

# 5 Projectors, Hubbard $U$ , Charge Self-Consistency, and Double-Counting

Tim Wehling

Institute for Theoretical Physics

Bremen Center for Computational Material Sciences

University of Bremen

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Correlated subspaces and projectors</b>	<b>3</b>
2.1	Quantum impurity problems . . . . .	4
2.2	Projector formalism in LDA+DMFT . . . . .	6
<b>3</b>	<b>Interaction terms: Hubbard <math>U</math> and beyond</b>	<b>10</b>
3.1	The constrained random phase approximation . . . . .	12
3.2	Non-local Coulomb interactions . . . . .	14
<b>4</b>	<b>Double-counting and charge self-consistency</b>	<b>17</b>
<b>5</b>	<b>Conclusions</b>	<b>21</b>

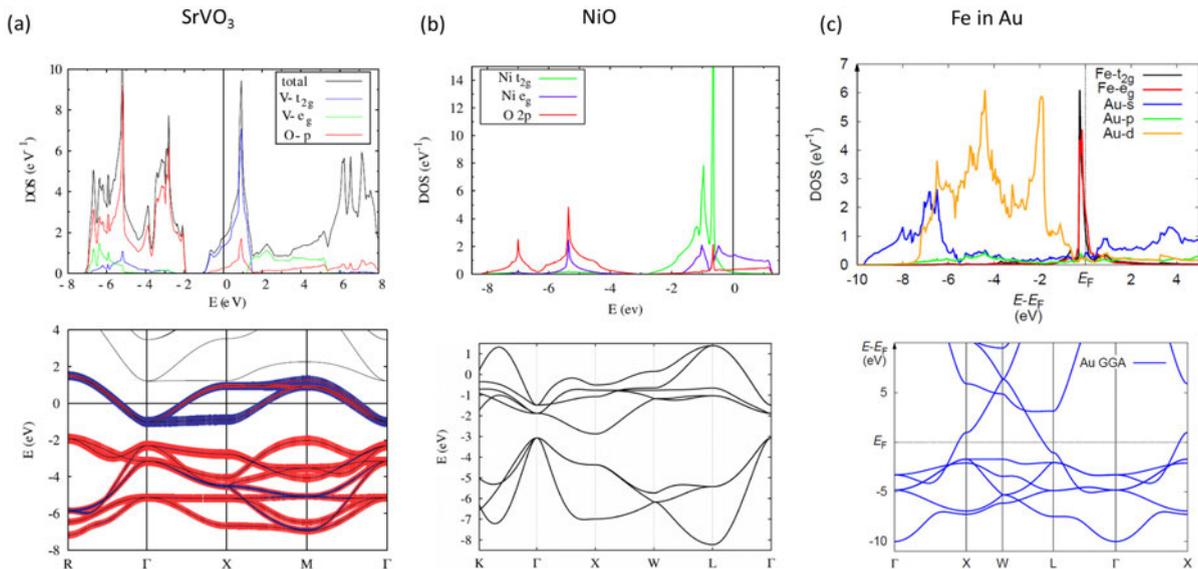
# 1 Introduction

The electronic structure problem in real materials involves two sources of complexity: First, on a single-particle level, a large Hilbert space can be required to describe how electronic wave functions adjust to a certain arrangement of nuclei. Second, electrons interact, and the dimension of the Fock space to describe a many-electron system grows exponentially with system size. There are two kinds of electronic structure approaches that circumvent either one of these problems.

So-called first-principles methods – most prominently density functional theory (DFT) – map the interacting-electron problem onto an auxiliary single-particle system but treat the full complexity of the single-particle wave functions. Thereby, DFT and related approaches provide material specific and atomistic descriptions of various extended systems (particularly *sp*-electron metals). These approaches fail, however, for strongly correlated materials such as transition-metal oxides or Kondo systems, where the electronic ground state requires a superposition of multiple Slater determinants and where excitations are governed by dynamic self-energy effects. The description of strongly correlated electron systems relies on model Hamiltonians, which operate on many-body Fock spaces constructed out of a reduced set of single-particle basis states. In this way, several correlation phenomena such as the Kondo effect, metal-insulator transitions, magnetism or unconventional superconductivity can be addressed but at the price that relations of models and materials are sometimes ambiguous due to *a priori* unknown model parameters.

Obviously, first-principles and model Hamiltonian approaches have complementary merits and shortcomings, so that their combination presents a promising route towards realistic, i.e., atomistic and material specific, descriptions of strongly correlated electron systems. The combination of density functional theory with dynamical mean field theory, termed LDA+DMFT, is one very successful example of such hybrid approaches and is introduced in the lecture of Alexander Lichtenstein. Generally, approaches which are based on combining DFT and model Hamiltonian approaches to correlated systems are referred to as DFT++ [1].

In this lecture, we will discuss how the realm of first-principles theories can be brought into model Hamiltonian approaches to strongly correlated electron systems. We will see in section 2 that a possible strategy to do so works as follows: A correlated subspace  $\mathcal{C}$ , i.e., a subset of single-particle states where electron correlations take place, is identified. Then, it will turn out that projectors from the full space of Kohn-Sham eigenstates onto the subspace  $\mathcal{C}$  provide a general way for linking DFT and model Hamiltonian approaches. Within the correlated subspace, the DFT band structure is augmented with interaction terms to generate the correlation effects missed in DFT. These interactions affect only a subset of orbitals and often only local interactions are considered, while clearly electrons in all orbitals and also at distant sites interact with each other. The question of how to determine meaningful interaction parameters entering the many-body models will be addressed in section 3. DFT includes already some (partly unknown portion of) interaction effects in a static mean field manner, which have to be accounted for. This leads to so-called double-counting corrections which are discussed in section 4. The inclu-



**Fig. 1:** *Electronic structure of three example systems. (a) SrVO<sub>3</sub>, (b) NiO, and (c) Fe impurities in Au. The electronic density of states (DOS) and band structures as obtained from DFT are shown. For Fe in Au the band structure of the host is shown. The systems illustrate different levels of complexity on the single-particle level. While for SrVO<sub>3</sub> the block of three  $t_{2g}$ -bands around the Fermi level turns out to control the low energy physics, NiO is experimentally known to be a charge transfer insulator and a description of elementary electronic excitations requires to consider both Ni-3d and O-2p states. For Fe impurities in Au, there is a continuum of  $sp$ -like host states together with impurity 3d-states at the Fermi level. Crystal momentum is no more a good quantum number of single-particle states in the the impurity system, which makes its description already involved on the single-particle level.*

sion of interactions in the correlated subspace can lead to a redistribution of charges between different orbitals of the system, which would in turn modify the mean field interaction terms contained in DFT and thereby included in the single-particle part of the DFT++ models. Resulting issues of charge self-consistency between the DFT and the many-body parts will finally be considered in section 4.

## 2 Correlated subspaces and projectors

We start our discussion with SrVO<sub>3</sub>, NiO, and magnetic Fe impurities in Au, which are examples of, respectively, correlated metals, charge transfer insulators and Kondo systems. Band structures and density of states as derived from DFT calculations are summarized in Fig. 1.

SrVO<sub>3</sub> has a block of three bands, the so-called  $t_{2g}$  bands, in the vicinity of the Fermi level. These bands have mainly V 3d character and it turns out that the low energy electronic structure can be understood in terms of these bands. In other words, a many-body Hamiltonian for the description of correlation effects in SrVO<sub>3</sub> could be obtained from these  $t_{2g}$  bands alone.

NiO is a so-called charge transfer insulator [2]. In contrast to Mott-Hubbard insulators, where the Hubbard  $U$  opens a charge gap within the transition-metal  $d$  bands, in charge-transfer systems the gap typically opens between ligand  $p$  bands and the upper Hubbard bands derived from

the transition-metal  $d$  states. Thus, the so-called charge transfer energy  $|\varepsilon_p - \varepsilon_d|$  determines the size of the gap. Quite often, ligand  $p$  bands and transition-metal  $d$  states mix through hybridization and it is thus the interplay of transition-metal  $d$  states and oxygen  $2p$  states that defines the fundamental electronic excitations. Paramagnetic DFT fails to reproduce this behavior; it predicts NiO to be metallic with Ni- $e_g$  bands close to the Fermi level,  $E_F = 0$ , and O  $2p$ -states more than 3eV below  $E_F$ . It reveals, nevertheless, hybridization between Ni and O-derived bands. A description of the electronic excitations of NiO should involve the intermixed O  $2p$  and Ni  $3d$ -bands, while we expect that electronic correlation effects result mostly from the partially filled transition-metal  $3d$  states. A natural many-body model would thus involve 8 bands, i.e., 3 carrying mainly O  $2p$ -weight and 5 bands derived from the Ni  $3d$  states. We would then end up with a correlated subspace of dimension 5 embedded into larger 8 dimensional space of single-particle Bloch states.

Finally, Fe in Au is an impurity problem, and there are no well defined bands since the crystal momentum  $k$  is not a good quantum number of the single-particle states any more. In practice, impurity problems are often modeled using supercells containing the transition-metal impurity atom and on the order of hundred atoms to mimic the host. Assuming that electron correlations mainly take place in the impurity  $d$ -orbitals one still has a correlated subspace of dimension 5 but embedded into single-particle Hilbert space spanned by a few hundred supercell bands. We will below see that coupling between the correlated and uncorrelated parts of the single-particle Hilbert space can be elegantly formulated based on projection operators. The discussion follows mainly Refs. [3–8].

