# Dynamical screening effects in correlated materials: Plasmon satellites and spectral weight transfers from a Green's function ansatz to extended dynamical mean field theory

Michele Casula,[1,2,3] Alexey Rubtsov,[4] and Silke Biermann[2,3]

[1]*CNRS and Institut de Minéralogie et de Physique des Milieux condensés, Université Pierre et Marie Curie, case 115, 4 place Jussieu, FR-75252, Paris Cedex 05, France*

[2]*Centre de Physique Théorique, Ecole Polytechnique, CNRS, FR-91128 Palaiseau Cedex, France*

[3]*Japan Science and Technology Agency, CREST, Kawaguchi 332-0012, Japan*

[4]*Department of Physics, Moscow State University, 119992 Moscow, Russia*

(Received 15 July 2011; published 18 January 2012)

Dynamical screening of the Coulomb interactions in correlated electron systems results in a low-energy effective problem with a dynamical Hubbard interaction $\mathcal{U}(\omega)$. We propose a Green's function ansatz for the Anderson impurity problem with retarded interactions, in which the Green's function factorizes into a contribution stemming from an effective static-$U$ problem and a bosonic high-energy part introducing collective plasmon excitations. Our approach relies on the scale separation of the low-energy properties, related to the instantaneous static $U$, from the intermediate- to high-energy features originating from the retarded part of the interaction. We argue that for correlated materials where retarded interactions arise from downfolding higher-energy degrees of freedom, the characteristic frequencies are typically in the antiadiabatic regime. In this case, accurate approximations to the bosonic factor are relatively easy to construct, with the most simple being the boson factor of the dynamical atomic-limit problem. We benchmark the quality of our method against numerically exact continuous-time quantum Monte Carlo results for the Anderson-Holstein model both at half- and quarter-filling. Furthermore, we study the Mott transition within the Hubbard-Holstein model within extended dynamical mean field theory. Finally, we apply our technique to a realistic three-band Hamiltonian for $SrVO_3$. We show that our approach reproduces both the effective mass renormalization and the position of the lower Hubbard band by means of a dynamically screened $U$, previously determined *ab initio* within the constrained random phase approximation. Our approach could also be used within schemes beyond dynamical mean field theory, opening a quite general way of describing satellites and plasmon excitations in correlated materials.

DOI: 10.1103/PhysRevB.85.035115 PACS number(s): 71.27.+a, 71.30.+h, 71.45.Gm, 71.10.−w

## I. INTRODUCTION

Over the last years, significant progress has been made in the modelization of strongly correlated materials. Such systems typically contain partially filled $d$ or $f$ orbitals,[1,2] which lie relatively close to the nuclei. Electronic Coulomb interactions can then induce substantial corrections to a Bloch one-particle picture, ranging from renormalization of effective parameters in the sense of Landau to full localization of the $d$ or $f$ degrees of freedom in the Mott insulator.

One of the difficulties in describing such correlation effects is to separate the usually rather small energy range of the correlated $d$ and $f$ orbitals from the larger energy scale of the itinerant degrees of freedom, e.g., the $p$ orbitals of ligand atoms, but also of higher- or lower-lying states of the transition-metal, rare-earth, or actinide atoms themselves. The former orbitals are mainly responsible for the low-energy physical properties of the compounds, while the latter act as a screening medium, setting in particular the actual value of the Coulomb repulsion of the correlated degrees of freedom.

Screening is a dynamical process, where the efficiency of the screening medium in reducing the effects of a perturbation depends on the energy scales involved.[3] The effective Coulomb interaction that enters a low-energy description of a solid is thus in general a frequency-dependent quantity. In particular, the local interaction matrix element ("Hubbard $U$") becomes a frequency dependent $U = U(\omega)$. In a realistic approach to strongly correlated materials, $U$ can be determined thanks to the constrained random phase approximation (cRPA),[4] once the bands and their eigenstates are computed by an *ab initio* calculation. The static value $U_0 = U(0)$, evaluated in this way,[5,6] has been recently used in the dynamical mean field theory (DMFT) calculations of materials (see, e.g., Ref. 7). The DMFT approach, combined with the density functional theory (DFT), is an extremely powerful tool to treat *ab initio* strongly correlated systems, once the low-energy model is determined.[8] However, very little is known on the impact of the frequency dependence of the interaction in the low-energy part of the spectrum. The hardest obstacle in order to include the dynamic $U$ into the DMFT framework has been the lack of reliable solver techniques for the quantum impurity problem with frequency-dependent Hubbard interactions.

The main difficulties come from two sources: the unscreened value of the Coulomb interaction for $3d$ electrons, e.g., is typically as large as several tens of eV, ruling out the application of traditional weak-coupling expansion methods. Furthermore, realistic $U(\omega)$ can be parametrized by bosonic (plasmon or particle-hole excitation) modes, but in general a representation in terms of multiple modes is required. This prevents the direct use of methods developed in the context of the Hubbard-Holstein Hamiltonian,[9–16] as many bosonic baths coupled to the fermion degrees of freedom will be necessary to fully resolve $U(\omega)$. Recently, this problem has been overcome by a continuous-time quantum Monte Carlo (CTQMC) solver proposed by Werner and Millis,[17,18] where a multiplasmon Lang-Firsov transformation[19] is treated exactly in the context of a hybridization expansion algorithm for the DMFT impurity Hamiltonian.[20,21] Also, the weak-coupling

 

CTQMC algorithm[22,23] by Rubtsov can treat generic retarded interactions, but it is limited to a not-so-large dynamic $U$ and not-so-large screening frequencies, and therefore it becomes prohibitively costly for realistic applications of dynamic screening interactions in a multiorbital context.

Another major problem still left (even if reliable Monte Carlo data are available) is the possibility of computing spectral properties, such as high-energy plasmon satellites. Indeed, the presence of screening modes leads usually to a quite complicated spectrum with a series of peaks located at high energies (at multiples of the plasma frequencies). Those features are hard to get by the usual maximum entropy (ME) methods[24,25] used to invert the noisy QMC data in the imaginary-time domain into the spectral properties at real frequencies. The ME methods are quite reliable at low energy, but usually are not capable to deal with high-frequency features.

In this paper, we present a DMFT approach based on a Bose factor ansatz (BFA) for the Green's function, which is able to handle a generic $U(\omega)$ interaction in a strong-coupling antiadiabatic regime, a typical situation in strongly correlated materials, and provides a robust and general way to compute the full spectrum of the frequency dependent (retarded) $U$, with an accuracy capable to resolve the high-energy satellites. Our method is based on the separation between the energy scales set by the screened value $U_0$ and treated using well-established solvers[17,22,26] and the dynamic part treated with various levels of approximation, the simplest and most insightful one taken from the dynamic atomic limit.

The paper is organized as follows. In Sec. II, we present the motivations for using the dynamically screened interactions in *ab initio* calculations. In Sec. III, we introduce the models used to parametrize the dynamic interactions in many-body realistic Hamiltonians. In Sec. IV, we describe the Green's function ansatz used in our method and its spectral properties. In Sec. V, we present the dynamic atomic limit approximation (DALA) to solve the Hubbard model with dynamically screened $U$. In Sec. VI, we show various ways to improve upon the DALA, and in Sec. VII we describe their performances. In Secs. VIII and IX, we report our results for a single-band lattice model at half-filling and a three-band model with the DFT density of state (DOS) of $SrVO_3$, respectively. Finally, Sec. X summarizes our findings.

## II. MOTIVATION: DYNAMICALLY SCREENED HUBBARD INTERACTIONS

### A. Dynamically screened interactions in lattice models for correlated fermions

At the time when Hubbard designed the model that carries nowadays his name,[27] the mere possibility of the *parameters* involved in this model becoming one day quantities that could be calculated from a microscopic theory would not even have been envisioned. Even though it was clear that the hopping integrals between different atomic lattice sites correspond to matrix elements of the coupling part of the Hamiltonian in appropriately localized Wannier functions in a tight-binding sense, it took several years, and the advent of modern electronic structure techniques[28] before calculations that extract this information explicitly from band structure[29] became standard tools of condensed matter physics. Determining the local interaction parameters that translate the Coulomb cost of creating a double occupancy on a given lattice site, the "Hubbard $U$", turned out to be an even harder issue. The "constrained local density approximation" (CLDA) tries to assess this quantity by calculating total energy differences for systems with modified particle numbers within DFT in the local density approximation. This route has met a certain success, making CLDA the most commonly used method for this purpose nowadays, even though ambiguities concerning the screening channels that are included persist. Indeed, it is clear on physical grounds that the Hubbard interactions depend on the choice of the one-particle part of the Hamiltonian, not only through the localization of the orbitals, but also through the number and nature of states included in the low-energy description. An elegant way to address this question is the "constrained random phase approximation" (cRPA), which calculates the Hubbard interactions as partially screened Coulomb interactions, that is, screening processes that are included in the low-energy model are suppressed in the calculation of $U$. The idea behind this scheme is in fact more general than its (now commonly used) name suggests, and is conceptually not restricted to a random phase approximation treatment. It rather relies on the fact that the electronic polarization $P$ of a solid can be divided into two parts, a part that is explicitly included in the low-energy model to be treated $P_d$ and the one stemming from the remaining screening channels $P_r = P - P_d$. In this framework, the bare interaction of the low-energy model $W_r$ can be identified to be the bare Coulomb interaction $v(r - r') = \frac{1}{|r-r'|}$ screened by $P_r$:

$$W_r = \frac{v}{1 - P_r v}. \tag{1}$$

This identification is motivated by the fact that further screening of $W_r$ by the low-energy screening channels, expressed by the polarization $P_d$, leads to the fully screened Coulomb interactions (the $W$ of Hedin's theory[30]) of the original full-scale model of the solid. The matrix elements of $W_r$ between localized orbitals used for the description of the low-energy effective model can therefore be identified as the Hubbard $U$ parameters of the model. While this construction directly translates the dependence of $U$ on the choice of the low-energy model in a conceptually satisfying manner, it reveals a further basic feature: screening is a dynamical process since the efficiency of a screening medium in responding to a perturbation depends on the energy scale of this perturbation, as translated by the frequency dependence of the polarization function or the corresponding dielectric response functions. The recognition of the Hubbard $U$ as a partially screened interaction thus leads in a natural way to a generalization of the original Hubbard model to a model with frequency-dependent Hubbard $U$. The natural formulation is thus no longer in the form of a static Hamiltonian, but rather in the framework of an action formalism, even though the additional frequency dependence of the Hubbard interaction can be parametrized in terms of bosonic degrees of freedom (particle-hole excitations and plasmon modes) opening the way to formulating the problem in terms of a Hubbard-Holstein model, where electronic degrees of freedom are coupled to bosonic modes. As alluded to above, it is the

goal of this paper to introduce a simple but accurate way to deal with such a frequency dependence in the context of many-body calculations. We address this problem in the dynamical mean field approximation, in which the lattice problem is mapped onto an auxiliary local problem of a single atomic site embedded in a bath. The bath is determined by a self-consistency condition imposing the local part of the lattice Green's function to coincide with that of the auxiliary problem. In the present framework, where the lattice model itself is dynamical (through the interaction), this amounts to solving a dynamical local impurity model, as in extended dynamical mean field theory,[31–35] but with a fixed dynamical interaction. A mathematical formulation of this strategy is given in Sec. III, before we introduce our dynamical impurity solver scheme in Sec. IV.

### B. Converting nonlocality into frequency dependence: The "*GW* + DMFT view" on dynamical impurity models

In the discussion above, the dynamical nature of the interaction was supposed to stem only from the downfolding of higher-energy degrees of freedom. It can then be directly assessed by a cRPA construction of a dynamical lattice model. The construction of the impurity model serves in this case merely as a tool to solve the dynamical lattice problem, with a fixed local dynamical interaction that is assumed to be the local part of the cRPA one. It is, however, possible to adopt a more general point of view, which also incorporates the contribution of nonlocal interactions and nonlocal screening effects in a solid, giving rise to an additional frequency dependence of the effective local interaction. The Hubbard interaction $\mathcal{U}$ in this case should no longer be interpreted as the local part of the physical Coulomb interaction of a downfolded model, but as an effective quantity that incorporates both the effects of screening by downfolding and by representing a lattice model by an impurity. This perspective goes beyond the cRPA view of the problem, but can be consistently formulated within extended DMFT.[31–35] It is also the starting point for the combined "$GW$ + DMFT" scheme,[36] which yields a prescription on how to calculate both the one-particle part of the Hamiltonian and the Coulomb interactions of a correlated material from first principles. The idea behind this method is to calculate the nonlocal parts of the self-energy and polarization to first order in the screened Coulomb interaction $W$ as in Hedin's $GW$ approximation,[30] and to combine them with the local part of these quantities as calculated from a dynamical impurity model. One thus *represents* two physical quantities, namely, the local Green's function $G$ of the solid and the local part of the fully screened Coulomb interaction $W$ by a local model, defined by some effective Weiss field $\mathcal{G}_0$ and the auxiliary Coulomb interaction $\mathcal{U}(\omega)$. The latter is constructed such that the solution of the impurity model yields the local part of $W$. This is akin in spirit to other theories in solid state physics, where a physical quantity is represented by the self-consistent solution of an effective auxiliary model, famous examples being density functional theory or DMFT itself. In DFT, the physical density of a system is represented by an auxiliary system in an effective one-particle (Kohn-Sham) potential; in DMFT, the local lattice Green's function is constructed from an impurity model with an effective Weiss field or, equivalently, a local self-energy. The auxiliary quantities such as the Kohn-Sham potential of DFT or the impurity self-energy acquire the role of Lagrange multipliers fixing the density (in DFT) or the local Green's function (in DMFT) to their physical values. In extended DMFT, a nonlocal interaction in the original lattice model gives rise to a *dynamical* impurity model representing the physical quantities of the model.

The question of how to solve dynamical impurity models in an efficient and accurate way is thus a crucial one, motivated by the dynamical nature of the effective local Coulomb interaction due to both downfolding of higher-energy degrees of freedom and nonlocal interactions. In the following, we will restrict the discussion to the case of Sec. II A, in which a dynamical lattice model is solved within a local approximation, keeping in mind, however, the more general usefulness of the designed approximations in the context of the $GW$ + DMFT scheme.

