Influence of Fock exchange in combined many-body perturbation and dynamical mean field theory

Thomas Ayral,1,2 Silke Biermann,3 Philipp Werner,4 and Lewin Boehnke4

1Institut de Physique Théorique (IPhT), CEA, CNRS, UMR 3681, 91191 Gif-sur-Yvette, France
2Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854, USA
3Centre de Physique Théorique, École Polytechnique, CNRS, Université Paris Saclay, 91128 Palaiseau, France
4Department of Physics, University of Fribourg, 1700 Fribourg, Switzerland

(Rceived 26 January 2017; revised manuscript received 5 May 2017; published 23 June 2017)

In electronic systems with long-range Coulomb interaction, the nonlocal Fock-exchange term has a band-widening effect. While this effect is included in combined many-body perturbation theory and dynamical mean field theory (DMFT) schemes, it is not taken into account in standard extended DMFT (EDMFT) calculations. Here, we include this instantaneous term in both approaches and investigate its effect on the phase diagram and dynamically screened interaction. We show that the largest deviations between previously presented EDMFT and GW+EDMFT results originate from the nonlocal Fock term, and that the quantitative differences are especially large in the strong-coupling limit. Furthermore, we show that the charge-ordering phase diagram obtained in GW+EDMFT methods for moderate interaction values is very similar to the one predicted by dual-boson methods that include the fermion-boson or four-point vertex.

DOI: 10.1103/PhysRevB.95.245130

Dynamical mean field theory (DMFT) [1] self-consistently maps a correlated Hubbard lattice problem with local interactions onto an effective impurity problem consisting of a correlated orbital hybridized with a noninteracting fermionic bath. If the bath is integrated out, one obtains an impurity action with retarded hoppings. Extended dynamical mean field theory [2–10] (EDMFT) extends the DMFT idea to systems with long-range interactions. It does so by mapping a lattice problem with long-range interactions onto an effective impurity model with self-consistently determined fermionic and bosonic baths, or, in the action formulation, an impurity model with retarded hoppings and retarded interactions.

While EDMFT captures dynamical screening effects and charge-order instabilities, it has been found to suffer from qualitative shortcomings in finite dimensions. For example, the charge susceptibility computed in EDMFT does not coincide with the derivative of the average charge with respect to a small applied field [11], nor does it obey local charge conservation rules relevant to an adequate description of collective modes such as plasmons.

The EDMFT formalism has an even more basic deficiency: since it is based on a local approximation to the self-energy, it does not include even the first-order nonlocal interaction term, the Fock term. The combined GW+EDMFT [13–15] scheme corrects this by supplanting the local self-energy from EDMFT with the nonlocal part of the GW diagram, where $G$ is the interacting Green’s function and $W$ the fully screened interaction. Indeed, the nonlocal Fock term $G_{\psi}^{\text{nonloc}}$ is included in the nonlocal $[GW]^{\text{nonloc}}$ diagram. As described in more detail in Ref. [15] (see also the appendix of Ref. [16]), the $GW+\text{EDMFT}$ method is formally obtained by constructing an energy functional of $G$ and $W$, the Almbladh [17] functional $\Psi$, and by approximating $\Psi$ as a sum of two terms, one containing all local diagrams (corresponding to EDMFT) and the other containing the simplest nonlocal correction (corresponding to the GW approximation [18]).

This functional construction rules out double counting of local terms in the self-energy and polarization [15,19]. Even though it has been introduced under the name $GW+\text{DMFT}$ [13] in the literature, we denote this full scheme by $GW+\text{EDMFT}$ to emphasize that it is based on the EDMFT formalism, and to distinguish it from simplified implementations without two-particle self-consistency, which have appeared in the literature (and which we denote in the following as $GW+\text{DMFT}$).

In a recent implementation of the $GW+\text{EDMFT}$ method (Ref. [15], and related papers [20,21]), the nonlocal Fock term was omitted [22]. Here, we explore and highlight the role of this term and its interplay with the local correlations. We quantify the band-widening effect of the Fock term and study the consequences of its presence or absence on various observables, and on the charge-order phase boundary. Our self-consistent implementation goes beyond previous studies of the effect of the Fock exchange in realistic calculations, where it was studied systematically within $GW$ [23] and $GW+\text{DMFT}$ [24], albeit not in a self-consistent way.

The paper is organized as follows: In Sec. I, we recap the $GW+\text{EDMFT}$ equations with special emphasis on the Fock term and make general statements about the expected impact. In Sec. II, we show explicit results for the effective renormalization of the band structure by the instantaneous Fock contribution within $GW+\text{EDMFT}$, followed by systematic comparisons with the results of simplified formalisms in Sec. III. Section IV discusses the role of the Fock term in the Mott-insulating phase, where it stays relevant up to very large values of the on-site interaction. Finally, in Sec. V, we compare our results with results obtained within the recent dual-boson method.

