Moment-functional based spectral density-functional theory

Frank Freimuth\textsuperscript{1,2}, Stefan Blügel\textsuperscript{1}, and Yuriy Mokrousov\textsuperscript{1,2}
\textsuperscript{1}Peter Grünberg Institut and Institute for Advanced Simulation, Forschungszentrum Jülich and JARA, 52425 Jülich, Germany
\textsuperscript{2}Institute of Physics, Johannes Gutenberg University Mainz, 55099 Mainz, Germany

We describe a density-functional method which aims at computing the ground state electron density and the spectral function at the same time. One basic ingredient of our method is the construction of the spectral function from the first four spectral moment matrices. The second basic ingredient is the construction of the spectral moment matrices from density functionals. We call our method moment-functional based spectral density-functional theory (MFbSDFT), because it is based on density-functionals for the spectral moments and because it allows us to compute the spectral function. If it is implemented in second variation our method consumes only a fraction more computer time than a standard DFT calculation with the PBE functional. We show that MFbSDFT captures correlation effects such as the valence-band satellite in Ni and the formation of lower and upper Hubbard bands in SrVO\textsubscript{3}. For the purpose of constructing the spectral function from the first four $N \times N$ spectral moment matrices we describe an efficient algorithm based on the diagonalization of one hermitean $2N \times 2N$ matrix.

I. INTRODUCTION

In density-functional theory (DFT) the ground state electron density is determined by minimizing the total energy functional \cite{2008_Sokol_Book}. While most contributions to the total energy functional, such as the Hartree energy, the exchange energy, and the correlation energy, can be expressed as functionals of the electron density, it is difficult to express the kinetic energy directly in this way. This is why within the most popular kind of DFT – the Kohn-Sham (KS) DFT – the KS-Hamiltonian \cite{1965_Kohn-Sham} is set up and solved with the main purpose to provide the kinetic energy.

However, the KS energy bands agree very often fairly well with photoemission data \cite{2010_Hanke} and the KS spectrum is therefore even used to compute response properties such as the anomalous Hall effect \cite{2011_Thiaville}, the Gilbert damping \cite{2012_Dybbs}, the direct and inverse spin-orbit torque \cite{2014_Rong}, and the inverse Faraday effect \cite{2015_Vogt} in metallic systems. These KS response functions are often in good agreement with the corresponding material property tensors measured experimentally.

Well-known deficiencies of this approach are the underestimation of the band gap, which may require the application of band shifts when computing optical responses such as photocurrents \cite{2010_Hanke} in semiconductors such as GaAs. Instead of shifting the bands to match the band gap known from experiments, one may use the GW approximation \cite{2010_Savrasov}, which is a parameter-free technique based on many-body perturbation theory and which often predicts gaps that are closer to experiment than KS-DFT. However, since one deals then directly with a many-body Hamiltonian, one forsakes the DFT idea of obtaining all properties as directly as possible from the ground state density in order to avoid the complexity and factorial growth of the many-body Hilbert space. Another short-coming of KS-spectra is the overestimation of the magnetic moment and the resulting overestimation of the exchange splitting of some weak itinerant ferromagnets such as MnSi, which requires us to reduce the exchange field by a scaling factor in order to compute the topological Hall effect in MnSi \cite{2016_Xu}.

Moreover, a well-known deficiency of the KS spectrum is the absence of the splitting of bands into lower and upper Hubbard bands due to strong electron correlations \cite{1982_Millis}. Such a splitting of the single-particle bands leads for example to the appearance of a satellite peak roughly 6 eV below the Fermi energy in Ni \cite{2012_Rong, 2014_Rong}. In order to compute the spectrum in such cases of strongly interacting electrons one often uses DFT only to obtain the KS wavefunctions of a small manifold corresponding to a small energy window around the Fermi energy and constructs an interacting Hamiltonian for this manifold, which one solves by dynamical mean-field theory (DMFT) \cite{2000_Savrasov, 2003_Maier} in order to obtain the spectral function. In other words, one remains within the DFT-framework in order to determine the ground state density, but similarly to GW one leaves this framework and directly solves an interacting many-electron Hamiltonian in order to obtain the spectrum of the correlated system instead of evaluating a density functional. However, one may also take a different viewpoint: The local spectral function of DMFT minimizes the effective action. In this sense DMFT is a spectral density-functional approach \cite{2003_Maier}.

Nevertheless, the question still poses itself if it is possible to obtain both ground state density and correlated spectral function within a density-functional approach which avoids the direct use of many-body techniques such as GW and DMFT. A hermitean $N \times N$ matrix has $N$ real-valued eigenvalues. This well-known fact from linear algebra is exploited in many electronic structure programs based on density-functional theory, where the KS equations are solved numerically by diagonalizing a hermitean matrix. The direct construction of $2N$ state
we present applications of
we explain how the MFbSDFT method may
VII
VIII
II. THEORY
A. The concept of moment functionals

The ground state charge density defines the Hamiltonian uniquely (up to a constant) \[1\]. Consequently, it determines also the spectral function uniquely. In order to write the spectral function in matrix form we need a suitable set of orthonormal basis functions \(\phi_n(r)\). Denoting the creation and annihilation operators corresponding to state \(\phi_n(r)|\sigma\rangle\) – where \(|\sigma\rangle\) is a spinor – by \(c^\dagger_{\sigma n}\) and \(c_{\sigma n}\), respectively, the matrix elements of the spectral function matrix are

\[
S_{\sigma nm}(E) = \frac{1}{2\pi} \int d t e^{iE t} \langle [c_{\sigma n}(t), c^\dagger_{\sigma m}]^+ \rangle.
\]
When periodic boundary conditions are used, the spectral function and the spectral moments acquire an additional \( k \)-index for the \( k \) point \( k \), which we often suppress in this manuscript for notational convenience.

The spectral moments may be obtained by plugging Eq. 2 into Eq. 1. Since the spectral function is uniquely determined by the ground state density, also the spectral moments are uniquely defined by it. The spectral moments may be expressed in terms of real-space coordinates:

\[
M^{(I)}_{\sigma}(r, r') = \frac{1}{R} \int dE \int \sum_{nm} S_{\sigma nm}(E) \phi_n(r) \phi_m^*(r').
\]

We may consider \( M^{(I)}_{\sigma}(r, r') \) as a non-local potential, from which we may obtain the spectral moment matrices by computing the matrix elements:

\[
M^{(I)}_{\sigma \sigma'} = \int d^3 r d^3 r' M^{(I)}_{\sigma}(r, r') \phi^*_n(r) \phi_m(r').
\]

According to our arguments above, the non-local potentials \( M^{(I)}_{\sigma}(r, r') \) are unique functionals of the electron density.

In KS-DFT the total energy functional is split into the kinetic energy, the Hartree energy, and the exchange-correlation energy \[2\]. The kinetic energy is computed from the KS-wavefunctions, the Hartree energy is computed from the charge density, and for the exchange-correlation energy one often uses analytical expressions in terms of the charge density, which have been derived for the uniform electron gas \[25\]. Similarly, the potentials \( M^{(I)}_{\sigma}(r, r') \) (\( I = 1, 2, 3 \) if we use the first four moments) contain contributions from the kinetic energy and from the Hartree term. In the following section we show that these contributions may be identified and separated from a remainder, which thus plays a similar role in MFBSDFT like the exchange-correlation potential does in KS-DFT. We expect that useful expressions for this remainder can be found by evaluating it for the uniform electron gas.

B. Explicit expressions for the moments

We consider the Hamiltonian

\[
H = \sum_{\sigma nm} T_{nm} c^\dagger_{\sigma n} c_{\sigma m} + \frac{1}{2} \sum_{\sigma \sigma' nm'n'} V_{nmn'm'} c^\dagger_{\sigma n'} c_{\sigma' m'} c_{\sigma m} c^\dagger_{\sigma n},
\]

where

\[
T_{nm} = \int d^3 r \phi^*_n(r) \left(-\frac{1}{2} \Delta + V(r)\right) \phi_m(r)
\]

and

\[
V_{nmn'm'} = \int d^3 r_1 d^3 r_2 \frac{\phi^*_n(r_1) \phi^*_m(r_2) \phi_n(r_2) \phi_m^*(r_1)}{|r_1 - r_2|},
\]

and \( V(r) \) is the lattice potential. Note that in the entire Sec. IIIB we use Hartree atomic units for notational convenience.

Many-body approaches such as LDA+DMFT often take into account the Coulomb matrix element \( V_{nmn'm'} \) only when all orbitals, i.e., \( n, m, n' \) and \( m' \), describe the same crystal lattice site. In the simplest approximation \( V_{nmn'm'} \) is described by a single parameter, the so-called Hubbard-\( U \). Components of \( V_{nmn'm'} \) that are neglected hereby are of course partly treated in LDA+DMFT, because the lattice potential \( V(r) \) is replaced by the KS potential in this case. Therefore, the Hubbard-\( U \) only describes the Coulomb interaction from strong localization of electrons. These effects are underestimated by KS-DFT and become important when \( U \) approaches or exceeds the bandwidth. In contrast, we do not restrict \( V_{nmn'm'} \) at this point, i.e., both local and non-local contributions are described by it in MFBSDFT and \( V(r) \) is the pure lattice potential without exchange-correlation terms.

The zeroth moment is given by

\[
M^{(0)}_{\sigma \sigma'} = \langle \{ c_{\sigma n} c^\dagger_{\sigma m} \} \rangle = \delta_{nm},
\]

where \( \{ \ldots \} \) denotes the anticommutator, and the first moment evaluates to

\[
M^{(1)}_{\sigma \sigma'} = \langle \{ c_{\sigma n}, H \} c^\dagger_{\sigma m} \rangle = T_{nm} + \sum_{n'n'm'} V_{nn'n'm'} c^\dagger_{\sigma n'} c_{\sigma m'}
\]

\[-\sum_{n'n'm'} V_{nn'n'm'} c^\dagger_{\sigma n'} c_{\sigma m'}.
\]

Defining the Hartree potential by

\[
V^H(r) = \sum_{\sigma' n'm'} \int d^3 r_2 \frac{\phi^*_n(r_1) \phi^*_m(r_2) \phi_{\sigma n}(r_2) \phi_{\sigma' m'}(r_1)}{|r_1 - r_2|} c^\dagger_{\sigma n} c_{\sigma' m'}
\]

and the non-local exchange potential by

\[
V^X_{\sigma}(r_1, r_2) = -\sum_{n'm'} \frac{\phi^*_n(r_2) \phi_m(r_1)}{|r_1 - r_2|} c^\dagger_{\sigma n} c_{\sigma' m'}
\]

we may write the first moment as

\[
M^{(1)}_{\sigma \sigma'} = T_{nm} + V^H_{nm} + V^X_{\sigma nm} = M^{HF}_{\sigma \sigma'},
\]

where \( V^H_{nm} \) and \( V^X_{\sigma nm} \) are the matrix elements of the Hartree potential and of the non-local exchange potential, respectively. Thus, one obtains a method of Hartree-Fock (HF) type if one considers only the first two moments. It differs from the exact Hartree-Fock method by the self-interaction error (SIE) \[28\] (see also Sec. VII).
for a brief discussion of SIE from the perspective of MFbSDFIT). Therefore, we introduced the alternative label \( \mathcal{M}^{\text{HF}}_{\sigma \sigma m} \) for the first moment, which expresses concisely what this first moment contains. Instead of using the non-local exchange potential Eq. (11) one may use the local exchange potential \( \mathcal{V}^{\text{loc X}}_{\sigma} = \frac{\partial}{\partial n(r)} [n(r) e^X(n(r))] \),

\[
V^{\text{loc X}}_{\sigma}(r) = \frac{\partial}{\partial n(r)} [n(r) e^X(n(r))], \tag{13}
\]

where

\[
n(r) = \sum_{\sigma nm} \phi^\dagger_\sigma(r)\phi_\sigma(r) \langle c_{\sigma n}^\dagger c_{\sigma m} \rangle \tag{14}
\]

is the electron density at position \( r \) and \( e^X(n(r)) \) is the exchange energy density for electron density \( n(r) \). The local potential Eq. (13) has the advantage that it is computationally very cheap to evaluate in contrast to the non-local version Eq. (11). However, hybrid density functionals, which admix exact exchange, are often more precise than density functionals that use only the local approximation Eq. (13). Fortunately, one may reduce the computational burden of exact exchange by screening the Coulomb potential \( \mathcal{P} \). In the numerical calculations in this work we will use only the local expression Eq. (13), but, similar to KS-DFT, we expect that the precision of the MFbSDFT approach can be increased by avoiding the approximation of non-local potentials by local potentials. We leave it for future work to explore how the MFbSDFT approach may be combined with non-local potentials.

