4 Hubbard Dimer in GW and Beyond

Pina Romaniello
Laboratoire de Physique Théorique
Université de Toulouse
18 Route de Narbonne
31062 Toulouse Cedex 4, France

Contents

E. Pavarini and E. Koch (eds.)
Simulating Correlations with Computers
Modeling and Simulation Vol. 11
Forschungszentrum Jülich, 2021, ISBN 978-3-95806-529-1
http://www.cond-mat.de/events/correl21

1 Introduction

Exactly solvable models are of paramount importance to test the quality of approximations used in various theoretical methods. Having at hand the exact solution of a given problem can allow one to observe possible problems of a given approximation, to get insights into the origin of them, and, ideally, to suggest possible solutions. Several models can be used for this purpose. Here we will concentrate on the Hubbard model. This model is widely used to deal with the physics of strongly correlated materials. Since the model can be solved exactly for small cluster sizes, it is very useful for theoretical investigations. Of course care must be taken when extrapolating conclusions to realistic systems. In this chapter we will use it to scrutinize approximations used in many-body perturbation theory based on Green functions. We will concentrate on the one-body Green function (1-GF) at zero temperature and at equilibrium, which corresponds to the electronic Hamiltonian $\hat{H} = \hat{T} + \hat{V} + \hat{W}$, with $\hat{T}$, $\hat{V}$, and $\hat{W}$ the kinetic-energy operator, the operator corresponding to an external time-independent potential (for example the external potential of the nuclei), and the electron-electron interaction operator, respectively. The 1-GF is a powerful quantity since it contains a wealth of information about a physical system, such as the expectation value of any single-particle operator over the ground state, the ground-state total energy and the spectral function (which is related to photoemission spectra). It can be obtained by solving a Dyson equation of the form $G = G_0 + G_0 \Sigma G$, where G_0 is the non-interacting one-body Green function and the self-energy Σ is an effective potential which contains all the many-body effects of the system under study. In practice Σ needs to be approximated and a well-known approximation is the so called GW approximation (GWA) in which the self-energy reads $\Sigma = v_H + iGW$, where v_H is the classical Hartree potential, and W the dynamically screened Coulomb potential. The GWA works well for many systems, but it also suffers from some shortcomings, such as self-screening (i.e. the fact that an electron screens itself, which is unphysical) and wrong atomic limit (i.e. the fact that the GWA does not capture the so-called strong correlation). These errors of the GWA and possible solutions will be illustrated using the Hubbard dimer.

Notation

We will use combined space-spin-time coordinates of the type $(1) \equiv (\mathbf{x}_1, t_1) \equiv (\mathbf{r}_1, s_1, t_1)$ and $(1^+) = (\mathbf{x}_1, t_1^+)$ with $t_1^+ = t_1 + \delta$ ($\delta \to 0^+$). Moreover, to keep a light notation, integration over indices not present on the left-hand side of an equation is implied.

2 Theoretical background: the GW approximation & beyond

Much of the success of many-body perturbation theory based on Green functions [1] is due to the simplicity to access photoemission spectra through the one-body Green function. As illustrated in Fig. 1 photoemission is about absorption and emission of single electrons and it is a unique source of information about electronic structure and excitations in materials. We will focus precisely on this observable in the rest of the chapter.