The first step of any DFT++ approach is to identify a correlated subspace  $\{|m\rangle\}$ , where the Kohn-Sham Hamiltonian  $H_K$  is augmented by interactions  $H_U$  and a double-counting correction  $H_{DC}$ . One thus arrives at a Hamiltonian

$$H = \underbrace{\sum_k \varepsilon_k c_k^\dagger c_k}_{H_K} - \underbrace{\mu_{DC} \sum_m d_m^\dagger d_m}_{H_{DC}} + \frac{1}{2} \underbrace{\sum_{m\dots m'''} U_{m\dots m'''} d_m^\dagger d_m^\dagger d_{m''} d_{m'''} }_{H_U}, \quad (1)$$

where the Kohn-Sham energies  $\varepsilon_k$  and the Kohn-Sham eigenstates  $|k\rangle$  obtained from DFT define the non-interacting starting point. To make calculations feasible one often assumes local interactions that couple only the correlated orbitals at the same site. In the general form of Eq. (1),  $H$  could be a multiband Hubbard or a multiorbital Anderson Impurity Model (AIM).

## 2.1 Quantum impurity problems

We start with the discussion of an impurity problem, as for instance realized by ad-atoms on surfaces or by magnetic dopants in bulk metals. We will show that projections  $\langle k|m\rangle$  of the Kohn-Sham states onto the states of the correlated subspace are sufficient to connect the DFT real material simulations with the multi-orbital AIM. In general, states  $|k\rangle$  and  $|m\rangle$  have a finite overlap  $\langle k|m\rangle \neq 0$ . i.e., they are non-orthogonal. If one constructs an orthonormalized basis of single-particle states which includes the states  $|m\rangle$  and an orthogonal set  $\{|\tilde{k}\rangle\}$  with

corresponding Fermi operators  $c_{\tilde{k}}$  one could rewrite the Hamiltonian in the form

$$H_{\text{AIM}} = \sum_{\tilde{k}} \varepsilon_{\tilde{k}} c_{\tilde{k}}^\dagger c_{\tilde{k}} + \sum_{\tilde{k}, m} (V_{\tilde{k}m} c_{\tilde{k}}^\dagger d_m + h.c.) + \sum_m (\varepsilon_m - \mu_{DC}) d_m^\dagger d_m + \frac{1}{2} \sum_{m \dots m'} U_{m \dots m'} d_m^\dagger d_{m'}^\dagger d_{m''} d_{m'''}. \quad (2)$$

Equivalently, the problem can be characterized through the corresponding action

$$S_{\text{AIM}}(c^*, c, d^*, d) = \int_0^\beta d\tau \sum_{\tilde{k}} c_{\tilde{k}}^*(\tau) \partial_\tau c_{\tilde{k}}(\tau) + \sum_m d_m^*(\tau) \partial_\tau d_m(\tau) + H_{\text{AIM}}(c^*, c, d^*, d)(\tau), \quad (3)$$

where the Fermi operators are replaced by Grassmann numbers, which leads to the partition function  $\mathcal{Z}$  via the imaginary time path integral

$$\mathcal{Z} = \int \mathcal{D}[c^*, c, d^*, d] e^{-S_{\text{AIM}}(c^*, c, d^*, d)}. \quad (4)$$

This integral is Gaussian in the Grassmann numbers  $c_{\tilde{k}}^*$  and  $c_{\tilde{k}}$ , i.e., the bath parts can be integrated out and we arrive at an effective action

$$S_{\text{eff}}(d^*, d) = - \int_0^\beta d\tau \int_0^\beta d\tau' \sum_{m, m'} d_m^*(\tau) (\mathcal{G}_0^{-1}(\tau - \tau')) d_m(\tau') + \int_0^\beta d\tau H_{\text{U}}(d^*, d)(\tau), \quad (5)$$

which specifies all local electronic properties of the AIM. The bath enters indirectly through the non-interacting Green function  $\mathcal{G}_0$  of the correlated orbitals.  $\mathcal{G}_0$  can be easily obtained from our first-principles calculations: As the Kohn-Sham eigenvalues and eigenstates *define* the non-interacting starting point of our DFT++ impurity treatment, the Kohn-Sham Green function

$$G_{KS}(i\omega_n) = \sum_k \frac{|k\rangle\langle k|}{i\omega_n + \mu - \varepsilon_k} \quad (6)$$

plays the role of the non-interacting Green function and the matrix elements entering Eq. (5) read

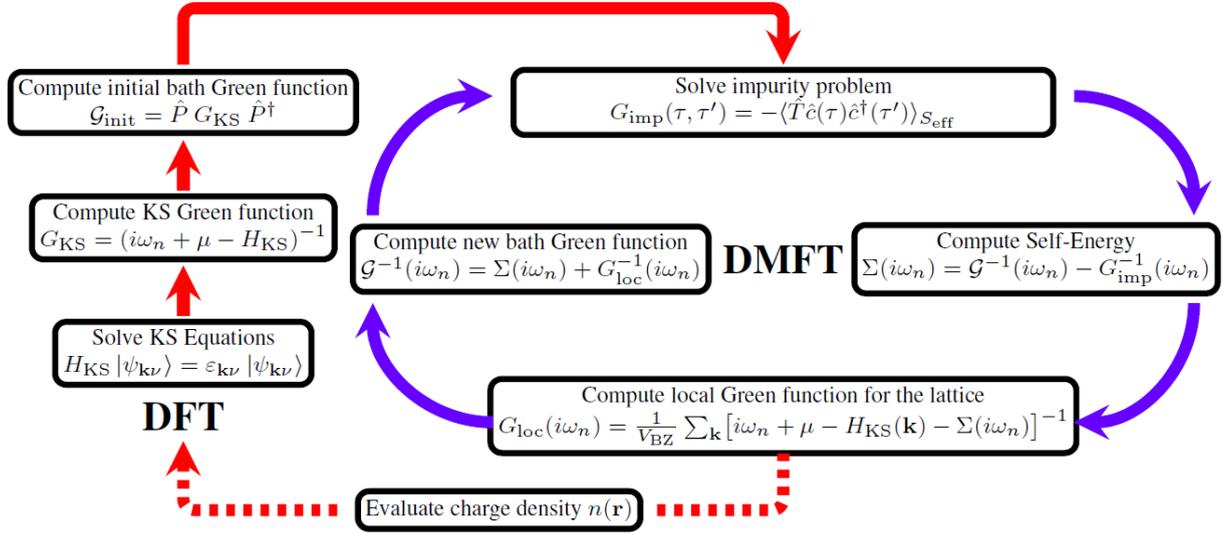
$$\mathcal{G}_0(i\omega_n)_{mm'} = \langle m | G_{KS}(i\omega_n) | m' \rangle = \sum_k \frac{\langle m | k \rangle \langle k | m' \rangle}{i\omega_n + \mu - \varepsilon_k}. \quad (7)$$

The single-particle part of the AIM is thus fully defined, once the DFT Kohn-Sham problem of the system of interest is solved and the projections  $\langle k | m \rangle$  of the Kohn-Sham eigenstates onto the basis states of the correlated subspace are known.

Defining the projector  $\mathcal{P}_{\mathcal{C}} = \sum_m |m\rangle\langle m|$  onto the correlated subspace  $\mathcal{C}$ ,  $\mathcal{G}_0(i\omega_n)$  from Eq. (7) can be represented in a compact matrix notation

$$\mathcal{G}_0(i\omega_n) = \mathcal{P}_{\mathcal{C}} G_{KS}(i\omega_n) \mathcal{P}_{\mathcal{C}}. \quad (8)$$

In this way, the single-particle terms entering the DFT++ quantum impurity model for a complex system like a magnetic impurity in a metal or metal-organic molecules on metal surfaces are defined from first-principles. Once the interaction (see section 3) and double-counting terms (see section 4) are also specified, the impurity model, Eq. (5), can be solved, e.g., by quantum Monte Carlo methods as explained in the lecture by Fakher Assaad.



**Fig. 2:** *DFT+DMFT self-consistency cycle. The algorithm starts with a DFT calculation, which yields the Kohn-Sham Hamiltonian and Kohn-Sham eigenstates  $|k\rangle \equiv |\psi_{\mathbf{k}\nu}\rangle$ . Then, the KS Green function and from it the starting value for the bath Green function  $\mathcal{G}_0$  are constructed and passed on to the DMFT loop. The DMFT loop involves unfolding of the self-energy, Eq. (15), and downfolding of the full Green function to the correlated subspace, Eq. (17). Both can be accomplished using projectors. A potential self-consistency over the charge density  $n(\mathbf{r})$  is also indicated. From Ref. [8].*

## 2.2 Projector formalism in LDA+DMFT

Dynamical mean field theory maps correlated lattice models such as the Hubbard model onto Anderson impurity models with a self-consistency condition, as explained in the lecture by Antoine Georges. In DMFT, the auxiliary impurity problems involve the correlated orbitals with local interaction, coupled to a self-consistent energy-dependent bath. In the effective action formulation, this takes again the form of Eq. (5), where only the meaning of  $\mathcal{G}_0$  changes. It is no longer the bare (i.e. non-interacting) local Green function of the DFT++ Hamiltonian (1), but rather the dynamical mean-field, which is determined self-consistently and which encodes the coupling of the embedded atom to the effective bath. In other words,  $\mathcal{G}_0$  is the analogue of the Weiss field in the mean field theory of classical magnets. Due to the self-consistency cycle of DMFT,  $\mathcal{G}_0$  depends on many-body effects in the material under consideration.

The central quantity in the DMFT formalism is the local Green function  $G_{\mathbf{R}}^{\text{loc}}(i\omega_n)$ , which contains simply those matrix elements of the full Green function  $G(i\omega_n)$  which belong to correlated orbitals from the same site  $R$ . We can thus use the projection operators  $\mathcal{P}_{\mathbf{R}}^{\text{C}} = \sum_m |\mathbf{R}m\rangle\langle\mathbf{R}m|$  to correlated orbitals  $|\mathbf{R}m\rangle$  at site  $\mathbf{R}$  to write the local Green function

$$G_{\mathbf{R}}^{\text{loc}}(i\omega_n) = \mathcal{P}_{\mathbf{R}}^{\text{C}} G(i\omega_n) \mathcal{P}_{\mathbf{R}}^{\text{C}}. \quad (9)$$

As before, we interpret the Green functions as operators acting on the space of single-particle states.