In Ref. [17] we explain that for independent electrons the spectral moment matrices commute, i.e.,

\[
[M^{(I)}_{\sigma}, M^{(J)}_{\sigma}]_\pm = 0 \tag{15}
\]

for all \( I \) and \( J \), and that the eigenvalues of the spectral moment matrix \( M^{(I)}_{\sigma} \) are simply the eigenvalues of the single-particle Hamiltonian raised to the \( I \)-th power, i.e., \( (E_{\sigma n})^I \). For correlated electrons this is not the case [17]. However, we may expect that the moment \( M^{(I)}_{\sigma} \) contains a term \( [M^{\text{HF}}_{\sigma}, \mathcal{P}]_I \), because there may be cases where Hartree-Fock provides an excellent description because correlation effects are small, and these special cases have to be accommodated by the general theory. We may therefore expect that the second moment should contain a term

\[
\mathcal{M}_{\sigma}^{\text{HF}} \mathcal{M}_{\sigma}^{\text{HF}} = TT + TV^H + TV^X + V^H T + V^H V^H + + V^H V^X + V^X T + V^X V^H + V^X V^X , \tag{16}
\]

which is indeed what we find. We may use this observation to split the second moment into the anticipated part \( \mathcal{M}_{\sigma}^{\text{HF}} \mathcal{M}_{\sigma}^{\text{HF}} \) plus additional new terms, which we denote by \( \mathcal{M}_{\sigma}^{(2+)} \), i.e.,

\[
\mathcal{M}_{\sigma}^{(2)} = \mathcal{M}_{\sigma}^{\text{HF}} \mathcal{M}_{\sigma}^{\text{HF}} + \mathcal{M}_{\sigma}^{(2+)} . \tag{17}
\]

In contrast to the single-band Hubbard model with onsite Coulomb interaction, where higher-order correlation functions appear in the third moment and in the higher moments, already the second moment \( \mathcal{M}_{\sigma}^{(2)} \) of the many-band case with the full Coulomb interaction contains the higher-order correlation function \( \langle c_{\sigma n}^\dagger c_{\sigma m} c_{\sigma' m'} c_{\sigma' m''} \rangle \). In order to identify the terms \( V^H V^H, V^H V^X, V^X V^H, \) and \( V^X V^X \) predicted by Eq. (10) we need to evaluate \( \langle c_{\sigma n}^\dagger c_{\sigma m} c_{\sigma n'} c_{\sigma' m'} \rangle \) in perturbation theory. In contrast, the terms \( TT, TV^H, TV^X, V^H T, \) and \( V^X T \) can be identified without using perturbation theory, because they appear with the correlation function \( \langle c_{\sigma n}^\dagger c_{\sigma m} c_{\sigma' m''} \rangle \). The term \( TT \) appears even without any correlation function. Therefore, we define

\[
\langle c_{\sigma n}^\dagger c_{\sigma m} c_{\sigma n'} c_{\sigma' m'} \rangle = \langle c_{\sigma n}^\dagger c_{\sigma m} c_{\sigma n'} c_{\sigma' m'} \rangle - \langle c_{\sigma n}^\dagger c_{\sigma m} c_{\sigma n'} \rangle \langle c_{\sigma' m'} \rangle , \tag{18}
\]

The idea behind Eq. (18) is that the diagrammatic expression of \( \langle c_{\sigma n}^\dagger c_{\sigma m} c_{\sigma n'} c_{\sigma' m'} \rangle \) as obtained within perturbation theory contains terms that may be written as \( \langle c_{\sigma n} c_{\sigma n'} \rangle \langle c_{\sigma' m'} \rangle \) and \( -\langle c_{\sigma n} c_{\sigma n'} \rangle \langle c_{\sigma' m'} \rangle \). Since these latter two terms occur sometimes in \( \mathcal{M}_{\sigma}^{\text{HF}} \mathcal{M}_{\sigma}^{\text{HF}} \) we introduce the notation of Eq. (18) in order to split \( \mathcal{M}_{\sigma}^{(2)} \) into \( \mathcal{M}_{\sigma}^{\text{HF}} \mathcal{M}_{\sigma}^{\text{HF}} + \mathcal{M}_{\sigma}^{(2+)} \). Using this notation we may write \( \mathcal{M}_{\sigma}^{(2+)} \) as a sum of 17 terms:

\[
M_{\sigma mn n' m'}^{(2+)} = \sum_{n'' m'' t' t'} V_{n n'' t' t''} V_{t' m' n''} \langle c_{\sigma n}^\dagger c_{\sigma m}^\dagger c_{\sigma' m'} \rangle , \tag{19}
\]

which is spin-independent,

\[
M_{\sigma nm n' m'}^{(2+)} = - \sum_{n'' m'' t'} V_{n n'' t' t''} V_{t' m' n''} \langle c_{\sigma n}^\dagger c_{\sigma m}^\dagger c_{\sigma' m'} \rangle , \tag{20}
\]

which does not depend on the spin,

\[
M_{\sigma nm n' m'}^{(2+)} = - \sum_{n'' m'' t' t''} V_{n n'' t' t''} V_{t' m' n''} \langle c_{\sigma n}^\dagger c_{\sigma m}^\dagger c_{\sigma' m'} c_{\sigma t} c_{\sigma' t'} \rangle , \tag{21}
\]

\[
M_{\sigma nm n' m'}^{(2+)} = \sum_{n'' m'' t' t''} V_{n n'' t' t''} V_{t' m' n''} \langle c_{\sigma n}^\dagger c_{\sigma m}^\dagger c_{\sigma' m'} c_{\sigma t} c_{\sigma' t'} \rangle , \tag{22}
\]

\[
M_{\sigma nm n' m'}^{(2+)} = \sum_{n'' m'' t' t''} V_{n n'' t' t''} V_{t' m' n''} \langle c_{\sigma n}^\dagger c_{\sigma m}^\dagger c_{\sigma' m'} c_{\sigma t} c_{\sigma' t'} \rangle , \tag{23}
\]

\[
M_{\sigma nm n' m'}^{(2+)} = \sum_{n'' m'' t' t''} V_{n n'' t' t''} V_{t' m' n''} \langle c_{\sigma n}^\dagger c_{\sigma m}^\dagger c_{\sigma' m'} c_{\sigma t} c_{\sigma' t'} \rangle , \tag{24}
\]

\[
M_{\sigma nm n' m'}^{(2+)} = - \sum_{n'' m'' t' t''} V_{n n'' t' t''} V_{t' m' n''} \langle c_{\sigma n}^\dagger c_{\sigma m}^\dagger c_{\sigma' m'} c_{\sigma t} c_{\sigma' t'} \rangle , \tag{25}
\]

\[
M_{\sigma nm n' m'}^{(2+)} = - \sum_{n'' m'' t' t''} V_{n n'' t' t''} V_{t' m' n''} \langle c_{\sigma n}^\dagger c_{\sigma m}^\dagger c_{\sigma' m'} c_{\sigma t} c_{\sigma' t'} \rangle . \tag{26}
\]
\[ M^{(2+9)}_{\sigma nm} = - \sum_{n'm't't'z} V_{nn't't'} V_{m't'zm} \langle \epsilon_{\sigma n}^t, \epsilon_{\sigma m}^t, \epsilon_{\sigma t'} c_{-\sigma z} \rangle, \] (27)

\[ M^{(2+10)}_{\sigma nm} = - \sum_{n'm't't'z} V_{nn't't'} V_{m't'zm} \langle \epsilon_{\sigma n}^t, \epsilon_{\sigma m}^t, \epsilon_{\sigma t'} c_{-\sigma z} \rangle, \] (28)

\[ M^{(2+11)}_{\sigma nm} = \sum_{n'm't't'z} V_{nn't't'} V_{m't'zm} \langle \epsilon_{\sigma n}^t, \epsilon_{\sigma m}^t, \epsilon_{\sigma t'} c_{-\sigma z} \rangle, \] (29)

\[ M^{(2+12)}_{\sigma nm} = \sum_{n'm't't'z} V_{nn't't'} V_{m't'zm} \langle \epsilon_{\sigma n}^t, \epsilon_{\sigma m}^t, \epsilon_{\sigma t'} c_{-\sigma z} \rangle, \] (30)

\[ M^{(2+13)}_{\sigma nm} = \sum_{n'm't't'z} V_{nn't't'} V_{m't'zm} \langle \epsilon_{\sigma n}^t, \epsilon_{\sigma m}^t, \epsilon_{\sigma t'} c_{-\sigma z} \rangle, \] (31)

\[ M^{(2+14)}_{\sigma nm} = - \sum_{\sigma' n'm't't'z} V_{nn't't'} V_{m't'zm} \langle \epsilon_{\sigma' n}^t, \epsilon_{\sigma m}^t, \epsilon_{\sigma t'}^c \rangle \langle \epsilon_{\sigma' m}^t, \epsilon_{\sigma t'}^c \rangle, \] (32)

which is spin-independent,

\[ M^{(2+15)}_{\sigma nm} = - \sum_{n'm't't'z} V_{nn't't'} V_{m't'zm} \langle \epsilon_{\sigma n}^t, \epsilon_{\sigma m}^t, \epsilon_{\sigma t'}^t \rangle \langle \epsilon_{\sigma m}^t, \epsilon_{\sigma t'}^t \rangle, \] (33)

\[ M^{(2+16)}_{\sigma nm} = \sum_{n'm't't'z} V_{nn't't'} V_{m't'zm} \langle \epsilon_{\sigma n}^t, \epsilon_{\sigma m}^t, \epsilon_{\sigma t'}^t \rangle \langle \epsilon_{\sigma m}^t, \epsilon_{\sigma t'}^t \rangle, \] (34)

and

\[ M^{(2+17)}_{\sigma nm} = \sum_{n'm't't'z} V_{nn't't'} V_{m't'zm} \langle \epsilon_{\sigma n}^t, \epsilon_{\sigma m}^t, \epsilon_{\sigma t'}^t \rangle \langle \epsilon_{\sigma m}^t, \epsilon_{\sigma t'}^t \rangle. \] (35)

In order to evaluate the contributions to \( M^{(2+)} \) in a way similar to Eq. (13), we suggest to consider the contractions

\[ C^{(2+)}_\sigma = \sum_{nm} M^{(2+)}_{\sigma nm} \langle \epsilon_{\sigma n}^t, \epsilon_{\sigma m}^t \rangle \] (36)

and to compute them for the uniform electron gas as a function of electron density. Similarly to Eq. (13) we assume that we may derive local potentials

\[ \Psi^{(2+)}_\sigma (r) = \frac{\partial}{\partial \phi_n^t (r)} \left[ C^{(2+)}_\sigma \right] \] (37)

from these contractions and compute the moments from these local potentials:

\[ M^{(2+)}_{\sigma nm} = \int d^3r \Psi^{(2+)}_\sigma (r) \phi_n^t (r) \phi_m^t (r). \] (38)

Many popular exchange-correlation potentials are constructed with the help of Green’s function Monte Carlo simulations of the energy of the uniform electron gas [25, 28], because the universality of the exchange correlation potential implies that it may be constructed from a uniform system. However, Green’s function Monte Carlo data are not yet available for our expressions Eq. (19) through Eq. (35). On the other hand, diagrammatic perturbation theory has been used to derive expressions for the energy of the uniform electron gas in the limit of low and high density and these results are considered in the construction of exchange correlation potentials as well [25, 28]. For the purpose of demonstrating the feasibility of the MFBSDFT approach it is sufficient to find simple approximate expressions for the contractions Eq. (36). Therefore, we evaluate these contractions for the uniform electron gas using perturbation theory in Appendix [13]. We leave it for future work to find accurate analytic representations of the contractions of Eq. (19) through Eq. (35) based on techniques such as Green’s function Monte Carlo simulations and diagrammatic expansions for the high-density limit.