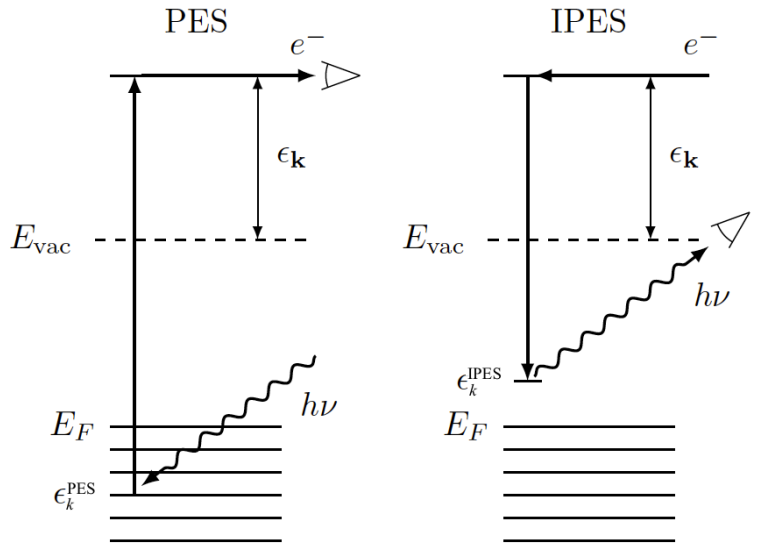

Fig. 1: *Schematic picture of direct (PES) and inverse photoemission spectroscopies (IPSE). In PES (IPES) one samples the occupied (unoccupied) energy levels* $\varepsilon_k^{PES} = E_0^N - E_k^{N-1}$ *(*$\varepsilon_k^{IPES} = E_k^{N+1} - E_0^N$*) of the system.*

Let us consider the following many-body Hamiltonian in second quantization

$$\hat{H} = \int d\mathbf{x}\, \hat{\psi}^\dagger(\mathbf{x}) h(\mathbf{r}) \hat{\psi}(\mathbf{x}) + \frac{1}{2} \iint d\mathbf{x} d\mathbf{x}'\, \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}^\dagger(\mathbf{x}') \hat{\psi}(\mathbf{x}') \hat{\psi}(\mathbf{x}) v_c(\mathbf{x}, \mathbf{x}'), \tag{1}$$

where $\hat{\psi}$ and $\hat{\psi}^\dagger$ are field operators, v_c is the Coulomb potential, and $h(x) = -\nabla_{\mathbf{r}}^2/2 + v_{ext}(\mathbf{r})$ is the one-body part of the Hamiltonian, which comprises the kinetic term and a local external potential. At zero temperature the time-ordered equilibrium 1-GF can be written in terms of its Lehmann representation in frequency space as

$$G(\mathbf{x}, \mathbf{x}'; \omega) = \lim_{\eta \to 0^+} \left(\sum_m \frac{B_m^A(\mathbf{x}, \mathbf{x}')}{\omega - (E_m^{N+1} - E_0^N) + \mathrm{i}\eta} + \sum_n \frac{B_n^R(\mathbf{x}, \mathbf{x}')}{\omega - (E_0^N - E_n^{N-1}) - \mathrm{i}\eta} \right), \tag{2}$$

where $B_m^A(\mathbf{x}, \mathbf{x}') = \langle \Psi_0^N | \hat{\psi}(\mathbf{x}) | \Psi_m^{N+1} \rangle \langle \Psi_m^{N+1} | \hat{\psi}^\dagger(\mathbf{x}') | \Psi_0^N \rangle$ and $B_n^R(\mathbf{x}, \mathbf{x}') = \langle \Psi_0^N | \hat{\psi}^\dagger(\mathbf{x}') \big| \Psi_n^{N-1} \rangle$ $\langle \Psi_n^{N-1} | \hat{\psi}(\mathbf{x}) \big| \Psi_0^N \rangle$, $|\Psi_0^N\rangle$ and E_0 are the many-body ground-state wavefunction and energy of the N-electron system, respectively, and $|\Psi_m^{N\pm1}\rangle$ and $E_m^{N\pm1}$ are the many-body wavefunctions and corresponding energies of the $(N \pm 1)$-electron system. This definition of the 1-GF makes clear its connection to the charged excitations of the system, which are measured in photoemission experiments. It is indeed convenient to define the spectral function in terms of the imaginary part of the 1-GF according to