By definition of the self-energy  $\Sigma(i\omega_n)$ , the full Green function of the lattice problem is given

by

$$G(i\omega_n)^{-1} = G_{KS}(i\omega_n)^{-1} - \Sigma(i\omega_n). \quad (10)$$

Since the interaction terms of the DFT++ Hamiltonian are restricted to the correlated subspace, the self-energy can be non-zero only within  $\mathcal{C}$ . In the DMFT approximation,  $\Sigma(i\omega_n)$  is local and obtained from the auxiliary impurity problem according to

$$\Sigma(i\omega_n) = \mathcal{G}_0^{-1}(i\omega_n) - G_{imp}^{-1}(i\omega_n), \quad (11)$$

where  $G_{imp}(i\omega_n)$  is the Green function of the auxiliary impurity problem defined through  $S_{\text{eff}}$ . From the DMFT self-consistency condition

$$G_{imp}(i\omega_n) = G_{\mathbf{R}}^{\text{loc}}(i\omega_n), \quad (12)$$

we obtain a prescription on how to construct the bath Green function:

$$\mathcal{G}_0^{-1}(i\omega_n) = \Sigma(i\omega_n) + (G_{\mathbf{R}}^{\text{loc}})^{-1}(i\omega_n). \quad (13)$$

Eqs. (11) and (13) define the DMFT self-consistency cycle in a basis-independent way as illustrated in Fig. 2. For actual computations, a basis  $\{|B_{\mathbf{k}\alpha}\rangle\}$  (often referred to as Bloch basis) has to be chosen to represent the Green functions explicitly as matrices. There are two natural choices: one could either use the Kohn-Sham eigenstates or any basis set (e.g. plane waves, projector augmented plane waves, or full potential linearized augmented plane waves) which is implemented in the DFT-code used.

To translate the equations defining the self-consistency cycle (Eqs. (9–13)) to this basis set dependent notation, we need the matrix-representations of the Kohn-Sham Hamiltonian

$$H_{\text{KS}}(\mathbf{k})_{\alpha\alpha'} = \sum_k \langle B_{\mathbf{k}\alpha} | k \rangle \varepsilon_k \langle k | B_{\mathbf{k}\alpha'} \rangle \quad (14)$$

and of the self-energy operator

$$\Sigma_{\alpha\alpha'}(\mathbf{k}, i\omega_n) = \langle B_{\mathbf{k}\alpha} | \Sigma(i\omega_n) | B_{\mathbf{k}\alpha'} \rangle = \langle B_{\mathbf{k}\alpha} | m \rangle \Sigma_{mm'}(i\omega_n) \langle m' | B_{\mathbf{k}\alpha'} \rangle. \quad (15)$$

These lead directly to the full Green function

$$G_{\alpha\alpha'}(\mathbf{k}, i\omega_n) = \{i\omega_n + \mu - H_{\text{KS}}(\mathbf{k}) - \Sigma(\mathbf{k}, i\omega_n)\}_{\alpha\alpha'}^{-1} \quad (16)$$

in the Bloch basis. Eq. (15) obviously unfolds the self-energy, which is obtained from the solution of the impurity problem in the localized basis of  $\mathcal{C}$ , to the full space of Bloch basis functions. Although  $\Sigma(i\omega_n)$  is purely local when expressed in the set of correlated orbitals, it acquires in general momentum dependence when expressed in an arbitrary basis set.

The bath Green function is provided to the impurity solver in the localized basis and thus requires (cf. Eq. (13)) the local Green function in the localized basis

$$G_{\mathbf{R}}^{\text{loc}}(i\omega_n)_{mm'} = \sum_{\mathbf{k}, \alpha, \alpha'} \langle \mathbf{R}m | B_{\mathbf{k}\alpha} \rangle G_{\alpha\alpha'}(\mathbf{k}, i\omega_n) \langle B_{\mathbf{k}\alpha'} | \mathbf{R}m' \rangle. \quad (17)$$

That is indeed all we need to implement the DFT+DMFT cycle as depicted in Fig. 2. The projections  $\langle \mathbf{R}m | B_{\mathbf{k}\alpha} \rangle$  facilitate both the upfolding of the self-energy to the Bloch basis and downfolding of the full Green function to the local Green function on the correlated subspace. To summarize, DFT+DMFT requires decisions on the following two issues [4]:

1. The local orbitals  $|\mathbf{R}m\rangle$  spanning the correlated subspace  $\mathcal{C}$  have to be chosen. Different definitions of  $\mathcal{C}$  can lead to different results and the quality of the DMFT approximation will in general depend on the choice of  $\mathcal{C}$ . One might want to define  $\mathcal{C}$  such that the DMFT approximation is best justified, which is intuitively associated with well-localized orbitals.
2. To keep the DFT+DMFT computationally tractable, the basis of Bloch states  $\{|B_{\mathbf{k}\alpha}\rangle\}$  should be chosen in such a way that the number of states that have non-zero overlap with the correlated subspace  $\langle B_{\mathbf{k}\alpha} | \mathcal{P}_{\mathbf{R}}^{\mathcal{C}} | B_{\mathbf{k}\alpha} \rangle > 0$  remains sufficiently small. Taking  $\{|B_{\mathbf{k}\alpha}\rangle\}$  to be the Kohn-Sham eigenstates is often a good choice in this respect while simple plane waves typically lead to too large Bloch spaces.

There are indeed several possibilities for the construction of local orbitals including different flavors of Wannier functions such as maximally localized Wannier functions or so-called  $N$ th order Muffin Tin Orbitals (NMO). A very practical way to construct a basis for  $\mathcal{C}$  is to use entities that are already existing in most of the common band-structure codes, namely, the decomposition of local atomic-like orbitals  $|\widetilde{\mathbf{R}m}\rangle$  in terms of Bloch basis functions [5]. Indeed, if the set of Bloch states (e.g. the Kohn-Sham eigenstates generated by the DFT code) were complete, we could simply take the set  $\{|\widetilde{\mathbf{R}m}\rangle\}$  as basis of the correlated subspace  $\mathcal{C}$ .

However, independently of the particular Bloch basis set which is chosen, one has to restrict practical DFT+DMFT calculations always to a finite space of  $N_B$  Bloch states. Those states span a finite subspace  $\mathcal{W}$  of the total Hilbert space. The local atomic-like states  $|\widetilde{\mathbf{R}m}\rangle$  will, in general, have a decomposition involving all Bloch bands. Projections of  $\{|\widetilde{\mathbf{R}m}\rangle\}$  onto  $\mathcal{W}$  can thus lead to a non-orthonormal set of localized states. The obvious way out is to reorthonormalize, which is easiest done in the following way:

We consider the Bloch transform of the local atomic-like orbitals,  $|\widetilde{\mathbf{k}m}\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\mathbf{R}} |\widetilde{\mathbf{R}m}\rangle$ , where  $N$  is the number of atoms in the crystal. The projections of  $|\widetilde{\mathbf{k}m}\rangle$  onto  $\mathcal{W}$  reads

$$|\widetilde{\mathbf{k}m}\rangle = \sum_{\alpha \in \mathcal{W}} |B_{\mathbf{k}\alpha}\rangle \langle B_{\mathbf{k}\alpha} | \widetilde{\mathbf{k}m}\rangle. \quad (18)$$

The  $\{|\widetilde{\mathbf{k}m}\rangle\}$  are not true Wannier functions as they are not orthonormal, i.e., their overlap matrix  $O_{mm'}(\mathbf{k}) = \langle \widetilde{\mathbf{k}m} | \widetilde{\mathbf{k}m'} \rangle$  is not the unit matrix. We arrive however at an orthonormal basis set of the correlated subspace by orthonormalizing according to

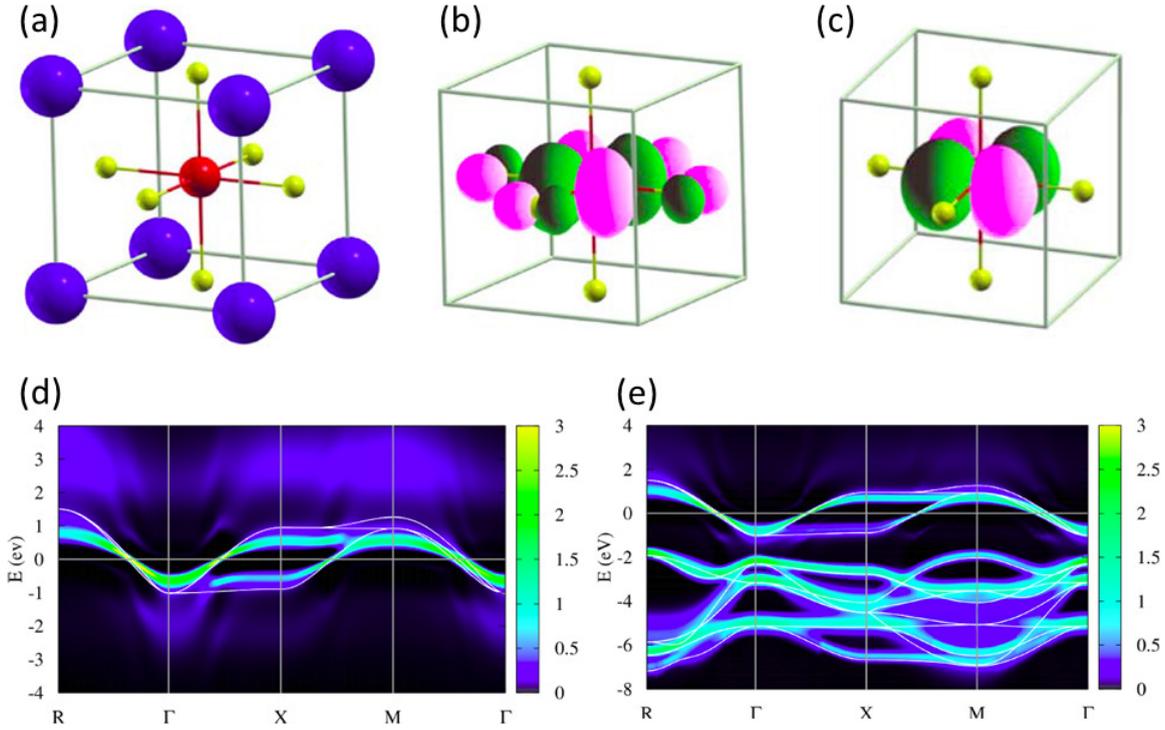
$$|\mathbf{k}m\rangle = \sum_{m'} O_{mm'}^{-1/2}(\mathbf{k}) |\widetilde{\mathbf{k}m'}\rangle. \quad (19)$$

Eqs. (18) and (19) indeed define a proper Wannier construction. We call the resulting Wannier functions "projector guided Wannier functions" (PWF). PWFs often serve as starting point of an MLWF localization procedure. The extent to which PWFs differ from MLWFs is system-specific. For rather localized states, like in transition-metal oxides, the differences are indeed small [6], while they are larger in materials with highly extended electronic states like  $p_z$  electrons in graphene.