Similar to Eq. (17), one may anticipate that the third moment may be decomposed as

\[ M^{(3)}_\sigma = M^{HF}_\sigma M^{HF}_\sigma M^{HF}_\sigma + M^{(3+)}_\sigma, \] (39)

where

\[ M^{HF}_\sigma M^{HF}_\sigma M^{HF}_\sigma = TTT + TTV^H + TTV^X + TV^HT + TV^HTV^H + TV^HTV^X + TV^XT + TV^XTV^X + TV^XTV^X + TV^XTV^X + TV^XTV^X + TV^XTV^X, \] (40)

which is indeed what we find: To identify \( TTT \) in \( M^{(3)}_\sigma \) one needs to check the terms without correlation functions. To find the terms that contain two factors of the matrix \( T \), i.e., the terms \( TV^HT, TV^HT, TV^HT, TV^HT, TV^HT \), \( TV^XT, TV^XT, TV^XT, TV^XT, TV^XT \), and \( TV^XT, TV^XT \), one needs to look out for the contributions to \( M^{(3)}_\sigma \) that contain the correlation function \( \langle \epsilon_{\sigma n}^t, \epsilon_{\sigma m}^t, \epsilon_{\sigma n'}^t \rangle \). To track down the terms that contain a single factor of the matrix \( T \), i.e., the terms \( TV^HT, TV^HT, TV^HT, TV^HT, TV^HT \), \( TV^XT, TV^XT, TV^XT, TV^XT, TV^XT \), and \( TV^XT, TV^XT \), one needs to find the contributions to \( M^{(3)}_\sigma \) that contain the correlation function \( \langle \epsilon_{\sigma n}^t \epsilon_{\sigma m}^t, \epsilon_{\sigma n'}^t \rangle \) and one has to evaluate this correlation function in perturbation theory. In order to identify all those terms in Eq. (40) that do not contain the matrix \( T \), i.e., the terms \( V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V^H + V}
and one has to evaluate this correlation function in perturbation theory.

When one evaluates the correlator \( \langle c_{\sigma m}^\dagger c_{\sigma' n}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \) in perturbation theory to extract the terms discussed above, e.g. \( V^H V T \), one may use Eq. \((18)\) like for the second moment. This procedure generates a group of terms in \( M_{3+} \) that contain \( \langle c_{\sigma m}^\dagger c_{\sigma' n}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \). Similarly, it is convenient to define

\[
\langle c_{\sigma m}^\dagger c_{\sigma' n}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle - \langle c_{\sigma m}^\dagger c_{\sigma' n}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \tag{41}
\]

and to use Eq. \((13)\) in order to replace the correlators of the type \( \langle c_{\sigma m}^\dagger c_{\sigma' n}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \) on the right-hand side of Eq. \((11)\) by \( \langle c_{\sigma m}^\dagger c_{\sigma' n}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \) and the simpler correlators \( \langle c_{\sigma m}^\dagger c_{\sigma' n}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \). When we use this procedure to express \( \langle c_{\sigma m}^\dagger c_{\sigma' n}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \) in terms of the correlators \( \langle c_{\sigma m}^\dagger c_{\sigma m'}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \) and thereby extract the terms discussed above, e.g. \( V^H V^H V^H \), we generate additional groups of terms in \( M_{3+} \), which contain \( \langle c_{\sigma m}^\dagger c_{\sigma' n}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \) or \( \langle c_{\sigma m}^\dagger c_{\sigma m'}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \).

The remaining contributions to \( M_{3+} \) may be split into groups of formally similar expressions. The first group of two terms in \( M_{3+} \) contains two matrices \( T \) and the correlation function \( \langle c_{\sigma m}^\dagger c_{\sigma' n}^\dagger \rangle \):

\[
M_{nm}^{3+} = \sum_{n' t' \sigma'} V_{nn't'} T_{tm} S_{\sigma' t' n'}, \tag{42}
\]

which is spin-independent, and

\[
M_{\sigma nm}^{3+} = - \sum_{n' t'} V_{nn't'} T_{tm} S_{\sigma' t' n'}, \tag{43}
\]

where

\[
S_{\sigma' t' n'} = \sum_{m'} \left[ T_{tm'} \langle c_{\sigma' n'}^\dagger c_{\sigma m'} \rangle - \langle c_{\sigma m}^\dagger c_{\sigma' n'} \rangle T_{m'n'} \right] \tag{44}
\]

is the commutator between the matrix \( T \) and the density matrix.

It is desirable to rewrite Eq. \((42)\) in a form that permits a numerically efficient evaluation of this term, because the direct computation of Eq. \((42)\) will often be numerically demanding due to the four indices of the Coulomb matrix element. We may exploit that the indices \( t' \) and \( n' \) couple to the Coulomb matrix element in a way that allows us to identify the density matrix. Therefore we may define

\[
N(r) = \sum_{n' t' \sigma'} \phi_{n'}^*(r) \phi_{t'}(r) S_{\sigma' t' n'}, \tag{45}
\]

from which we compute the Hartree-type integral

\[
\mathcal{F}(r) = \int d^3 r_2 \frac{N(r_2)}{|r - r_2|}, \tag{46}
\]

Eq. \((42)\) may now be written as

\[
M_{nm}^{3+} = \sum_t F_{nt} T_{tm}, \tag{47}
\]

where \( F_{nt} \) are the matrix elements of the Hartree-type potential \( \mathcal{F}(r) \).

Similarly, in order to evaluate Eq. \((43)\) we may exploit that the indices \( t' \) and \( n' \) couple to the Coulomb matrix element in a way that allows us to identify a non-local potential. Therefore we may define

\[
L_{\sigma}(r_1, r_2) = \sum_{n' t'} \frac{\phi_{n'}^*(r_1) \phi_{t'}(r_2)}{|r_1 - r_2|} S_{t' n'}. \tag{48}
\]

Eq. \((43)\) may now be written as

\[
M_{\sigma nm}^{3+} = - \sum_t L_{\sigma t} T_{tm}, \tag{49}
\]

where \( L_{\sigma t} \) are the matrix elements of the non-local Fock-type potential \( L_{\sigma}(r_1, r_2) \).

The next group of two terms in \( M_{3+} \) does not contain the matrix \( T \), but it contains the correlation function \( \langle c_{\sigma m}^\dagger c_{\sigma m'} \rangle \):

\[
M_{nm}^{3+} = \sum_{\sigma' n' t' z' z''} V_{n'n't'} V_{t't'z'z''} V_{zz'm'm'} \langle c_{\sigma' n'}^\dagger c_{\sigma' m'} \rangle, \tag{50}
\]

which does not depend on the spin, and

\[
M_{\sigma nm}^{3+} = - \sum_{n' t' z' z''} V_{n'n't'} V_{t't'z'z''} V_{zz'm'm'} \langle c_{\sigma m}^\dagger c_{\sigma m'} \rangle. \tag{51}
\]

In Appendix A we show that the sum of Eq. \((50)\) and Eq. \((51)\) turns into the very simple result \( U_{\langle n_\sigma m \rangle} \) in the case of the single-band Hubbard model. However, for realistic many-band systems, the direct evaluation of these expressions may be numerically demanding due to the four indices of the Coulomb matrix element. Therefore, we may try to use concepts from DFT to simplify the calculations. In order to approximate these contributions by local functionals, we may use the recipe described above in Eq. \((46)\), Eq. \((47)\), and Eq. \((48)\). In Appendix C we evaluate the corresponding contractions for the uniform electron gas.

\( M_{3+} \) contains several additional groups of terms that we have not discussed yet. One group of terms contains three factors of the Coulomb matrix element and the correlator \( \langle c_{\sigma m}^\dagger c_{\sigma m'}^\dagger c_{\sigma m'} c_{\sigma' n'} \rangle \), but the indices are not connected in a way that terms such as \( HHH \) or \( HXXH \) arise, which we have already discussed above.
An example from this group of terms is
\[
M^{(3+5)}_{\sigma\sigma' m} = - \sum_{n't'\ u'w' z' m'} \sum_{n' w' z' m'} V_{m' n't' t'} V_{m' u' w' v' z' m'} \times \langle c_{\sigma n\tau} c_{\sigma' z' \sigma'_t} c_{\sigma u\sigma w'} \rangle.
\]
(52)

The first index \( t' \) of \( V_{n't'u'w'} \) is shared with \( V_{n't'u'w'} \), while the second index \( m' \) of \( V_{n'm'w'} \) is shared with \( V_{z'z'm'} \). The last two indices, \( u' \) and \( w' \) are contracted with the correlation function \( \langle c_{\sigma n\tau} c_{\sigma' z' \sigma'_t} c_{\sigma u\sigma w'} \rangle \). In this term, it is therefore not possible to express \( V_{n't'u'w'} \) through the matrices \( H \) or \( X \) when perturbation theory is used.

Other terms in this group differ from Eq. (52) for example due to different spin quantum numbers in the correlation function, e.g. \( \langle c_{-\sigma n\tau} c_{\sigma' z' \sigma'_t} c_{\sigma u\sigma w'} \rangle \), or they differ due to different indices of the Coulomb matrix elements, e.g. \( V_{n't'u'w'} \times V_{n'm't'w'} \).

There is a second group of terms that contains three factors of the Coulomb matrix element as well. However, it contains the correlator \( \langle c_{\sigma n\tau} c_{\sigma' z' \sigma'_t} c_{\sigma u\sigma w'} \rangle \) instead. An example from this group of terms is
\[
M^{(3+6)}_{\sigma\sigma' m} = - \sum_{n't'\ u'w' z' m'} \sum_{n' w' z' m'} V_{m' n't' t'} V_{n' u' w' v' z' m'} \times \langle c_{\sigma n\tau} c_{\sigma' z' \sigma'_t} c_{\sigma u\sigma w'} \rangle.
\]
(53)

Since \( V_{n't'z'} \) couples to the correlation function only through the index \( z \), it cannot be expressed through the matrices \( H \) or \( X \) when perturbation theory is used. Similar to the previous group, the other members in this group differ from this example due to different spin quantum numbers in the correlation function, or due to different indices of the Coulomb matrix elements.