$$\begin{aligned} A(\mathbf{x}, \mathbf{x}'; \omega) &= \frac{1}{\pi} \mathrm{sign}(\mu - \omega)\, \mathrm{Im}\, G(\mathbf{x}, \mathbf{x}'; \omega) \\ &= \sum_m B_m^A(\mathbf{x}, \mathbf{x}') \delta\big(\omega - (E_m^{N+1} - E_0^N)\big) + \sum_n B_m^R(\mathbf{x}, \mathbf{x}') \delta\big(\omega - (E_0^N - E_n^{N-1})\big), \end{aligned} \tag{3}$$

where μ is the chemical potential. Of course the definition (2) is not very useful for determining the 1-GF, since it relies on the knowledge of many-body wavefunctions, which is precisely what

one wants to avoid when using theories based on simpler physical quantities, such as the 1-GF. Therefore the development of approximate methods to calculate the 1-GF has been an active research topic in many-body physics since the sixties, and many routes have been explored in order to find increasingly accurate GFs. A very popular class of methods is based on the solution of an integral equation for the 1-GF, namely,

$$G(1,2) = G_0(1,2) + G_0(1,3)\Sigma(3,4)G(4,2)\,, \tag{4}$$

where G_0 is the noninteracting 1-GF and Σ the self-energy. A good starting point to make approximations to the self-energy is to use Hedin's equations, which are often solved within the so-called GW approximation (GWA) to the self-energy [2], where G is the one-particle Green function and W the dynamically screened Coulomb potential. Within this approximation one neglects so-called vertex corrections, which take into account the fermionic nature of the system, and one treats the system and its response classically (besides exchange). This becomes clear by starting from the following exact expression for the self-energy [3, 4]

$$\Sigma(11') = v_H(1)\delta(11') + \Sigma_x(11') + iv_c(1^+2)G(13)\Xi(35;1'4)L(42;52^+), \tag{5}$$

with $v_H(1) = -i\int d2\, v_c(12)G(22^+)$ the Hartree potential, $\Sigma_x(11') = iv_c(1^+1')G(11')$ the exchange contribution to the self-energy, $\Xi(35;1'4) = \frac{\delta\Sigma(31')}{\delta G(45)}$ the effective interaction, and $L(42;52) = \frac{\delta G(45)}{\delta U_{ext}(2)}\Big|_{U_{ext}=0}$ the time-ordered "response" of the system to an external perturbation U_{ext}. This way to write the self-energy directly displays the physics behind it, i.e., the description of a particle interacting with the system: the particle can scatter against the density of the system (Hartree term), it can exchange with another particle of the system (exchange term), it can do something to the system (last term), i.e., it can have an effective interaction with the system (Ξ), the system responds (L), and the particle feels this response through the Coulomb interaction (v_c). There are two crucial ingredients in Eq. (5), namely the response of the system L and the effective interaction Ξ. Combining approximations to these two quantities, various approximations to the self-energy can be created. In situations where the screening is important one can treat L accurately and use a rough approximation to the induced potential Ξ, whereas in situations where the quantum nature of the interaction is important[1] one would concentrate on Ξ, although L and Ξ are of course in principle linked through the Bethe-Salpeter equation [5] and one might wish to keep them approximately consistent.

2.1 GW

Neglecting the variation of $\Sigma_{xc} = \Sigma - v_H$ in Ξ, i.e., keeping only the classical interaction v_c, one obtains $\Sigma_{xc}(11') = \Sigma_x(11') + iv_c(12)G(11')v_c(1'4)\chi(42)$, with $\chi(42) = -iL(42;42)$ the time-ordered response function. Hence one gets a screening contribution with respect to Σ_x: this is the GW form, with $W = v_c + v_c\chi v_c$, which can also be written as $W = v_c + v_cPW$

[1] The atomic limit of the Hubbard molecule is an example where the correlation part of the interaction is crucial; see later.

