sigma}$  and  $\sum_q w_q^\alpha \langle \phi_q^\alpha \bar{\rho}^\alpha \rangle = -\sum_\omega \tilde{\Lambda}_\omega^\alpha \chi_\omega^\alpha$  have been obtained from functional derivatives in the previous section. To use the latter in Eq. (D3), one has to account for the difference in  $\rho$  and  $\bar{\rho} = \rho - \langle \rho \rangle$ . Back in Eq. (D3) we have

$$\begin{aligned} -\sum_{\nu\sigma} (i\nu + \tilde{\mu} - \Delta_{\nu\sigma}) g_{\nu\sigma} e^{i\nu 0^+} &= -2\delta(0^+) - 2\tilde{U}d + \sum_{\alpha\omega} \tilde{\Lambda}_\omega^\alpha \chi_\omega^\alpha - \tilde{\Lambda}_{\omega=0}^\alpha \langle \rho^\alpha \rangle^2. \end{aligned} \quad (\text{D4})$$

We recognize  $g^0 = [i\nu + \tilde{\mu} - \Delta]^{-1}$  on the LHS and use Dyson's equation  $g/g^0 = 1 + \Sigma g$ .  $\sum_{\nu\sigma} e^{i\nu 0^+} = 2\delta(0^+)$  on the LHS cancels, leaving us with a modified Migdal-Galitzkii formula

$$d = \frac{1}{2\tilde{U}} \sum_{\nu\sigma} g_{\nu\sigma} \Sigma_{\nu\sigma} e^{i\nu 0^+} + \frac{1}{2\tilde{U}} \sum_{\alpha\omega} \tilde{\Lambda}_\omega^\alpha (\chi_\omega^\alpha - \beta \langle \rho^\alpha t \rangle^2 \delta_\omega). \quad (\text{D5})$$

The factor  $\beta$  on the RHS accounts for the factor  $\beta^{-1}$  implied in the frequency summation. The above formula does not express the double occupancy  $d$  in terms of 1P quantities, due to the contribution of  $\chi$  on the RHS. However, bringing this contribution to the LHS, one still obtains a relation between 2P and 1P quantities, which justifies the label  $d_{\text{imp}}^{\text{1P}}$  used in the main text.

To accurately calculate the sum  $\sum_{\nu\sigma} g_{\nu\sigma} \Sigma_{\nu\sigma} e^{i\nu 0^+}$ , one needs to separate the constant Hartree part from the self-energy  $\Sigma_\nu = \Sigma'_\nu + \Sigma^{\text{H}}$  and to treat the asymptotes of  $g_\nu = 1/i\nu + \dots$  and  $\Sigma'_\nu = c_1/i\nu + \dots$  analytically. Then, the equal-time limit can be taken safely:

$$\sum_{\nu\sigma} g_{\nu\sigma} \Sigma_{\nu\sigma} e^{i\nu 0^+} = \sum_{\nu\sigma} g_{\nu\sigma} (\Sigma'_{\nu\sigma} + \Sigma^{\text{H}}) e^{i\nu 0^+} \quad (\text{D6})$$

$$= \sum_{\nu\sigma} [g_{\nu\sigma} \Sigma'_{\nu\sigma} - c_1/(i\nu)^2] - c_1\beta/2 + \langle n \rangle \Sigma^{\text{H}}. \quad (\text{D7})$$

### APPENDIX E: WARD IDENTITIES AND ASYMPTOTES IN THE BOSE-FERMI-KONDO MODEL

We derive the Ward identities of the Bose-Fermi-Kondo model (C2) and establish a relation to the susceptibility asymptotes derived in Appendix C2.