Another group of terms contains the matrix \( T \) once, the Coulomb matrix elements twice, and the correlation function \( \langle c_{\sigma n\tau} c_{\sigma' z' \sigma'_t} \rangle \). An example from this group of terms is
\[
M^{(3+7)}_{\sigma\sigma' m} = - \sum_{n't'} \sum_{n' m' z} V_{n'n't'} V_{m't' m} T_{m' n'} \langle c_{\sigma n\tau} c_{\sigma' z} \rangle.
\]
(54)

Similar to the previous two groups, the other members in this group differ from the example of Eq. (54) due to different spin quantum numbers in the correlator, i.e., \( \langle c_{\sigma n\tau} c_{\sigma' z} \rangle \), and due to different indices in the Coulomb matrix elements.

A similar group of terms contains the correlator \( \langle c_{\sigma n\tau} c_{\sigma' z} c_{\sigma' t} c_{\sigma u} \rangle \) instead of \( \langle c_{\sigma n\tau} c_{\sigma' z} \rangle \). An example is given by
\[
M^{(3+8)}_{\sigma\sigma' m} = \sum_{n't'} \sum_{n' m' z} V_{n'n't'} V_{n' m' z} T_{n' m'} \langle c_{\sigma n\tau} c_{\sigma' z} c_{\sigma' t} c_{\sigma u} \rangle.
\]
(55)

### III. Construction of the Spectral Function from the Spectral Moments

In Ref. [17] we have shown that the spectral function of 4 spectral moment matrices of size \( N \times N \) may be obtained by solving \( 4N^2 \) coupled non-linear equations. While this approach is efficient for small \( N \), it may become inefficient for large \( N \). The reason may be understood from the amount of computer memory needed to store the Jacobian of the system of non-linear equations. The size of the Jacobian scales like \( 16N^4 \). In contrast, the size of the KS Hamiltonian matrix used in DFT codes scales like \( N^2 \) with the number \( N \) of basis functions. Therefore, we describe an alternative algorithm in this section, which is more efficient than solving systems of coupled non-linear equations when \( N \) is large. Since we discuss in Ref. [17] that finding the spectral function from non-commuting spectral matrices can be interpreted as a generalization of matrix diagonalization, it is perhaps not surprising that the new algorithm that we describe in this section uses such concepts.

In the following we describe the algorithm to construct the spectral function from the spectral moment matrices \( M^{(1)} \), \( M^{(2)} \), \( M^{(3)} \), where we assume that the zeroth spectral moment matrix is simply the unit matrix. Only the final result is described here, while the detailed proof is given in Appendix [D]. First, construct the hermitean \( N \times N \) matrix
\[
M^{(2+)} = M^{(2)} - M^{(1)} M^{(1)\dagger}.
\]
(56)

Next, diagonalize \( M^{(2+)} \):
\[
M^{(2+)} = U D U \dagger,
\]
(57)

where \( U \) is a unitary matrix and \( D \) is a diagonal matrix. Using \( D \) and \( U \) construct the matrix
\[
B_1 = U \sqrt{D}.
\]
(58)

Employ the inverse of its hermitean adjoint together with the moment matrices to compute the matrix
\[
B_2 = [M^{(3)} - M^{(2)} M^{(1)\dagger}] [B_1^{-1}]^{-1}.
\]
(59)

Use it to obtain the matrix
\[
D_1 = B_1^{-1} [B_2 - M^{(1)\dagger} B_1] [B_1^{-1}]^{-1}.
\]
(60)

Finally, take \( B_1, D_1 \) and the first moment matrix to construct the \( 2N \times 2N \) matrix
\[
B^{(1)} = \begin{pmatrix} M^{(1)} & B_1 \\ B_1 D_1 \end{pmatrix}
\]
(61)

and diagonalize it:
\[
B^{(1)} = U D U \dagger.
\]
(62)

The unitary matrix \( U \) contains the normalized eigenvectors of \( B^{(1)} \) as its columns. Compute the spectral weight of state \( j \) from
\[
a_j = \sum_{i=1}^{N} U_{ij} [U_{ij}]^*.
\]
(63)
Note that \( a_j \) may be smaller than one, because the summation over the index \( i \) goes only from 1 to \( N \) and not from 1 to \( 2N \). Therefore, spectral weights smaller than 1 may occur when bands split into lower and upper Hubbard bands.

Construct the \( N \times 2N \) matrix \( \mathbf{V} \) according to

\[
V_{ij} = \frac{U_{ij}}{\sqrt{a_j}}. \tag{64}
\]

Note that \( i = 1, ..., N, \) i.e., only the first \( N \) entries of the \( j \)-th column of \( \mathbf{U} \) are used, while every column of \( \mathbf{U} \) has of course \( 2N \) entries in total.

The spectral function is given by

\[
S_{ij}(E) = \frac{2N}{\hbar} \sum_{l=1}^{2N} a_l V_{il} V_{lj} \delta(E - E_l), \tag{65}
\]

where \( E_l \) is the \( l \)-th diagonal element of \( \mathbf{D} \), i.e., \( E_l = D_{ll} \). Here, \( 1 \leq i, j \leq N \), because in Eq. (64) we utilize only the first \( N \times 2N \) block of the \( 2N \times 2N \) matrix \( \mathbf{U} \) to construct the matrix \( \mathbf{V} \).

Note that in this paper we do not use the grand canonical Hamiltonian \( \mathcal{H} = H - \mu N \), where \( \mu \) is the chemical potential, but instead we use \( H \), because most DFT codes do not work with \( \mathcal{H} = H - \mu N \), as the chemical potential \( \mu \) is typically re-determined only before the end of every iteration in the self-consistency loop to achieve matching between electronic and nuclear charge, i.e., charge neutrality. When comparing our result Eq. (65) to the literature, one therefore needs to be aware of this difference by \( \mu \) in the expressions for the spectral function.

### IV. CHOICE OF THE MOMENT FUNCTIONALS

In Appendix B we have shown that

\[
\psi^{(2+)}_\sigma(r) = \frac{c^{(2+)}_\sigma}{[r_s(r)]^2} + \ldots \tag{66}
\]

and in Appendix C we have found

\[
\psi^{(3+)}_\sigma(r) = \frac{c^{(3+)}_\sigma}{[r_s(r)]^3} + \ldots, \tag{67}
\]

where

\[
r_s(r) = \frac{1}{a_B} \left( \frac{9\pi}{4|k_F(r)|^3} \right)^{\frac{1}{2}} = \left( \frac{3}{4\pi n(r)} \right)^{\frac{1}{2}} \tag{68}
\]

is the dimensionless density parameter. The corresponding matrix elements of the moments are obtained from these potentials according to

\[
M^{(2+)}_{\sigma nm} = \int d^3 r \psi^{(2+)}_\sigma(r) \phi^*_n(r) \phi_m(r) \tag{69}
\]

\[
M^{(3+)}_{\sigma nm} = \int d^3 r \psi^{(3+)}_\sigma(r) \phi^*_n(r) \phi_m(r). \tag{70}
\]

While it might be tempting to use these expansions, Eq. (66) and Eq. (67), to compute the moment functionals it is instructive to recall first the parameterization of the correlation energy of the uniform electron gas.

In order to construct an accurate analytic representation of the correlation energy of the uniform electron gas one considers the high-density expansion, the low-density expansion, and Green’s-function Monte Carlo data [29]. In the low-density expansion, the leading order is \( r_s^{-1} \) for the exact correlation energy. In the high-density limit one considers instead the parameterization \( c_0(\zeta) \ln r_s - c_1(\zeta) + c_2(\zeta) r_s \ln r_s \). Since these two functional forms for the low and high density limits differ considerably, we cannot expect good results, if we construct moment functionals only based on the parameterizations Eq. (66) and Eq. (67), which describe the case of low density. To give an impression of the deviation of the correlation energy density \( \epsilon_c \) from the low-density behavior \( \propto r_s^{-1} \), we plot in Fig. 1 the quantity \( |V_c r_s|^2 \), where

\[
V_c = \frac{d(\epsilon_c n)}{dn}. \tag{71}
\]

In order to take into account Monte Carlo simulations in the construction of the moment functionals, we would need such calculations for correlation functions such as Eq. (19) through Eq. (64) for the uniform electron gas. Since these data are currently not available in the literature, we nevertheless use the parameterization Eq. (66) in our applications below. As in this paper we present our first tests of the MFbSDFT-method, this slightly crude approach is justified, because the development of accurate moment functionals will probably take similarly long
as the development of the modern functionals used in KS-DFT calculations. Therefore, it is important to demonstrate the feasibility of the method before developing accurate moment functionals.

Additionally, we test the following strategy to find more elaborated moment functionals: Eq. (66) suggests that the leading order at low density is $r_s^{-2}$. Since the leading order of the correlation energy is $r_s^{-3}$ in this limit, we try to replace $r_s^{-2}$ in Eq. (66) by the square of the correlation potential $V_c$, i.e.,

$$\psi_{\sigma}^{(2+)}(r) = d_{\sigma}^{(2+)} |V_c(r_s)|^2$$  \hfill (72)

and similarly

$$\psi_{\sigma}^{(3+)}(r) = d_{\sigma}^{(3+)} |V_c(r_s)|^3.$$  \hfill (73)

This strategy should yield better results, because the $r_s^{-2}$ and $r_s^{-3}$ of the low-density expansion are thereby replaced by a more realistic functional form at high density.

In Appendix B we have estimated Eq. (19) through Eq. (35) based on zeroth order perturbation theory. According to this estimate, the prefactor of $r_s^{-2}$ is of the order of $10^{-3}$ [Ry]. The prefactor of $r_s^{-1}$ in the low-density expansion of the correlation potential is of the order of $1.2$ [Ry]. When we use the square of the correlation energy we have to choose the prefactor $d_{\sigma}^{(2+)}$ of the square of the correlation energy so that $d_{\sigma}^{(2+)}(1.2r_s^{-1})^2$ becomes comparable to $10^{-2}$. We therefore expect $d_{\sigma}^{(2+)}$ to be of the order of 10. At this order of magnitude of $d_{\sigma}^{(2+)}$, we indeed find a strong satellite peak in Ni (see Sec. VI).

V. SECOND VARIATION APPROACH

In this section we describe the implementation of our MFiSDFT method within a second variation approach. By second variation we mean that first a standard KS Hamiltonian is diagonalized at a given $k$-point and only part of its eigenvectors are used to compute the matrix elements of the moment functionals. The computation of the state vector matrix $\mathbf{V}$ and of the energies $E_l = \mathcal{D}_l$ may therefore be considered as a second variation step.

The size of the KS Hamiltonian matrix depends on the number of basis functions $N_B$. We do not compute all eigenvectors, but only as many eigenvectors as we need to describe the occupied bands and a fraction of the unoccupied bands. We call this number $N \ll N_B$. At a given $k$ point we additionally compute the $N_B \times N_B$ matrices $M^{(2+)}$ and $M^{(3+)}$ and project them onto the $N$ eigenstates. By $M^{(2+)}$ and $M^{(3+)}$ we denote these projections:

$$\tilde{M}^{(2+)} = \mathbf{U}^\dagger M^{(2+)} \mathbf{U}$$ \hfill (74)

and

$$\tilde{M}^{(3+)} = \mathbf{U}^\dagger M^{(3+)} \mathbf{U},$$ \hfill (75)

where $\mathbf{U}$ is a $N_B \times N$ matrix, which holds the $N$ eigenvectors in its $N$ columns.

The implementation of the moments $M^{(2+)}$ and $M^{(3+)}$ is easy to do: In the subroutines computing the standard KS-Hamiltonian one needs to switch off the kinetic energy contribution such that only the computation of the matrix elements of the potential remains. If one additionally replaces the exchange-correlation potential by the moment functional potential for $M^{(2+)}$ or $M^{(3+)}$, the subroutine computes the corresponding moment matrix.