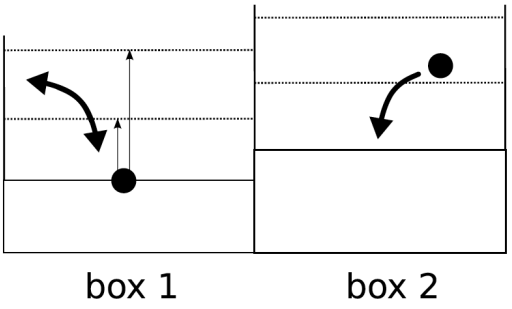

Fig. 2: *Model system for removal and addition energies from Ref. [10]: first an electron is added to/removed from box 1 occupied by zero/one electron; in a second step an electron is added to box 2 in presence of the electron in box 1 (see text).*

with the polarizability $P = \chi(1 + v_c\chi)^{-1}$. At this stage it has not been specified yet how to calculate the screening: different approximations to the screening will give the various GW flavors, e.g. GW^{RPA}, in which $P = -iGG$, and beyond.[2] If one keeps an approximate Σ_{xc} in Ξ, one goes beyond GW and includes so-called vertex corrections. The GW approximation works well for systems where the screening is important; instead in systems which show atomic-like physics, such as transition-metal oxides, the GWA shows difficulties or even failures [3, 7–15]. More accurate approximations are then needed. One way to go beyond the standard methods is to try to correct some basic shortcomings that plague them. Here we will focus on the following two errors GW suffers from:

- **Self-screening error**
  This error [16, 10, 17] can be illustrated by considering an empty box with a potential V_0, as depicted in Fig. 2. The problem of adding a particle to the empty box can be solved exactly by the independent-particle Schrödinger equation

$$\left(-\frac{\nabla^2}{2} + V_0(x_1)\right)\phi_1(x_1) = \varepsilon_1\phi_1(x_1). \tag{6}$$

  The total energy difference between the system with $N = 0$ and $N+1 = 1$ particles is given by the eigenvalue ε_1. This situation is described exactly by the GWA (the same holds for any other commonly used approach, like DFT or Hartree-Fock, since the density of the system is zero). Problems arise when looking at the reverse problem, namely to extract the particle from the box, since now there is a non-vanishing charge density in the box. Again, Eq. (6) correctly describes the total energy difference between $N = 1$ and $N-1 = 0$ particles. Also Hartree-Fock (HF) is exact, as one can easily verify, since there is the exact cancellation between the Hartree and the Fock terms because the density is built with just the state ϕ_1 one is looking at. It then becomes clear that GW does not give the exact result, since the exchange term that in HF cancelled the Hartree self-interaction is now screened, and the cancellation is no longer perfect. If we

[2] In this case the nature of the screening is always test-charge-test-charge [6].

had $W = v_c$ (no screening) we would be back to the exact HF result. However, the screening is in general non-zero, since it depends on the density of the system, which is non-zero for a particle in a box. Therefore, it becomes clear that GW *suffers from a self-screening problem*: the particle that is extracted from the system screens itself, which makes the approach non-symmetric with respect to the case where the particle is added to the empty box. In other words, when the quasiparticle eigenvalue is used to represent a total energy difference, the energy of the system has changed after one has first added and then extracted one electron.

- **Atomic limit**:
  Another very important diagnostic tool is the study of the atomic limit. This limit corresponds to pulling apart the atoms of a system, so that the overlap between wavefunctions on different atoms is negligible. In particular H_2^+ and H_2 dissociations have been extensively studied in order to trace errors in the approximations used in DFT and MBPT. The study of the H_2 dissociation within GW shows that this approximation is not size-consistent, i.e., the total energy in the atomic limit is not equal to the sum of the total energies of the two isolated atoms [8, 9]. Also the description of the spectrum can be expected to be problematic, and this can be understood as follows: the GW approximation takes into account only the Hartree, exchange, and induced Hartree potentials to describe an interacting system and its response to an additional electron or hole, i.e., besides exchange it gives a classical description of the system and of its response. This means for the example H_2^+ that an additional electron sees half an electron on each atom at any distance between the two atoms; this results in only one type of addition energy, as we will later show using the Hubbard dimer at 1/4 filling. In reality, instead, one has the same probability of finding the whole electron on one or the other atom, resulting in two types of addition energies depending if the additional electron is added to an empty or to an occupied atom.