## III. MODELS

### A. From dynamical Coulomb interactions to the Hubbard-Holstein model

Putting the arguments of Sec. II A about the dynamical nature of the partially screened Coulomb interactions into mathematical language, we replace the general many-body Hamiltonian by a low-energy effective description. This amounts to constructing the effective action for the low-energy degrees of freedom $\Psi$:

$$S = \int d\mathbf{r}\, \Psi^\dagger(\mathbf{r})\,(\partial\tau - \mu + H_0)\,\Psi(\mathbf{r}) + \int d\mathbf{r} \int d\mathbf{r}'\, W_r(r,r',\tau-\tau')\Psi^\dagger(\mathbf{r})\Psi(\mathbf{r})\Psi^\dagger(\mathbf{r}')\Psi(\mathbf{r}'), \tag{2}$$

where $\mathbf{r} = (r,\tau)$ is a generalized space imaginary-time coordinate, $\mu$ is the chemical potential, and the "bare interaction" of the low-energy model is now a dynamical quantity, as a consequence of the downfolding.

Using a standard procedure for expressing the continuous model in a basis of localized Wannier orbitals $\phi_{RL}(r)$ (with $R$ the atomic site and $L$ some orbital index) leads then to the action of a Hubbard model with dynamical interaction matrix elements $\mathcal{U}_{L_1,L_2,L_3,L_4}(\tau-\tau')$. Such an interaction term can be parametrized by a set of bosonic degrees of freedom. In the following, we write out this relationship explicitly for the two cases that we consider in this paper: the one-orbital model and a multiorbital model with local density-density interactions.

#### 1. Single-band model

Let us take into account the single-band Hubbard-Holstein Hamiltonian $H = H_{\text{kin}} + H_{\text{plasmons}} + H_{\text{int}}$, defined as

$$H_{\text{kin}} = \sum_{\mathbf{k},\sigma} [\epsilon(\mathbf{k}) - \mu] d^\dagger_{\mathbf{k}\sigma} d_{\mathbf{k}\sigma},$$

$$H_{\text{plasmons}} = \sum_{i,\nu} \omega_\nu b^\dagger_{i\nu} b_{i\nu} + \sum_{i,\nu} \lambda_\nu (b^\dagger_{i\nu} + b_{i\nu}) \sum_\sigma n_{i\sigma},$$

$$H_{\text{int}} = \sum_i U_\infty n_{i\uparrow} n_{i\downarrow},$$

where $d^\dagger_\sigma$ ($d_\sigma$) is a creation (annihilation) operator for an electron with spin $\sigma$, $\epsilon(\mathbf{k})$ is the band in $\mathbf{k}$ space with chemical potential $\mu$, $b^\dagger_\nu$ ($b_\nu$) is a creation (annihilation) operator for a bosonic mode $\nu$ of frequency $\omega_\nu$, which couples locally to the total local electronic density $\sum_\sigma n_{i\sigma} \equiv \sum_\sigma d^\dagger_{i\sigma} d_{i\sigma}$ on the $i$th site. The Hamiltonian in Eq. (3) can be written in the action (finite-temperature) formalism as follows:

$$\begin{aligned}\mathcal{S} = &- \sum_{ij\sigma} \int_0^\beta d\tau \int_0^\beta d\tau'\, d^\dagger_{i\sigma}(\tau)\mathcal{G}^{-1}_{ij\sigma}(\tau-\tau')d_{j\sigma}(\tau') \\ &+ \sum_i \Bigg( \int_0^\beta d\tau\, n_{i\uparrow}(\tau)U_\infty n_{i\downarrow}(\tau) \\ &+ \int_0^\beta d\tau \int_0^\beta d\tau'\, n_{i\uparrow}(\tau)U_{\rm ret}(\tau-\tau')n_{i\downarrow}(\tau')\Bigg), \qquad (3)\end{aligned}$$

where $\beta$ is the inverse temperature, $i$ and $j$ are lattice sites, and $\mathcal{G}_{ij\sigma}$ is such that $\mathcal{G}^{-1}_\sigma(i\omega,\mathbf{k}) = i\omega - \epsilon(\mathbf{k}) + \mu$. The last interacting term is the screening part, which is always negative, retarded (nonlocal in time), and comes from the integration of the bosonic operators. The full time-dependent interaction is therefore $U(\tau) = U_\infty\delta(\tau) + U_{\rm ret}(\tau)$, and the unscreened instantaneous $U_\infty$ is reduced by the regular part. It turns out that the retarded interaction reads as

$$U_{\rm ret}(\tau) = -\sum_\nu \lambda^2_\nu \cosh[(\tau-\beta/2)\omega_\nu]/\sinh[\omega_\nu\beta/2], \qquad (4)$$

a relation that can be obtained by using a Hubbard-Stratonovich transformation in the bosonic fields.[10,21] Therefore, the generic screened interaction $U(\tau)$ can be represented by a set of local (in space) bosonic operators $b_\nu$, couplings $\lambda_\nu$, and frequencies $\omega_\nu$. In the case of a single mode $\nu = 0$ with frequency $\omega_0$ and coupling $\lambda_0$, Fourier transforming Eq. (4) will lead to the well-known Matsubara representation of $U(\tau)$ for the single-boson (either plasmon or phonon) Hubbard-Holstein model, namely, $U(i\nu_n) = U_\infty - 2\lambda_0^2\omega_0/(\nu_n^2+\omega_0^2)$, with $\nu_n = 2n\pi/\beta$. We define the fully screened interaction as the static limit of $U(i\nu)$. Therefore, in the latter case, the fully screened interaction is $U_0 = U_\infty - 2\lambda_0^2/\omega_0$, while in the most general case, it reads as

$$U_0 = U_\infty - 2\sum_\nu \lambda^2_\nu/\omega_\nu. \qquad (5)$$

### 2. *Three-band model*

Now, let us generalize the single-band case presented above to a multiband model. Here, we take explicitly into account a three-band model with $t_{2g}$ symmetry valid for $SrVO_3$.[37] The Hamiltonian is again constituted by three parts, $H = H_{\rm kin} + H_{\rm plasmons} + H_{\rm int}$, defined analogously to the one-band case as

$$\begin{aligned}H_{\rm kin} &= \sum_{\mathbf{k},m,\sigma} [\epsilon_m(\mathbf{k}) - \mu] d^\dagger_{\mathbf{k}m\sigma} d_{\mathbf{k}m\sigma}, \\ H_{\rm plasmons} &= \sum_{i,\nu} \omega_\nu b^\dagger_{i\nu} b_{i\nu} + \sum_{i,\nu} \lambda_\nu (b^\dagger_{i\nu} + b_{i\nu}) \sum_{m,\sigma} n_{im\sigma}, \qquad (6) \\ H_{\rm int} &= \frac{1}{2}\sum_i \Bigg( U_\infty \sum_{m,\sigma} n_{im\sigma} n_{im-\sigma} + (U_\infty - 2J) \sum_{\substack{m,m',\sigma \\ m\neq m'}} n_{im\sigma} n_{im'-\sigma} + (U_\infty - 3J) \sum_{\substack{m,m',\sigma \\ m\neq m'}} n_{im\sigma} n_{im'\sigma} \Bigg),\end{aligned}$$

where $n_{m\sigma}$=$d^\dagger_{m\sigma} d_{m\sigma}$ is the usual density operator, with $m$,$\sigma$ denoting orbital and spin indexes, and $d^\dagger_{m\sigma}$, $d_{m\sigma}$ representing the $t_{2g}$ localized orbitals. In Eq. (6), we have introduced the Hund's coupling $J$, which takes into account spin effects in an effective density-density interaction, and fulfills the $t_{2g}$ symmetry. As in the one-band model, the bosonic modes couple to the total charge on a site, with the orbital indexes summed over. Therefore, by going to the finte-temperature action formalism, the bosonic degrees of freedom can be integrated out again, thanks to the same Hubbard-Stratonovich transformation used for the single-band model. The resulting effective contribution screens the bare $U_\infty$, which is the density-density total coupling, while $J$, which acts selectively on the spin and orbital character of the electrons, is left unscreened by the multiboson Holstein bath. The resulting action for the $d$ electrons is the following:

$$\begin{aligned}\mathcal{S} = &- \sum_{ijmm'\sigma} \int_0^\beta d\tau \int_0^\beta d\tau'\, d^\dagger_{imm'\sigma}(\tau)\mathcal{G}^{-1}_{ijm\sigma}(\tau-\tau')d_{jm'\sigma}(\tau') \\ &+ \frac{1}{2}\sum_i \int_0^\beta d\tau \Bigg( -2J \sum_{\substack{m,m',\sigma \\ m\neq m'}} n(\tau)_{im\sigma} n(\tau)_{im'-\sigma} - 3J \sum_{\substack{m,m',\sigma \\ m\neq m'}} n(\tau)_{im\sigma} n(\tau)_{im'\sigma} + \int_0^\beta d\tau' \sum_{\substack{m,m',\sigma,\sigma' \\ (m\sigma)\neq(m'\sigma')}} n(\tau)_{im\sigma} U(\tau-\tau') n(\tau')_{im'\sigma'} \Bigg), \qquad (7)\end{aligned}$$

where $U(\tau) = U_\infty\delta(\tau) + U_{\rm ret}(\tau)$, as in the one-band model, with $U_{\rm ret}(\tau)$ given by Eq. (4).

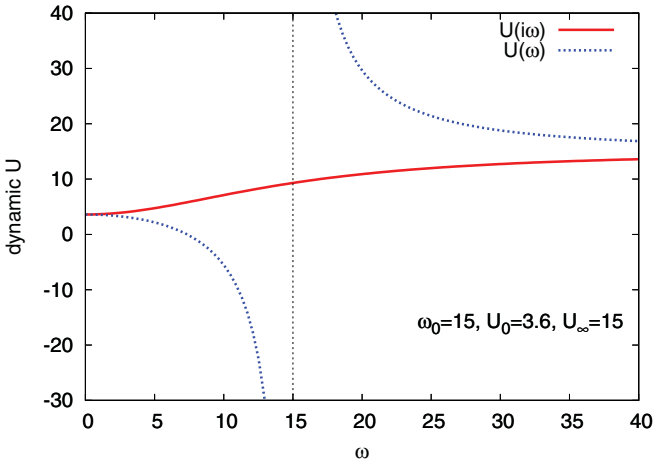


FIG. 1. (Color online) Dynamic $U$ for a single plasmon mode with $\omega_0 = 15$, $U_0 = 3.6$, and $U_\infty = 15$. This is the single plasmon modelization of the dynamic $U$ for $SrVO_3$, determined *ab initio* by constrained RPA (Ref. 5). The function $U(z)$ is plotted on both the real ($z = \omega$) and imaginary ($z = i\omega$) frequency domains. The vertical asymptote is shown at $\omega = \omega_0$ for $U(\omega)$, where the function has a single pole. In the most general case, there will be as many poles as plasmon modes.

### B. Dynamical Hubbard interactions: General features

Thanks to the cRPA, it has become possible to explicitly compute the frequency-dependent effective interaction $U(\omega)$ for specific materials.[5,6,38] The resulting functions share some common features: The unscreened $U_\infty$ value is in general about an order of magnitude larger than the screened $U_0$, and $U(i\omega)$ varies, e.g., from roughly 3 to 30 eV for nickel, or from 4 to 15 eV in $SrVO_3$. In oxides,[5–7] a jump is generally observed at a quite well-defined plasma frequency $\omega_0$, that is, the frequency-dependent interaction is well parametrized by a single bosonic mode coupling to the electrons. The plasma frequency is typically much larger than the bandwidth ($\omega_0$ is of the order of 15 eV in $SrVO_3$ for example), thus placing the system in the antiadiabatic regime. In transition metals or the recentpnictide systems,a more complex frequency dependenceof $U(\omega)$ requires parametrization by a continuumof bosonic modes. Nevertheless,the energy scale atwhich the frequency-dependent interaction finallygoes up to the unscreened value is typicallyof the order of 20 to 30 eV.A general $U(\omega)$can be resolved into many plasmon contributions,whichcharacterize the screening process. More explicitly,the coupling $\lambda_\omega$ thatcouples the mode of frequency $\omega$ to the electronicsystem can be related to the spectral representation ofthe interaction by$\lambda_\omega = -\frac{1}{\pi}\mathrm{Im}U_{\mathrm{ret}}(\omega)$. This quantity corresponds to the $\alpha^2 F$ of electron-phonon models.

In Fig. 1, we illustrate a typical form of the frequency-dependent interaction used in this work, on both the real and imaginary frequency axes.

### C. From lattice to the impurity models: Dynamical mean field approximation

In order to solve the Hubbard-Holstein models in Eqs. (7) and (3), we are using the dynamical mean field theory (DMFT), which maps the full lattice problem into a single-site Anderson impurity problem coupled to an effective bath. The bath is determined self-consistently by requiring that the impurity Green's function equals the onsite Green's function on the lattice.[1] Therefore, computing in the most efficient way the Green's function of the Anderson model (AM) is of key importance to have a feasible DMFT scheme. Here, as we have seen in Sec. II, we have the additional complication that the effect of screening makes the onsite Hubbard interaction retarded.

In the action formalism, the Anderson model reads as in Eqs. (3) and (7), except that the site dependence is replaced by an effective local (but dynamic) hybridization term $\Delta$, which is included in $\mathcal{G}$, namely,

$$\mathcal{S}_{\mathrm{AM}} = -\sum_{mm'\sigma}\iint d\tau\, d\tau' d^\dagger_{m\sigma}(\tau)\mathcal{G}^{-1}_{mm'\sigma}(\tau-\tau')d_{m'\sigma}(\tau') + \frac{1}{2}\sum_{\substack{m,m',\sigma,\sigma' \\ (m\sigma)\neq(m'\sigma')}}\iint d\tau d\tau' n_{m\sigma}(\tau)V^{\sigma\sigma'}_{mm'}(\tau-\tau')n_{m'\sigma'}(\tau'). \tag{8}$$

In the DMFT scheme, $\mathcal{G}^{-1}_{mm'\sigma}(i\omega) = i\omega + \mu - \Delta_{mm'\sigma}(i\omega)$, where $\Delta$ has to be determined self-consistently. The Hamiltonian representation of the action in Eq. (8) with $\mathcal{G}^{-1}_{mm'\sigma}$ and $V$ (which includes both $J$ and dynamic $U$) is nothing but the Anderson-Holstein model, which this work is primarily concerned with.