The first and zeroth moments in the basis of the $N$ eigenstates are diagonal matrices:

$$\tilde{M}^{(1)}_{nm} = E_{nm}^\text{HF} \delta_{nm}$$ \hfill (76)

and

$$\tilde{M}^{(0)}_{nm} = \delta_{nm}.$$ \hfill (77)

Note that in contrast to a standard KS-DFT calculation, the KS-Hamiltonian used in the first variation step does not use the full exchange-correlation potential, but only the local or non-local first-order exchange, i.e., either Eq. (11) or Eq. (13). Therefore, we denote the band energies from the first variation step by $E_{nm}^\text{HF}$ in Eq. (76). Moments and band energies depend additionally on the $k$-point if periodic boundary conditions are used, but we suppress again the $k$ index in the moments and also in the band energy, i.e., instead of $E_{nm}^\text{HF}$ we write $E_{nm}^\text{HF}$.

The second and third moments are given by

$$\tilde{M}^{(2)} = \tilde{M}^{(1)} \tilde{M}^{(1)} + \tilde{M}^{(2+)}$$ \hfill (78)

and

$$\tilde{M}^{(3)} = \tilde{M}^{(1)} \tilde{M}^{(1)} \tilde{M}^{(1)} + \tilde{M}^{(3+)}$$ \hfill (79)

respectively.

The size of the matrices $\tilde{M}^{(0)}$, $\tilde{M}^{(1)}$, $\tilde{M}^{(2)}$, and $\tilde{M}^{(3)}$ is $N \times N$ and typically $N \ll N_B$. Therefore, the second variation approach is fast.

Close to the end of the selfconsistency cycle the Fermi energy is determined such that the total electronic charge compensates the nuclear charge. Typically, the subroutine computing the Fermi energy makes use of the eigenvalues and of weights, which are determined by the multiplicities of the $k$ points, when symmetries are used. In order to include the spectral weights Eq. (66) into the calculation of the Fermi energy, one only needs to multiply the $k$-point weights with these spectral weights. Similarly, the spectral weights need to be considered when computing the charge density from the matrix $\mathbf{V}$ of state vectors, Eq. (63), according to

$$n(r) = \sum_{\sigma nmj} \phi_{nm}^*(r) \phi_m(r) a_{j\sigma} V_{nj\sigma} V_{mj\sigma} f(E_{j\sigma}).$$ \hfill (80)
which may be derived from Eq. (14) by using the spectral theorem (17) to express the correlator \( \langle c_{\sigma n}^\dagger c_{\sigma m} \rangle \) in terms of the spectral function.

We illustrate the selfconsistency loop by the flowchart in Fig. 2. All results presented in Sec. IV have been obtained according to the flowchart in Fig. 2.

KS-DFT is so constructed that it may be used to obtain the total energy and the charge density in principle exactly. However, in practice the exact exchange correlation potential is not known and therefore the charge density computed in KS-DFT is an approximation. Since considerable progress has been made in the construction of exchange correlation potentials, the KS charge density is a very good approximation in many cases. Whenever the KS charge density is sufficiently correct, one may run MFbSDFT in a simplified mode: The converged KS charge density is used as starting density in Fig. 2 and only one iteration is performed, i.e., the output charge density is not computed but instead the results are calculated immediately from the state vector matrix \( \mathbf{V} \), from the spectral poles, and from the spectral weights.

VI. APPLICATIONS

In this section we apply the MFbSDFT method to several well-studied materials that show genuine many-body effects such as satellite peaks. According to the literature, the details of the spectral function of these materials depend strongly on the theoretical model used to study them. According to our discussion in Sec. IV the parameterizations that we use for the moment functionals should be considered only as a first step towards the development of accurate moment functionals. Consequently, if the results shown below are more similar to one theoretical model than they are to another one this does not imply at all that MFbSDFT confirms one particular theoretical model, because accurate moment functionals remain to be developed. The main purpose of this section is therefore to show that MFbSDFT is able to reproduce spectral features qualitatively that have been identified as genuine correlation effects before.

However, beyond validating the concept of MFbSDFT, the results shown here hint at a practical perspective for MFbSDFT already at this early stage of its development: Since the MFbSDFT reproduces spectral features of correlated materials, it may be used to compute response properties (17) such as the anomalous Hall effect, which is likely to require considerably less computer time than LDA+DMFT. For such an application one would fine-tune the parameters in the parameterizations of the moment functionals to match the spectral function known from LDA+DMFT or photoemission. While this approach is not parameter-free, it is similar to many applications of LDA+U, where the \( U \) parameter is chosen to reproduce a material property.

A. fcc Ni

The DOS obtained in KS-DFT with the PBE functional is shown in Fig. 3. The valence DOS starts to become significant starting from 5 eV below the Fermi energy and the exchange splitting is around 0.75 eV. In contrast, the width of the main bands found experimentally is significantly smaller than 5 eV, namely only 3 eV. Additionally, a much smaller exchange splitting of around 0.3 eV is found in photoemission experiments (32, 33). Moreover, the satellite peak observed in experiments at around 6 eV below the Fermi energy is absent in the KS-DFT spectrum.

Next, we discuss the MFbSDFT-spectrum obtained with \( \nu_{\sigma}^{(2+)}(r) = 15 \zeta_{\sigma}^{(5/3)} [V_{\sigma}(r_s)]^2 \) and \( \nu_{\sigma}^{(3+)}(r) = 0 \). Here \( \zeta_{\sigma} = (1 - \sigma(n_\uparrow - n_\downarrow)/n) \). We use \( N = 36 \). With this choice of parameters the magnetic moment computed self-consistently in MFbSDFT is 0.58 \( \mu_B \). The resulting DOS is presented in Fig. 4. The exchange splitting of around 0.3 eV is strongly reduced compared to the KS-DFT calculation and close to the experiments (32, 33). Additionally, the main bands are much narrower than in KS-DFT and therefore in much better agreement with experiments. Moreover, satellite peaks are found close to 6 eV. However, the spectral weight of these satellite peaks is smaller than what is found in experiments and in LDA+DMFT calculations (see e.g. Fig. 9 in Ref. 32, Fig. 2 in Ref. 14, and Fig. 2 in Ref. 33).

Finally, we discuss the MFbSDFT-spectrum obtained with \( \nu_{\sigma}^{(2+)}(r) = 0.015 \zeta_{\sigma}^{(7/3)} r_s^{-2} \) [Ry] \( ^2 \) and \( \nu_{\sigma}^{(3+)}(r) = -0.00472 \zeta_{\sigma}^{(1/3)} r_s^{-3} \) [Ry] \( ^3 \). We use \( N = 36 \). With this choice of parameters the magnetic moment is 0.63 \( \mu_B \). Fig. 5 shows the density of states (DOS) of Ni in the ferromagnetic state as computed selfconsistently in MFbSDFT. While the exchange splitting is similar to KS-DFT, the width of the main bands is reduced, leading to a slightly better agreement with experiment. Around 6 eV below the Fermi energy satellite peaks appear with a spectral weight of a similar order of magnitude like in experiment and LDA+DMFT. However, the spin polarization of the satellite peak structure differs from both experiment and LDA+DMFT, which both predict the minority satellite to be strongly suppressed (see e.g. Fig. 9 in Ref. 32, Fig. 2 in Ref. 14, and Fig. 2 in Ref. 33). In contrast, in Fig. 5 the satellites of the majority and minority band are comparable in magnitude and only shifted in energy.

As discussed in Sec. IV the parameters employed in \( \nu_{\sigma}^{(2+)}(r) = 15 \zeta_{\sigma}^{(5/3)} [V_{\sigma}(r_s)]^2 \) (used to generate the DOS shown in Fig. 4) are of the order of magnitude expected from the estimate given in Sec. IV. In contrast, we determined the parameters employed in \( \nu_{\sigma}^{(2+)}(r) = 0.015 \zeta_{\sigma}^{(7/3)} r_s^{-2} \) [Ry] \( ^2 \) (used to generate the DOS shown in Fig. 5) only based on try-out, because it is unclear how to renormalize the parameters of the low-density expansion.
Input Charge Density \( n_{\text{inp}}(r) \)

Hartree and Exchange potentials \( V^H(r), V^X(r) \) (Eq. (10), Eq. (11), Eq. (13))

Diagonalize \( H = -\frac{\hbar^2}{2m} \Delta + V(r) + V^H(r) + V^X(r) \)

Moment potentials \( V^{(2+)}(r), V^{(3+)}(r) \) (Eq. (66), Eq. (67), Eq. (72), Eq. (73))

Compute moments \( \bar{M}^{(1)}, \bar{M}^{(2)}, \bar{M}^{(3)} \) (Eq. (76), Eq. (78), Eq. (79))

Compute spectral poles \( E_l = D_{ll} \) (Eq. (62))

Compute state vectors \( V \) (Eq. (64))

Compute spectral weights \( a_j \) (Eq. (63))

Compute output charge density \( n_{\text{out}}(r) \) (Eq. (80))

Mix charges to obtain new charge density for next iteration

FIG. 2. Flowchart of the MFbSDFT selfconsistency cycle.

FIG. 3. DOS of Ni vs. energy \( E \) in the ferromagnetic state as obtained in KS-DFT. \( E_F \) is the Fermi energy.

FIG. 4. DOS of Ni vs. energy \( E \) in the ferromagnetic state as obtained in MFbSDFT when the moment functional is constructed according to Eq. (72). \( E_F \) is the Fermi energy.

FIG. 5. DOS of Ni vs. energy \( E \) in the ferromagnetic state as obtained in MFbSDFT when the moment functional is constructed according to Eq. (66) and Eq. (67). \( E_F \) is the Fermi energy.

so that it effectively describes the regimes of intermediate and high densities as well. However, Fig. 5 is useful nevertheless, because it shows that satellite peaks with the correct order of magnitude of spectral weight can be produced by MFbSDFT. Taken together with the result of Fig. 4 which shows that band widths, exchange splittings and location of the satellite peaks are predicted very well if the correlation potential is used to construct the moment functionals, the overall conclusion from this finding is that it is likely that assessing Eq. (19) through Eq. (35) in the low-density and the high-density regimes and using Monte Carlo results to interpolate between these limits will allow us to formulate a moment functional that predicts the spectral features in Ni quite well.

B. SrVO\(_3\)

In Fig. 6 we show the contributions of the V-3d \( e_g \) and \( t_{2g} \) states to the DOS of SrVO\(_3\), as obtained with KS-DFT using the PBE functional. The KS spectrum is not
In good agreement with experiment. It has been shown that the agreement with experiment is improved significantly \cite{11}, when DMFT is used to supplement these bands with correlation effects. The detailed rearrangement of the spectral features obtained from LDA+DMFT depends on the details of the modelling of the correlation effects by the Hubbard model. Ref. \cite{11} includes only the \( t_{2g} \) states into the Hubbard model. In this case the DOS of the \( t_{2g} \) states obtained from LDA+DMFT is distributed into three pronounced spectral peaks: A dominant central peak roughly 0.5 eV above the Fermi energy, a lower Hubbard band around 2 eV below the Fermi energy and an additional upper Hubbard band around 3 eV above the Fermi energy. The position of these peaks is in good agreement with experiments, which find peaks roughly at -1.7 eV, 0.5 eV, and 2.4 eV \cite{31,32}. These spectral features are also observed in Ref. \cite{36}, which includes the \( e_g \) states, however they strongly depend on the parameters, and the intensities of the lower and upper Hubbard bands are much smaller for some parameters. Additionally, the intensities of the lower and upper Hubbard bands depend strongly on the double counting correction.