2.2 Vertex corrections

We can now formulate vertex corrections with the aim to correct the problems of self-screening and wrong atomic limit that GW suffers from. If one keeps an approximate Σ_{xc} in Ξ of Eq. (5) one goes beyond GW and includes vertex corrections. The question is which approximate Σ_{xc} to consider.

2.2.1 Vertex corrections from DFT

By approximating Σ_{xc} by the exchange-correlation potential of DFT, v_{xc}, one gets the expression $\Sigma_{xc}(1,1') = \Sigma_x(1,1') + iv_c(1,2)G(1,1')\left[v_c(1',4) + f_{xc}(1',4)\right]\chi(4,2)$, where $f_{xc} = \frac{\delta v_{xc}}{\delta \rho}$; this leads to $\Sigma_{xc}{=}iG\tilde{W}$, with a modified screened Coulomb interaction $\tilde{W}{=}v_c{+}(v_c{+}f_{xc})\chi v_c$ [6]. This is a test-charge test-electron screened Coulomb interaction (TC-TE) instead of a test-charge test-charge W, and it expresses the fact that an additional particle in the system can-

not be described as a classical charge. This approximate xc self-energy can be rewritten as $\Sigma_{xc} = iGW\Gamma$ with $\Gamma = 1 + f_{xc}P$ a two-point vertex and P the irreducible polarizability.[3] The f_{xc} that appears makes the one-electron case exact, since for this case $f_{xc} = -v_c$ from which $\Sigma_{xc} = iGv_c$, which exactly cancels the Hartree term for one electron. In the simple case of one electron, hence, this simple vertex correction removes the self-screening and yields the exact result for the highest occupied state. It seems reasonable to generalize these intuitive findings to the case of more bands. In fact, we know that the KS potential yields often a quite good description for all valence bands, not only the highest occupied one. In these cases one should expect that the test-charge test-electron approximation for the self-energy derived above is the method of choice to describe valence bands. Of course, the exact KS potential and kernel are not known, but one might use LDA as a first approximation. For conduction states that are spatially distinct from valence states or with different spin, which are modelled by a second electron added to a second separate box in Fig. 2, one can argue that one should rather use GW [10]. The effectiveness of this approach has yet to be investigated and it is an interesting outlook. Preliminary results on silicon indicate that this kind of vertex corrections appear to affect just the quasiparticle energies [18].

2.2.2 Vertex corrections from the T-matrix approximation

As we will show using the Hubbard dimer, the vertex corrections which cure the self-screening problem are not sufficient to improve the description of the atomic limit, precisely because the two problems have a different nature; therefore, a more complex vertex is needed to fix both. For the spectral function in the atomic limit what matters is that the electron in the system and the electron added to the system "see" each other. In other words it is the quantum nature of the electron interaction that is important. Starting from the exact expression for the self-energy, Eq. (5), one could use the rough approximation $L(4,2;5,2) \simeq L_0(4,2;5,2)$ for the response of the system but concentrate on a clever approximation for Ξ. To this end one can introduce an effective 4-point interaction T such that, similar to GW, $\Sigma(1,1') = iG(4',2')\,T(1,2';1',4')$. T is linked to Ξ through the functional derivative of the self-energy as $\Xi(3,5;1',4) = iT(3,5;1',4) + iG(4,2)\frac{\delta T(3,2;1',4)}{\delta G(4,5)}$. Neglecting $\frac{\delta T(3,2;1',4)}{\delta G(4,5)}$, in analogy with what one usually does in the framework of Bethe-Salpeter calculations based on GW [19–21], Eq. (5) becomes an integral equation for GT

$$\begin{aligned}\Sigma(1,1') &= iG(4',2')\,T(1,2';1',4') \\ &= v_H(1)\,\delta(11') + \Sigma_x(11') - v_c(12)\,G(1,3)\,G(42)\,G(25)\,T(3,5;1',4). \end{aligned} \tag{7}$$