#### 1. Ward identities

We can follow the derivation for the Ward identities on the lattice in Appendix A which led to (A3). Omitting momentum indices, and inserting impurity instead of lattice quantities,  $G_k \rightarrow g_v$ ,  $G_{kk'}^{(2),\alpha} \rightarrow g_{vv'\omega}^{(2),\alpha}$ ,  $X_{kk'q}^\alpha \rightarrow \chi_{vv'\omega}^\alpha$ ,  $\Gamma_{kk'q}^\alpha \rightarrow \gamma_{vv'\omega}^\alpha$ ,  $H \rightarrow H_{\text{imp}}$ , we obtain in frequency space

$$\frac{1}{2} \sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha \left\{ -i\omega \langle c_{v\sigma} c_{v+\omega,\sigma'}^\dagger \rho_\omega^\alpha \rangle + \langle c_{v\sigma} c_{v+\omega,\sigma'}^\dagger [\rho_\omega^\alpha, H_{\text{imp}}] \rangle \right\} = g_{v+\omega} - g_v. \quad (\text{E1})$$

The impurity Hamiltonian  $H_{\text{imp}}$  is defined in (C2), where  $[\rho_\omega^\alpha, H_{\text{at}}] = [\rho_\omega^\alpha, H_\Delta^0] = [\rho_\omega^\alpha, H_\Lambda^0] = 0$ . We treat the remaining contributions from  $H_\Delta$  and  $H_\Lambda$  separately using  $[\rho^\alpha, H_\Delta] = \sum_{k\sigma\sigma'} s_{\sigma'\sigma}^\alpha (v_k c_{\sigma'}^\dagger f_{k\sigma} - v_k^* f_{k\sigma}^\dagger c_\sigma)$ . We need to calculate the following correlation function in Eq. (E1):

$$\begin{aligned} & \frac{1}{2} \sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha \langle c_{v\sigma} c_{v+\omega,\sigma'}^\dagger [\rho_\omega^\alpha, H_\Delta] \rangle \\ &= \frac{1}{2} \sum_{\sigma_1\sigma_1'\sigma_2\sigma_2'} s_{\sigma_1'\sigma_1}^\alpha s_{\sigma_2'\sigma_2}^\alpha \left\{ \sum_{k\nu'} v_k \langle c_{v\sigma_1} c_{v+\omega,\sigma_1'}^\dagger c_{\nu'\sigma_2'}^\dagger f_{k\nu'+\omega,\sigma_2} \rangle \right. \\ & \quad \left. - \sum_{k\nu'} v_k^* \langle c_{v\sigma_1} c_{v+\omega,\sigma_1'}^\dagger f_{k\nu'\sigma_2'}^\dagger c_{\nu'+\omega,\sigma_2} \rangle \right\}. \quad (\text{E2}) \end{aligned}$$

As demonstrated in Appendices C2 and D, the bath operators in the impurity averages can be transferred to Grassmann numbers in the path-integral formalism and can then be integrated out. This leaves convolutions with the hybridization function  $\Delta_{v\sigma} = \sum_k |v_k|^2 \mathcal{G}_{k\sigma\nu}$  [cf. (C4)]. We ascertain that the following replacements are valid: “ $\sum_k v_k f_{v\sigma} \rightarrow \Delta_{v\sigma} c_{v\sigma}$ ” and “ $\sum_k v_k^* f_{v\sigma}^* \rightarrow \Delta_{v\sigma} c_{v\sigma}^*$ .” Performing these replacements in Eq. (E2) and identifying the four-point correlation function

$$g_{vv'\omega}^{(2),\alpha} = -\frac{1}{2} \sum_{\sigma_1\sigma_1'\sigma_2\sigma_2'} s_{\sigma_1'\sigma_1}^\alpha s_{\sigma_2'\sigma_2}^\alpha \langle c_{v\sigma_1} c_{v+\omega,\sigma_1'}^\dagger c_{\nu'+\omega,\sigma_2} c_{\nu'\sigma_2'}^\dagger \rangle$$

and the generalized susceptibility  $\chi_{vv'\omega}^\alpha = g_{vv'\omega}^{(2),\alpha} + 2\beta g_v g_{v'} \delta_\omega \delta_\alpha$  of the impurity, one can separate the noninteracting part of (E1) as

$$\begin{aligned} g_{v+\omega} - g_v &= \sum_{\nu'} \chi_{vv'\omega}^\alpha [\Delta_{\nu'+\omega} - \Delta_{\nu'} - i\omega] \\ & \quad + \frac{1}{2} \sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha \langle c_{v\sigma} c_{v+\omega,\sigma'}^\dagger [\rho_\omega^\alpha, H_\Lambda] \rangle. \quad (\text{E3}) \end{aligned}$$