I. FORMALISM

We aim at solving the extended Hubbard model on the two-dimensional square lattice by constructing an effective impurity problem that gives the local part of the self-energy $\Sigma$ and polarization $P$, and a diagrammatic expansion in their nonlocal components. The model is defined by the

$$\cdots$$
Hamiltonian
\[
H = -\sum_{ij} t_{ij} c_i^\dagger c_j + \frac{1}{2} \sum_{ij} v_{ij} n_i n_j - \mu \sum_i n_i. \tag{1}
\]

Here, \( t_{ij} \) are the real-space hopping matrix elements, \( c_i^\dagger \) the electronic annihilator (creator) on site \( i \), \( v_{ij} \) the Coulomb interaction, and \( n_i = c_i^\dagger c_i \). We restrict ourselves to models with hoppings and interactions between nearest neighbors and next-nearest neighbors only,
\[
t_{ij} = t \delta_{ij} + t' \delta_{ij}, \tag{2}
\]
\[
v_{ij} = U \delta_{ij} + V \delta_{ij} + V' \delta_{ij}, \tag{3}
\]
where \( \delta_{ij} \) is the usual Kronecker delta, \( \delta_{ij} \) (respectively \( \delta_{ij} \)) is 1 for \( i \) and \( j \) nearest neighbors (respectively next-nearest neighbors, along the diagonal of the square lattice) and zero otherwise. This results in the Fourier transforms
\[
e_k = 2\{\cos(k_x) + \cos(k_y)\} + 2t'\{\cos(k_x + k_y) + \cos(k_x - k_y)\}, \tag{4}
\]
and
\[
v_q = U + 2V[\cos(q_x) + \cos(q_y)] + 2V'[\cos(q_x + q_y) + \cos(q_x - q_y)]. \tag{5}
\]

The full expression for the self-energy in the GW+EDMFT approximation is
\[
\Sigma(k,i\omega_n) = \Sigma_{\text{imp}}(i\omega_n) + \Sigma_{G_W}(k,i\omega_n) + \Sigma_F(k). \tag{6}
\]
The last two terms correspond to the nonlocal part of the \( G \) self-energy. They can be expressed as a function of imaginary time \( \tau \) and momentum \( k \) as follows:
\[
\Sigma_{G_W}^{\text{nonloc}}(k,\tau) = -\sum_q G_{q+k}(\tau) W_q^c(\tau)
\]
\[
+ \left[ \sum_q G_{q+k}(\tau) W^c_q(\tau) \right]_\text{loc}, \tag{7}
\]
\[
\Sigma_F^{\text{nonloc}}(k) = -\sum_q G_{q+k}(\tau = 0^+) v_q
\]
\[
+ \left[ \sum_q G_{q+k}(\tau = 0^+) v_q \right]_\text{loc}. \tag{8}
\]

Fourier transformations between \( \tau \) and fermionic [bosonic] Matsubara frequencies \( i\omega_n = i(2n + 1)\pi T \) \([i\nu_m = i2m\pi T]\) are assumed where needed (\( \beta \) denotes the inverse temperature). The “loc” suffix denotes a sum over the first Brillouin zone. The interacting lattice Green’s function is
\[
G_k(i\omega_n) = [i\omega_n + \mu - \Sigma(k,i\omega_n)]^{-1}, \tag{9}
\]
and \( W^c \) is defined as
\[
W^c_q(i\nu_m) \equiv \frac{v_q}{1 - v_q P_q(i\nu_m)} - v_q, \tag{10}
\]
with \( P \) the polarization function. All results are given in units of \( D = 4|t| \) (which is the half bandwidth when \( t' = 0 \)), and the momentum discretization is \( N_k = 32 \times 32 \) points in the first Brillouin zone, unless otherwise stated. We use the original formulation of the GW+EDMFT scheme [13], corresponding—within a functional formulation—to a Hubbard-Stratonovich decoupling of the full interaction term, dubbed “HS-UV decoupling” in Ref. [15]. As argued there, this choice has the advantage that it treats local and nonlocal interactions on the same footing.