### D. Atomic limit

The atomic limit is reached when the hybridization $\Delta_{mm'\sigma}(i\omega) = 0$, and each single site is independent of the rest, acting as an isolated atom. The importance of the atomic limit is given by the fact that usually high-energy excitations are well described in that limit. Therefore, it is a test for the quality of local approximations in the high-energy regime. In this section, we show that an exact solution is known for the atomic limit with retarded interactions, as proved in Ref. 39 for a single-orbital case. Here, we sketch that proof for a generic multiorbital model, where a retarded interaction $U(\tau)$ couples to the total local charge $N(\tau) = \sum_{m,\sigma} n_{m\sigma}(\tau)$. In this case, the action is

$$\mathcal{S} = \int_0^\beta d\tau \sum_{m\sigma} d^\dagger_{m\sigma}(\tau)(\partial_\tau - \mu)d_{m\sigma}(\tau) + \frac{1}{2}\int_0^\beta d\tau \int_0^\beta d\tau'\, N(\tau)U(\tau-\tau')N(\tau'), \tag{9}$$

while the Green's function is defined as $G_{m\sigma}(\tau) = \langle \mathcal{T} d_{m\sigma}(\tau) d^\dagger_{m\sigma}(0)\rangle$, where the time-ordered thermal average is weighted by the action in Eq. (9). Now, we introduce a scalar Hubbard-Stratonovich field $\phi(\tau)$ with Fourier components $\phi_n$, which decouples the interaction term, and leads to the following action:

$$\mathcal{S} = \int_0^\beta d\tau \sum_{m\sigma} d^\dagger_{m\sigma}(\tau)[\partial_\tau - \mu + i\phi(\tau)]d_{m\sigma}(\tau) + \frac{1}{2}\int_0^\beta d\tau \int_0^\beta d\tau'\, \phi(\tau)U^{-1}(\tau-\tau')\phi(\tau'), \tag{10}$$

where we have to integrate both over the $d_{m\sigma}$ and $\phi_n$ components to compute the thermal average. It is now possible to carry out analytically both functional integrations (see Ref. 39 for details), and the calculation yields for the Green's function

$$G_{m\sigma}(\tau) = \mathcal{F}_{m\sigma}(\tau,\mu,U_0)\exp\left(\frac{1}{\beta}\sum_{n\neq 0}\frac{U_n}{\nu_n^2}(e^{i\nu_n\tau}-1)\right), \tag{11}$$

where $U_n = U(i\nu_n)$ with $\nu_n = 2n\pi/\beta$ bosonic frequencies, and $\mathcal{F}_{m\sigma}$ is a complex function we can leave undetermined here, for the sake of readability. For a static (instantaneous) interaction, $U_n = U_{\rm static}\forall n$ and, therefore, Eq. (11) implies that

$$G_{m\sigma}(\tau) = G^{\rm static}_{m\sigma}(\tau,\mu,U_{\rm static}) \times \exp\left(\frac{1}{\beta}\sum_{n\neq 0}\frac{U_n - U_{\rm static}}{\nu_n^2}(e^{i\nu_n\tau}-1)\right), \tag{12}$$

where $G^{\rm static}_{m\sigma}(\tau)$ is the usual atomic-limit Green's function of static models, multiplied by an extra factor, which takes into account the dynamic nature of the interaction. This result holds for an arbitrary filling and number of orbitals, their dependence being included only in $G^{\rm static}_{m\sigma}(\tau)$.

Equation (12) reveals an important physical outcome reported in this paper. In the atomic limit, the exact solution of the dynamic Anderson model is the factorization between a static Green's function (depending only on the instantaneous interaction) and a bosonic factor, which is a functional of the retarded (frequency dependent) $U$. The rest of the paper is devoted to the extension of this result and its consequences to the most general case, with $\Delta \neq 0$.

Equation (12) raises also the question about which instantaneous part of $U$ one has to take explicitly. There are two possible physical choices. Indeed, either the explicit instantaneous part is the unscreened $U_\infty$, in which case $U(\tau) = U_\infty\delta(\tau) + U_{\rm ret}(\tau)$, as written in Sec. III A, or the explicit instantaneous part is the fully screened $U_0$, and $U(\tau) = U_0\delta(\tau) + \bar{U}(\tau)$. Both choices are equivalent, in the sense that the resulting $U(\tau)$ must be the same. The former is usually done in dealing with electron-phonon models, as $U_{\rm ret}$ has a natural interpretation in terms of the $\alpha^2F$. Here and thereafter, we are going to take the latter convention, as in the context of plasmon screening it is more natural to deal with the fully screened interaction $U_0$ [the static limit of $U(i\nu)$, which sets the low-energy behavior] as the instantaneous part, while the dynamic part $\bar{U}(\tau)$ is such that $\int_0^\beta d\tau\,\bar{U}(\tau) = 0$, and links the screened interaction to the bare $U_\infty$. Therefore, in our conventions, $U_{\rm static} = U_0$ and $G^{\rm static} = G_0$, i.e., the Green's function with instantaneous interaction $U_0$.

## IV. GREEN's FUNCTION BOSE FACTOR ANSATZ (BFA)

### A. Bose factor ansatz

We are mainly interested in evaluating the Green's function $G(\tau) = \langle \mathcal{T} d(\tau)d^\dagger(0)\rangle$ and its spectral properties for the model in Eq. (8). In analogy with the atomic limit of Eq. (12), we are going to rewrite it in the form

$$G(\tau) = G_0(\tau)F(\tau), \tag{13}$$

where we dropped all orbital and spin indexes to simplify the notation. As reported in Sec. III D, $G_0(\tau)$ is the Green's function for the model in Eq. (8), with instantaneous onsite repulsion, namely, $U(\tau) = U_0\delta(\tau)$, with $U_0$ the static limit of $U(i\nu)$. We highlight that the above factorization is defined in the *time domain*, a feature that is borrowed from the dynamic atomic solution of the problem, the form of which is known analytically, as explained in Sec. V. In that limit, the Green's function assumes exactly the form in Eq. (13), with $G_0$ the instantaneous $U_0$ atomic Green's function. The static model is much easier than the dynamic one since it contains only the energy scales set by the screened $U(\ll U_\infty)$ and the Kondo resonance with the bath, and it can be solved by means of various techniques,[17,22,26] which are usually very robust and efficient in this case. On the other hand, $F(\tau)$ is a Bose factor, which is a functional of $\bar{U}(\tau)$, and it is not known *a priori*. However, we will present various approximations where the function $F(\tau)$ is derived. It contains the information of the plasmon (or phonons) excitations and the plasmon (or phonons) satellites.

### B. Spectral properties

A great advantage of dealing with the Green's function ansatz in Eq. (13) is the possibility to compute very accurate spectral functions over the whole energy range, including the intermediate- to high-energy plasmon satellites. Indeed, since the Bose factor $F(\tau)$ can be estimated analytically by means of some approximation, its numerical value is known at machine precision, and its spectral function $B(\omega)$ can be obtained via a Padé approximant[40] in an accurate and robust way. On the other hand, a ME approach has to be used to find the spectral function $A_0(\omega)$ of the static Green's function $G_0(\omega)$. However, this does not pose any particular problem since there are no high-energy features in $A_0(\omega)$, and its energy range is set by $U_0$, where the ME is reliable in the presence of data with good statistics.[25] The spectral function $A(\omega)$ of the full Green's function $G(\tau)$ expressed as a functional of $B$ and $A_0$ reads as

$$A(\omega) = \int_{-\infty}^{\infty} d\epsilon\, B(\epsilon)\frac{1+e^{-\beta\omega}}{(1+e^{\beta(\epsilon-\omega)})(1-e^{-\beta\epsilon})}A_0(\omega-\epsilon). \tag{14}$$

The spectral functions obtained in this way are reported for instance in Fig. 3, which corresponds to the Green's functions plotted in Fig. 2. The quality of the satellite resolution is striking, much higher than that usually obtained with ME methods,[25] particularly at energies far away from the Fermi level.

We would like to stress that the spectral convolution in Eq. (14) is general and can be used not only for the approximated Green's functions we are going to derive in Secs. V and VI. For instance, our approach to compute the spectral properties can be applied to the Green's function obtained by means of the algorithm in Ref. 21. Given the full Green's function $G(\tau)$ of the dynamic impurity problem of Eq. (8), one defines an auxiliary Green's function $G_{\rm aux}(\tau)$ as $G(\tau)/F_{\rm DALA}(\tau)$, with $F_{\rm DALA}$ taken from the atomic limit as described in Sec. III D. This is an effective way to exploit the separation of the low-energy properties, kept in $G_{\rm aux}$, from the high-frequency features correctly reproduced by $F_{\rm DALA}$. At this point, one computes

$A_0$, the spectral representation of $G_{\rm aux}$, by using ME and $B$, the spectrum of $F_{\rm DALA}$, by means of the Padé approximant, and evaluates the full spectral function in Eq. (14).

## V. DYNAMIC ATOMIC-LIMIT APPROXIMATION

In order to find a way to determine $F(\tau)$ in Eq. (13), we invert it and get

$$F(\tau) = \frac{G(\tau)}{G_0(\tau)} \approx \left(\frac{G(\tau)}{G_0(\tau)}\right)\Big|_{\Delta=0}, \tag{15}$$

where the rightmost-hand side of the above equation for $F(\tau)$ is taken in the dynamic atomic-limit approximation (DALA), when the hybridization $\Delta$ is zero. As we have seen in Sec. III D, the Green's function in the DALA is analytically solvable by means of a Hubbard-Stratonovich transformation,[39] and therefore $G/G_0$ is exactly known in a close analytic form, such that

$$F_{\rm DALA}(\tau) = \exp\left(\frac{1}{\beta}\sum_{n\neq 0}\frac{U(i\nu_n)-U_0}{\nu_n^2}(e^{i\nu_n\tau}-1)\right), \tag{16}$$

where $\nu_n = 2n\pi/\beta$ are bosonic Matsubara frequencies, with $n$ relative integer.

Aside from the atomic limit, this approximation is exact in the static and the noninteracting limits [in both cases $F_{\rm DALA}(\tau) = 1$]. Notice that it retains all the nonperturbative character of $G_0(\tau)$. To have a better idea on the quality of this approximation, we are going to test it for the dynamic $U$ with a single plasmon mode $\omega_0$, which is equivalent to the Anderson-Holstein model with $U(i\nu) = U_\infty - 2\lambda^2\omega_0/(\nu^2+\omega_0^2)$ and the electron-"phonon" coupling given by $\lambda = \sqrt{(U_\infty - U_0)\omega_0/2}$. We use the CTQMC algorithm by Rubtsov,[22,23] which can handle retarded interactions and yield the exact Green's function in a weak-coupling regime, to benchmark our approximation for the particle-hole symmetric system with $U_0 = 2$, $\beta = 10$, and few values of $\omega_0$ and $U_\infty$. The energy units are expressed in terms of the half bandwidth ($D/2 = 1$) of the semicircular DOS.

As one can see from Fig. 2, the DALA works very well for the cases analyzed since it gives a $G_{\rm DALA}(\tau)$, which almost coincides with the numerically exact $G(\tau)$ given by the CTQMC algorithm. The accuracy is particularly impressive in the case with $\beta = 10$, $U_0 = 2$, $U_\infty = 6$, and $\omega_0 = 5$, where the impact of the dynamic part is reduced by the larger $\omega_0 (\gg U_0)$, and a smaller $U_\infty$, namely, when the the energy scales of the static part set by $U_0$ are well separated from the dynamic contributions in $U(i\nu)$. Moreover, from Fig. 2 it is apparent that the low-energy properties of the system are strongly renormalized by the effect of the high-energy components of $U$. This is quite insightful on the importance of the dynamic screening effects in the treatment of more realistic models that we are going to tackle in Sec. VIII. This can be noted also from the spectral properties reported in Fig. 3, which correspond to the Green's functions $G_{\rm DALA}(\tau)$ plotted in Fig. 2. A spectral weight transfer from the low-frequency spectrum to the high-energy satellites is clearly visible in that figure.

Since the DALA is obtained in the $\Delta = 0$ limit, it works well in the intermediate- to strong-coupling regime, with $U_0$ and the dynamic part large, as we have seen in the cases

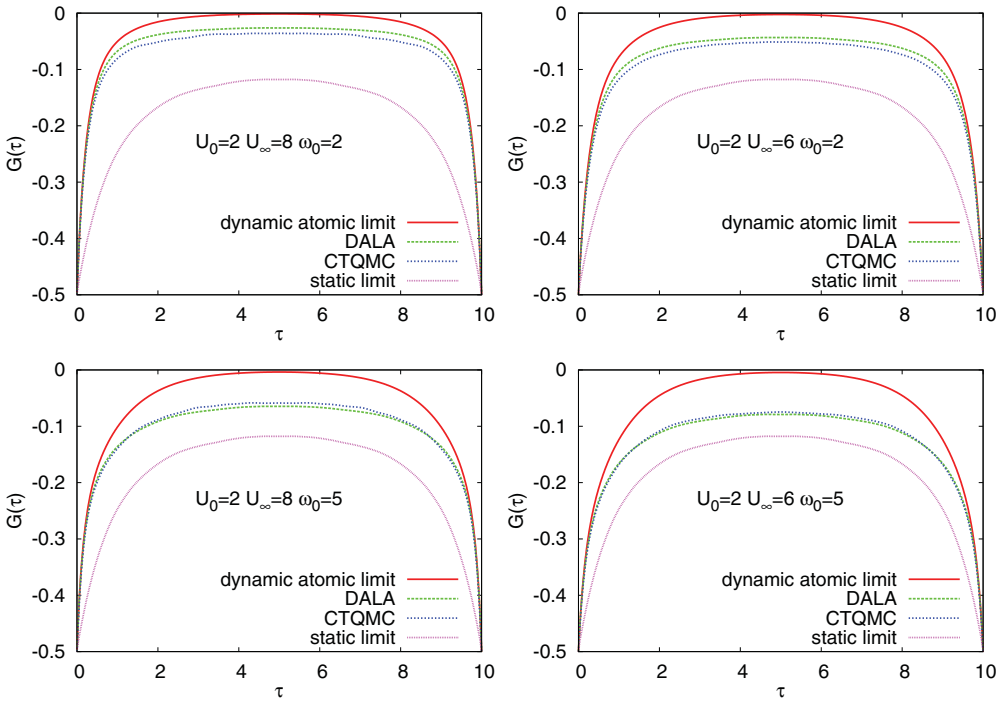


FIG. 2. (Color online) Green's functions for a half-filled Anderson model with dynamically screened $U$ and a semicircular density of states at $\beta = 10$. The $G_{\rm DALA}$ obtained by the method proposed in Sec. V is plotted (green long-dashed line) and compared to the exact Rubtsov's CTQMC numerical result (blue dotted line). Also, the dynamic atomic limit (red solid line) and the static one (pink dotted-dashed line) are reported. The system is computed at a quite large temperature such that an accurate benchmark against the numerically exact CTQMC is still possible, even for quite large values of $U$.