We are left with the commutator  $[\rho^\alpha, H_\Lambda]$  on the RHS with the retarded interactions  $H_\Lambda = \sum_\alpha H_\Lambda^\alpha$ . We recall that  $H_\Lambda = \sum_{\alpha q} w_q^\alpha \bar{\rho}^\alpha \phi_q^\alpha$  and recognize that  $[\rho^0, H_\Lambda] = 0$ , that is, the retarded interactions do not contribute to the charge current.  $\Lambda^{x,y,z}$  on the other hand do contribute to the spin currents due to the commutation relations for spin operators  $[\rho^\alpha, \rho^\beta] = 2i \sum_\gamma \varepsilon_{\alpha\beta\gamma} \rho^\gamma$ ,

$$[\rho^\alpha, H_\Lambda] = 2i \sum_{\beta\gamma} \varepsilon_{\alpha\beta\gamma} \rho^\gamma \sum_q w_q^\beta \phi_q^\beta. \quad (\text{E4})$$

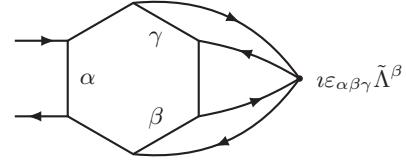


FIG. 9. Symbolic representation of the second line of Eq. (E5). One obtains a representation of the RHS of Eq. (E6) by tapering the open Green's function lines (implying summation over  $\alpha = x, y, z$ ).

Inserting this relation into Eq. (E3), one can transfer the bosonic operators  $\phi^\beta$  into complex variables in the path integral and integrate these out. As exercised in Appendix C2, this leads to a replacement rule “ $\sum_q w_q^\beta \phi_q^\beta \rightarrow \tilde{\Lambda}_\omega^\beta \rho_\omega^\beta$ ,” and hence  $\phi^\beta$  give rise to the retarded spin-spin interaction  $\tilde{\Lambda}^\beta$ . We obtain a six-point correlation function on the RHS of Eq. (E3), leaving us with the Ward identities of the BFK:

$$\begin{aligned} g_{v+\omega} - g_v &= \sum_{\nu'} \chi_{vv'\omega}^\alpha [\Delta_{\nu'+\omega} - \Delta_{\nu'} - i\omega] \\ & \quad + \frac{1}{2} \sum_{\sigma\sigma'} s_{\sigma'\sigma}^\alpha \sum_{\omega'\beta\gamma} 2i \varepsilon_{\alpha\beta\gamma} \tilde{\Lambda}_{\omega'}^\beta \langle c_{v\sigma} c_{v+\omega,\sigma'}^\dagger \rho_{\omega'}^\beta \rho_{\omega-\omega'}^\gamma \rangle. \quad (\text{E5}) \end{aligned}$$

The six-point correlation function in the second line of Eq. (E5) contains essentially three-particle-irreducible contributions (which cannot be broken up into parts by cutting one, two, or three fermion lines) (see Fig. 9). Hence, the Ward identities of the BFK (E5) cannot be recast into a relation which features only the one- and two-particle-irreducible vertices  $\Sigma$  and  $\gamma^\alpha$ .