More severe than differences between PWFs and MLWFs are differences associated with different choices of the Bloch space  $\mathcal{W}$ . This shall be illustrated with the example of  $\text{SrVO}_3$ , which is a metal with one electron in the  $t_{2g}$  bands and empty  $e_g$  bands (see Fig. 1).  $\text{SrVO}_3$  is a good test-case for DFT+DMFT calculations because excitation spectra and thermodynamic properties hint at correlation effects taking place: First, direct comparison of the photoemission spectra with the one-particle band-structures, e.g. from DFT-LDA, yields poor agreement. Moreover, the linear coefficient in the temperature-dependent specific heat is twice larger than estimated from DFT-LDA, which suggests correlation induced mass enhancement [4].

For an DFT+DMFT description of  $\text{SrVO}_3$ , one could make two rather different choices of  $\mathcal{W}$ : First, we could focus on a very limited set of low-energy Bloch bands, such as the three  $t_{2g}$  bands in the vicinity of the Fermi level (cf. Fig. 1) and generate  $\mathcal{W}$  just from the three corresponding Bloch functions. In this case, we would have  $\mathcal{W} = \mathcal{C}$ . Since now the Bloch bands span a narrow energy window, the Wannier functions defining  $\mathcal{C}$  will be rather spatially extended: As Fig. 3b shows, they are centered on vanadium atoms but also have sizable weight on neighboring oxygen atoms, which reflects the hybrid character of the low energy  $t_{2g}$ -type states. This first choice of  $\mathcal{W} = \mathcal{C}$  is of course appealing since it involves a minimal number of bands but comes at the expense that the investigation of the indirect effects of correlations on bands other than the  $t_{2g}$  ones are out of scope.

Alternatively, we could define  $\mathcal{W}$  from a larger energy window of Bloch bands including all bands associated with the O  $2p$  and all V  $3d$  states. Then, the indirect impact of electron correlations on largely O  $2p$ -derived bands can also be addressed. Having an enlarged Bloch space  $\mathcal{W}$  means also that the basis orbitals  $|\mathbf{R}m\rangle$  of the correlated subspace can involve Bloch states from a correspondingly wider energy range (cf. Eq. (19)) and will be more spatially localized. As Fig. 3c shows, they are now indeed much closer to vanadium atomic-like orbitals. A comparison of  $\mathbf{k}$ -resolved spectral functions as obtained from DFT+DMFT simulations of  $\text{SrVO}_3$  with a Bloch space made up by the  $t_{2g}$  bands only and involving additionally the oxygen  $2p$  states is given in Fig. 3. In both cases, we find a narrowing of the  $t_{2g}$ -like quasi-particle states close to the Fermi level with renormalization factors  $Z \approx 0.6$ , which is in line with experimentally found band narrowing [6]. Also, independently of the choice of  $\mathcal{W}$ , we see spectral weight being transferred from the  $t_{2g}$  bands to lower and upper Hubbard bands. The major effect of different choices of  $\mathcal{W}$  concerns, as expected, the oxygen  $2p$  derived states. For the larger Bloch space, hybridization with the correlated subspace leads to lifetime broadening of the oxygen  $2p$  bands. While for  $\text{SrVO}_3$  both models yield a reasonable description of the low energy physics, there are many materials for which the inclusion of bands beyond the correlated subspace is absolutely required. For example, in charge-transfer insulators like NiO



**Fig. 3:** (a)  $\text{SrVO}_3$  structure with Sr large (blue), V (red), and O small (yellow). (b,c) Perspective view of Maximally Localized Wannier functions of the  $t_{2g}$  states as obtained from the three  $t_{2g}$ -like bands around the Fermi level (b) and from a 14 band calculation, i.e., involving three  $2p$ -like states from three oxygen ions and five  $3d$ -like states from the vanadium ion. From Ref. [4]. Momentum resolved spectral functions as calculated within DFT+DMFT using models involving three Bloch bands only (d) and a Bloch space involving also the oxygen derived states (e). The DFT bands are indicated as white lines. From Ref. [6].

(see section 4), the inclusion of oxygen  $2p$  states in the DFT+DMFT procedure is necessary to describe the fundamental electronic excitations.

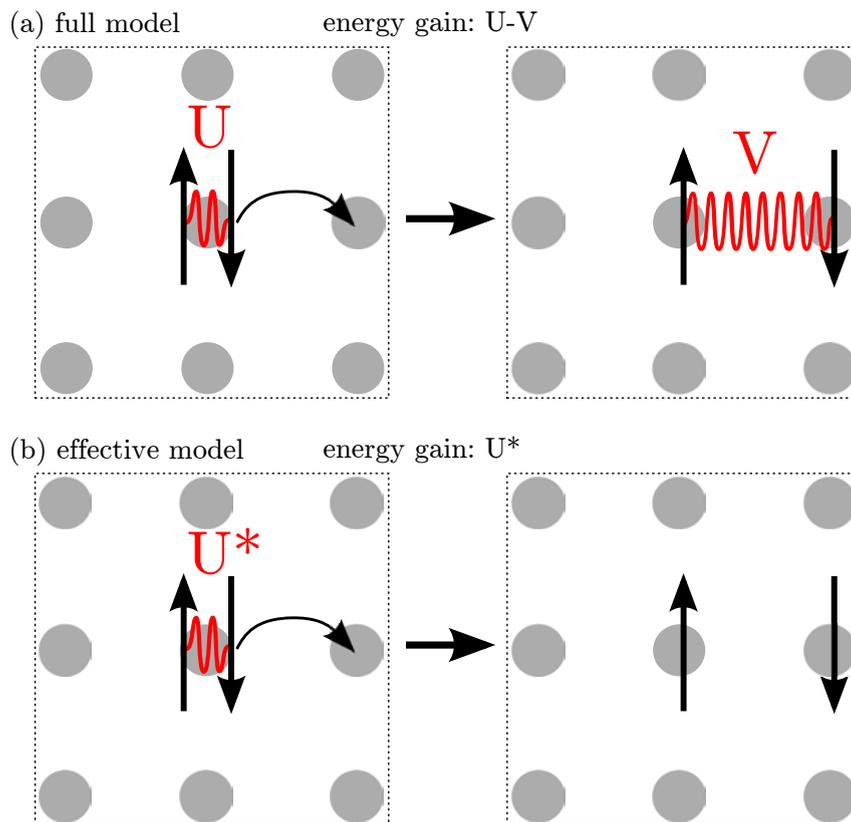
Using the projector formalism outlined so far, we can in principle fix all terms entering the DFT++ Hamiltonian, Eq. (1), apart from the double-counting shifts and the interaction matrices. We will discuss these terms in the following two sections.

### 3 Interaction terms: Hubbard $U$ and beyond

Correlation effects are generated by the interaction terms in the DFT++ Hamiltonian, Eq. (1). What are these interaction terms? Naively, one might guess that they could be obtained as matrix elements of the *bare* Coulomb interaction

$$U_{m\dots m'''} \stackrel{?}{=} \langle \mathbf{R}m | \langle \mathbf{R}m' | \frac{e^2}{|\hat{\mathbf{r}} - \hat{\mathbf{r}}'|} | \mathbf{R}m'' \rangle | \mathbf{R}m''' \rangle.$$