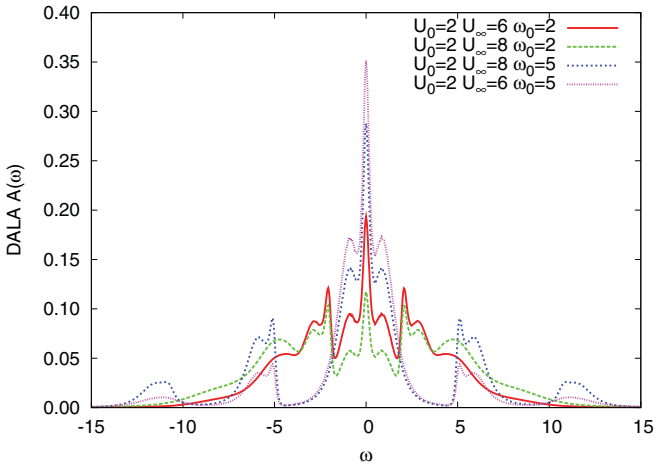


FIG. 3. (Color online) Spectral function of the $G_{\mathrm{DALA}}$ Green's functions reported in Fig. 2.

analyzed in Fig. 2, where $U_0$ was quite close to the critical $U_{c2}$ ($\approx$2.6, see Ref. 41) for the Mott transition of the static Hubbard model. However, it deteriorates as $U_0$ is getting smaller and the hybridization $\Delta$ becomes important to set the low-energy properties of the system. To show this, let us take into account the Anderson-Holstein model with $U_0 = 2, \omega_0 = 5$, and $U_\infty = 6$ at $\beta = 10$, for which the DALA gives a result very close to the exact one. Now, let us keep $\omega_0$ and $U_\infty - U_0$ fixed, such that the DALA Bose factor [Eq. (16)] is unchanged, while we vary $U_0$ from strongly to more weakly correlated values. The resulting Green's functions are plotted in Fig. 4. It is apparent that the DALA accuracy reduces as $U_0$ decreases, and the resulting bias is more pronounced in the low-energy part of the Green's function. Indeed, the discrepancy is larger around $\beta/2$ in $G(\tau)$ [Fig. 4(a)], which corresponds to a larger difference at small Matsubara frequencies in $G(i\omega_n)$ [Fig. 4(b)]. On the other hand, the high-energy tails of $G(\tau)$ are very well reproduced by the DALA, as it is confirmed also by the inspection of $G(i\omega)$ at large $\omega_n(>4)$.

The correct high-energy asymptotics of the DALA is a non-trivial property of this approximation, which is borrowed from the atomic limit exactly built in. In order to further analyze this important feature, we take into account the temperature dependence of the DALA in the symmetric Anderson model with a dynamic interaction given by $U_0 = 1.25$, $\omega_0 = 2$, and $U_\infty = 2.5$. In this not-so-correlated case, the model can be solved exactly down to low temperatures ($\beta = 160$), even by the CTQMC algorithm, to benchmark the temperature dependence of our approximation. Results are plotted in Fig. 5. As we already found in the previous analysis, at large Matsubara frequencies, the dependence of the Green's function is correctly given by the DALA, which in this case becomes almost indistinguishable from the exact CTQMC result for $\omega_n > 7$. We note that the DALA is capable to reproduce the decay of the imaginary part of $G(i\omega)$ well beyond the $1/i\omega$ term, as it is apparent from Fig. 5(b). It is also worth noting that the relative accuracy of the approximation increases with the temperature, as it is shown in Fig. 5(d). Indeed, the exact Bose factor $F(\tau)$ is getting closer to $F_{\mathrm{DALA}}(\tau)$ as the temperature increases. At low temperatures, it is the $F(\beta/2)$ value that is poorly reproduced by the DALA. Again, this is related to the

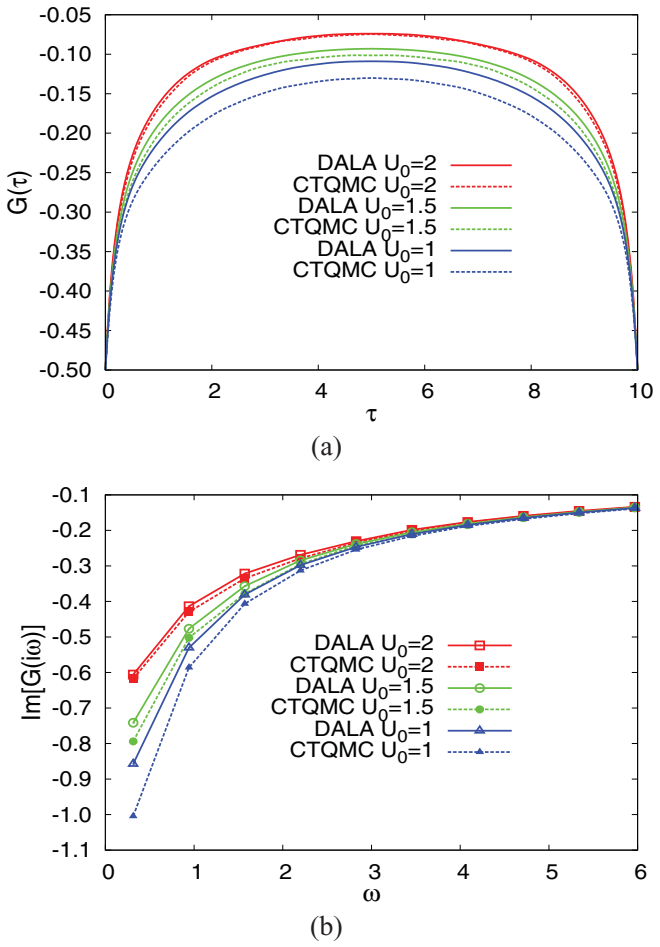


FIG. 4. (Color online) Half-filled Anderson impurity model with retarded screened interaction: $\omega_0 = 5$, $U_\infty - U_0 = 5$, $\beta = 10$, for various values of $U_0$. (a) Numerically exact Rubtsov's CTQMC and DALA $G(\tau)$; (b) CTQMC and DALA imaginary part of $G(i\omega_n)$.

roughness of the approximation at low frequency, which does not describe accurately the low-energy excitations around and below the coherent temperature. Indeed, the Friedel sum rule, fulfilled by Fermi liquids with local self-energies, is clearly violated, as one can see in Fig. 5(a), where the condition $G''(i0^+) = -4/D$ valid at half-filling and pinned by the value of the density of states at the Fermi level is not met by the DALA.

Therefore, going beyond the DALA is needed to capture the low-energy low-temperature features of the spectral function, while its high-energy properties, as the plasmon satellites, can be successfully taken into account at this level of approximation.

## VI. BEYOND THE DYNAMIC ATOMIC LIMIT

### A. DALA Lang-Firsov approximation

A way to improve the dynamic atomic limit approximation in both the intermediate- and low-energy regimes is provided by the Lang-Firsov (LF) approach. This approximation has been widely used in the literature to tackle electron-phonon models in the antiadiabatic limit, when the electron-phonon

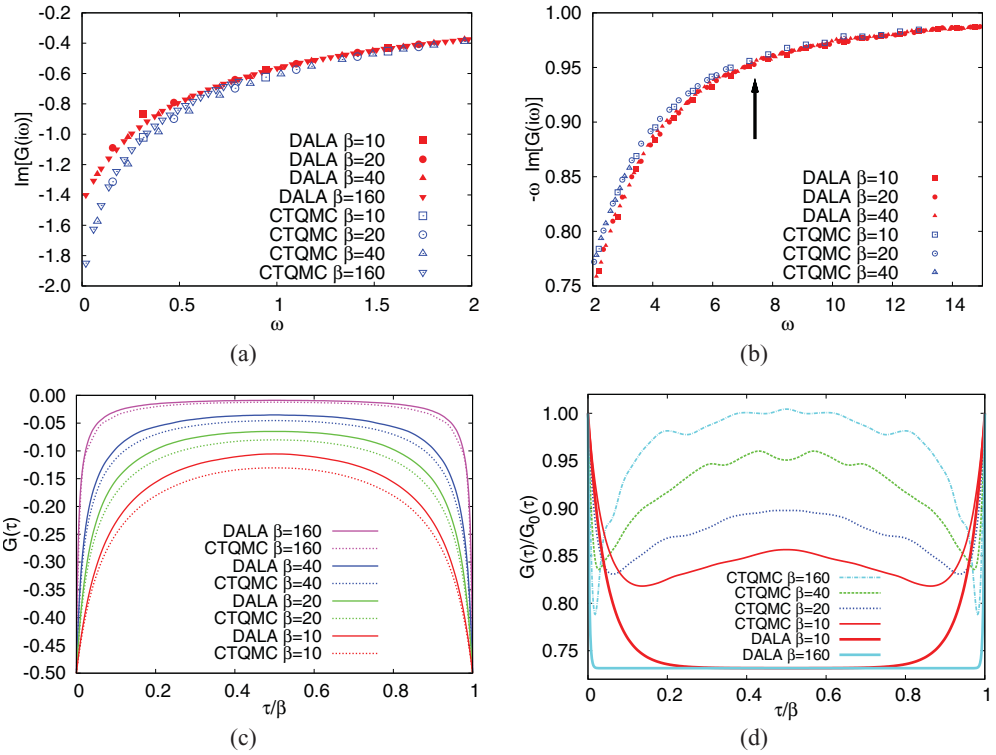


FIG. 5. (Color online) Symmetric Anderson model with retarded screened interaction: $U_0 = 1.25, \omega_0 = 2, U_\infty = 2.5$, at different $\beta$. (a) Imaginary part of the numerically exact Rubtsov's CTQMC and DALA Green's functions for small Matsubara frequencies $\omega_n$; (b) $-\omega_n \mathrm{Im}[G(i\omega_n)]$ in the intermediate frequency range. The arrow indicates the frequency when the DALA and the CTQMC Green's functions become practically indistinguishable. The DALA reproduces correctly the intermediate- to high-energy behavior of the Green's function well beyond the "trivial" $1/i\omega$ term; (c) $G(\tau)$; (d) the CTQMC and DALA Bose factor $F(\tau)$. The thicker lines are for the DALA. The discrepancy between the DALA and the exact factors is getting smaller as the temperature increases. The wiggles are due to the stochastic noise of the data.

coupling is $\lambda \ll \omega_0$, with $\omega_0$ the phonon frequency. The same applies to models where the interaction is retarded by screening plasmons. In the latter case, the antiadiabatic regime is met more often since the plasmon frequencies are larger than the phonon ones, usually by an order of magnitude. However, it should be noted that in the realistic retarded Hubbard $U$, also the electron-plasmon coupling is stronger, as $\lambda = \sqrt{(U_\infty - U_0)\omega_0/2}$, and $U_\infty$ is an order of magnitude larger than $U_0$. In any case, if $U_\infty - U_0 < \omega_0$, the Lang-Firsov approach describes well the low-energy properties of the system. Here, we use the Lang-Firsov approximation as a low-frequency correction to the dynamic factor of our Green's function ansatz.

The factorization in the $\tau$ space implies a convolution in the $\omega_n$ space

$$G(i\omega_n) = \frac{1}{\beta} \sum_m G_0(i\omega_m) F(i\omega_n - i\omega_m), \tag{17}$$

where $F(i\nu_n)$ are the Matsubara components of the Bose factor. A way to improve upon the DALA is to choose the $F$ factor such that the Lang-Firsov behavior is obtained at low frequency. This can be done by introducing an enhanced $F$ defined as

$$F_{\mathrm{DALA+LF}}(i\nu_n) = \begin{cases} a & \text{if } \nu_n = 0, \\ \left[1 - b \exp(-\frac{\nu_n}{c})\right] F_{\mathrm{DALA}}(i\nu_n) & \text{elsewhere,} \end{cases} \tag{18}$$

where $a$, $b$, and $c$ are parameters determined by the following conditions: $F(0) = F(\beta) = 1$, which keeps the correct number of particles provided by $G_{\mathrm{static}}$; $G(i\omega_l) = G_{\mathrm{LF}}(i\omega_l)$ for $l = 0$ (the first Matsubara frequency), which gives the correct Friedel sum rule fulfilled by the Lang-Firsov approximation (and broken by the DALA); $c$ is the crossover frequency between the LF behavior at low energy and the atomic DALA behavior at high frequency, and its optimal value is $\approx \omega_0/10$, with $\omega_0$ the lowest Holstein frequency.

The LF Green's function is given by the usual expression

$$G_{\mathrm{LF}}(i\omega) = \frac{\exp\left(-\frac{\lambda^2}{\omega_0^2}\right)}{i\omega + \mu - \lambda^2/\omega_0 - \exp\left(-\frac{\lambda^2}{\omega_0^2}\right)\Delta(i\omega) - \Sigma[U_0](i\omega)}, \tag{19}$$

which clearly implies that an Anderson model with static Hubbard $U_0$ must be solved, hybridized via a renormalized bath $\exp(-\lambda^2/\omega_0^2)\Delta(i\omega)$. Therefore, to get the DALA + LF Green's function, one has to solve two static models (with regular and renormalized bath) and mix them together by using our definition in Eq. (18).