Assuming an isotropic retarded interaction  $\tilde{\Lambda}^{x,y,z} = \tilde{\Lambda}'$ , we can sum the impurity Ward identities (E5) over  $\nu$  and  $\alpha$ , use the definition of the spin-density operators  $\rho^{x,y,z} = 2S^{x,y,z} = \sum_{\sigma\sigma'} c_\sigma^\dagger s_{\sigma\sigma'}^\alpha c_{\sigma'}$ , and the definition of the vector product  $(\mathbf{A} \times \mathbf{B})^\alpha = \sum_{\beta\gamma} \varepsilon_{\alpha\beta\gamma} A^\beta B^\gamma$  to see that  $\tilde{\Lambda}'$  couples to the time-dependent spin chirality [61]

$$\begin{aligned} & \sum_{\nu\nu'\alpha} \chi_{vv'\omega}^\alpha [\Delta_{\nu'+\omega} - \Delta_{\nu'} - i\omega] \\ &= 8i \sum_{\omega'} \tilde{\Lambda}'_{\omega'} \langle \mathbf{S}_{-\omega} (\mathbf{S}_\omega \times \mathbf{S}_{\omega-\omega'}) \rangle. \quad (\text{E6}) \end{aligned}$$

Writing the RHS in imaginary time,  $\int_0^\beta \tilde{\Lambda}'_{\tau_2-\tau_3} \langle T_\tau [\mathbf{S}_{\tau_1} (\mathbf{S}_{\tau_2} \times \mathbf{S}_{\tau_3})] \rangle d\tau_2$ , the spin chirality obviously vanishes when two of its time indices are equal. Hence, this contribution arises exclusively in presence of a time-dependent spin-spin interaction.

In the Anderson impurity model, we have  $\Lambda = 0$  and the interaction  $U n_\uparrow n_\downarrow$  does not contribute to the currents. In this case, the six-point correlation function drops out of the impurity Ward identities (E5). Then, analogous to the lattice Ward identities (A7), one can recast the impurity Ward identities into the form  $\Sigma_{v+\omega} - \Sigma_v = -\sum_{\nu'} \gamma_{vv'\omega}^\alpha [g_{\nu'+\omega} - g_{\nu'}]$ , i.e., Eq. (7) in the main text.

#### 2. Relation to susceptibility asymptotes

We prove that the Ward identities of the AIM (7) determine the  $(i\omega)^{-2}$  coefficient of the impurity susceptibility. Consequently, finding a different coefficient proves a violation of Eq. (7), which is used in Sec. III. In the AIM we have  $\tilde{\Lambda} = 0$ . Then, similar considerations as in Appendix B1

[cf. Eq. (B5)] show that the impurity Ward identities (E5) imply the following high-frequency asymptote of the susceptibility:

$$\begin{aligned} \lim_{\omega \rightarrow \infty} (i\omega)^2 \chi_\omega^\alpha &= \lim_{\omega \rightarrow \infty} 2 \sum_\nu (g_\nu - g_{\nu+\omega}) [\Delta_{\nu+\omega} - \Delta_\nu] \\ &= -4 \sum_\nu g_\nu \Delta_\nu. \end{aligned} \quad (\text{E7})$$

In the last step it was used that  $g_{\nu+\omega}, \Delta_{\nu+\omega}$  vanish at  $\omega \rightarrow \infty$ . Equation (E7) is in agreement with the asymptotes of  $\chi$  in the BFK for  $\tilde{\Lambda}^{x,y,z} = 0$  [see Eq. (C18)]. For  $\tilde{\Lambda} \neq 0$  we find by comparison of Eq. (E7) with the asymptote of the exact solution (C18) that the six-point correlation function in the impurity Ward identities (E5) contributes to the asymptote of the spin susceptibility  $\chi^{\alpha=x,y,z}$  with  $4 \sum_{\omega', \beta \neq \alpha} \tilde{\Lambda}_{\omega'}^\beta \chi_{\omega'}^\beta$ .