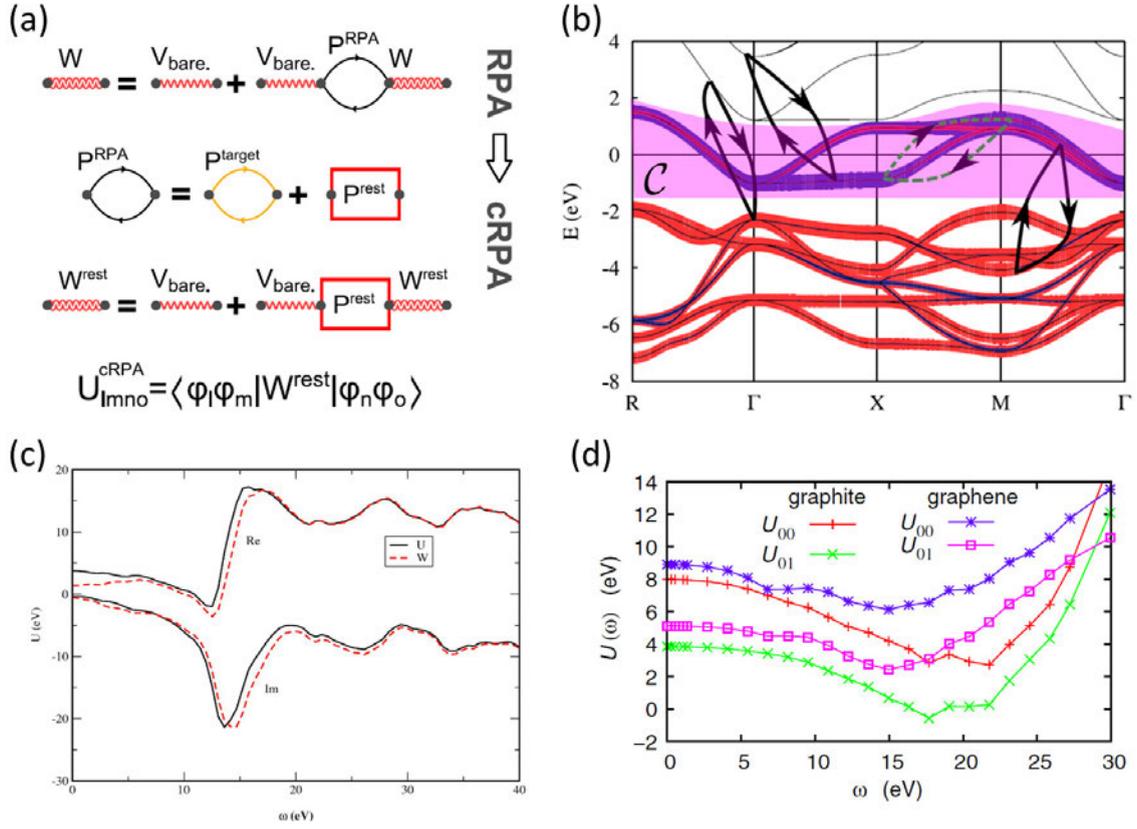
Indeed, that is not the case, since the DFT++ Hamiltonian restricts interactions to the correlated subspace  $\mathcal{C}$  and also assumes that interactions are local. The *Gedankenexperiment* depicted in Fig. (4) illustrates the problem.



**Fig. 4:** Interaction energies in models with non-local (a) and purely on-site interactions (b). Wavy lines illustrate Coulomb interactions. (a) An electron in the extended Hubbard model hopping from a doubly occupied site to an empty one, gains an energy  $U - V$ . (b) The same situation in the model with purely local interactions leads to an energy gain of  $U^*$ . From Ref. [9].

In a model with on-site repulsion  $U$  and nearest neighbor interaction  $V$ , an electron hopping from a doubly occupied site to a neighboring empty site gains interaction energy  $U - V$ , while in a model with purely local Hubbard repulsion  $U^*$  the corresponding energy gain is  $U^*$ . A similar argument applies for electrons moving from a correlated orbital of an atom, e.g., the  $3d$ -orbital of a transition-metal atom like Fe or Ni, to an uncorrelated  $4p$ -orbital at the same site. Here, one would have to account for neglecting the repulsion  $U_{pd}$  between  $p$  and  $d$  electron when determining some Hubbard interaction which is restricted to the space of  $d$ -orbitals. Thus far, our discussion involved only two electrons. Obviously in a real solid, all surrounding electrons rearrange in response to the Coulomb potential generated by a charge moving through the material and Coulomb interactions will be screened. The DFT++ model Hamiltonian involves explicitly a limited set of Coulomb processes and associated screening effects. It thus accounts for some amount of electronic screening but many screening channels are indeed neglected. The Coulomb matrix elements entering the DFT++ model should therefore be derived from an appropriately partially screened interaction.

One approach to this problem is the so-called constrained local-density approximation (cLDA) [10]: The Hubbard interaction  $U$ , e.g. between  $3d$ -electrons at the same site, contributes to the



**Fig. 5:** Screened interaction in the random phase approximation (a). (From Ref. [12]) The constrained random phase approximation (cRPA) excludes bubbles which are entirely within the correlated subspace  $\mathcal{C}$  as illustrated here with the example of  $\text{SrVO}_3$ . If we choose  $\mathcal{C}$  to be the  $t_{2g}$  block around  $E_F$ , the bubbles (dashed) involving  $t_{2g}$  to  $t_{2g}$  transitions are excluded (b). Frequency dependent effective interactions in  $\text{SrVO}_3$  (c) (from Ref. [13]) and graphene (d) (from Ref. [14]) as obtained from cRPA.

total energy a term of the form  $E = \frac{1}{2}U n_{3d}(n_{3d} - 1)$ . The second derivative of  $E$  with respect to the occupancy thus yields the Hubbard interaction

$$U = \frac{\partial^2 E}{\partial n_{3d}^2}. \quad (20)$$

This approach is within the DFT reasoning and has been implemented in several DFT codes.

### 3.1 The constrained random phase approximation

More recently, a method implementing the idea of partial screening in a diagrammatic language – the so-called constrained random phase approximation (cRPA) – has been proposed [11] and is widely used to date. Let us assume that we have a solid with a well-defined correlated subspace of flat bands near the Fermi level, e.g. transition-metal  $3d$  bands, with the remainder including, for instance,  $4s$  or  $4p$  bands. For simplicity, we assume that there is no hybridization between  $3d$  and  $4sp$  states, i.e., we can span our Bloch space by Kohn-Sham eigenstates  $|\psi_d\rangle$  and  $|\psi_r\rangle$  referring to  $3d$ -states and the rest, respectively.

The idea is now to construct a partially screened interaction, which accounts for all screening processes except for those involving  $3d$ -to- $3d$ -transitions since the latter processes will be contained in our DFT++ Hamiltonian and should not be counted twice. We can implement this idea within the random phase approximation if we include all bubble diagrams except for those involving  $3d$ -to- $3d$ -transitions in the expansion of the polarization function (cf. Fig. 5). The sum of bubble diagrams leads to the polarization operator

$$\hat{P}(i\Omega_m) = -\frac{1}{\beta} \sum_{i\omega_n} \hat{G}(i\omega_n + i\Omega_m) \hat{G}(i\omega_n). \quad (21)$$

The screened interaction  $W$  in RPA is given by

$$\hat{W} = \hat{v} + \hat{v}\hat{P}\hat{v} + \hat{v}\hat{P}\hat{v}\hat{P}\hat{v} + \dots = \left[1 - \hat{v}\hat{P}\right]^{-1} \hat{v}. \quad (22)$$

Taking  $G$  to be the Kohn-Sham Green function, the evaluation of the Matsubara sum, (21), leads to a sum over transitions between occupied and empty states which reads in position space representation as

$$P(\mathbf{r}, \mathbf{r}'; i\Omega_m) = \sum_i^{\text{occ}} \sum_j^{\text{empty}} \psi_i(\mathbf{r}) \psi_i^*(\mathbf{r}') \psi_j(\mathbf{r}') \psi_j^*(\mathbf{r}) \left\{ \frac{1}{i\Omega_m + \varepsilon_i - \varepsilon_j} - \frac{1}{i\Omega_m + \varepsilon_j - \varepsilon_i} \right\}. \quad (23)$$

We split  $\hat{P}$  into  $\hat{P} = \hat{P}_d + \hat{P}_r$ , where  $\hat{P}_d$  includes only  $3d$  to  $3d$  transitions (i.e, restricting the sums in Eq. (23) to  $i, j \in \mathcal{C}$ ), and  $P_r$  be the rest of the polarization. The screened interaction in RPA can then be expressed as

$$\begin{aligned} \hat{W} &= \left[1 - \hat{v}\hat{P}_r - \hat{v}\hat{P}_d\right]^{-1} \hat{v} \\ &= \left[(1 - \hat{v}\hat{P}_r) \{1 - (1 - \hat{v}\hat{P}_r)^{-1} \hat{v}\hat{P}_d\}\right]^{-1} \hat{v} \\ &= \left[1 - (1 - \hat{v}\hat{P}_r)^{-1} \hat{v}\hat{P}_d\right]^{-1} (1 - \hat{v}\hat{P}_r)^{-1} \hat{v} \\ &= \left[1 - \hat{W}_r \hat{P}_d\right]^{-1} \hat{W}_r, \end{aligned} \quad (24)$$

where

$$\hat{W}_r(i\Omega_n) = (1 - \hat{v}\hat{P}_r(i\Omega_n))^{-1} \hat{v} \quad (25)$$

is the partially screened interaction we were searching for [11]. It describes the interaction between the electrons of the correlated subspace and accounts for screening by the rest of the system in RPA. The matrix elements of  $\hat{W}_r$  can therefore be used obtain the interaction terms entering the DFT++ model from first principles

$$U_{m\dots m'''}(i\Omega_n) = \langle \mathbf{R}m | \langle \mathbf{R}m' | \hat{W}_r(i\Omega_n) | \mathbf{R}m'' \rangle | \mathbf{R}m''' \rangle. \quad (26)$$

From Eq. (25) we see that, the frequency dependence in the polarization function leads to a frequency dependent, i.e. retarded, interaction which in general carries a real and an imaginary part. This effective DFT++ theory will thus not take a Hamiltonian form and the interaction

$\hat{W}_r(i\Omega_n)$  is, in general non-local. We analyze two example cases, SrVO<sub>3</sub> and graphene, to illustrate this point.

For SrVO<sub>3</sub>, cRPA yields in the static limit a value of  $U = 3.5$  eV [13], which is indeed close to the value of  $U \approx 4$  eV used in many DFT+DMFT calculations which include only the  $t_{2g}$  bands in the Bloch space and the correlated subspace (cf. Refs. [4–6]). With this value, DFT+DMFT yields quasi-particle weights and the location of the Hubbard bands in reasonable agreement with experiments [4]. The interpretation of the static limit of the cRPA interaction as the interaction to include in the DFT++ Hamiltonian thus seems to be a reasonable approximation and that is indeed what is often done in practice. It should, however, be noted that this approximation neglects several physical processes that affect electronic excitation spectra in real materials. Plasmons, manifesting for instance as poles/resonances in the screened interaction, i.e., dynamic plasmon effects associated with the uncorrelated states, will not enter the DFT++ model if the static limit of the RPA interaction is chosen. For Ni, it has been shown that these high-energy plasmons can affect the low-energy spectra by spectral weight transfer to higher-energy plasmon satellites and a concomitant reduction of the quasi particle weight [13]. It has been suggested that this spectral weight transfer can be accounted for through a renormalization of the hopping and hybridization parameters [15].