One should note that the Lang-Firsov Green's function in Eq. (19) has been written for a single plasmon (or phonon) Anderson-Holstein model, which gives a rough representation of the dynamically screened $U(i\nu)$ present in *ab initio* models. For instance, the cRPA approximation for $U(i\nu)$ leads in some

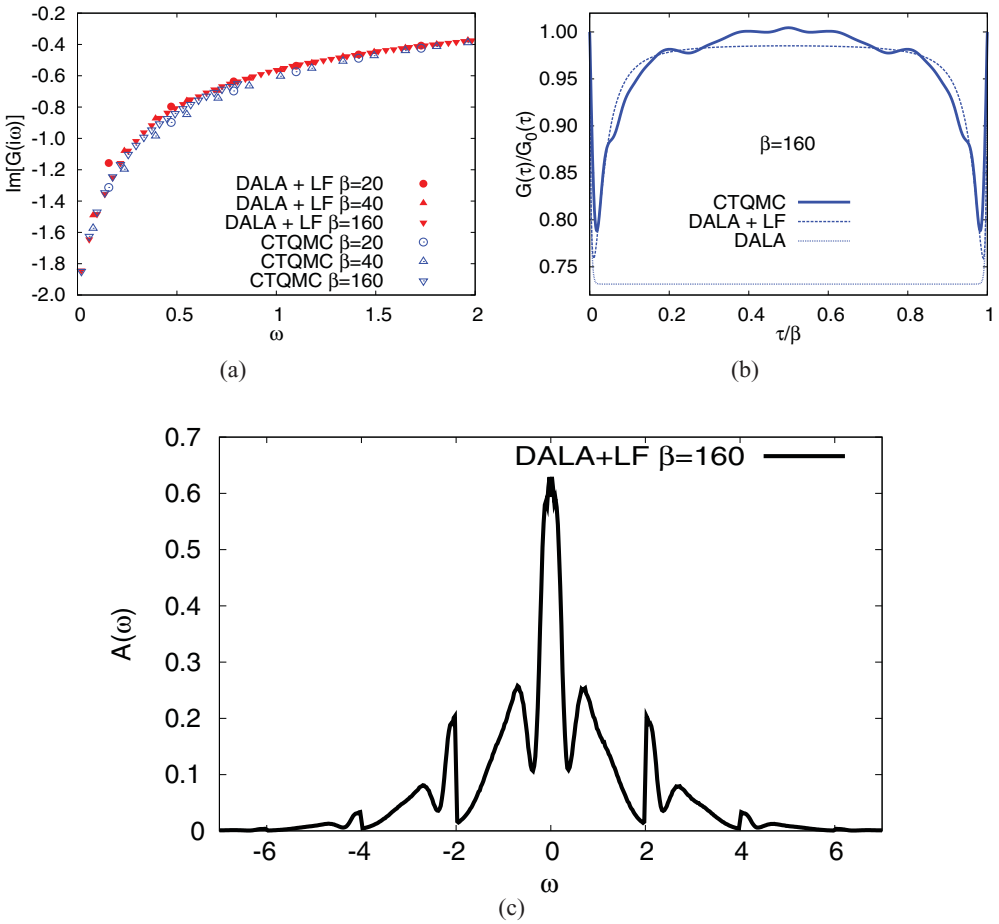


FIG. 6. (Color online) Symmetric Anderson model with retarded screened interaction as in Fig. 5: $U_0 = 1.25, \omega_0 = 2, U_\infty = 2.5$. (a) Imaginary part of the numerically exact Rubtsov's CTQMC and DALA + LF Green's functions for small Matsubara frequencies $\omega_n$ at different $\beta$. (b) The CTQMC, DALA, and DALA + LF Bose factor $F(\tau)$ at $\beta = 160$. The improvement of the $F(\tau)$ provided by the Lang-Firsov correction is apparent. The wiggles are due to the interpolation of noisy QMC data. (c) Spectral representation of the DALA + LF Green's function in (a).

cases to a quite broad spectrum of screening plasmons, which is difficult or impossible to fit accurately by a single frequency model, as stated in Sec. III B. However, a generalization of Eq. (19) can be easily done by following the same lines as in Ref. 21. If the LF renormalization factor $\exp(-\lambda^2/\omega_0^2)$ is replaced by

$$\exp\left(\sum_\nu \frac{1}{\pi}\frac{\mathrm{Im}\bar{U}(\omega_\nu)}{\omega_\nu^2}\right), \tag{20}$$

where $\bar{U}(\omega) = U(\omega) - U_0$, and $\omega_\nu$ are real frequencies, all the screening plasmons $\nu$ are treated on the same footing. Although the LF approximation is accurate only for large $\omega$, Eq. (20) is a good approximation if $\mathrm{Im}\bar{U}(\omega)/\omega^2$ goes rapidly to zero for small $\omega$. For the Holstein single-mode interaction, $\mathrm{Im}\bar{U}(\omega) = -\pi\lambda^2[\delta(\omega-\omega_0) - \delta(\omega+\omega_0)]$, and one recovers the standard LF renormalization factor.

The improvement of the DALA + LF correction with respect to the simpler DALA factor is apparent in Fig. 6(a), to be compared with Fig. 5(a). In particular, at low frequency, the LF is capable to recover the pinning condition [$G''(i0^+) = -2$ for the semicircular density of states at half-filling], violated by the DALA at low temperatures. This is reflected in the behavior of the Bose factor $F_{\mathrm{DALA+LF}}$ drawn in Fig. 6(b), the value at $\beta/2$ of which goes correctly to 1 for $\beta \to \infty$, in contrast to the $F_{\mathrm{DALA}}$, which does not change with temperature. Indeed, the exact $F$ factor shows a quite strong temperature dependence, as reported in Fig. 5(d). In Fig. 6(c), we plot the spectral representation of the DALA + LF Green's function at $\beta = 160$ obtained by the DALA factor assisted analytic continuation described in Sec. IV B.

### B. Diagrammatic first-order expansion: Gaussian cumulants (GC)

Another way to improve upon the DALA is to rely on the cumulant expansion of the interaction. This can be done in various ways. Here, we took two routes: making a cumulant Green's function expansion in the full $U(i\nu)$ and its "alternating" part $\bar{U}(i\nu) = U(i\nu) - U_0$. In the former approach, dubbed "Gaussian cumulants" (GC), the perturbation theory is built on the Gaussian action containing the hybridized $\mathcal{G}_0$ as the bare Green's function. In the latter method, named "instantaneous bold cumulants" (IBC), the perturbation expansion requires the calculation of density-density correlators within

the static $U_0$ model by means of the CTQMC algorithm (or other algorithms suitable for static interacting models).[1]

Let us first consider how a perturbation $\delta U$ affects the Green's function in general. Taylor expansion in perturbation up to the first order gives

$$G(\tau) = G_0(\tau) + \langle \delta U d_\tau d_0^\dagger \rangle_0 - \langle \delta U \rangle_0 G_0(\tau), \quad (21)$$

where $G_0$ is a Green's function of the unperturbed system, and $\langle \cdots \rangle_0$ is an average over this system. Note that so far the unperturbed system is not supposed to be Gaussian, and Wick's theorem is not employed to obtain the averages in Eq. (21).

In the spirit of the DALA, we would like to draw an expression of the type $G(\tau) = G_0(\tau) \exp[\mathcal{F}(\tau;[\delta U])]$. The advantage of using an exponential form for $\mathcal{F}(\tau;[\delta U])$ (or in other words a cumulant expansion) is that the exact dynamic atomic limit can be obtained by just the first two Taylor terms. Rewriting Eq. (21) as an exponent, one obtains (again, up to the first order)

$$G(\tau) = G_0(\tau) \exp(\langle \delta U d_\tau d_0^\dagger \rangle_0 G_0^{-1}(\tau) - \langle \delta U \rangle_0). \quad (22)$$

Now, let us turn into the specific case of retarded density-density interactions. An explicit form $\delta U = \frac{1}{2}\int\int U(\tau - \tau')N(\tau)N(\tau')d\tau\, d\tau'$, with $N(t) = \sum_\sigma n_\sigma(t)$ being the spin (and orbital) integrated density, leads to the expression

$$G(\tau) = G_0(\tau) \exp\left(-\frac{1}{2}\int\int_0^\beta dt\, dt' \chi(t,t',\tau) U(t-t')\right), \quad (23)$$

where

$$\chi(t,t',\tau) = \frac{\langle \mathcal{T} d_\tau d_0^\dagger N(t) N(t') \rangle_0}{G_0(\tau)} - \langle \mathcal{T} N(t) N(t') \rangle. \quad (24)$$

We note that the second term in Eq. (24) is $\tau$ independent and so contributes to an overall prefactor in Eq. (23), which can be determined from the normalization condition, as both the perturbed $G$ and unperturbed $G_0$ are normalized. Therefore, the second term in Eq. (24) does not need to be calculated explicitly.

If the perturbation theory is built with respect to the full $U(\tau)$ (GC approximation), the unperturbed system is noninteracting, and so $G_0^{-1}(i\omega) = \mathcal{G}_0^{-1}(i\omega) = i\omega - \mu - \Delta(i\omega)$. This implies that Wick's theorem is valid for calculating $\langle \ldots \rangle_0$, and one can check that only connected parts enter in the first term of Eq. (24). Therefore, by reminding that $F \equiv G_U(\tau)/G_{U_0}(\tau)$, the resulting expression for the Bose factor $F$ in the GC approximation is

$$F_{\rm GC}(\tau) = F_0 \exp\left(-\frac{1}{2}\int\int_0^\beta dt\, dt' \chi_{\rm GC}(t,t',\tau) \bar{U}(t-t')\right), \quad (25)$$

with $\bar{U}(t) = U(t) - U_0\delta(t)$, $F_0$ the normalization factor, and $\chi_{\rm GC}$ given by the following equation:

$$\chi_{\rm GC}(t,t',\tau) = \frac{\mathcal{G}_0(t)\mathcal{G}_0(t'-t)\mathcal{G}_0(\tau - t) + \{t \leftrightarrow t'\}}{\mathcal{G}_0(\tau)}. \quad (26)$$

The right-hand side of the above equation is conveniently computed in the frequency domain, where the singularities in $\mathcal{G}_0$ are properly taken into account. Since we integrate over $t$ and $t'$ in (23) and $U(t)$ is an even function, the two terms in the nominator give identical contributions. The normalization factor $F_0$ is obtained from the condition $F_{\rm GC}(0) = F_{\rm GC}(\beta) = 1$.

The expression in Eq. (25) for the dynamic factor has the advantage to be very accurate at both small $U$ since the interaction is treated perturbatively, and strong coupling, as the density-density correlator, will factor up by giving the exact dynamic atomic limit (see the Appendix). As already highlighted in the derivation, this is the reason why the cumulant expansion is more effective than the regular perturbation theory for this case.

### C. Diagrammatic first-order expansion: Instantaneous bold cumulants (IBC)

One can go beyond the approximation for $F$ in Eq. (25) and compute the density-density correlator $\chi$ by using the interacting instantaneous $U_0$ Green's function $G_0(\tau)$ as propagator instead of the hybridized Green's function $\mathcal{G}_0(\tau)$. This requires a correlated method, as the corresponding action is no longer Gaussian and the Wick theorem can not be applied. The QMC methods can compute $\chi$ directly in the interacting systems.[1] This development represents a consistent diagrammatic first-order expansion in $\bar{U}(i\nu)$ where the reference theory is the static $U_0$ model. The main difference with respect to the previous method reported in Sec. VI B stems from the fact that not only hybridization effects, but also the impact of the instantaneous $U_0\delta(t-t')$ interaction, are included in $\chi$.

This approximation gives the following Bose factor $F$:

$$F_{\rm IBC}(\tau) = F_0 \exp\left(-\frac{1}{2}\int\int_0^\beta dt\, dt' \chi_{\rm IBC}(t,t',\tau) \bar{U}(t-t')\right), \quad (27)$$

where $F_0$ is the normalization constant, and $\chi_{\rm IBC}$ is given by

$$\chi_{\rm IBC}(t,t',\tau) = \frac{\langle \mathcal{T} d_\tau d_0^\dagger N(t) N(t') \rangle_0}{G_0(\tau)}. \quad (28)$$

In the following, we report a comparison on the performance of the different approaches proposed in this work. As a common benchmark, we chose to compute the Green's function of the Anderson-Holstein impurity problem both in the particle-hole symmetric and asymmetric cases, and check the BFA solutions against the numerically exact one provided by Rubtsov's CTQMC algorithm.

## VII. OVERVIEW ON THE PERFORMANCE OF THE PROPOSED FACTORIZATION APPROXIMATIONS

The factorization introduced in Eq. (13) is an extremely useful Green's function ansatz to compute thermal and spectral properties coming from a generic retarded interaction in the multiband Anderson model and also the Hubbard model at the DMFT level. The product in the time domain between a Bose factor $F$ embedding the dynamic properties of the interaction and an auxiliary Green's function $G_0$ including instantaneous interactions and low-energy features leads to a deconvolution of the spectrum into low- and high-frequency contributions. Its low-frequency part, depending on $G_0$, can be easily obtained by available and well-developed ME methods, while the high-energy features, as the plasmon satellites, difficult or impossible to obtain by standard analytical continuation, are

directly given by the analytically known Bose factor $F$, which can be inverted in an accurate way by the Padé approximants.

The Green's function Bose factor is a very general ansatz. However, the factor $F$ is not exactly known in the generic case and needs some approximations. The most practical and physically insightful one is borrowed from the dynamic atomic limit, where the factorization is exactly given by a $G_0$, depending only on the instantaneous $U_0$ times a factor that depends only on the retarded part $\bar{U}(i\nu)$. The DALA proposed in Sec. V consists of keeping the atomic dynamic factor $F$ and taking the $G_0$ from the exact numerical solution of the instantaneous $U_0$ model. As reported in Figs. 2 and 7(b), the DALA performs well in the antiadiabatic regime when $\omega_0 > U_\infty - U_0$, and in the strong-coupling regime with a large $U_0$. The quality of the low-energy part of the DALA Green's function worsens as the interaction becomes weaker and $\omega_0$ gets smaller, while its high-energy tails are well reproduced even in the intermediate coupling. The major failure of the DALA is the breaking of the Friedel sum rule, which is apparent at low temperature (below $\beta = 40$) or away from the antiadiabatic regime. That is not surprising since the DALA is built upon the atomic limit. In general, the DALA works when the high-energy (unscreened) part of the retarded interaction is well separated (in frequency) from the low-energy (screened) part, which is the most common situation in the case of realistic Hamiltonians, with $U(i\nu)$ determined *ab initio* by the cRPA approach.[4]

An improvement upon the DALA is represented by the Lang-Firsov approximation, which fulfills the Friedel sum rule and cures the low-energy low-temperature behavior. The way to incorporate the LF into the DALA factor, described in Sec. VI A, clearly improves the low-energy features of the DALA Green's function, as one can see in Figs. 7 and 8, where the value of $G(\beta/2)$ yielded by the DALA + LF approach is closer to the exact numerical result in both the half-filling and quarter-filling Anderson-Holstein impurity models. The DALA + LF factor provides an overall better agreement with the numerically exact Green's function computed by the CTQMC method. The LF correction turns out to be important particularly in the weak and intermediate couplings and at low temperature, while in other cases, the simple DALA approximation is already good enough. For instance, in Fig. 7(b), when $U_0$ is strong and the temperature not so low ($\beta = 10$), the DALA Green's function coincides with the DALA + LF one, and both are on the top of the exact numerical solution. A common limitation of the DALA and DALA + LF approximations is that they become inaccurate when $\omega_0 < U_\infty - U_0$, while they work very well in the antiadiabatic regime.