- 
- [1] E. Dagotto, *Rev. Mod. Phys.* **66**, 763 (1994).  
[2] M. Imada, A. Fujimori, and Y. Tokura, *Rev. Mod. Phys.* **70**, 1039 (1998).  
[3] E. G. C. P. van Loon, H. Hafermann, A. I. Lichtenstein, and M. I. Katsnelson, *Phys. Rev. B* **92**, 085106 (2015).  
[4] M. Aichhorn, E. Arrigoni, M. Potthoff, and W. Hanke, *Phys. Rev. B* **74**, 024508 (2006).  
[5] G. Baym and L. P. Kadanoff, *Phys. Rev.* **124**, 287 (1961).  
[6] G. Baym, *Phys. Rev.* **127**, 1391 (1962).  
[7] H. Hafermann, E. G. C. P. van Loon, M. I. Katsnelson, A. I. Lichtenstein, and O. Parcollet, *Phys. Rev. B* **90**, 235105 (2014).  
[8] N. Mermin and H. Wagner, *Phys. Rev. Lett.* **17**, 1133 (1966).  
[9] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).  
[10] T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, *Rev. Mod. Phys.* **77**, 1027 (2005).  
[11] M. Potthoff, *Condens. Matter Phys.* **9**, 557 (2006).  
[12] A. Toschi, A. A. Katanin, and K. Held, *Phys. Rev. B* **75**, 045118 (2007).  
[13] G. Rohringer and A. Toschi, *Phys. Rev. B* **94**, 125144 (2016).  
[14] A. N. Rubtsov, M. I. Katsnelson, and A. I. Lichtenstein, *Phys. Rev. B* **77**, 033101 (2008).  
[15] G. Rohringer, A. Toschi, H. Hafermann, K. Held, V. I. Anisimov, and A. A. Katanin, *Phys. Rev. B* **88**, 115112 (2013).  
[16] T. Ayrál and O. Parcollet, *Phys. Rev. B* **92**, 115109 (2015).  
[17] T. Ayrál and O. Parcollet, *Phys. Rev. B* **93**, 235124 (2016).  
[18] C. Taranto, S. Andergassen, J. Bauer, K. Held, A. Katanin, W. Metzner, G. Rohringer, and A. Toschi, *Phys. Rev. Lett.* **112**, 196402 (2014).  
[19] A. N. Rubtsov, M. I. Katsnelson, and A. I. Lichtenstein, *Ann. Phys.* **327**, 1320 (2012).  
[20] N. E. Bickers, D. J. Scalapino, and S. R. White, *Phys. Rev. Lett.* **62**, 961 (1989).  
[21] Y.M. Vil'k and A.-M.S. Tremblay, *J. Phys. I (France)* **7**, 1309 (1997).  
[22] G. Rohringer, H. Hafermann, A. Toschi, A. A. Katanin, A. E. Antipov, M. I. Katsnelson, A. I. Lichtenstein, A. N. Rubtsov, and K. Held, [arXiv:1705.00024](https://arxiv.org/abs/1705.00024).  
[23] J. Gukelberger, E. Kozik, and H. Hafermann, *Phys. Rev. B* **96**, 035152 (2017).  
[24] S. Isakov, A. E. Antipov, and E. Gull, *Phys. Rev. B* **94**, 035102 (2016).  
[25] A. M. Sengupta and A. Georges, *Phys. Rev. B* **52**, 10295 (1995).  
[26] J. Otsuki and Y. Kuramoto, *Phys. Rev. B* **88**, 024427 (2013).  
[27] S. Sachdev and J. Ye, *Phys. Rev. Lett.* **70**, 3339 (1993).  
[28] Q. Si and J. L. Smith, *Phys. Rev. Lett.* **77**, 3391 (1996).  
[29] H. Kajueter, Ph.D. thesis, Rutgers University, 1996.  
[30] J. L. Smith and Q. Si, *Phys. Rev. B* **61**, 5184 (2000).  
[31] R. Chitra and G. Kotliar, *Phys. Rev. B* **63**, 115110 (2001).  
[32] P. Sun and G. Kotliar, *Phys. Rev. B* **66**, 085120 (2002).  
[33] T. Ayrál, P. Werner, and S. Biermann, *Phys. Rev. Lett.* **109**, 226401 (2012).  