The nonlocal Fock term of Eq. (6), which is real valued and instantaneous, renormalizes the bandwidth. It can become quite large and momentum dependent. This is illustrated in Fig. 1 for the parameters \( U = 2 \) and \( V = 0.4 \) (and \( t' = 0 \) and \( V' = 0 \), which is assumed in the following if not explicitly stated otherwise). The figure also indicates that for the case of nearest-neighbor hopping and interaction, the nonlocal Fock term can be exactly absorbed into the bare dispersion Eq. (4) by defining a \( U \)- and \( V \)-dependent hopping
\[
\tilde{t}(t',U,V,V') = t + \delta t(t',U,V,V'). \tag{11}
\]
This can be understood by looking at the real-space representation of the Fock term, Eq. (8):
\[
\Sigma^{\text{nonloc}}_{ij} = -G_{ij}(\tau = 0^+) v_{ij} + G_{ij}(\tau = 0^+) v_{ij}\delta_{ij}
\]
\[
= -G_{ij}(\tau = 0^+) V \delta_{ij} - G_{ij}(\tau = 0^+) V' \delta_{ij}, \tag{12}
\]
where the notation \( (ij) \) (\( ((ij)) \)) denotes a restriction to nearest-neighbor (next-nearest-neighbor) interactions.

Thus, for the case of Fig. 1, where \( V' = 0 \), the Fock term enters Eq. (9) as a renormalization of the nearest-neighbor hopping by
\[
\delta t(t',U,V,V') = -G_{ij}(\tau = 0^+) V. \tag{13}
\]

The Green’s function factor, which is closely related to the occupation number, has an implicit dependence on all parameters of the lattice problem. In the presence of a next-nearest-neighbor interaction, \( t' \) also gets renormalized according to
\[
\delta t'(t',U,V,V') = -G_{ij}(\tau = 0^+) V'. \tag{14}
\]

Note that such a term breaks the particle-hole symmetry of the lattice. In the particle-hole-symmetric case with \( t' = 0 \)
indicates the phase transition to the charge-ordered phase and the line half filling (top) and half filling, the next-nearest-neighbor Green’s function factor vanishes: 
\[ G_{ij}(\tau = 0^+) = 0. \]

**II. EFFECTIVE BAND STRUCTURE**

In the simplest case of nearest-neighbor interaction \( V \), the nearest-neighbor hopping renormalization \( \delta t \) determines the band widening [see Eq. (13)]. Since the half bandwidth is \( D = 4|t| \), the widening will be \( \delta D = 4\delta t \). Figure 2 illustrates this effect throughout the homogeneous part of the phase diagram for the particle-hole-symmetric \( (t' = 0) \) half-filled case. The most obvious feature is the increase with \( V \), which is expected from Eq. (13), and the decrease close to the Mott-insulating phase. \( \delta t \) nonetheless remains significant even at very high values of \( U \), a property that is further investigated in Sec. IV.

Away from half filling, where the Mott-insulating phase does not exist, the corresponding suppression of \( \delta t \) disappears, but otherwise the dependence on \( V \) and \( U \) is very similar to the half-filled case; see the bottom panel of Fig. 2.

To study a model with broken particle-hole symmetry, we introduce a nearest-neighbor hopping \( t' = t/\sqrt{2} \) and fix the filling at \( n = 0.8 \) as well as the nearest-neighbor interaction \( V \). We then calculate \( \delta t \) as a function of the next-nearest-neighbor interaction \( V' \). As shown in Fig. 3, the main effect on the hopping renormalization comes from the (essentially linear) dependence on \( V \). The qualitative effect of \( V' \) is to slightly reduce the renormalization.

In order to make the connection to realistic electronic structure calculations, we note as a side remark that there the situation is slightly more subtle. The band-widening effect is indeed relative to the reference point. Let us consider three reference Hamiltonians: (i) the Kohn-Sham Hamiltonian \( H_{KS} \) of density functional theory (DFT); (ii) the Hartree Hamiltonian 
\[ H_0 = H_{KS} - V [v_{xc}(\mathbf{r})], \]
where \( v_{xc}(\mathbf{r}) \) denotes the Kohn-Sham exchange-correlation potential, which is local in the electronic structure sense [that is, “local” denotes a function depending only on one space variable \( f(\mathbf{r}) \), while “nonlocal” denotes a function depending on two variables \( f(\mathbf{r}, \mathbf{r}') \)]; and (iii) the nonlocal exchange Hamiltonian 
\[ H_{xc} = H_0 + V [v_{xc}^{\prime}(\mathbf{r}, \mathbf{r}')] \]
where \( v_{xc}^{\prime}(\mathbf{r}, \mathbf{r}') \) denotes an exchange-correlation potential including the “nonlocal” Fock exchange. Then the hierarchy of the bandwidths in a metallic system is
\[ H_0 > H_{xc}^{\prime} > H_{KS}. \]

Thus, \( H_{xc}^{\prime} \) indeed widens the band with respect to density functional theory calculations. The question of the relative bandwidth changes thus implies a question on the starting band structure. We refer the interested reader to Ref. [25] for a systematic construction of explicit low-energy many-body Hamiltonians.