A different route to determine $F$ is built on diagrammatic expansion techniques. We have proven that the cumulant expansion of the Green's function up to the first order in the interaction provides the exact connection to the dynamic atomic limit (see Appendix). Indeed, the $F$ factor computed in this way [see Eq. (23)] has the exact atomic behavior in the zero hybridization limit. We propose two flavors for the cumulant expansion. In both cases, one has to evaluate the density-density correlator $\langle \mathcal{T} c_\tau c_0^\dagger N(t) N(t') \rangle_0$. In the first one (the Gaussian cumulants or "GC"), the thermal average is computed with a Gaussian action based on the hybridized noninteracting Green's function, while in the other one (the instantaneous bold cumulants or "IBC"), the brackets are computed with the instantaneous $U_0$ Green's function. The GC works well in the weak interacting regime (small $U_0$) and it worsens as $U_0$ is getting larger, as shown in Fig. 7, as expected for a perturbative expansion. Away from half-filling, the GC Green's function is reliable in the "empty" part of the spectrum (for $\tau < \beta/2$) and where the dynamic part is more relevant in setting the tails of the Green's function. On the other hand, the IBC performs well everywhere in all cases when $\omega_0$ is large, as seen in Figs. 7 and 8, and it is supposed to work quite well even away from the antiadiabatic regime since it retains the feedback of the instantaneous interaction on the dynamic part via the $\chi$ correlator. The price to pay in the latter case is the statistical uncertainty of $F_{\rm IBC}$, as $\chi$ must be evaluated by means of Monte Carlo techniques, in contrast to the GC approach where $\chi$ is known up to machine precision thanks to Wick's theorem. This could lead to some inaccuracy in the analytical continuation of the IBC Bose factor.

If the goal is to compute the spectral representation of a Green's function with retarded interactions, we found that

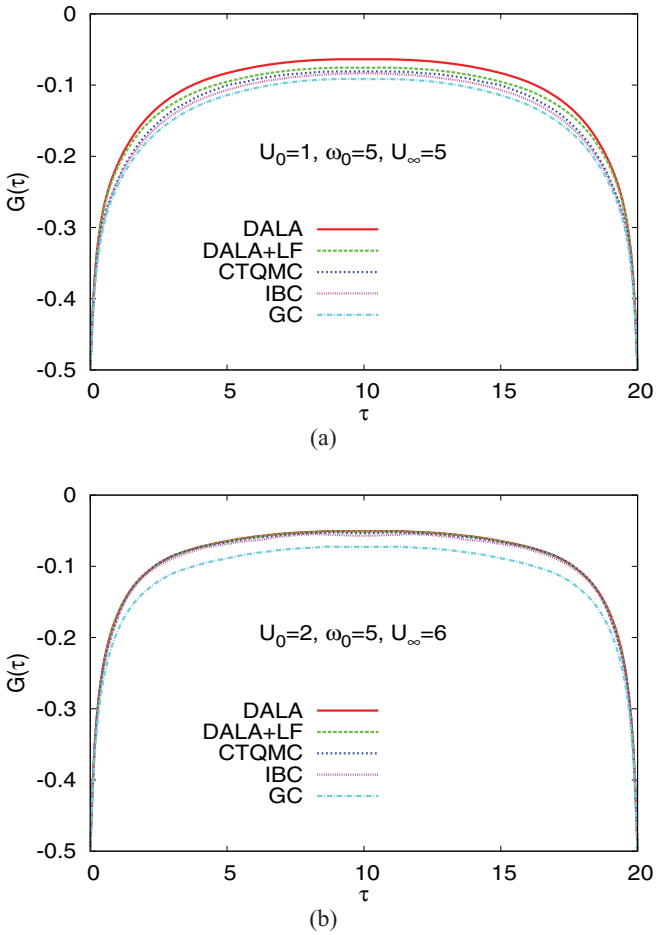


FIG. 7. (Color online) Symmetric Anderson model with retarded screened interaction as in Fig. 4: $\omega_0 = 5$, $U_\infty - U_0 = 5$, $\beta = 10$, for $\beta = 20$ and two values of $U_0$. (a) $U_0 = 1$; (b) $U_0 = 2$. The numerically exact Rubtsov's CTQMC, DALA, DALA + LF, GC, and IBC Green's functions are reported.

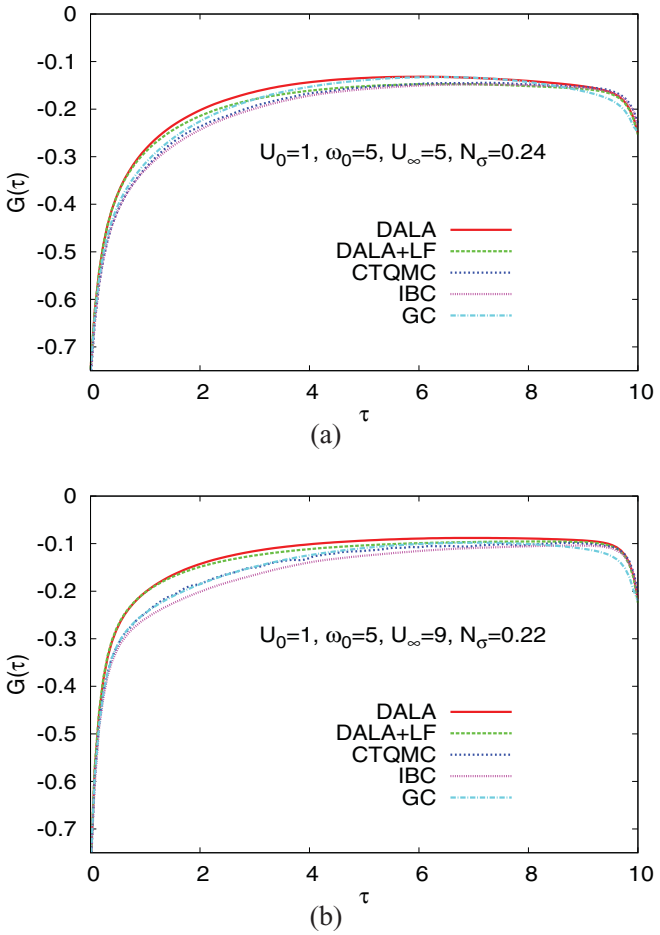


FIG. 8. (Color online) Asymmetric spin unpolarized Anderson model with retarded screened interaction as in Fig. 7: $\omega_0 = 5$, $U_\infty - U_0 = 5$, for $\beta = 20$ and two values of $U_0$. (a) $U_0 = 1$; (b) $U_0 = 2$. In this case, the number of particles $N_\sigma$ was set to be around the quarter-filling. The numerically exact Rubtsov's CTQMC, DALA, DALA + LF, GC, and IBC Green's functions are reported.

the most effective way is to use the Bose factor taken from the dynamic atomic limit. As already reported in Sec. IV, one has to define an auxiliary Green's function $G_{\rm aux}(\tau)$ as $G(\tau)/F_{\rm DALA}(\tau)$, and then use Eq. (14) to compute the spectrum. The DALA factor is the simplest to invert, and its bosonic spectral representation is directly related to the "physical" density of plasmonic (or phononic) modes, whereas the other factors beyond the DALA include also some low-energy contributions. Therefore, the DALA factor is the most recommended in the assisted analytical continuation.

## VIII. APPLICATION TO THE SINGLE-BAND HUBBARD-HOLSTEIN MODEL

In the following, we are going to present applications where our BFA approach is used as solver of the Anderson impurity problem resulting from the DMFT self-consistency conditions. Therefore, in contrast to what has been shown in the methodological sections, the following results are the converged solutions of the DMFT equations for the full lattice Hamiltonian, which incorporates the retarded $U(i\nu)$ as the onsite interaction.

TABLE I. Upper and lower critical $U_\infty$ values obtained at $\beta = 40$ by different BFA methods for the Mott transition in the single-band Hubbard-Holstein model on the Bethe lattice at half-filling, once $U_0(=2)$ and $\omega_0(=10)$ are fixed. The values are expressed in the half-bandwidth units. By our BFA, we always get a first-order phase transition, with a hysteresis effect, in agreement with what was reported in literature (Ref. 16).

| Method | $U_\infty^{c1}$ | $U_\infty^{c2}$ |
|---|---|---|
| GC | 12.3 | 12.5 |
| DALA | 8.2 | 8.4 |
| DALA + LF | 4.9 | 5.2 |

The first application is on the single-band half-filled Hubbard-Holstein model solved by the DMFT on the Bethe lattice. To study the impact of the retarded screened interaction on the Mott transition, we analyze the half-filled model at $\beta = 40$ with different $U_0$, $U_\infty$, and $\omega_0$ parameters. We choose to work always with $U_\infty > U_0$, a "physical" condition that states that the unscreened $U$ is larger than the screened one. The screening frequency is taken such that we are in the antiadiabatic regime, and we use our various Bose factor approximations to predict the critical value of $U_\infty$ for the Mott transition, once the other parameters are fixed, such that $\omega_0 = 10$ and $U_0 = 2$. As reported in literature,[15,16] the metal-Mott insulator transition in the Hubbard-Holstein model is first order. In Table I, we present our lower and upper critical values obtained for $\beta = 40$ by means of the BFA at various levels of approximation. As one can see, the actual values at criticality depend quite strongly on the approximation used, with the DALA + LF giving the results aligned with the numerically exact CTQMC method by Werner and Millis[21] (see Fig. 9). In order to have a close comparison to the data published in Ref. 21, we carried out some calculations also for $\beta = 100$. The agreement between the DALA + LF and the Monte Carlo predictions is quite remarkable. This highlights again the importance of the LF correction at low energy in order to accurately predict the physical properties.

Figure 9 summarizes the results for the upper critical line for different sets of $\omega_0$ and $U_0$, the values of which are taken not so far from the critical $U_{c2} \approx 2.6$ of the static model computed at $\beta = 40$.[41,42] It is clear that the fully retarded model with $U_\infty > U_0$ is more correlated than the static one with the same instantaneous $U_0$. The Mott transition happens at values of screened $U_0$ *lower* than the critical static $U_{c2}$ for any finite $\omega_0$. The dependence of the critical parameters on $\omega_0$ is also clear. At fixed screened $U_0$, a smaller $\omega_0$ corresponds to a smaller unscreened $U_\infty$ at which the Mott transition is reached. Indeed, if the frequency of the screening plasmon is closer to the Fermi level, it is easier for the unscreened part to induce the Mott transition at low energy. The same conclusions were reached by Werner and Millis by means of their numerically exact CTQMC algorithm. This shows the impact of the dynamic screening features on the low-energy properties of the model. In order to get the same effective low-energy parameters, the "effective" instantaneous $U_0$ can be up to 20%–25% larger than the true screened value even for plasmon frequencies $\omega_0 (\gg U_0, \gg D)$ in the antiadiabatic regime.

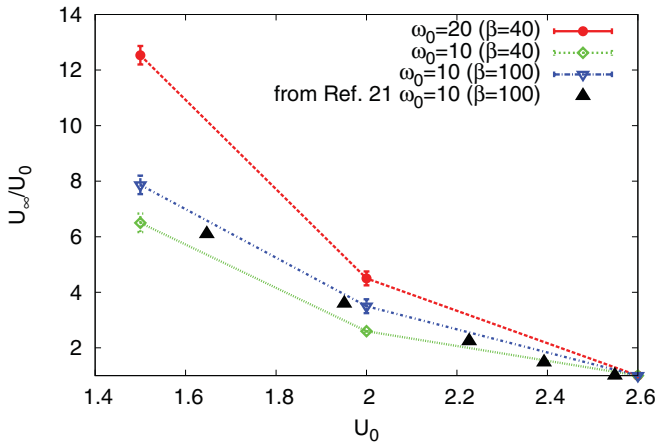


FIG. 9. (Color online) Critical $U_{c2}$ line of the Mott transition calculated at $\beta = 40$ and half-filling on a Bethe lattice for the screened instantaneous interaction $U_0$ as a function of $U_\infty/U_0$ and $\omega_0 = 10,20$ (green diamonds and red circles, respectively). The critical values are obtained from the Green's function computed by the DALA + LF ansatz. The behavior of the critical screened $U$ is in agreement with the one computed by Werner and Millis with their CTQMC algorithm at $\omega_0 = 10$ and $\beta = 100$ (black upper triangles). Indeed, the critical line calculated by means of our DALA + LF approach (blue lower triangles) is on top of their CTQMC points (Ref. 21). This highlights the accuracy of our approximated DALA + LF method if compared to the exact numerical result. Note the quite strong temperature dependence of the critical $U_{c2}$ in going from $\beta = 40$ to 100. This is a quite interesting effect, which deserves further analysis.

In Fig. 10, we report the spectral representation of a half-filled Holstein-Hubbard model with $\omega_0 = 10$, $U_0 = 2$, $U_\infty = 6.5$, and temperature $\beta = 100$, quite close to the Mott transition. The spectral function has been obtained with the help of the DALA factor by the method in Sec. IV B. It describes with high accuracy not only the low-energy features,

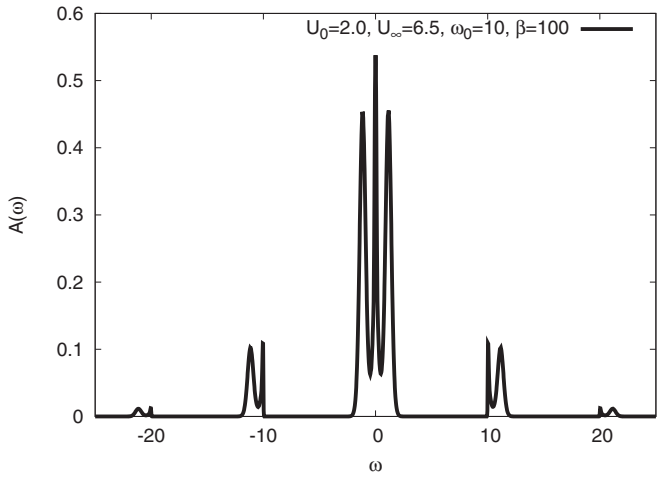


FIG. 10. Spectral representation of the Green's function for a Holstein-Hubbard model with $\omega_0 = 10$, $U_0 = 2$, $U_\infty = 6.5$, and $\beta = 100$, close to the blue Mott transition line of Fig. 9. The analytic continuation has been assisted by the DALA Bose factor, which allows one to describe accurately the plasmon satellites centered around frequencies multiple of $\omega_0$.

but also the plasmon satellites in the antiadiabatic regime, typical of realistic materials.