### 3.2 Non-local Coulomb interactions

So far, we considered only local interaction terms. One reasoning behind this is that in transition-metals like Fe or Ni, the uncorrelated  $sp$ -electron bands provide efficient screening such that the non-local terms are small [16]. This is not necessarily always the case, as can be seen from the the partially screened interaction  $\hat{W}_r$  of the  $p_z$ -electrons in graphene and graphite in Fig. 5. We see that graphene hosts both sizable on-site repulsion  $U_{00} = 9.3$  eV  $\approx 3.3t$  and nearest neighbor interaction  $U_{01} = 5.5$  eV  $\approx 2t$ , which exceed the nearest neighbor hopping  $t$  and are both on the order of the electronic band width  $D = 6t$  [14]. This coexistence of local- and non-local interaction terms is typical for effective models of low dimensional materials and has also been found for two-dimensional superstructures of ad-atoms on semiconductor surfaces [17]. We will therefore discuss the example of graphene a bit closer.

Quantum Monte Carlo simulations of the Hubbard model on the honeycomb lattice have indicated many-body instabilities from a Dirac material towards gapped phases for interaction strengths  $U \gtrsim 3.5t$ . At large interactions  $U > 4.5t$  the formation of an antiferromagnetic insulator appears well established, while there is controversy about intermediate interaction strengths  $3.5t < U < 4.5t$ , where Ref. [18] argues for the presence of a spin-liquid phase. More recent calculations question this formation of a spin-liquid but find an antiferromagnetic insulator for  $U > 3.9t$  [19]. Thus, taking the cRPA local Coulomb interaction  $U_{00} \approx 3.3t$  and neglecting all other terms would put graphene close to an instability towards an insulating phase driven by local correlations. This appears surprising, since graphene is indeed one of the best known electric conductors and electrons in graphene are generally assumed to be rather delocalized.

So, what is wrong, here? Obviously, the non-local interaction terms have been neglected or equivalently have been only included on a Hartree level.<sup>1</sup> From the illustration of Fig. 4, it is however clear that non-local terms could indeed weaken the effective local interactions. If we are interested in thermodynamic instabilities (e.g. transitions between a Dirac material and an antiferromagnetic insulator), the following variational approach provides a connection between models with strictly local and non-local interactions [9]:

The starting point is the extended Hubbard model

$$H = - \sum_{i,j,\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} + \frac{1}{2} \sum_{\substack{i \neq j \\ \sigma, \sigma'}} V_{ij} n_{i\sigma} n_{j\sigma'}, \quad (27)$$

where  $t_{ij}$  are the hopping matrix elements and  $U$  and  $V_{ij}$  are the local and nonlocal Coulomb matrix elements, respectively. The goal is to map the Hamiltonian (27) onto the effective model

$$H^* = - \sum_{i,j,\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U^* \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (28)$$

The effective on-site interaction  $U^*$  shall be chosen such that the canonical density operator  $\rho^* = 1/Z^* e^{-\beta H^*}$  of the auxiliary system, where  $Z^* = \text{Tr} \{ e^{-\beta H^*} \}$  is the partition function, approximates the exact density operator  $\rho$  derived from  $H$  as close as possible. This requirement leads to the Peierls-Feynman-Bogoliubov variational principle [20–22] for the functional

$$\tilde{\Phi}[\rho^*] = \Phi^* + \langle H - H^* \rangle^*, \quad (29)$$

where  $\Phi^* = -\frac{1}{\beta} \ln Z^*$  is the free energy of the auxiliary system.  $\langle \dots \rangle^*$  denotes thermodynamic expectation values with respect to the auxiliary system:  $\langle H - H^* \rangle^* = \text{Tr} \rho^* (H - H^*)$ . In the case of  $\rho^* = \rho$  the functional  $\tilde{\Phi}[\rho^*]$  becomes minimal and coincides with the free energy. The optimal  $U^*$  is thus obtained for minimal  $\tilde{\Phi}[\rho^*] = \tilde{\Phi}[U^*]$ :

$$\partial_{U^*} \tilde{\Phi}[U^*] = 0. \quad (30)$$

By evaluating Eq. (30) one finds

$$U^* = U + \frac{1}{2} \sum_{\substack{i \neq j \\ \sigma, \sigma'}} V_{ij} \frac{\partial_{U^*} \langle n_{i\sigma} n_{j\sigma'} \rangle^*}{\sum_l \partial_{U^*} \langle n_{l\uparrow} n_{l\downarrow} \rangle^*}. \quad (31)$$

This rule quite closely resembles the *Gedankenexperiment* depicted in Fig. 4: Increasing the on-site term  $U^*$  reduces the double occupancy  $\langle n_{i\uparrow} n_{i\downarrow} \rangle^*$  and pushes away electrons approaching an already occupied site  $i = 0$  to neighboring sites. In the case of purely local Coulomb interactions, there is a Coulomb energy gain of  $U^*$  upon suppressing the double occupancy (Fig. 4b). However, when there are nonlocal Coulomb interactions with the surrounding lattice sites  $j$ , the displaced electrons raise the energy of the system by terms proportional to  $V_{0j}$ . This

<sup>1</sup>Neglecting the non-local interactions or inclusion on a Hartree level are equivalent in the model of Eq. (27) if translation invariance is assumed due to cancellation of Hartree terms with the positive charge background stemming from the nuclei.

process is depicted in Fig. 4a for the simple case of two electrons on one site. In this case, it is obvious that the Coulomb energy gain due to the electron displacement in the full and the auxiliary model become energetically equivalent for  $U^* = U - V$ . In general, this energy gain depends both on the sites to which the charge density is displaced due to the local Coulomb interaction and on how strong the nonlocal Coulomb terms are.

For a translationally invariant system, the local part of the interaction  $U$  is reduced according to  $U^* = U - \bar{V}$ , where

$$\bar{V} = - \sum_{\substack{j \neq 0 \\ \sigma'}} V_{0j} \frac{\partial_{U^*} \langle n_{0\uparrow} n_{j\sigma'} \rangle^*}{\partial_{U^*} \langle n_{0\uparrow} n_{0\downarrow} \rangle^*}. \quad (32)$$

The conservation of the total electron number  $N$  leads to the sum rules  $\sum_{j\sigma} \langle n_{0\uparrow} n_{j\sigma} \rangle^* = \text{const.}$  and  $\partial_{U^*} \langle n_{0\uparrow} n_{0\downarrow} \rangle^* = - \sum_{j \neq 0, \sigma} \partial_{U^*} \langle n_{0\uparrow} n_{j\sigma} \rangle^*$ . Thus,  $\bar{V}$  is a weighted average of the nonlocal Coulomb interactions. Under the assumption that an increasing  $U^*$  displaces electrons only to next neighbors, we find  $\partial_{U^*} \langle n_{0\uparrow} n_{0\downarrow} \rangle^* = -N_n \partial_{U^*} \sum_{\sigma} \langle n_{0\uparrow} n_{1\sigma'} \rangle^*$ , where  $N_n$  is the coordination number. Equation (31) then yields

$$U^* = U - V_{01}, \quad (33)$$

which is exactly the situation depicted in Fig. 4.

It is reasonable that  $\bar{V}$  is positive (repulsive) in most situations that correspond to real materials. Then, the nonlocal Coulomb interaction reduces the effective on-site interaction and therefore stabilizes the Fermi sea against transitions e.g. to a Mott insulator. This is indeed what happens also in graphene where an evaluation of Eq. (32) using correlation functions  $\partial_{U^*} \langle n_{0\uparrow} n_{j\sigma'} \rangle^*$  obtained by means of lattice QMC calculations yields  $U^* \approx 1.6t$  [9]. The effective local interaction is thus reduced due to the non-local Coulomb terms by more than a factor of two and the Dirac electron phase in graphene is correspondingly stabilized against transitions into an antiferromagnetic insulating phase. The example of graphene thus shows that treatments of non-local interactions beyond the Hartree approximation can be very important to assess phase transitions in strongly correlated electron systems.

The approach discussed here is variational and comes with the simplicity that the auxiliary system (which we solve numerically) involves only local interactions and can thus be treated e.g. by standard DMFT. It is also possible to account for non-local interactions diagrammatically, as for instance in the GW+DMFT approach [23]. In GW+DMFT non-local interactions and related electronic correlation effects are included on an RPA level. Regarding the effect of non-local interactions on boundaries between metallic and Mott insulating phases in low dimensional correlated materials, GW+DMFT also predicts that non-local interactions can stabilize the metallic phase [17]. The pictures emerging from GW+DMFT and the above explained variational approach are thus consistent.

## 4 Double-counting and charge self-consistency

The central idea of the DFT++ Hamiltonian, Eq. (1), is to introduce interaction terms  $U_{m\dots m'}$  within the correlated subspace to account for dynamic correlation effects. However, the Kohn-Sham energies from DFT already include interaction effects through the Hartree and exchange-correlation terms. Without correction, some interaction contributions would thus be counted twice in DFT++. Thus, some double-counting correction  $H_{DC}$  has to be included. One often assumes a form like

$$H_{DC} = \mu_{DC} \sum_m d_m^\dagger d_m. \quad (34)$$

The major problem is that widely used exchange-correlation functionals such as LDA or GGA are non-linear, do not have a diagrammatic representation, and most-importantly do not allow one to judge which portion of exchange- and correlation entering the Kohn-Sham eigenvalues is associated with the interactions added in DFT++ within the correlated subspace. Double-counting problems are typical for electronic structure methods where semilocal approximate DFT functionals are augmented with additional interaction terms and also occur in approaches like LDA+U. There is no universal solution to this problem, and the following discussion of practical ways to deal with double-counting will be quite empirical.