Here, we only comment on the specific point of the Hartree and Fock terms, in order to put our work on the extended Hubbard Hamiltonian into perspective with respect to realistic electronic structure calculations. Indeed, as argued in Ref. [25], in realistic electronic structure calculations one needs to avoid double counting of interactions at the one- and two-particle level. Let us consider first the case of the Hartree terms: Standard electronic structure techniques (e.g., a DFT calculation) produce a band structure including the Hartree contribution. This one-body potential contribution is then already part of the effective hopping parameter determined from this band structure. Reference [25] explains how to avoid double counting by including—at the level of the many-body calculation—only terms beyond Hartree. Here, we do not need to address this point in detail, since a Hartree term included in the model calculation would cancel out with the corresponding shift of the chemical potential, since the particle number eventually determines the energetic level of the single orbital included in the present model.

Let us now move to the analogous question for the Fock term: One may examine the relevance of excluding it at the level of the many-body calculation, and keeping it at
the level of the electronic structure calculation instead. The answer is based on several elements: The first point to note is that standard DFT calculations do treat exchange in a local approximation (where “local” here means again “local in the electronic structure sense”; see above), which relies on an error-cancellation effect with part of the correlation contribution (see, e.g., Ref. [26]) and is not relevant here. The next question is, therefore, Why not start from a Hartree-Fock calculation in the continuum in the full energy range of the Coulomb Hamiltonian? Such a treatment would neglect the crucial screening of the bare interaction by high-energy degrees of freedom (typically, matrix elements of the bare Coulomb interaction in the relevant Wannier functions are of the order of several tens of electron volts, while the effective Hubbard interactions are usually a few electron volts). Therefore, what is relevant here is indeed the exchange term calculated using the effective bare interaction of the low-energy Hamiltonian. For realistic electronic structure calculations, this interaction should correspond to a partially screened interaction, where screening by high-energy degrees of freedom is taken into account (as done, e.g., in the screened exchange + DMFT scheme [27,28]). We refer the interested reader to Refs. [25–27] for details.

III. SIMPLIFIED VARIANTS OF GW + EDMFT

In the following, we study the effect of this $V$-dependent bandwidth renormalization on local observables as well as the critical value of the nearest-neighbor repulsion for the transition into the charge-ordered phase. We call $GW_c$ + EDMFT the formula implemented in Ref. [15] (which contains only the $GW_c$ term; see Ref. [22]), and $GW$ + EDMFT the formula with the self-energy expression (6). For comparison, we also show results for $Gv$ + EDMFT, a scheme where $\Sigma$ is the sum of the impurity self-energy and of the nonlocal Fock term only [the first and third terms of Eq. (6)]. In all three schemes, the polarization is the sum of the impurity polarization with the nonlocal part of the interacting bubble (as described in steps (5)(a) and (5)(b) of Sec. V in Ref. [15]).

In Fig. 4, we plot the self-energy and polarization obtained from the three schemes at different momenta, and we compare the results to the local EDMFT self-energy and polarization. All these results are for half filling. One can observe the following trends:

(i) While the imaginary part of the self-energy in $GW_c$ + EDMFT is larger than in EDMFT, the opposite is true for $GW$ + EDMFT; i.e., the $GW$ + EDMFT self-energy is less correlated than the EDMFT self-energy.

(ii) The $GW$ + EDMFT result is more strongly correlated than $Gv$ + EDMFT.

(iii) At small Matsubara frequencies, the polarization is overall larger in the $GW$/$Gv$ + EDMFT method than in the $GW_c$ + EDMFT.

The trend in the self-energy (i.e., less correlated in $GW$ + EDMFT than EDMFT) can be understood easily from the broadening effect of the nonlocal Fock term on the band: When the (effective) bandwidth gets larger, so does the polarization $P$, and hence screening effects are more important, interactions are more screened, and, as a result,

![Figure 4](https://example.com/figure4)

**FIG. 4.** Im$\Sigma$($k, i\omega$) (left) and Re$\Pi(q, i\nu)$ (right) (black lines for EDMFT, magenta lines for $GW$ + EDMFT, blue lines for $Gv$ + EDMFT). Circles, squares, and triangles show results for the $(0,0)$, $(0,\pi)$, and $(\pi,\pi)$ points, respectively. Top, $U = 2.0$ and $V = 0.4$; bottom, $U = 3.0$ and $V = 1.0$ ($\beta = 100, n = 1$).

the imaginary part of the Matsubara self-energy is smaller in absolute magnitude.

Less trivial is the comparison between $GW$ + EDMFT and $Gv$ + EDMFT. Here, the band-widening effect is included in both calculations, and it turns out that the additional nonlocal $GW$ contributions to the self-energy lead to stronger correlations. This is consistent with the conclusions of Ref. [15], which compared $GW_c$ + EDMFT to EDMFT.