## IX. REALISTIC APPLICATION TO $SrVO_3$

Our second application is on a realistic Hamiltonian for $SrVO_3$, a very well-studied material, which represents a benchmark for theories describing strongly correlated compounds. $SrVO_3$ is the prototype of a correlated metal where the many-body treatment of correlation in the $d$ manifold is important to explain the spectral properties.[2] Therefore, it has been the subject of intensive studies[7,43–47] applying DMFT in the context of realistic strongly correlated Hamiltonians. Indeed, its band structure is relatively simple due to its undistorted perovskite geometry, resulting in the occupation of one electron in three degenerate $t_{2g}$ bands crossing the Fermi level. The $p$ oxygen ligands are quite well separated from the $d$ levels, such that the definition of a low-energy $t_{2g}$ Hamiltonian is unambiguous. Thus, $SrVO_3$ has been the testing case for many new DFT-DMFT implementations.[7,48–51] On the other hand, $SrVO_3$ has been the subject of intensive experimental activity, with magnetic, electrical, and optical measurements,[52–54] and by means of photoemission spectroscopy[55–58] (PES) and angled resolved PES (ARPES).[59–63]

Here, we consider a model where only the $t_{2g}$ electron is retained, and all the others contribute to screen the local lattice interaction. The DFT band structure has been calculated with the linear muffin-tin orbital (LMTO) framework in the atomic sphere approximation (ASA), which allows one to work with a native $d$ projected and localized orbital representation. The realistic retarded $U$ for this compound has been computed in Ref. 5 based on the cRPA construction.

The low-energy Hamiltonian we are going to work with consists of the LDA $t_{2g}$ Hamiltonian introduced in Eq. (6). For the $t_{2g}$ orbitals of $SrVO_3$, the screened value of the interaction $U_0 = 3.6$ eV, and the Hund's coupling $J = 0.68$ eV. The additional retarded interaction, which is included in our model, couples to the total charge of the system, as described by Eq. (7). There is no need for an explicit double-counting term since such a correction is absorbed into the effective chemical potential fixing the particle number to one.

To study the impact of $U_{\rm ret}$ on the low-energy properties of the model, we took into account different Holstein single-plasmon $U(i\nu)$'s and compared them with the corresponding static model. We used our BFA approach in its simplest DALA formulation, and computed the spectrum at $\beta = 10$ and $20\ {\rm eV}^{-1}$. Indeed, it turns out that at those temperatures, the LF correction is irrelevant, and so the DALA performs here at best even in terms of efficiency. In Fig. 11, we report the spectra computed with the instantaneous static $U_0(=3.6$ eV) interaction, and two retarded interactions with the same unscreened $U_\infty(=7$ eV), but different screening $\omega_0$'s (one at 5 eV, the other at 15 eV). The static model has a quasiparticle peak at the Fermi level, typical of strongly correlated compounds, with a lower and upper shoulder reminiscent of the lower and upper Hubbard bands. In the dynamic model, the effect of the 15-eV plasmon is to renormalize the quasiparticle spectral weight, transferred at higher energies, while the shape of the low-energy spectrum is almost unchanged. The effect of the 5-eV plasmon, closer to the low-energy sector, is more

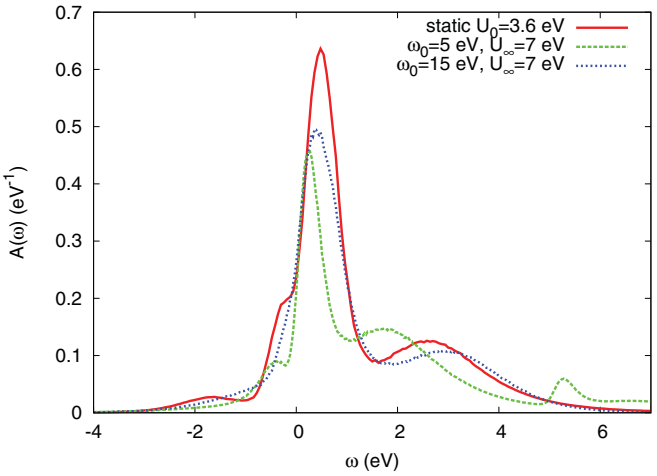


FIG. 11. (Color online) Spectral function of one electron in 3 $t_{2g}$ bands for the $SrVO_3$ obtained by the DFT in the local density approximation. The DMFT calculations are done at $\beta = 10$ eV$^{-1}$ with the DALA Bose factor solver for the retarded $U$ and the Hirsch-Fye QMC solver for the instantaneous interaction. Three types of interactions are reported: the purely static one, with $U_0 = 3.6$ eV, a dynamic Holstein one with the same instantaneous interaction but with the unscreened $U_\infty = 7$ eV, and a plasmon frequency $\omega_0 = 5$ eV, another Holstein interaction with the same $U_0$ and $U_\infty$ but different $\omega_0$(=15 eV). Note the displacement of the upper Hubbard shoulder shifted at lower energy by the presence of the plasmon at $\omega_0 = 5$ eV.

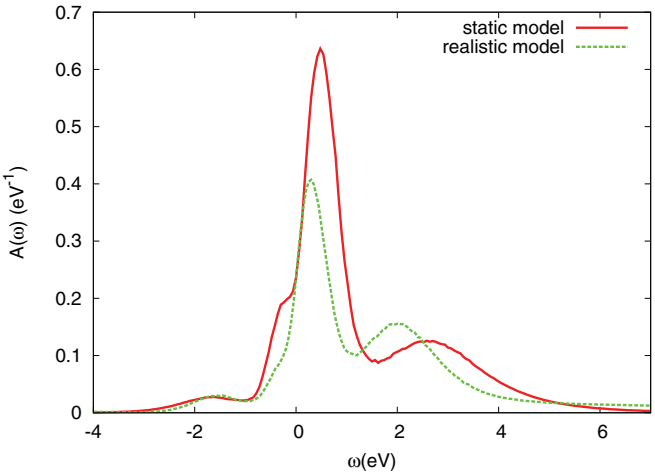


FIG. 12. (Color online) Spectral function at $\beta = 10$ eV$^{-1}$ of one electron in 3 $t_{2g}$ bands for the $SrVO_3$ obtained by the DFT in the local density approximation. The DMFT calculations are carried out with the DALA Bose factor solver for the retarded $U$ and the Hirsch-Fye QMC solver for the instantaneous interaction. The result of the realistic Green's function is reported, showing a more correlated behavior than the corresponding static model with the same instantaneous $U_0$(=3.6 eV). The quasiparticle peak is smaller than the corresponding static model, with a spectral weight transfer at higher energies.

remarkable. Aside from a stronger spectral weight reduction, there is a shift of the upper Hubbard shoulder to lower energies than in the corresponding static model. This effect is certainly due to the interplay between the plasmon satellite at 5 eV, visible in Fig. 11, and the low-energy features of the spectrum.

The realistic retarded $U(i\nu)$ is characterized by two main screening frequencies at 5 and 15 eV,[5,6] while the unscreened value of $U$ is 16 eV, much larger than the one in the models considered so far. By applying our Bose factor ansatz to this problem, we found the spectral function in Fig. 12, reported together with the instantaneous interaction with the same screened $U_0$(= 3.6eV). Also, here we note the quasiparticle weight reduction, and the shift of the upper Hubbard shoulder to lower frequencies, as in the model $U$ analyzed before, while the lower Hubbard band is almost unchanged with respect of the static model. Its maximum corresponds to the position found in PES measurements, and well documented in previous DMFT studies on $SrVO_3$. The weight reduction of the quasiparticle peak coincides with a smaller value of $Z = 1/(1 - \partial\Sigma''(i\omega)/\partial\omega|_{\omega=0})$, namely, a larger value of the effective mass $m^*$. In particular, for the realistic dynamic model, we found a $Z \approx 0.5$, which gives an effective mass renormalized by 2 with respect to the DFT band structure, while for the corresponding static model, we obtained a value of $Z \approx 0.7$. Recent ARPES data yielded an effective mass $m^* \approx 2m_0$,[59,61,63] which is in a good agreement with our findings for the realistic retarded interaction. On the other hand, the static model with the same screened $U_0$ underestimates the correlation by a factor of 1.4, and so the value of $m^*$. In the literature, a static model with a larger instantaneous $U_0$(≈5 eV) has been used to find the experimental mass renormalization.[44–47] Such a larger value of $U$ could be justified by the constrained LDA method[64,65] used to determine *a priori* the onsite interaction, but known to overestimate its strength. The difficulty here is to reproduce by the same model both the effective mass and the position of the lower Hubbard band, which turns out to be shifted at lower energies (≈ −2 eV) by a stronger $U_0$. This is the reason why in Ref. 49 the authors made the choice to work with $U_0 = 4$ eV, a slightly weaker effective static interaction, which gives the lower Hubbard band correctly peaked at −1.5 eV. Some cluster calculations give the same position for the lower Hubbard band, with the interpretation that its peak depends on the hybridization and screening properties provided by the ligands.[66,67] With our dynamically screened model, we describe correctly both the mass renormalization ($2m_0$) and the lower Hubbard band position peaked at −1.5 eV, as apparent in Fig. 12. This analysis highlights the importance of including the proper *retarded* interaction to have a reliable and fully *ab initio* description of the correlation in these materials.

We note, however, that the present description includes only the $t_{2g}$ orbitals, so that the above conclusions are valid in an energy range where no other orbital contributions are present.

## X. CONCLUSIONS

We introduced a factorized form $G = G_0F$ for the Green's function of the Anderson model with generic retarded (dynamic) interaction, dubbed "Bose factor ansatz." We proposed various approximations for the Green's function Bose factor $F$, the most practical and effective one borrowed directly from the dynamic atomic limit (DALA), the form of which is analytically known. The DALA provides an accurate way to compute the Green's function in the antiadiabatic limit *and* evaluate the spectral properties in the full frequency range

by means of an improved analytical continuation method. In practice, the inversion from the imaginary to the real frequency domain is assisted by the factor $F$, which retains the main information on the position and strength of the plasmon satellites, and enters in the analytical continuation as a convolution with the instantaneous part $G_0$. We carefully analyzed the pros and cons of the DALA, and found various ways to improve $F$ by either using the Lang-Firsov approximation at low frequency or resorting to diagrammatic techniques. Finally, we applied our approach to lattice problems in the context of the DMFT formalism. We took into account the Hubbard-Holstein Hamiltonian at half-filling and on the Bethe lattice, and we studied the Mott transition in the spirit of looking at the retarded interaction as resulting from "realistic screening" of the bare $U$ by a single plasmon. The second application has been for the realistic $SrVO_3$ Hamiltonian, where the $t_{2g}$ electrons interact via a retarded onsite $U$ previously determined *ab initio* at the cRPA level.[4] In both cases, it turns out that it is important to retain the retarded features of the local interaction resulting from the dynamic screening in order to have a reliable *ab initio* description of correlated materials. Also, our approach could be useful to determine whether some spectral signatures at intermediate energy ($\approx 10$ eV) seen in a broad class of correlated materials[68] come from a bulk dynamic screening. In perspective, more work has to be done theoretically to rationalize the effects of the screening on both the ground- and excited-state properties of correlated compounds. Moreover, by means of the same formalism, one can study Jahn-Teller models to describe the impact of the electron-phonon coupling to the spectral function of the distorted compound. Other electron-phonon coupling models in strongly correlated systems[69] could be solved by our ansatz, if they are in the antiadiabatic regime. Last but not least, dealing with a frequency dependent $U$ is an essential step toward the implementation of the $GW$ + DMFT framework,[36] where the screening resulting from the $GW$ polarization has to be included consistently in the low-energy correlated model solved at the DMFT level. Therefore, behind this work, there are important experimental, theoretical, and methodological implications that one can now start to take into account.

## ACKNOWLEDGMENTS

This work was supported by IDRIS/GENCI under Grant No. 101393, and by the French ANR under project SURMOTT. A.R. acknowledges support from the Dynasty Foundation, Russia. We thank F. Aryasetiawan, A. I. Lichtenstein, A. J. Millis, T. Miyake, J. M. Tomczak, and P. Werner for stimulating discussions.

## APPENDIX: FROM THE FIRST-ORDER CUMULANT EXPANSION TO THE DYNAMIC ATOMIC LIMIT

In this Appendix, we prove that the Bose factor $F$ reported in Eq. (25) fulfills the exact atomic limit, i.e., it equals the one in Eq. (16) for $\Delta = 0$. This is a nontrivial property, which guarantees the GC approximation to provide a reasonable description of the insulating phase (and the metal-to-insulator transition) in the DMFT framework, where the hybridization function of the Anderson impurity assumes a crucial frequency dependence in order to represent the coupling with the self-consistent bath. In the strong-coupling case, the hybridization $\Delta(i\omega_n)$ goes to zero for small $\omega_n$. Thus, in the DMFT language, this phase is mapped into an Anderson impurity problem close to the atomic limit, and its accurate solution is required around $\Delta = 0$.