In Fig. 7 we present the contributions of the V-3d \( e_g \) and \( t_{2g} \) states to the DOS, as obtained with MFlbSDFT when we use Eq. (72) and Eq. (73), where we set \( d^{2+}_{t_2} = 100 \) and \( d^{3+}_{t_2} = -200 \). We use \( N = 200 \). The total V-d DOS is in good agreement with both the experimental spectrum and the LDA+DMFT spectrum (See e.g. Fig. 7 in Ref. \cite{11} for comparison. Ref. \cite{11} uses a broadening of 0.36 eV in order to reproduce the experimental resolution. We use 0.36 eV in our Fig. 7 as well.). However, in our case the peak between 2 eV and 2.5 eV stems mainly from the \( e_g \) states, which are not included into the Hubbard model in Ref. \cite{11}. Ref. \cite{12} includes the \( e_g \) states, but still finds a small peak from the upper Hubbard band for the \( t_{2g} \) states at around 3 eV above the Fermi energy. Such a small peak is consistent with our Fig. 7, where a shoulder in the V-3d(\( t_{2g} \)) is clearly visible between 2 eV and 2.5 eV. Moreover, Ref. \cite{12} finds a large contribution from the \( e_g \) states to the DOS at this energy as well. In this regard, our Fig. 7 resembles closely Fig. 8 in Ref. \cite{12} when the energy is above the Fermi energy. However, Fig. 8 in Ref. \cite{12} does not find a strong V-d DOS at around 2 eV below the Fermi energy. In contrast, we find a strong V-d DOS at around 2 eV below the Fermi energy with a dominant part from the \( e_g \) states and a small contribution from the \( t_{2g} \) states.

In Fig. 8 we show the contributions of the V-3d \( e_g \) and \( t_{2g} \) states to the DOS, as obtained with MFlbSDFT when we take \( N = 200 \) and use Eq. (72) and Eq. (73), where we set \( c^{(2+)}_{t_2} = 1.1 [\text{[Ry]}^3] \), and \( c^{(3+)}_{t_2} = -1.5 [\text{[Ry]}^3] \). The peak between 2 eV and 2.5 eV is very pronounced and both \( e_g \) and \( t_{2g} \) bands contribute to it. In contrast, the peak around -2 eV in Fig. 7 is significantly smaller in Fig. 8 and shifted to lower energy between -2.5 eV and -3 eV. Several main features of the \( t_{2g} \) band in Fig. 8 resemble those obtained from a LDA+DMFT calculation with a Hubbard \( U \) of 6 eV (see Fig. 8 in Ref. \cite{36}). Notably, the \( t_{2g} \) band, which ends around 2 eV in Fig. 8 is expanded to higher energies like in LDA+DMFT. Overall, the total V-d DOS in Fig. 8 is qualitatively similar to the one in Fig. 8 of Ref. \cite{36}, which displays pronounced peaks close to 1 eV and close to 2.5 eV, while the V-d DOS close to -2 eV is small, in agreement with our result in Fig. 8. However, the peak close to 2.5 eV is much larger in our Fig. 8.

We may reduce the intensity of this peak at 2.5 eV.
VII. DISCUSSION AND OUTLOOK

In the previous section we have shown that MFbSDFT reproduces features such as satellite peaks and spectral weight shifts which are usually obtained by solving the correlated electron problem directly, e.g. by means of LDA+DMFT. While these first MFbSDFT results look therefore very promising there is a large number of open questions and obvious possibilities to improve this method further.

First, accurate moment functionals are required. The construction of accurate moment functionals should be possible based on Monte Carlo data of correlation functions such as those given in Eq. (19) through Eq. (35).

Second, it is desirable to derive gradient approximations for the moment functionals. Using local functionals that depend only on the spin densities for MFbSDFT misses effects related to their spatial inhomogeneity. Like GGA is an improvement over LDA, we expect that MFbSDFT will become more accurate by adding gradient corrections to the functionals.

Third, in this work we do not explore the calculation of total energies and atomic forces for structural relaxation. However, since force calculations in correlated materials are possible within LDA+DMFT we expect that total energies and forces may also be obtained within our MFbSDFT approach.

Fourth, one may use more than the first 4 moments. While the first four moments are sufficient to reproduce the quasi-particle band structure qualitatively correctly in the strong-correlation regime [37], the precision of MFbSDFT is expected to increase with the number of moments used. The moments become increasingly more complicated with increasing order. However, one may use computer algebra systems in order to derive the higher-order moments and to assess them for the uniform electron gas. This seems feasible since the complexity is probably comparable to higher-order perturbation theory in QED, where high-order contributions have to be tackled by computer algebra. Assuming that computer algebra systems can manage the complexity of the higher-order moments the question remains if the spectral function can be found for more than 2 or 4 moments. In Ref. [17] we give an argument that the spectral function may be found from the first 4 moments, which is based on counting the number of available equations and the number of parameters that determine the spectral function and showing that these numbers match. In the present paper we have explicitly constructed the spectral function from the first 4 moments in Sec. III. We may generalize the argument given in Ref. [17] and show that from the first 2P moments (P = 1, 2, . . .) one may construct the spectral function. This generalization is discussed in App. E. If one uses only Delta-functions in the expression for the spectral function (as in Eq. (55))

\[ \text{DOS} \]
one misses lifetime effects, which may be accommodated by employing Gaussians instead \(35\). When 4 moments are required to put the spectral peaks, such as satellite peaks, at the right energies, it is clear that more than 4 moments are generally required to correct the spectral widths of these spectral features by lifetime effects.

Fifth, perhaps the method of spectral moments may contribute to the understanding of SIE, because it is remarkable that setting \(\mu_s(r) = 0\) in Eq. (2.22) of Ref. \[2\] leads to a HF-type method that suffers from the same SIE and is equivalent to the method of spectral moments derived from the first two moments (see Sec. \[IV\]). We suspect that increasing the number of moments used will ultimately eliminate the SIE. However, it is an open question, how the self-interaction correction (SIC) takes place exactly within the method of spectral moments. Within the KS-DFT framework the explanation of SIC is that \(\mu_s(r)\) in Eq. (2.22) of Ref. \[2\] has to eliminate SIE when the exact exchange correlation functional is used. However, within the spectral moment method a valid explanation of SIE seems to be that using only the first two moments produces an error, which may be eliminated by using more moments. Of course, the precise moment functionals are expected to be necessary in order to remove SIE. However, Eq. \[19\] through Eq. \[55\] provide explicit expressions, which may be used for the construction of the moment functionals.

Sixth, it is an important open question how to extend the MFbSDFT approach to finite temperatures. In Ref. \[17\] we have shown how to generalize the spectral moment method so that it can be applied to many-band Hamiltonians. Since the method of Ref. \[17\] computes the correlation functions from the spectral theorem, which involves the Fermi function and the actual excitation energies, it naturally includes finite temperature effects. As the spectral theorem is not used for the higher-order correlation functions in MFbSDFT, which are obtained from moment functionals, it is currently unknown how to accommodate finite temperatures accurately in this method.

While accurate moment functionals are currently not available yet, the MFbSDFT method may also be used in practice in a way similar to LDA+\(U\): In LDA+\(U\) the \(U\) and \(J\) parameters are usually chosen for a given material in order to add correlation effects that are not described by LDA. Similarly, one may use parameterizations of the moment functionals similar to the ones that we discussed in Sec. \[IV\] and choose the coefficients in the functional in order to optimize spectral features.

**VIII. SUMMARY**

We describe the concept of moment functionals, which allow us to obtain the spectral moments from functionals of the charge density. These functionals play a similar role in MFbSDFT as the exchange correlation functional does in KS-DFT. We derive explicit expressions for the moment functionals and use perturbation theory to investigate their scaling with the charge density. We describe an efficient algorithm to obtain the spectral function from the first four spectral moments. We demonstrate that MFbSDFT allows us to reproduce spectral features such as satellite peaks in Ni and lower and upper Hubbard bands in SrVO\(_3\). At this stage of its development, MFbSDFT may be used in a way similar to LDA+\(U\): The parameters in the moment functionals are chosen such that spectral features found in experiments are reproduced.

**ACKNOWLEDGMENTS**

We acknowledge funding by SPP 2137 “Skyrmionics” of the DFG and by Sino-German research project DIS-TOMAT (DFG project MO 1731/10-1). We gratefully acknowledge financial support from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (Grant No. 856538, project “3D MAGiC”). The work was also supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – TRR 288 – 422213477 (project B06). We also gratefully acknowledge the Jülich Supercomputing Centre and RWTH Aachen University for providing computational resources under project No. jiff40.

**Appendix A: Comparison of the spectral moments of the many-band case to the spectral moments of the single-band Hubbard model**

The spectral moments of the many-band case contain many new contributions that do not have counterparts in the single-band Hubbard model. In this appendix we discuss which of the many-band terms have a correspondence in the single-band Hubbard model. In the single band Hubbard model the Coulomb matrix element is simply

\[
V_{nn'\tau\tau'} = U \delta_{nn'} \delta_{\tau\tau'} \delta_{nt},
\]

where \(U\) is the Hubbard-\(U\). For the single band Hubbard model the first four moments are:

\[
\tilde{M}^{(0)}_{k\sigma} = \frac{1}{N} \sum_{ij} e^{ik \cdot (R_i - R_j)} \langle [c_{\sigma i}, c_{\sigma j}^\dagger]_+ \rangle = 1,
\]

\[
\tilde{M}^{(1)}_{k\sigma} = \frac{1}{N} \sum_{ij} e^{ik \cdot (R_i - R_j)} \langle [c_{\sigma i}, H]_{-}, c_{\sigma j}^\dagger \rangle + e(k) + U \langle n_{-\sigma} \rangle,
\]
Here, \( N \) is the number of \( k \) points.

Clearly, Eq. (5) and Eq. (11) turn into Eq. (A2) and Eq. (A3), respectively, when one evaluates them for the single-band Hubbard model and performs a Fourier transformation.

The following contributions to \( M^{(2+)} \) are zero for the single-band Hubbard model: \( M^{(2+),3} \), \( M^{(2+),4} \), \( M^{(2+),5} \), \( M^{(2+),6} \), \( M^{(2+),7} \). The sum \( M^{(2+),1} + M^{(2+),2} \) turns into \( U^2 \langle n_{-\sigma} \rangle \) in the single-band case, which is the last term in Eq. (A3). The sum \( TV^H + TV^X + V^H T + V^X T \), which contributes to Eq. (10), evaluates to \( 2U \langle n_{-\sigma} \rangle \epsilon(k) \) in the case of the single-band Hubbard model. This is the middle term in Eq. (A3).

For the single-band Hubbard model the sum of \( M^{(3+),3} \) (Eq. (15)) and \( M^{(3+),4} \) (Eq. (16)) is \( U^3 \langle n_{-\sigma} \rangle \), which is the last term in the third line of Eq. (A3).

Appendix B: Evaluation of \( M^{(2+),j} \)

In order to keep the notation simple, we discuss the contractions \( C_{\sigma}^{(2+),j} \), Eq. (39), for the uniform electron gas without spin-polarization.