In the top panels of Fig. 5, we replot Figs. 15(a) and 15(b) of Ref. [15], and we show, in the bottom panels, the same observables for $U = 3$ and $V = 1$. One sees that the deviation of the $GW$ + EDMFT and $Gv$ + EDMFT results from the EDMFT result is very small for $U = 2$ and $V = 0.4$, but sizable for larger interaction values ($U = 3$ and $V = 1$).

In Fig. 6, we plot the corresponding spectral functions (the EDMFT and $GW_c$ + EDMFT results are identical to Fig. 2(b) of Ref. [20]). As a logical consequence of the above observations, the $GW$ + EDMFT and $Gv$ + EDMFT spectra are close to each other and slightly less correlated than the EDMFT spectrum, in the sense that the integrated weight of the quasiparticle peak is larger in those methods.

We next consider the phase diagram in the $U$-$V$ plane. In Fig. 7, we plot the dependence of the inverse charge susceptibility $\chi^{-1}_{q=\pi,\pi}(i\nu_m = 0)$ on the nearest-neighbor repulsion $V$, with $\chi$ defined by

$$\chi_q(i\nu_m) = \frac{-\Pi_q(i\nu_m)}{1 - \Pi_q(i\nu_m)}.$$ (16)

When the inverse susceptibility vanishes, the charge susceptibility diverges, signaling a transition to a charge-ordered
phase with a checkerboard pattern. The corresponding phase diagram is shown in Fig. 8, where we plot the results from Fig. 5 of Ref. [15] together with the phase boundaries for the \(GW+EDMFT\) and \(G_{\nu}+EDMFT\) methods.

At low and intermediate \(U\), \(GW+EDMFT\) and \(GW_{\nu}+EDMFT\) yield quantitatively similar critical nonlocal interactions \(V_c\) for the transition to the charge-ordered phase over a wide range of the local interactions. More importantly, they capture the expected \(GW\) behavior at low \(U\) that EDMFT misses due to its local self-energy. In the strong-coupling limit, the value of \(V_c\) is substantially reduced (middle and bottom panels) when going from EDMFT to \(GW+EDMFT\) or even only \(G_{\nu}+EDMFT\) (\(GW_{\nu}+EDMFT\) is very close to EDMFT).

**IV. Inside the Mott Phase**

In the large-interaction regime of the phase diagram (Fig. 8), the nonlocal Fock term has a significant effect. The schemes
which lack this instantaneous contribution, EDMFT and \( G_W + \text{EDMFT} \), yield a larger and steeper phase boundary than the schemes that take the Fock term into account (\( G_V + \text{EDMFT} \) and \( G_W + \text{EDMFT} \)).

The exact phase boundary in the Mott phase is difficult to predict \textit{a priori}. It can be computed in the classical (\( t_{ij} \to 0 \)) and zero-temperature limit of the extended Hubbard model by exact Monte Carlo simulations, as, e.g., in Ref. [31], and is given by the analytical expression

\[
V_c^\text{int} = \frac{U}{4},
\]

where 4 corresponds to the number of nearest neighbors. This is plotted as a dashed grey line in Fig. 8. This result can be obtained by a simple comparison between the interaction energies of the Mott-insulating phase and of the checkerboard phase. In the full-fledged model, finite temperature and quantum tunneling have to be taken into account. In the quantum case, the bandwidth renormalization (proportional to \( \delta t \)) stays relevant in the vicinity of the charge-ordering transition even at large \( U \).

\[ \delta t = \frac{t V}{2U - V}. \]

Thus \( \delta t \) may become arbitrarily large if the nonlocal interaction coefficient exceeds twice the value of the local one. However, even disregarding the fact that the generic case is certainly the opposite one (local interactions in general exceed nonlocal ones), one should be aware of the fact that in that case the Hubbard-I approximation, which is justified in the strong-coupling limit, would no longer be appropriate.

By inspecting the phase diagram in Fig. 8, we can parametrize the phase boundary in the large-\( U \) limit as a constant slope, i.e., \( V_c = aU + b \). [Within the \( U \) range that we can simulate (we performed measurements up to \( U = 8.0 \)), we can estimate \( a = 1.25 \) and \( b = -2.5 \).] With this parametrization, we obtain

\[
\delta t_c = \frac{t(aU + b)}{2U - aU - b}. \]

Hence, the bandwidth renormalization (proportional to \( \delta t \)) stays relevant in the vicinity of the charge-ordering transition even at large \( U \).

\section{V. Beyond GW+EDMFT: Comparison to Dual Bosons and Trilex}

As mentioned in the introduction, the EDMFT formalism suffers from certain conceptual problems, such as the lack of thermodynamic consistency and an unreliable description of collective modes. These shortcomings are alleviated in the recently developed dual-boson (DB) method [34], which, in its full-fledged implementation [35,36], computes the susceptibility [11,12] after resumming an infinite number of ladder diagrams built from local impurity four-leg vertices.