We start by noting that the density-density correlator $\chi$ factorizes in the atomic limit as

$$\chi(t,t',\tau) \underset{\Delta=0}{=} \chi^1(t,\tau)\chi^1(t',\tau), \tag{A1}$$

where $\chi^1(t,\tau) = \langle \mathcal{T} d_\tau d_0^\dagger N(t)\rangle_0 / \mathcal{G}_0(\tau)$. In the most general case, $\chi$ can be resolved into its spin and orbital components by defining $\chi_{\sigma\sigma'}(t,t',\tau) = \langle \mathcal{T} d_\tau d_0^\dagger N_\sigma(t) N_{\sigma'}(t')\rangle_0 / \mathcal{G}_0(\tau)$, with $N_\sigma(t) = n_\sigma(t) - 1/2$. The factorization in Eq. (A1) holds also for the spin-resolved quantities: $\chi_{\sigma\sigma'} = \chi^1_\sigma \chi^1_{\sigma'}$. Once the Wick theorem is applied and the former correlators are written in terms of the $\mathcal{G}_0(\tau)$'s, it is straightforward to prove the spin-resolved identity and Eq. (A1) by using the atomic-limit expression for $\mathcal{G}_0(\tau)$. For instance, in the atomic limit and for a generic $\mu$, the connected part of $\chi^1$ reads as

$$\chi^1(t,\tau) = \begin{cases} -e^{-\mu\beta}/\left(1+e^{-\mu\beta}\right) & \text{for} \quad \tau < t < \beta, \\ 1/\left(1+e^{-\mu\beta}\right) & \text{for} \quad 0 \leqslant t \leqslant \tau. \end{cases} \tag{A2}$$

Now, we use a Hubbard-Stratonovich (HS) transformation to rewrite the $F_{GC}$ factor in a form that includes $\chi^1$ in linear terms only. We obtain the following identity:

$$\begin{aligned} &\exp\left(-\frac{1}{2}\int\!\!\int_0^\beta dt\, dt'\, \chi^1(t,\tau)\bar{U}(t-t')\chi^1(t',\tau)\right) \\ &\quad = \int \mathcal{D}\phi\, \exp\left(-i\int_0^\beta dt\, \chi^1(t,\tau)\phi(t)\right. \\ &\qquad \left. -\frac{1}{2}\int\!\!\int_0^\beta dt\, dt'\, \phi(t)\bar{U}^{-1}(t-t')\phi(t')\right), \end{aligned} \tag{A3}$$

where $\int \mathcal{D}\phi$ is the functional integral over the complex HS field $\phi(t)$. By exploiting the atomic-limit expression of $\chi^1$ in Eq. (A2), one can compute the integral involving the product of $\phi$ and $\chi^1$, which gives the result

$$\begin{aligned} \int_0^\beta dt\, \chi^1(t,\tau)\phi(t) &= \left(\frac{\tau}{\beta} - \frac{e^{-\mu\beta}}{1+e^{-\mu\beta}}\right)\phi_0 \\ &\quad + \frac{i}{\beta}\sum_{n\neq 0}\frac{\phi_n}{\nu_n}(e^{-i\nu_n\tau}-1), \end{aligned} \tag{A4}$$

with $\phi_n = \int_0^\beta dt\, \phi(t) e^{i\nu_n t}$ the Fourier components of the HS field. By integrating the functional integral in Eq. (A3) in the Fourier space, one gets the final expression for the exponent of the GC factor:

$$\begin{aligned} &-\frac{1}{2}\int\!\!\int_0^\beta dt\, dt'\, \chi(t,t',\tau)\bar{U}(t-t') \\ &\quad \underset{\Delta=0}{=} -\frac{1}{2}\beta\bar{U}(i\nu_0)\left(\frac{e^{-\mu\beta}}{1+e^{-\mu\beta}} - \frac{\tau}{\beta}\right)^2 \\ &\qquad + \frac{1}{\beta}\sum_{n\neq 0}\frac{\bar{U}(i\nu_n)}{\nu_n^2}(e^{i\nu_n\tau}-1). \end{aligned} \tag{A5}$$

We note here that $\bar{U}(i\nu_0) = 0$, as $\bar{U}(i\nu_n) = U(i\nu_n) - U_0$. Therefore, the first term in the right-hand side of the above equation vanishes, and we are left with the last term in Eq. (A5),

which is exactly equal to the exponent of the DALA factor. Thus, we have proven that

$$F_{\mathrm{GC}}(\tau) \underset{\Delta=0}{=} F_{\mathrm{DALA}}(\tau) \tag{A6}$$

for an arbitrary $\mu$ and $\bar{U}$.

To conclude, the fact that the first-order cumulant expansion fulfills the exact dynamic atomic limit justifies the use of cumulants instead of the standard first-order developments, and validates also the cumulant of the instantaneous bold factor in Eq. (27).


[1] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).
[2] M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998).
[3] D. C. Langreth, Phys. Rev. B **1**, 471 (1970).
[4] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. I. Lichtenstein, Phys. Rev. B **70**, 195104 (2004).
[5] F. Aryasetiawan, K. Karlsson, O. Jepsen, and U. Schönberger, Phys. Rev. B **74**, 125106 (2006).
[6] T. Miyake and F. Aryasetiawan, Phys. Rev. B **77**, 085122 (2008).
[7] M. Aichhorn, L. Pourovskii, V. Vildosola, M. Ferrero, O. Parcollet, T. Miyake, A. Georges, and S. Biermann, Phys. Rev. B **80**, 085101 (2009).
[8] T. Miyake, F. Aryasetiawan, and M. Imada, Phys. Rev. B **80**, 155134 (2009).
[9] A. C. Hewson, *The Kondo Problem to Heavy Fermions, Cambdridge Studies in Magnetism* (Cambridge University Press, Cambridge, UK, 1993).
[10] A. C. Hewson and D. Meyer, J. Phys.: Condens. Matter **14**, 427 (2002).
[11] G. Sangiovanni, M. Capone, C. Castellani, and M. Grilli, Phys. Rev. Lett. **94**, 026401 (2005).
[12] G. Sangiovanni, M. Capone, and C. Castellani, Phys. Rev. B **73**, 165123 (2006).
[13] P. Barone, R. Raimondi, M. Capone, C. Castellani, and M. Fabrizio, European Phys. Lett. **79**, 47003 (2007).
[14] M. Hohenadler and W. von der Linden, Phys. Rev. B **71**, 184309 (2005).
[15] W. Koller, D. Meyer, and A. C. Hewson, Phys. Rev. B **70**, 155103 (2004).
[16] G. S. Jeon, T.-H. Park, J. H. Han, H. C. Lee, and H.-Y. Choi, Phys. Rev. B **70**, 125114 (2004).
[17] P. Werner and A. J. Millis, Phys. Rev. B **74**, 155107 (2006).
[18] P. Werner, A. Comanac, L. dé Medici, M. Troyer, and A. J. Millis, Phys. Rev. Lett. **97**, 076405 (2006).
[19] I. G. Lang and Y. A. Firsov, Sov. Phys. JETP **16**, 1301 (1963) [Zh. Eksp. Teor. Fiz. **43**, 1843 (1962)].
[20] P. Werner and A. J. Millis, Phys. Rev. Lett. **99**, 146404 (2007).
[21] P. Werner and A. J. Millis, Phys. Rev. Lett. **104**, 146401 (2010).
[22] A. N. Rubtsov and A. I. Lichtenstein, JETP Lett. **80**, 61 (2004) [Pis'ma v ZhETF **80**, 67 (2004)].
[23] A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, Phys. Rev. B **72**, 035122 (2005).
[24] J. E. Gubernatis, M. Jarrell, R. N. Silver, and D. S. Sivia, Phys. Rev. B **44**, 6011 (1991).
[25] M. Jarrell and J. E. Gubernatis, Phys. Rep. **269**, 133 (1996).
[26] J. E. Hirsch and R. M. Fye, Phys. Rev. Lett. **56**, 2521 (1986).
[27] J. Hubbard, Proc. R. Soc. London, Ser. A **276**, 238 (1963).
[28] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[29] V. I. Anisimov, J. Zaanen, and O. K. Andersen, Phys. Rev. B **44**, 943 (1991).
[30] L. Hedin, Phys. Rev. **139**, A796 (1965).
[31] Q. Si and J. L. Smith, Phys. Rev. Lett. **77**, 3391 (1996).
[32] H. Kajueter, Ph.D. thesis, Rutgers University, 1996.
[33] P. Sun, G. Kotliar, Phys. Rev. B **66**, 085120 (2002).
[34] R. Chitra and G. Kotliar, Phys. Rev. B **63**, 115110 (2001)
[35] Th. Ayral, S. Biermann, and Ph. Werner (unpublished).
[36] S. Biermann, F. Aryasetiawan, and A. Georges, Phys. Rev. Lett. **90**, 086402 (2003).
[37] C. Castellani, C. R. Natoli, and J. Ranninger, Phys. Rev. B **18**, 4945 (1978).
[38] L. Vaugier, Ph.D. thesis, Ecole Polytechnique, Palaiseau, France, 2011.
[39] Serge Florens, Ph.D. Thesis, Ecole Normale Superieure, Paris, France, 2003.
[40] H. J. Vidberg and J. W. Serene, J. Low Temp. Phys. **29**, 179 (1977).
[41] M. J. Rozenberg, G. Kotliar, and X. Y. Zhang, Phys. Rev. B **49**, 10181 (1994).
[42] A. Georges and W. Krauth, Phys. Rev. B **48**, 7167 (1993).
[43] A. Liebsch, Phys. Rev. Lett. **90**, 096401 (2003).
[44] E. Pavarini, S. Biermann, A. Poteryaev, A. I. Lichtenstein, A. Georges, and O. K. Andersen, Phys. Rev. Lett. **92**, 176403 (2004).
[45] E. Pavarini, A. Yamasaki, J. Nuss, and O. K. Andersen, New J. Phys. **7**, 188 (2005).
[46] I. A. Nekrasov, G. Keller, D. E. Kondakov, A. V. Kozhevnikov, Th. Pruschke, K. Held, D. Vollhardt, and V. I. Anisimov, Phys. Rev. B **72**, 155106 (2005).
[47] I. A. Nekrasov, K. Held, G. Keller, D. E. Kondakov, Th. Pruschke, M. Kollar, O. K. Andersen, V. I. Anisimov, and D. Vollhardt, Phys. Rev. B **73**, 155112 (2006).
[48] V. I. Anisimov, D. E. Kondakov, A. V. Kozhevnikov, I. A. Nekrasov, Z. V. Pchelkina, J. W. Allen, S.-K. Mo, H.-D. Kim, P. Metcalf, S. Suga, A. Sekiyama, G. Keller, I. Leonov, X. Ren, and D. Vollhardt, Phys. Rev. B **71**, 125119 (2005).
[49] F. Lechermann, A. Georges, A. Poteryaev, S. Biermann, M. Posternak, A. Yamasaki, and O. K. Andersen, Phys. Rev. B **74**, 125120 (2006).
[50] B. Amadon, F. Lechermann, A. Georges, F. Jollet, T. O. Wehling, and A. I. Lichtenstein, Phys. Rev. B **77**, 205112 (2008).
[51] G. Trimarchi, I. Leonov, N. Binggeli, Dm. Korotin, V. I. Anisimov, J. Phys.: Condens. Matter **20**, 135227 (2008).
[52] M. Onoda, H. Ohta, and H. Nagasawa, Solid State Commun. **79**, 281 (1991).
[53] I. H. Inoue, O. Goto, H. Makino, N. E. Hussey, and M. Ishikawa, Phys. Rev. B **58**, 4372 (1998).
[54] H. Makino, I. H. Inoue, M. J. Rozenberg, I. Hase, Y. Aiura, and S. Onari, Phys. Rev. B **58**, 4384 (1998).

[55]A. Fujimori, I. Hase, H. Namatame, Y. Fujishima, Y. Tokura, H. Eisaki, S. Uchida, K. Takegahara, and F. M. F. de Groot, Phys. Rev. Lett. **69**, 1796 (1992).

[56]I. H. Inoue, I. Hase, Y. Aiura, A. Fujimori, Y. Haruyama, T. Maruyama, and Y. Nishihara, Phys. Rev. Lett. **74**, 2539 (1995).

[57]K. Maiti, D. D. Sarma, M. Rozenberg, I. Inoue, H. Makino, O. Goto, M. Pedio, and R. Cimino, Europhys. Lett. **55**, 246 (2001).

[58]A. Sekiyama, H. Fujiwara, S. Imada, S. Suga, H. Eisaki, S. I. Uchida, K. Takegahara, H. Harima, Y. Saitoh, I. A. Nekrasov, G. Keller, D. E. Kondakov, A. V. Kozhevnikov, Th. Pruschke, K. Held, D. Vollhardt, and V. I. Anisimov, Phys. Rev. Lett. **93**, 156402 (2004).

[59]T. Yoshida, K. Tanaka, H. Yagi, A. Ino, H. Eisaki, A. Fujimori, and Z.-X. Shen, Phys. Rev. Lett. **95**, 146404 (2005).

[60]H. Wadati, T. Yoshida, A. Chikamatsu, H. Kumigashira, M. Oshima, H. Eisaki, Z. X. Shen, T. Mizokawa, and A. Fujimori, Phase Transitions **79**, 617 (2006).

[61]R. Eguchi, T. Kiss, S. Tsuda, T. Shimojima, T. Mizokami, T. Yokoya, A. Chainani, S. Shin, I. H. Inoue, T. Togashi, S. Watanabe, C. Q. Zhang, C. T. Chen, M. Arita, K. Shimada, H. Namatame, and M. Taniguchi, Phys. Rev. Lett. **96**, 076402 (2006).

[62]M. Takizawa, M. Minohara, H. Kumigashira, D. Toyota, M. Oshima, H. Wadati, T. Yoshida, A. Fujimori, M. Lippmaa, M. Kawasaki, H. Koinuma, G. Sordi, and M. Rozenberg, Phys. Rev. B **80**, 235104 (2009).

[63]T. Yoshida, M. Hashimoto, T. Takizawa, A. Fujimori, M. Kubota, K. Ono, and H. Eisaki, Phys. Rev. B **82**, 085119 (2010).

[64]O. Gunnarsson, O. K. Andersen, O. Jepsen, and J. Zaanen, Phys. Rev. B **39**, 1708 (1989).

[65]M. Cococcioni and S. de Gironcoli, Phys. Rev. B **71**, 035105 (2005).

[66]R. J. O. Mossanek, M. Abbate, and A. Fujimori, Phys. Rev. B **74**, 155127 (2006).

[67]R. J. O. Mossanek, M. Abbate, T. Yoshida, A. Fujimori, Y. Yoshida, N. Shirakawa, H. Eisaki, S. Kohno, P. T. Fonseca, and F. C. Vicentin, Phys. Rev. B **79**, 033104 (2009).

[68]M. Taguchi, A. Chainani, N. Kamakura, K. Horiba, Y. Takata, M. Yabashi, K. Tamasaku, Y. Nishino, D. Miwa, T. Ishikawa, S. Shin, E. Ikenaga, T. Yokoya, K. Kobayashi, T. Mochiku, K. Hirata, and K. Motoya, Phys. Rev. B **71**, 155102 (2005).

[69]O. Gunnarsson and O. Rosch, J. Phys.: Condens. Matter **20**, 043201 (2008).