Several schemes to fix the double-counting terms have been put forward. All of them are based on some assumption either on how exchange and correlation effects within the correlated subspace are included in a functional like LDA or on some quantity which is assumed to be correctly obtained already from the DFT and which should not change when adding correlations within DFT++.

Since the double-counting correction is intrinsically an impurity quantity and not a global quantity (like the chemical potential  $\mu$ ) it appears natural to use intrinsic quantities of the impurity like the impurity self-energy or the impurity Green function to fix it. One physically intuitive assumption is to require that the electronic charge computed from the local noninteracting Green function and the one computed from the interacting impurity Green function are identical [5]

$$\text{Tr} G_{mm'}^{imp}(\beta) \stackrel{!}{=} \text{Tr} G_{mm'}^{0,loc}(\beta). \quad (35)$$

Alternatively one can also use the Weiss field  $\mathcal{G}_{mm'}$  instead of the local noninteracting Green function in the above equation. Both versions of the method give very similar results and work very well in metallic systems [5], since in a metal the total particle number of the system  $N$  and of the impurity  $n_{imp}$  are both very sensitive to small variations in  $\mu$  and  $\mu_{dc}$ .

One possible ansatz using the impurity self-energy  $\Sigma_{mm'}^{imp}$  is to constraint the high energy tails in the real part of the self-energy to sum up to zero

$$\text{Re Tr} (\Sigma_{mm'}^{imp}(i\omega_N)) \stackrel{!}{=} 0. \quad (36)$$

Here,  $\omega_N$  is the highest Matsubara frequency included in the computation. Physically this amounts to the requirement that the shift in the centroid of the impurity orbitals contains no

static component, i.e., that static mean field components of the self-energy are correctly provided by LDA. This criterion is sometimes used in insulating materials [24]. In metals it is otherwise reasonable to assume that the exchange correlation potential yields a good approximation of the self-energy at the Fermi level and thus to require that  $\text{Re Tr}(\Sigma_{mm'}^{imp}(i2\pi/\beta)) \stackrel{!}{=} 0$ . The criteria (35) and (36) define the double-counting correction implicitly. There are also two widely used schemes on how to fix the double-counting explicitly from occupation numbers termed “around mean field” (AMF) [25] and “fully localized limit” (FLL) [26]. AMF bases on the idea that exchange and correlation effects are included in LDA but only in a spherically and thus orbitally averaged mean field manner. The resulting double-counting potential is

$$\mu_{dc}^{AMF} = \sum_{m'} U_{mm'} n^0 + \sum_{m', m' \neq m} (U_{mm'} - J_{mm'}) n^0, \quad (37)$$

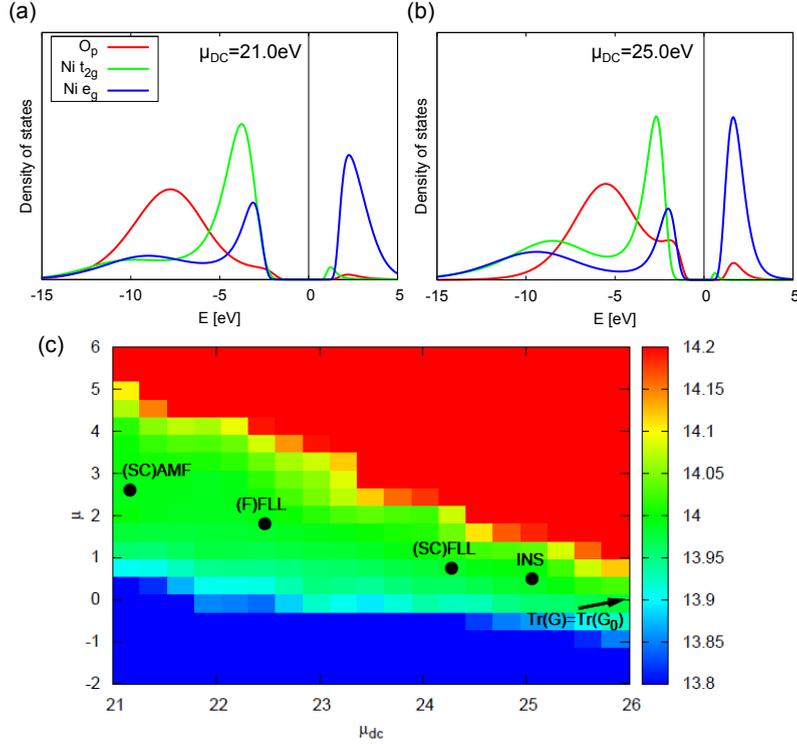
where  $n^0 = \frac{1}{2(2l+1)} \sum_{m,\sigma} n_{m\sigma}$  is the average occupancy. FLL is based on essentially the *opposite* idea. It assumes that total energies for fully localized atomic systems are rather well represented in functionals like LDA, i.e., that LDA (or LSDA) total energies are reliable if orbital occupation numbers  $n_{m\sigma}$  are either 1 or 0. While the LDA total energies are assumed to be appropriate in this fully localized case, for non-integer occupations LDA is known to be problematic since it does not correctly reproduce the derivative discontinuity of the exact density functional: it is known that the Kohn-Sham energies (which are derivatives of the total energy with respect to orbital occupations) in LDA do not jump discontinuously as they should for the exact density functional. One can combine the observation of good total energies at integer occupancy but lacking derivative discontinuity into the following prescription for the double-counting potential:

$$\mu_{dc}^{FLL} = U(N_{\text{imp}} - 1/2) + J(N_{\text{imp}}^{\sigma} - 1/2), \quad (38)$$

with  $N_{\text{imp}}^{\sigma} = \sum_m n_{m\sigma}$  being the total occupancy of the spin  $\sigma$ -component and  $N_{\text{imp}} = \sum_{\sigma} N_{\text{imp}}^{\sigma}$  being the total occupancy.

Indeed, different prescriptions for the double-counting can lead to different predictions regarding material properties like excitation spectra, as can be seen for the example system of NiO. The double-counting potential  $\mu_{dc}$  has profound impact on the density of states  $N_i(\omega) = -\frac{1}{\pi} \text{Im} G_i(\omega)$  shown in Fig. 6. In the LDA+DMFT study of Ref. [24], the double-counting potential has been treated as an adjustable parameter and has been varied between 21 eV and 26 eV.<sup>2</sup> The most prominent effects of the double-counting on the spectral properties are the shift of the oxygen  $p$  bands with respect to the nickel  $d$  bands, as well as the variation in gap size. Plainly speaking, the double-counting correction allows for a tuning of the spectral properties from a large gap Mott-Hubbard insulator to a metal. The regime of the charge transfer insulator, the expected physical state of NiO, lies somewhere in between. The calculated LDA+DMFT(QMC) spectral functions shown in Fig. 6 reveal basically the two different physical situations of a Mott-Hubbard, Fig. 6a, and a charge-transfer insulator, Fig. 6b, mentioned

<sup>2</sup>These values already contain the intrinsic shift due to the energy of the particle-hole symmetry in the Hirsch-Fye QMC method that amounts to 34 eV with our values of  $U$  and  $J$ .



**Fig. 6:** (a,b) Spectral functions of NiO obtained with LDA+DMFT (QMC) at inverse temperature  $\beta = 5\text{eV}^{-1}$  for different values of the double-counting  $\mu_{dc}$ . (c) Number of particles  $N$  per unit cell (color coded) as function of the chemical potential  $\mu$  and the double-counting potential  $\mu_{dc}$  as obtained with LDA+DMFT (QMC). Ni- $d$  and O- $p$  states are included in the calculation, which yields  $N = 14$  electrons per unit cell. The green plateau corresponds to a particle number very close to the desired value of  $N = 14$ . Values below are encoded in blue, values above in red. Additionally the results produced by different prescriptions to fix the double-counting are indicated. For the AMF and FLL functionals SC or F in parentheses indicates, that the occupancies from the DMFT or the formal occupancies have been used, respectively. INS refers to the double correction of  $\mu_{DC} = 25 \text{ eV}$ , where best agreement of ARPES spectra and LDA+DMFT simulations is achieved. From Ref. [24].

above, which are realized depending on the double-counting correction  $\mu_{DC}$ . The characteristic feature of a charge-transfer system, the strongly hybridized ligand  $p$  and transition-metal  $d$  character of the low-energy charge excitations, is only present in the spectrum in Fig. 6b. The spectrum in Fig. 6a is missing this feature almost completely and shows Mott-Hubbard behavior. This difference underscores the importance of the double-counting correction. A detailed comparison of calculated bandstructures with experiments shows that the choice of  $\mu_{dc} = 25 \text{ eV}$  yields best agreement of LDA+DMFT and the experimental data [24].

The pronounced impact of the double-counting correction can be further seen from the plot of the total number of electrons per unit cell on the chemical potential  $\mu$  and the double-counting correction  $\mu_{DC}$  in Fig. 6c.  $\mu_{DC}$  directly affects the  $pd$ -charge transfer energy and controls thereby the gap of the system which can be inferred from the  $N = 14$  plateau region. Where would the above explained prescriptions for fixing the double-counting correction lead to? Both the AMF and the self-energy criterion, Eq. (36), would lead to  $\mu_{DC} \approx 21 \text{ eV}$  and thus predict NiO to be basically a Mott-Hubbard but not a charge transfer system. I.e., these criteria are

not in agreement with experiments. FLL yields  $\mu_{DC} \approx 24$  eV, which comes closer to the experimental situation of a charge transfer insulator. Finally, criteria trying to fix the occupation of the correlated subspace to some value provided by LDA or the bath Green function, cf. Eq. (35), drives the system towards a metallic state at double-counting  $\mu_{DC} = 26.5$  eV indicated by the arrow pointing out of Fig. 6c. Here, FLL appears to describe the system best, which turns out to be often the case for insulators. On the other hand, metals are often well described by double-counting corrections based on traces of the Green function as in Eq. (35). So, the choice of an appropriate double-counting is rather empirical. It is sometimes beneficial to treat the double-counting correction as adjustable parameter and to study the dependence of LDA+DMFT prediction on the choice of the double-counting correction.