These four-leg vertices, which are also central to the dynamical vertex approximation [37–42] (which was recently shown to be a simplified version of the QUADRUply Irreducible Local EXpansion (QUADRILEX), a method consisting of an atomic approximation of the four-particle irreducible functional [43]), can nonetheless only be obtained at a considerable computational expense and require a proper parametrization and treatment of their asymptotic behavior [44–47]. Consequently, it is not possible to use them routinely in multiorbital calculations (see, however, Refs. [48,49]), and lightweight improvements on EDMFT, especially with realistic applications in mind, are
Let us start with the small-$U$ limit. For $U < 1.5$, the phase boundaries obtained in all the $GW+\text{EDMFT}$ as well as $GW$ alone are almost indistinguishable. The $GW$ transition is a straight line for all shown values, and the variants with an impurity polarization have varying degrees of upward curvature, with the $GW+\text{EDMFT}$ line in between the $GW_c+\text{EDMFT}$ and the $GV+\text{EDMFT}$ line. For $GV+\text{EDMFT}$ it is notable that the $U \to 0$ limit does not reproduce the $GW$ value.

The dual-boson lines start with a similar upwards trend, with the exception of the $DB-GW$ line, which follows essentially the $\text{HS-V}$ variant of $GW+\text{EDMFT}$ [more properly denoted as $GD + \text{-second-order perturbation theory (SOPT) + EDMFT}$] as shown in Ref. [52], and discussed in detail at the end of this section. Yet, the dual-boson variants start out with a lower slope, indicating stronger ordering tendencies already for the lowest values of $U$, while all $GW+\text{EDMFT}$ variants follow the slope of the “weak-coupling” $GW$ boundary in the vicinity of $U = 0$. Interestingly, a very recent cluster-EDMFT study [57] reports a similarly reduced slope in the weak-$U$ regime. We also note that only the full $DB$ critical line is above the $U/4$ line (dashed grey line, discussed in Sec. IV), while the (non-self-consistent) $DB-GW$ and $DB-GW_\gamma$ results are not (the latter only slightly so). Further comparisons of $DB$ with $GW+\text{EDMFT}$ (in the $\text{HS-V}$ decoupling) can be found in Ref. [52].

For $U > 2$, the phase boundary for $GW+\text{EDMFT}$ is below the dual-boson phase boundary, while $GW$ is even a bit lower. For $GW_c+\text{EDMFT}$, the $U = 2.5$ point already falls in the Mott-insulating phase and is not shown in the comparison. The too-low Mott-transition line for $GW_c+\text{EDMFT}$ is not surprising, since it lacks the band widening of the nonlocal Fock term.

Two possible decouplings were previously discussed in the literature: “HS-UV” (giving rise to $GW+\text{EDMFT}$) and “HS-V” (resulting in a combined “GD+SOPT+EDMFT” scheme, where $D$ is the screened nonlocal interaction; see Ref. [15]). We emphasize that, contrary to the $\text{HS-UV}$ variant and, e.g., the random-phase approximation (RPA), the $\text{HS-V}$ variant does not resum, in the nonlocal part of the self-energy, the local ($U$) and nonlocal ($V$) parts of the interaction to the same order. (They are resummed, respectively, to second and infinite order.) This arbitrary inconsistency raises questions concerning the soundness of the $\text{HS-V}$ scheme, as already pointed out in Ref. [15].

Recent works have indeed confirmed the deficiency of $\text{HS-V}$. For instance, all the $GW+\text{EDMFT}$–related variants shown in Fig. 7 of Ref. [52], which were obtained in an $\text{HS-V}$ flavor, yield phase boundaries which are much lower than either $DB$ or the $GW+\text{EDMFT}$ results in Fig. 9 (which correspond to the $\text{HS-UV}$ variant of the decoupling of the interaction). The only additional outlier is the simplest $DB$ type of approximation, the $DB-GW$ variant, for which Ref. [52] showed that it is formally similar to an $\text{HS-V}$ calculation. More recently, Ref. [57] used cluster dynamical mean field theory to study the extended Hubbard model, which allows a control on errors by increasing the size of the cluster (but neglecting intercluster interactions, which in the case of $\text{EDMFT}$ and $GW+\text{EDMFT}$ are treated via the retarded impurity interactions). These cluster results were shown to be in poor agreement with $\text{HS-V}$, but very close to the full-fledged $DB$ method. In the lower panel of Fig. 9, we show

1Note that the dual-boson variants and $\text{HS-V}$ calculations of Ref. [52] have been executed as a single lattice self-consistency iteration on top of the converged $\text{EDMFT}$ solution.
that the \(GW+\)EDMFT (HS-\(U/V\)) method yields a critical \(V_c\) in agreement (with a 20% accuracy or better) with the cluster results, a remarkable result in view of the reduced numerical cost of this method compared to cluster DMFT. Comparisons for the larger \(U\) values in Fig. 9 would be of great interest.