We evaluate the contraction of Eq. (20) by transforming it into the momentum representation, where we obtain at zero temperature

\[
C_{\sigma}^{(2+),2} = -\int d^3q d^3k d^3k_5 \frac{\Theta(k_F - |k_5 - q|) \Theta(k_F - |k_4 + q|)}{q^2|k_4 - k_5 + q|^2},
\]

(B1)

Here, \( k_F = (3\pi^2 n)^{1/3} \) (B2) is the Fermi wave number, \( \Theta(k) \) is the Heaviside step function,

\[
v(q) = \frac{8\pi}{V} |\text{Ry}| a_B \frac{1}{q^2} \quad (B3)
\]

is the Coulomb potential expressed in terms of the Bohr radius \( a_B \), \( \text{Ry} = 13.6 \) eV, and

\[
A_2 = \frac{(8\pi)^2}{(2\pi)^9} |\text{Ry}|^2 |a_B|^2 \frac{1}{n} \quad (B4)
\]

Scaling all momenta in Eq. (B1) by the factor \( \xi \), we observe that this integral is proportional to \( \xi^2 \), i.e., it is proportional to \( k_F^2 \). In this scaling analysis we took into account that \( n \) depends on \( k_F \) as well: \( n = k_F^2/(3\pi^2) \). It is convenient to express \( C_{\sigma}^{(2+),2} \) in terms of the dimensionless density parameter

\[
r_s = \frac{1}{a_B} \left( \frac{9\pi}{4k_F^2} \right)^{\frac{1}{3}}. \quad (B5)
\]

According to the scaling analysis above, it is sufficient to evaluate the integral for a single density parameter, e.g. \( r_s' = 1 \), because

\[
C_{\sigma}^{(2+),2}(r_s') = C_{\sigma}^{(2+),2}(r_s') \quad (B6)
\]

The integral can be performed numerically using the VEGAS [39, 40] package for the Monte-Carlo integration of high-dimensional integrals. We obtain

\[
C^{(2+),2}(r_s') = -\frac{2.57}{r_s'^2} |\text{Ry}|^2. \quad (B7)
\]

Next, we consider \( C^{(2+),16} \). This integral is given by

\[
C^{(2+),16} = \int d^3q d^3k_5 \frac{\Theta(q)}{(2\pi)^9 nV^{-2} v(q)} \times v(k_4 - k_5 + q) n_{k_5 - q} n_{k_4 + q}
\]

\[
= -A_2 \int d^3q d^3k_4 d^3k_5 \frac{\Theta(k_F - |k_5 - q|) \Theta(k_F - |k_4 + q|)}{q^2|k_4 - k_5 + q|^2}, \quad (B8)
\]

which also scales like \( r_s'^{-2} \), which is easy to see with a scaling analysis. Evaluating this integral with the VEGAS [39, 40] package gives

\[
C^{(2+),16}(r_s') = \frac{1.73}{r_s'^2} |\text{Ry}|^2. \quad (B9)
\]
For the contraction $C^{(2+,17)}$ we need to compute the integral
\[
C^{(2+,17)} = \int \frac{d^3q d^3k_4 d^3k_5}{(2\pi)^9 n V^2} v(q) \times v(k_4 - k_5 + q) n_{k_5} q_n k_4 + q n_{k_4}
\]
\[= A_2 \int d^3q d^3k_4 d^3k_5 \Theta(k_F - |k_4|) \times \frac{\Theta(k_F - |k_5 - q|) \Theta(k_F - |k_4 + q|)}{q^2 |k_4 - k_5 + q|^2}.
\] (B10)
Using a scaling analysis, we find that this integral also scales like $r_s^{-2}$. Employing the VEGAS package yields
\[C^{(2+,17)}(r_s) = \frac{1.73}{r_s^2} [\text{Ry}^2]. \] (B11)

While the contractions $C^{(2+,2)}$, $C^{(2+,16)}$, and $C^{(2+,17)}$ above are straightforward to evaluate with VEGAS, the contributions $C^{(2+,1)}$, $C^{(2+,6)}$, $C^{(2+,8)}$, $C^{(2+,12)}$, and $C^{(2+,15)}$ require more care, because their integrands contain factors $(v(q))^3$, which lead to a strong divergence of the integrands in the limit $q \to 0$. In contrast, the integrals of the contractions $C^{(2+,2)}$, $C^{(2+,16)}$, and $C^{(2+,17)}$ contain only a single factor $v(q)$, which does not produce a divergence, because it is compensated by the $q^2$ of $d^3q = q^2 \sin(\theta) d\theta d\phi$. However, these contributions may be grouped into pairs of two, where the two partners in a pair differ in sign. When we replace the Coulomb potential by
\[v(q) = \frac{8\pi}{V} [\text{Ry}] |a_B| \frac{1}{q^2 + \eta^2}, \] (B12)
we observe that the limit $\eta \to 0$ is finite for the pair, while both partners in a pair differ in this limit.

Consider for example the pair composed of $C^{(2+,1)}$ and $C^{(2+,14)}$. Evaluating the integrals
\[C^{(2+,1)} = 2 \int \frac{d^3q d^3k_4 d^3k_5}{(2\pi)^9 n V^2} v(q) v_n(q) n_{k_5 - q} n_{k_4 + q}
\]
\[= 2A_2 \int d^3q d^3k_4 d^3k_5 \Theta(k_F - |k_4 + q|) \times \frac{\Theta(k_F - |k_5 - q|)}{[q^2 + \eta^2]^2}
\] (B13)
and
\[C^{(2+,14)} = -2 \int \frac{d^3q d^3k_4 d^3k_5}{(2\pi)^9 n V^2} v(q) v_n(q) n_{k_5 - q} n_{k_4 + q} n_{k_4}
\]
\[= 2A_2 \int d^3q d^3k_4 d^3k_5 \Theta(k_F - |k_5|) \times \frac{\Theta(k_F - |k_5 - q|) \Theta(k_F - |k_4 + q|)}{[q^2 + \eta^2]^2}
\] (B14)
with the VEGAS package we obtain
\[
\lim_{\eta \to 0} \left[ C^{(2+,1)}(r_s) + C^{(2+,14)}(r_s) \right] = \frac{2.1051}{r_s^2} [\text{Ry}^2]. \] (B15)
We explicitly left the spin-degeneracy factor 2 in this equation.

Appendix C: Evaluation of $M_{\sigma \tau m \nu}$

In order to keep the notation simple, we discuss the contractions $C^{(3+,\nu)}$ for the uniform electron gas without spin-polarization.

Transforming Eq. (B11) into the momentum representation, we obtain the following expression for $C^{(3+,4)}$ in terms of an integral:
\[C^{(3+,4)} = \int \frac{d^3q d^3q' d^3k_4 d^3k_5}{(2\pi)^9 n V^2} v(q) v(q') v(k_4 - k_5 + q - q')
\times n_{k_5 - q} n_{k_4 + q}
\]
\[= A_3 \int d^3q d^3q' d^3k_4 d^3k_5 \Theta(k_F - |k_5 - q|) \Theta(k_F - |k_4 + q|)
\times \frac{\Theta(k_F - |k_4 - q|)}{|k_4 - k_5 + q - q'|^2}, \] (C1)
where
\[A_3 = \frac{(8\pi)^3}{(2\pi)^9 [\text{Ry}]^2 |a_B|^3} \frac{1}{n}. \] (C2)
Scaling all momenta in Eq. (C1) by the factor $\xi$ one may easily find that $C^{(3+,4)} \propto k_F^2 \propto r_s^{-3}$. Using VEGAS we obtain
\[C^{(3+,4)}(r_s) = -\frac{9.81}{r_s^3} [\text{Ry}]^3. \] (C3)

Appendix D: Algorithm to construct the spectral function from non-commuting spectral moment matrices

In this section we provide the derivation of the algorithm described in Sec. III for the construction of the spectral function from the first four $N \times N$ spectral moment matrices, $M^{(0)}$, $M^{(1)}$, $M^{(2)}$, and $M^{(3)}$, where $M^{(0)}$ is the unit matrix.

Assume that we manage to find hermitean $2N \times 2N$ matrices
\[B^{(1)} = \begin{pmatrix} M^{(1)} & B_1 \\ B_1^T & D_1 \end{pmatrix}, \] (D1)
\[B^{(2)} = \begin{pmatrix} M^{(2)} & B_2 \\ B_2^T & D_2 \end{pmatrix}, \] (D2)
and
\[
\mathbf{B}^{(3)} = \begin{pmatrix} M^{(3)} & B_3 \\ B_3^* & D_3 \end{pmatrix},
\] (D3)

which mutually commute, i.e.,
\[
[B^{(1)}, B^{(2)}] = 0,
\]
\[
[B^{(1), B^{(3)}]} = 0,
\]
and
\[
[B^{(2), B^{(3)}]} = 0,
\] (D5)

and
\[
\mathbf{B}^{(3)} = \mathbf{B}^{(1)} \mathbf{B}^{(2)} = \mathbf{B}^{(1)} \mathbf{B}^{(2)} \mathbf{B}^{(1)}.
\] (D6)

Note that Eq. (D4) is satisfied if Eq. (D5) and Eq. (D6) are satisfied. We will therefore solve only Eq. (D5) and Eq. (D6) below. The matrices \(M^{(1)}\), \(B_i\) and \(D_i\) have the size \(N \times N\). The matrices \(M^{(1)}\) are the given hermitean spectral moment matrices, while \(B_i\) and \(D_i\) are matrices that need to be determined such that Eq. (D5) and Eq. (D6) are satisfied. While \(D_i\) is required to be hermitean, \(B_i\) is not.

If we manage to find these matrices \(B^{(1)}\), \(B^{(2)}\), and \(B^{(3)}\), we know that they possess a common system of eigenvectors, i.e., they may be diagonalized by the same unitary transformation, because they are hermitean and they commute mutually. Consequently, we may find a unitary transformation \(U\) so that
\[
\mathbf{B}^{(1)} = \mathbf{UDU}^\dagger,
\] (D7)

where \(\mathbf{D}\) is a diagonal matrix. Using \(U\) and \(\mathbf{D}\) we may write
\[
\mathbf{B}^{(2)} = \mathbf{UD}^2 \mathbf{U}^\dagger
\] (D8)

and
\[
\mathbf{B}^{(3)} = \mathbf{UD}^3 \mathbf{U}^\dagger.
\] (D9)

In Ref. [17] we have shown that the eigenvalue problems Eq. (D7), Eq. (D8), and Eq. (D9) may be rewritten in the form
\[
\mathbf{W} \mathbf{A}^{(I)} = \mathbf{B}^{(I)},
\] (D10)

where \(I = 1, 2, 3\) (see Eq. (13) in Ref. [17]). When we denote the representation of the unit matrix as a column vector by \(\mathbf{B}^{(0)}\), we may combine Eq. (D7), Eq. (D8), and Eq. (D9) into the compact expression
\[
\mathbf{WA} = \mathbf{B},
\] (D11)

where
\[
\mathbf{A} = [\mathbf{A}^{(0)}, \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}]
\] (D12)

and
\[
\mathbf{B} = [\mathbf{B}^{(0)}, \mathbf{B}^{(1)}, \mathbf{B}^{(2)}, \mathbf{B}^{(3)}].
\] (D13)

Next, we rewrite \(\mathbf{B}\) as
\[
\mathbf{B} = \begin{pmatrix} \mathbf{M} \\ \mathbf{B}_{\text{Low}} \end{pmatrix}
\] (D14)

and
\[
\mathbf{W} = \begin{pmatrix} \mathbf{W} \\ \mathbf{W}_{\text{Low}} \end{pmatrix},
\] (D15)

where \(\mathbf{M}\) and \(\mathbf{W}\) are the matrices defined in Ref. [17] (see Eq. (7) and Eq. (8) in Ref. [17]). \(\mathbf{M}\) is a \(N^2 \times 4\) matrix, \(\mathbf{B}_{\text{Low}}\) is a \(3N^2 \times 4\) matrix, \(\mathbf{W}\) is a \(N^2 \times 2N\) matrix, and \(\mathbf{W}_{\text{Low}}\) is a \(3N^2 \times 2N\) matrix. Thus, we may rewrite Eq. (D11) as two equations:
\[
\mathbf{WA} = \mathbf{M}
\] (D16)

and
\[
\mathbf{W}_{\text{Low}} \mathbf{A} = \mathbf{B}_{\text{Low}}.
\] (D17)

Eq. (D16) is identical to the Eq. (9) in Ref. [17], which needs to be solved to obtain the spectral function. Thus, we may solve Eq. (D16) by determining the matrices \(B_1\) and \(D_1\), and by diagonalizing the matrix \(B^{(1)}\).