A promising way to circumvent double-counting issues are fully diagrammatic approaches like GW+DMFT [23]. These are currently being under development and first GW+DMFT studies of example materials like SrVO<sub>3</sub> [27–29] have been reported. GW+DMFT comes, however, at the expense of considerably higher computational demands than DFT+DMFT.

The DFT++ Hamiltonian in the form of Eq. (1) implicitly includes interactions between electrons in the correlated subspace and the rest of the system through the Hartree as well as the exchange correlation potential from DFT. As soon as the many-body part of DFT++ redistributes electrons between correlated and uncorrelated orbitals or also between different sites there will be associated Hartree (as well as possible exchange or correlation) energies and the DFT++ Hamiltonian should be correspondingly updated. In general, it is obviously problematic to obtain the update of the DFT++ Hamiltonian simply from a double-counting correction applied to the correlated subspace only. This can be better achieved by including self-consistency over the charge-density in the DFT++ approach.

To this end, one calculates the electron density of the DFT++ system,

$$n(r) = \frac{1}{\beta} \sum_{\mathbf{k}, \alpha, \alpha', n} \langle r | B_{\mathbf{k}\alpha} \rangle G_{\alpha\alpha'}(\mathbf{k}, i\omega_n) \langle B_{\mathbf{k}\alpha'} | r \rangle, \quad (39)$$

which includes corrections due to dynamic self-energy effects within the correlated subspace. With this density  $n(r)$  one can recalculate the DFT potential and solve the resulting Kohn-Sham Hamiltonian, which then reenters the non-interacting part of the DFT++ Hamiltonian, Eq. (1). In this way, a charge self-consistent DFT++ scheme is obtained, see Fig. 2, which includes interactions between electrons of the correlated subspace and the rest in a fully self-consistent static mean-field manner. Several implementations of charge self-consistent of DFT+DMFT have been reported, e.g. Refs. [30–32], based on projector formalisms similar to Sec. 2.

It is intuitively clear that the Hartree terms occurring within DFT++ charge self-consistency counteract large charge redistributions. In other words, ambiguities stemming for instance from the unknown double-counting potential can be expected to be less severe in charge self-consistent DFT++ calculations as compared to one-shot calculations. This has been explicitly demonstrated, e.g., for the Matsubara self-energies in the iron pnictide superconductor LaFeAsO, where the discrepancy between FLL and AMF approaches is significantly reduced in the fully charge self-consistent scheme [30].

## 5 Conclusions

The combination of first-principles and model Hamiltonian approaches termed DFT++ presents a promising route towards realistic and material specific descriptions of strongly correlated electron systems. The number of adjustable parameters normally present in models of strongly correlated materials like transition-metal compounds or impurities on surfaces can be indeed largely reduced by deriving them from *ab-initio* calculations. Thereby, realistic studies of ever more complex correlated electron systems are coming into reach. At the same time, the model Hamiltonian level involved in DFT++ offers the chance to study how material properties depend, e.g., on the strength of Coulomb interactions by deliberately treating them as adjustable parameters. With the projector formalism, DFT++ can in principle be applied to arbitrarily complex systems. Many developments in this direction are being pursued throughout the last few years. Naturally, this lecture covered only a very limited amount of these activities, as readers familiar with the subject of DFT++ will have noticed and as becomes clear from a deeper look into the literature referenced here. As has already become clear in the discussions of double-counting issues or frequency dependent and non-local interactions a lot of method development at the interface of first-principles and model based approaches remains still to be done. This includes both the further development of, e.g., DFT+DMFT to a point where it can be as widely and routinely applied as LDA+U, developments in the combination of diagrammatic *ab-initio* approaches with model based approaches such as GW+DMFT, descriptions of non-local correlation effects or also the coupling of correlated electrons and bosonic modes such as plasmons, phonons or magnons.

The DFT++ model Hamiltonians discussed in this lecture have put an emphasis on correlation effects due to local Coulomb interactions, which are indeed essential for various magnetic phenomena or Mott metal insulator transitions. Other many-body phenomena can rely on different kinds of interactions. Wigner crystallization of electrons, exciton binding, or plasmon modes are often controlled by non-local Coulomb interaction terms, and appropriate models for such phenomena will naturally have to include different interaction terms. In other words: any DFT++ modeling requires an idea on which interactions form the basis of the many-body problem to be described. It can of course be very challenging to identify the essential interactions responsible for an unknown phenomenon or to determine whether some observed effect is a many-body phenomenon or not.

## Acknowledgments

Support by Deutsche Forschungsgemeinschaft through FOR 1346 is gratefully acknowledged.

## References

- [1] A.I. Lichtenstein and M.I. Katsnelson, *Phys. Rev. B* **57**, 6884 (1998)
- [2] M. Imada, A. Fujimori, and Y. Tokura, *Rev. Mod. Phys.* **70**, 1039 (1998)
- [3] V.I. Anisimov, D.E. Kondakov, A.V. Kozhevnikov, I.A. Nekrasov, Z.V. Pchelkina, J.W. Allen, S.-K. Mo, H.-D. Kim, P. Metcalf, S. Suga, A. Sekiyama, G. Keller, I. Leonov, X. Ren, and D. Vollhardt, *Phys. Rev. B* **71**, 125119 (2005)
- [4] F. Lechermann, A. Georges, A. Poteryaev, S. Biermann, M. Posternak, A. Yamasaki, and O.K. Andersen, *Phys. Rev. B* **74**, 125120 (2006)
- [5] B. Amadon, F. Lechermann, A. Georges, F. Jollet, T.O. Wehling, and A.I. Lichtenstein, *Phys. Rev. B* **77**, 205112 (2008)
- [6] M. Karolak, T.O. Wehling, F. Lechermann, and A.I. Lichtenstein, *J. Phys.: Condens. Matter* **23**, 085601 (2011)
- [7] T.O. Wehling: *Impurities and Inhomogeneities in Nanoelectronic Systems* (PhD thesis, University of Hamburg, 2010)
- [8] M. Karolak: *Electronic Correlation Effects in Transition Metal Systems: From Bulk Crystals to Nanostructures* (PhD thesis, University of Hamburg, 2013)
- [9] M. Schüler, M. Rösner, T.O. Wehling, A.I. Lichtenstein, and M.I. Katsnelson, *Phys. Rev. Lett.* **111**, 036601 (2013)
- [10] O. Gunnarsson, O.K. Andersen, O. Jepsen, and J. Zaanen, *Phys. Rev. B* **39**, 1708 (1989)
- [11] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A.I. Lichtenstein, *Phys. Rev. B* **70**, 195104 (2004)
- [12] P. Hansmann, L. Vaugier, H. Jiang, and S. Biermann, *J. Phys.: Condens. Matter* **25**, 094005 (2013)
- [13] F. Aryasetiawan, K. Karlsson, O. Jepsen, and U. Schönberger, *Phys. Rev. B* **74**, 125106 (2006)
- [14] T.O. Wehling, E. Şaşıoğlu, C. Friedrich, A.I. Lichtenstein, M.I. Katsnelson, and S. Blügel, *Phys. Rev. Lett.* **106**, 236805 (2011)
- [15] M. Casula, P. Werner, L. Vaugier, F. Aryasetiawan, T. Miyake, A.J. Millis, and S. Biermann, *Phys. Rev. Lett.* **109**, 126408 (2012)
- [16] J. Hubbard, *Proc. Roy. Soc. (London) A* **276**, 238 (1963)

- [17] P. Hansmann, T. Ayrat, L. Vaugier, P. Werner, and S. Biermann, Phys. Rev. Lett. **110**, 166401 (2013)
- [18] Z.Y. Meng, T.C. Lang, S. Wessel, F.F. Assaad, and A. Muramatsu, Nature **464**, 847 (2010)
- [19] S. Sorella, Y. Otsuka, and S. Yunoki, Sci. Rep. **2**, 992 (2012)
- [20] R.E. Peierls, Phys. Rev. **54**, 918 (1938)
- [21] N.N. Bogoliubov., Dokl. Akad. Nauk SSSR **119**, 244 (1958)
- [22] R.P. Feynman: *Statistical Mechanics* (Benjamin, Reading Mass., 1972)
- [23] S. Biermann, F. Aryasetiawan, and A. Georges, Phys. Rev. Lett. **90**, 086402 (2003)
- [24] M. Karolak, G. Ulm, T. Wehling, V. Mazurenko, A. Poteryaev, and A.I. Lichtenstein, Journal of Electron Spectroscopy and Related Phenomena **181**, 11 (2010)
- [25] V.I. Anisimov, J. Zaanen, and O.K. Andersen, Phys. Rev. B **44**, 943 (1991)
- [26] M.T. Czyżyk and G.A. Sawatzky, Phys. Rev. B **49**, 14211 (1994)
- [27] J.M. Tomczak, M. Casula, T. Miyake, F. Aryasetiawan, and S. Biermann, Europhys. Lett. **100**, 67001 (2012)
- [28] R. Sakuma, P. Werner, and F. Aryasetiawan, Phys. Rev. B **88**, 235110 (2013)
- [29] C. Taranto, M. Kaltak, N. Parragh, G. Sangiovanni, G. Kresse, A. Toschi, and K. Held, Phys. Rev. B **88**, 165119 (2013)
- [30] M. Aichhorn, L. Pourovskii, and A. Georges, Phys. Rev. B **84**, 054529 (2011)
- [31] D. Grieger, C. Piefke, O.E. Peil, and F. Lechermann, Phys. Rev. B **86**, 155121 (2012)
- [32] B. Amadon, J. Physics: Condens. Matter **24**, 075604 (2012)