We end this section by examining two further questions, namely the influence of spin fluctuations and the impact of local vertex corrections. One can expect that neither is important for the charge-ordering instability under study, since (i) this is an instability in the charge channel, not the spin channel, and (ii) as \(V\) increases towards charge ordering, the effective static interaction \(U(\omega = 0)\) decreases to zero [15], making the system behave more and more like a weakly correlated metal, where vertex corrections are expected to be small.

In all previous implementations of the \(GW+\)EDMFT method, the interaction was formally decoupled in the charge channel only, neglecting the possible influence of spin fluctuations. Furthermore, in \(GW+\)EDMFT, the influence of the local vertex on the nonlocal self-energy is included only through the nonlocal Green’s function. In the TRILEX approximation, both charge and spin fluctuations are taken into account, as well as local vertex corrections to the nonlocal self-energy.

We can thus answer both questions of interest by implementing the TRILEX method for the extended Hubbard model. We refer the reader to Refs. [50,51] for implementation details. The only difference with respect to the application to the Hubbard model is that Eq. (41) of Ref. [51] must be modified to also describe nonlocal interactions, which means that Eq. (61b) of that publication becomes

\[
W^\eta(q,i\Omega) = \frac{v^\eta(q)}{1 - v^\eta(q)P^\eta(q,i\Omega)},
\]

where \(\eta\) denotes the charge (ch) or spin (sp) channel and

\[
v^{\text{ch}}(q) = U^{\text{ch}} + 2V[\cos(q_x) + \cos(q_y)],
\]

\[
v^{\text{sp}}(q) = U^{\text{sp}},
\]

and the bare on-site interactions in the charge and spin channels are parametrized, in the so-called Heisenberg decoupling [51], by a parameter \(\alpha\):

\[
U^{\text{ch}} = (3\alpha - 1)U, \quad U^{\text{sp}} = (\alpha - 2/3)U.
\]

In Fig. 10, we show TRILEX results for two characteristic points of the phase diagram, namely \(U = 1\) (characteristic of the metallic phase) and \(U = 3\) (characteristic of the Mott phase). First, we observe that the critical \(V_c\) (computed by looking for a vanishing inverse static susceptibility, shown in the left panels) is quite close to that of \(GW+\)EDMFT, justifying our \textit{a priori} intuition. This agreement is quite remarkable, since \(GW+\)EDMFT has only charge fluctuations, while TRILEX has both charge and spin fluctuations. Second, \(V_c\) only mildly depends on the ratio of the charge to spin fluctuations, as can be seen in the right panels, where quite large variations of \(U^{\text{ch}}\) (and correspondingly \(U^{\text{sp}}\)) lead to comparatively small variations in \(V_c\).

Taking inspiration from the comparison of the cluster extension of TRILEX with exact benchmark results for the two-dimensional Hubbard model [58] (there, one observes that whenever the TRILEX solution is close to the exact solution, the dependence on the decoupling is weak), this stability (compared to charge-only \(GW+\)EDMFT, and with respect to \(\alpha\)) can be used as a proxy for the quantitative robustness of the present \(GW+\)EDMFT results.

VI. CONCLUSION

In conclusion, we have shown that the nonlocal Fock term has a significant influence on the description of the charge fluctuations in the \(GW+\)EDMFT method, especially in the strong-coupling limit. By effectively enhancing the bandwidth, it lowers the critical value of the nonlocal interaction for the charge-ordering transition. We have also shown that the differences between the EDMFT and \(GW+\)EDMFT phase diagrams are to a large extent a consequence of the nonlocal Fock term, which is not included in EDMFT.

Another interesting result is that the simple extension from EDMFT to a \(Gv+\)EDMFT formalism yields results similar to the full-fledged \(GW+\)EDMFT method. This suggests the possibility of studying complex multiband materials, where a full \(GW+\)EDMFT computation would be too costly, using techniques in the spirit of the recent screened exchange + dynamical DMFT (SEx+DDMFT) method [26,27,59]. In realistic materials, the simple single-band description is not sufficient, and substantial screening effects resulting from the presence of higher-energy degrees of freedom must be taken into account [60–62].

Performing a self-consistent calculation of the screening by these higher-energy states is, however, computationally expensive, even within a multiterm approach [56], where the updates are restricted to an intermediate energy window.
A scheme which combines a properly renormalized band structure with a self-consistent treatment of screening effects within the low-energy subspace may provide a good basis for tractable, but still accurate, first-principles electronic structure methods for correlated electron materials.