Therefore, in order to prove the algorithm in Sec. [11] it remains to show that the matrices \(B_i\) and \(D_i\) may be found by solving Eq. (D5) and Eq. (D6). From Eq. (D5) we obtain the following equation for \(B_1\):
\[
B_1 B_1^\dagger = M^{(2)} - M^{(1)} M^{(1)}.
\] (D18)

Since \(B_1 B_1^\dagger\) is a hermitean matrix, it may be diagonalized:
\[
B_1 B_1^\dagger = \mathbf{UDU}^\dagger,
\] (D19)

where \(U\) is a unitary matrix and \(\mathbf{D}\) is a diagonal matrix. If \(B_1 B_1^\dagger\) is positive definite, we obtain
\[
B_1 = \mathbf{U} \sqrt{\mathbf{D}},
\] (D20)

which is Eq. (58) in the main text. If \(B_1 B_1^\dagger\) is not positive definite, the algorithm described in this section cannot be used. However, in all applications discussed in this paper, \(B_1 B_1^\dagger\) is positive definite. We suspect that the reason for this is that \(M^{(2+)}\) is generally positive definite.

From Eq. (D6) we obtain the following equation for \(B_2\):
\[
B_2 = \left[ M^{(3)} - M^{(2)} M^{(1)} \right] B_1^\dagger.
\] (D21)

This is Eq. (59) in the main text.
From Eq. (D5) we obtain the following equation for $D_1$: \[
D_1 = B_1^{-1} \left[ B_2 - M^{(1)} B_1 \right]. \tag{D22}
\]
This is Eq. (60) in the main text. $D_1$ is required to be hermitean, which is not directly obvious from Eq. (D22). However, making use of Eq. (D21) and Eq. (D18) it is straightforward to show that \[
D_1 - D_1^\dagger = 0. \tag{D23}
\]
At this point we have completely determined the matrix $B^{(1)}$, from which the spectral function may be constructed using its eigenvalues, which are contained in the diagonal matrix $D_1$ and the unitary transformation $U$ defined in Eq. (D7). However, it remains to show that all those additional equations that follow from Eq. (D5) and Eq. (D6) but that we did not use to derive the expressions for $B_1$ and $D_1$ can be satisfied as well. From Eq. (D22) we obtain \[
D_2 = B_1^\dagger B_1 + D_1 D_1, \tag{D24}
\]
and from Eq. (D6) we obtain \[
B_3 = M^{(1)} B_2 + B_1 D_2. \tag{D25}
\]
and \[
D_3 = B_2^\dagger B_2 + D_1 D_2. \tag{D26}
\]
$D_2$ as given by Eq. (D24) is hermitean, because $D_1$ is hermitean according to Eq. (D23). Thus, it does not violate any of the equations above. $B_3$ as given by Eq. (D25) does not violate any of the equations above either. $D_3$ as given by Eq. (D26) should be hermitean, which is not directly obvious. However, using Eq. (D24), Eq. (D22), and Eq. (D23), it is straightforward to show that \[
D_3 - D_3^\dagger = 0. \tag{D27}
\]
Eq. (65) in the main text follows from Eq. (D7), Eq. (D16), and Ref. [17].

Appendix E: Generalization to more moments

We may generalize the argument given in Ref. [17] and show that from the first $2P$ moments ($P = 1, 2, \ldots$) one may construct the spectral function: We may map each moment $M^{(I)}$ (where $M^{(I)}$ denotes the moment computed from the nested commutator expression – as opposed to the moment obtained from the explicit energy integration) onto an $N^2$-dimensional real-valued vector $M^{(I)}$, because $N^2$ real-valued parameters fully define a hermitean $N \times N$ matrix. We introduce the $N^2 \times 2P$ matrix $\mathcal{M}$ by \[
\mathcal{M} = [\mathcal{M}^{(0)}, \ldots, \mathcal{M}^{(2P-1)}]. \tag{E1}
\]
We try to approximate the spectral function by \[
\frac{S_{\alpha\beta}(E)}{h} = \sum_{\gamma=1,\ldots,N} \sum_{\gamma'=1}^{N} a_{\gamma\gamma'} V_{\alpha\gamma'} V_{\beta\gamma'}^* \delta(E - E_{\gamma'}) \tag{E2}
\]
because we expect that $PN$ bands can be computed from the first $2P$ spectral moment matrices. Inserting this approximation into Eq. (I) yields \[
M^{(I)}_{\alpha\beta} = \sum_{\gamma=1,\ldots,N} \sum_{\gamma'=1}^{N} a_{\gamma\gamma'} W_{\alpha\beta\gamma'} [E_{\gamma'}]^I, \tag{E3}
\]
where we defined $W_{\alpha\beta\gamma'} = V_{\alpha\gamma'} V_{\beta\gamma'}^*$. We may consider $W_{\alpha\beta\gamma'}$ as the row-$\alpha$ column-$\beta$ element of a hermitean matrix $W_{\gamma'}$. Since $\gamma = 1, \ldots, N$ and $p = 1, \ldots, P$, there are $PN$ such matrices. As the hermitean $N \times N$ matrix $W_{\gamma'}$ is equivalent to a $N^2$-dimensional real-valued vector $W_{\gamma'}$, we define the $N^2 \times PN$ matrix $W = [W_{11} \ldots W_{NP}]$. Additionally, we construct the $PN \times 2P$ matrix $\mathcal{A}$ by setting the element $A_{\gamma\gamma'}$ in row $(\gamma, p)$ and column $m$ to $a_{\gamma\gamma'} (E_{\gamma'})^{m-1}$. The requirements $\mathcal{M}^{(I)} = \mathcal{M}^{(I)}$ with $I = 0, 1, \ldots, 2P - 1$ (where $\mathcal{M}^{(I)}$ are the moments computed from the nested commutator expressions) can now be formulated in compact form by \[
\mathcal{W} \mathcal{A} = \mathcal{M}. \tag{E4}
\]
This is the generalization of Eq. (9) in Ref. [17] for the first $2P$ moments. The form of the equation is the same, only the sizes of the matrices are different. Since the matrix $\mathcal{M}$ contains $2PN^2$ elements, Eq. (E4) defines $2PN^2$ nonlinear equations. Each vector $V_{\gamma'}$ has $N$ components and there are $PN$ such vectors. $V_{\gamma'}$ is required to be normalized and the gauge-transformation $V_{\gamma'} \rightarrow e^{i\theta} V_{\gamma'}$ does not affect $W_{\alpha\beta\gamma'} = V_{\alpha\gamma'} V_{\beta\gamma'}^*$. Thus, every $V_{\gamma'}$ is determined by $2(N - 1)$ real-valued unknowns, i.e., $2P(N^2 - N)$ unknown coefficients need to be found to determine all vectors $V_{\gamma'}$. Additionally, we need to find the $PN$ energies $E_{\gamma'}$ as well as the $PN$ spectral weights $a_{\gamma\gamma'}$. Consequently, Eq. (E4) is a system of $2PN^2$ nonlinear equations for $2PN^2$ unknowns. Thus, one may expect that it should be possible to compute $PN$ bands from the first $2P$ spectral moment matrices of size $N \times N$, because the number of unknowns matches the number of available nonlinear equations.
[26] O. K. Kalashnikov and E. S. Fradkin, The spectral density method applied to systems showing phase transitions, Physica status solidi (b) 59, 9 (1973).

[27] S. R. White, Spectral weight function for the two-dimensional Hubbard model, Phys. Rev. B 44, 4670 (1991).

[28] V. Turkowski and J. K. Freericks, Nonequilibrium sum rules for the retarded self-energy of strongly correlated electrons, Phys. Rev. B 77, 205102 (2008).

[29] W. Nolting and W. Brewer, Fundamentals of Many-body Physics: Principles and Methods (Springer Berlin Heidelberg, 2009).

[30] J. P. Perdew and Y. Wang, Accurate and simple analytic representation of the electron-gas correlation energy, Phys. Rev. B 45, 13244 (1992).

[31] J. P. Perdew and A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, Phys. Rev. B 23, 5048 (1981).

[32] J. Heyd, G. E. Scuseria, and M. Ernzerhof, Hybrid functionals based on a screened coulomb potential, The Journal of Chemical Physics 118, 8207 (2003).

[33] D. M. Ceperley and B. J. Alder, Ground state of the electron gas by a stochastic method, Phys. Rev. Lett. 45, 566 (1980).

[34] M. Gell-Mann and K. A. Brueckner, Correlation energy of an electron gas at high density, Phys. Rev. 106, 364 (1957).

[35] W. Eberhardt and E. W. Plummer, Angle-resolved photoemission determination of the band structure and multielectron excitations in ni, Phys. Rev. B 21, 3245 (1980).

[36] D. E. Eastman, F. J. Himpsel, and J. A. Knapp, Experimental band structure and temperature-dependent magnetic exchange splitting of nickel using angle-resolved photoemission, Phys. Rev. Lett. 40, 1514 (1978).

[37] A. K. See and L. E. Klebanoff, Spin-resolved x-ray photoemission from ferromagnetic nickel, Phys. Rev. B 51, 11002 (1995).

[38] A. I. Lichtenstein, M. I. Katsnelson, and G. Kotliar, Finite-temperature magnetism of transition metals: An ab initio dynamical mean-field theory, Phys. Rev. Lett. 87, 067205 (2001).

[39] I. Inoue, I. Hase, Y. Aiura, A. Fujimori, K. Morikawa, T. Mizokawa, Y. Haruyama, T. Maruyama, and Y. Nishihara, Systematic change of spectral functions observed by controlling electron correlation in casio3 with fixed 3d1 configuration, Physica C: Superconductivity 235, 1007 (1994).

[40] A. Sekiyama, H. Fujiwara, S. Imada, S. Suga, H. Eisaki, S. I. Uchida, K. Takahara, H. Harima, Y. Saitoh, I. A. Nekrasov, G. Keller, D. E. Kondakov, A. V. Kozhevnikov, T. Pruschke, K. Held, D. Vollhardt, and V. I. Anisimov, Mutual experimental and theoretical validation of bulk photoemission spectra of r1−xca2vo3, Phys. Rev. Lett. 93, 156402 (2004).

[41] B. Amadon, F. Lechermann, A. Georges, F. Jollet, T. O. Wehling, and A. I. Lichtenstein, Plane-wave based electronic structure calculations for correlated materials using dynamical mean-field theory and projected local orbitals, Phys. Rev. B 77, 205112 (2008).

[42] M. Potthoff, T. Herrmann, T. Wegner, and W. Nolting, The moment sum rule and its consequences for ferromagnetism in the Hubbard model, Physica status solidi (b) 210, 199 (1998).
[38] W. Nolting and A. M. Oles, Spectral density approach for the quasiparticle concept in the s-f model (ferromagnetic semiconductors), *Journal of Physics C: Solid State Physics* **13**, 2295 (1980).

[39] G. Peter Lepage, A new algorithm for adaptive multidimensional integration, *Journal of Computational Physics* **27**, 192 (1978).

[40] G. Peter Lepage, Adaptive multidimensional integration: vegas enhanced, *Journal of Computational Physics* **439**, 110386 (2021).