[34] E. G. C. P. van Loon, A. I. Lichtenstein, M. I. Katsnelson, O. Parcollet, and H. Hafermann, *Phys. Rev. B* **90**, 235135 (2014).  
[35] E. G. C. P. van Loon, H. Hafermann, A. I. Lichtenstein, A. N. Rubtsov, and M. I. Katsnelson, *Phys. Rev. Lett.* **113**, 246407 (2014).  
[36] E. A. Stepanov, E. G. C. P. van Loon, A. A. Katanin, A. I. Lichtenstein, M. I. Katsnelson, and A. N. Rubtsov, *Phys. Rev. B* **93**, 045107 (2016).  
[37] E. G. C. P. van Loon, F. Krien, H. Hafermann, E. A. Stepanov, A. I. Lichtenstein, and M. I. Katsnelson, *Phys. Rev. B* **93**, 155162 (2016).  
[38] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, *Rev. Mod. Phys.* **83**, 349 (2011).  
[39] K. Haule, *Phys. Rev. B* **75**, 155113 (2007).  
[40] V. Galitskii and A. Migdal, *Zh. Eksp. Teor. Fiz.* **34**, 139 (1958) [*JETP* **7**, 96 (1958)].  
[41] V. Janiš, A. Kauch, and V. Pokorný, *Phys. Rev. B* **95**, 045108 (2017).  
[42] T. Ayrál and O. Parcollet, *Phys. Rev. B* **94**, 075159 (2016).  
[43] E. G. C. P. van Loon, M. Schüler, M. I. Katsnelson, and T. O. Wehling, *Phys. Rev. B* **94**, 165141 (2016).  
[44] J. Otsuki, *Phys. Rev. B* **87**, 125102 (2013).  
[45] H. Rostami, M. I. Katsnelson, G. Vignale, and M. Polini (unpublished).  
[46] A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, *Phys. Rev. B* **72**, 035122 (2005).  
[47] P. Werner and A. J. Millis, *Phys. Rev. Lett.* **99**, 146404 (2007).  
[48] P. Werner and A. J. Millis, *Phys. Rev. Lett.* **104**, 146401 (2010).  
[49] H. Hafermann, P. Werner, and E. Gull, *Comput. Phys. Commun.* **184**, 1280 (2013).  
[50] B. Bauer, L. D. Carr, H. G. Evertz, A. Feiguin, J. Freire, S. Fuchs, L. Gamper, J. Gukelberger, E. Gull, S. Guertler *et al.*, *J. Stat. Mech.* (2011) P05001.  
[51] H. Hafermann, K. R. Patton, and P. Werner, *Phys. Rev. B* **85**, 205106 (2012).  
[52] H. Hafermann, *Phys. Rev. B* **89**, 235128 (2014).  
[53] E. M. Lifshitz and L. P. Pitaevskii, *Statistical Physics: Theory of the Condensed State: 009 (Course of Theoretical Physics Vol. 9)* (Butterworth-Heinemann, Oxford, 2013).  
[54] J. R. Schrieffer, *Theory of Superconductivity*, Advanced Book Program (Avalon Publishing, New York, 1983).  
[55] D. Vollhardt and P. Wölfle, *Phys. Rev. B* **22**, 4666 (1980).



- [56] M. Revzen, T. Toyoda, Y. Takahashi, and F. C. Khanna, [Phys. Rev. B \*\*40\*\*, 769 \(1989\)](#).
- [57] A.-M. Daré, L. Chen, and A.-M. S. Tremblay, [Phys. Rev. B \*\*49\*\*, 4106 \(1994\)](#).
- [58] T. Toyoda and M. Okada, [Phys. Rev. B \*\*58\*\*, 1210 \(1998\)](#).
- [59] V. Janiš, [Phys. Rev. B \*\*64\*\*, 115115 \(2001\)](#).
- [60] A. Masumizu and K. Sogo, [Phys. Rev. B \*\*72\*\*, 115107 \(2005\)](#).
- [61] X. G. Wen, F. Wilczek, and A. Zee, [Phys. Rev. B \*\*39\*\*, 11413 \(1989\)](#).
- [62] M. H. Hettler, M. Mukherjee, M. Jarrell, and H. R. Krishnamurthy, [Phys. Rev. B \*\*61\*\*, 12739 \(2000\)](#).