ACKNOWLEDGMENTS

We acknowledge useful discussions with Y. Nomura, A. I. Lichtenstein, E. Stepanov, and E. van Loon. We thank A. Huber and A. I. Lichtenstein for providing us the dual-boson data for Fig. 9, as well as H. Terletska for providing us the DCA data useful insights into the nontrivial nature of the large-$U$ local self-energy and any nonlocal screening, but they provide with tractable, but still accurate, within the low-energy subspace may provide a good basis for methods for correlated electron materials.

APPENDIX: ESTIMATION OF THE BANDWIDTH WIDENING WITH COMBINED HUBBARD-I AND FOCK TREATMENT

In this Appendix we discuss a simple Hubbard-I (plus Fock)--type treatment of the $U$-$V$ model. These arguments are not meant to be exact or comprehensive; most notably we ignore the effect of the nearest-neighbor interaction on the local self-energy and any nonlocal screening, but they provide useful insights into the nontrivial nature of the large-$U$ and large-$V$ limit.

We start with an approximation to the self-energy which follows the spirit of the Hubbard-I approximation by taking the atomic $\frac{U^2}{4\varepsilon}$ self-energy locally, but goes beyond it by taking also the instantaneous nonlocal Fock contribution into account:

$$\Sigma(k,z) = \frac{U^2}{4\varepsilon} + 2\delta t (\cos k_x + \cos k_y).$$

The corresponding Green’s function reads

$$G(k,z) = \frac{1}{z - \tilde{\varepsilon}_k - \frac{U^2}{4\varepsilon}}$$

with $\tilde{\varepsilon}_k$ denoting the effective dispersion including the Fock term,

$$\tilde{\varepsilon}_k \equiv 2(t + \delta t)(\cos k_x + \cos k_y).$$  \hspace{1cm} (A1)

Thus, we can write

$$G(k,z) = \frac{z}{[z - z_+(k)][z - z_-(k)]}$$  \hspace{1cm} (A2)

with

$$z_{\pm}(k) = \frac{\tilde{\varepsilon}_k \pm \sqrt{\tilde{\varepsilon}_k^2 + U^2}}{2}.$$  \hspace{1cm} (A3)

As expected, in the atomic limit ($\tilde{\varepsilon}_k \rightarrow 0$), the function has two peaks at $\pm U/2$, corresponding to the two Hubbard bands.

We can decompose the expression of Eq. (A2) as

$$G(k,z) = \frac{A_+(k)}{z - z_+(k)} + \frac{A_-(k)}{z - z_-(k)}$$  \hspace{1cm} (A4)

with

$$A_+(k) = \frac{1}{2} \left( 1 \pm \frac{\tilde{\varepsilon}_k}{\sqrt{\tilde{\varepsilon}_k^2 + U^2}} \right).$$  \hspace{1cm} (A5)

$A_+$ and $A_-$ are the weights of the upper and lower Hubbard bands, respectively. Using Eq. (A4), one writes the spectral function as

$$A(k,\omega) = A_+(k) \pi \delta(\omega - z_+(k)) + A_-(k) \pi \delta(\omega - z_-(k)).$$

Under the assumption that the Hubbard bands are well separated ($U$ large enough), only the lower Hubbard band contributes to the occupancy (at $T = 0$ for simplicity):

$$n_k = \int_0^0 \frac{d\omega}{\pi} A(k,\omega) \approx A_-(k) \approx \frac{1}{2} \left( 1 - \frac{\tilde{\varepsilon}_k}{U} \right).$$  \hspace{1cm} (A6)

In the second equality, we have again used the fact that $U$ is large enough (to neglect $\tilde{\varepsilon}_k^2$ in the square root).

On the other hand, the occupancy is related to $G(k,\tau = 0^+)$ in the following way:

$$n_k = 1 + G_k(\tau = 0^+).$$  \hspace{1cm} (A7)

Lastly, $\delta t$ [in Eq. (A1)] is also known, in the Fock approximation, as a function of $G_{ij}(\tau = 0^+)$ [see Eq. (13)]:

$$\delta t = -\frac{t^2}{V} G_{ij}(\tau = 0^+).$$  \hspace{1cm} (A8)

Putting Eqs. (A6)–(A8) together and Fourier transforming, one gets

$$G_{ij}(\tau = 0^+) = \frac{1}{U} \left( -\frac{t}{V} - \frac{t V G_{ij}(\tau = 0^+)}{2U - V} \right).$$  \hspace{1cm} (A9)

Solving for $G_{ij}(\tau = 0^+)$, one gets

$$G_{ij}(\tau = 0^+) = -\frac{t}{2U - V}.$$  \hspace{1cm} (A10)

and

$$\delta t = \frac{t V}{2U - V}.$$  \hspace{1cm} (A11)

This expression is used in Sec. IV.