Since $G(4',2')\,T(1,2';1',4')$ cannot be inverted to find T, several choices of T make the correct Σ. A possible solution is

$$\begin{aligned}T^{pp}(1,2;1',4) = &- v_c(1,2)\,\delta(1,1')\,\delta(4,2) + v_c(1,2)\,\delta(2,1')\,\delta(4,1) \\ &+ iv_c(1,2\,G(1,3)\,G(2,5)\,T^{pp}(3,5;1',4), \end{aligned} \tag{8}$$

[3] This can be obtained by considering that $W = v_c + v_c\chi v_c$ and $\chi v_c = PW$.

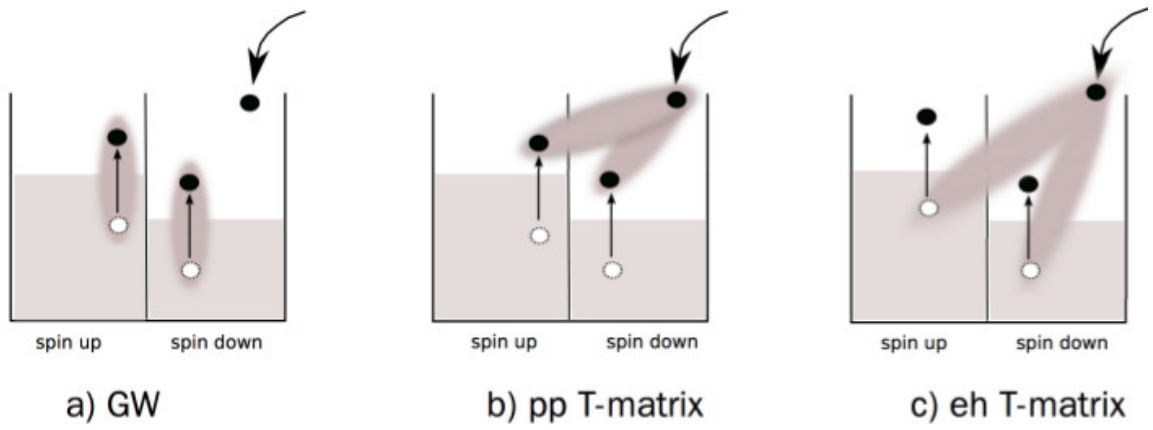

Fig. 3: *Schematic representation of the physical contents of GW (a), pp T-matrix (b), and eh T-matrix for particles with collinear spins from Ref. [3].*

which defines the particle-particle (because the kernel $G(1,3)\,G(2,5)$ in the correlation part describe two propagating electrons or holes) T-matrix [22, 23]. One can, moreover, decompose T^{pp} into a direct term (Hartree-like contribution) $T_1^{pp} = -v_c + iv_c\,G\,G\,T_1^{pp}$ and an exchange term (exchange-like contribution) $T_2^{pp} = v_c + iv_c\,G\,G\,T_2^{pp}$. Another possible solution of Eq. (7) is the electron-hole T-matrix, defined as

$$\begin{aligned} T^{eh}(1,5;1',2) = &- v_c(1',2)\,\delta(1,1')\,\delta(5,2) + v_c(1',2)\,\delta(1,2)\,\delta(1',5) \\ &+ iv_c(1,2)\,G(1,3)\,G(4,2)\,T^{eh}(3,5;1',4), \end{aligned} \tag{9}$$

where now the kernel $G(1,3)\,G(4,2)$ in the correlation part describes an electron and a hole. The structure of the T-matrix self-energy is hence very close to the structure of GW: one has a "particle-particle-screened" interaction T (or an "electron-hole-screened" interaction T in case of the electron-hole T-matrix, but where the electron-hole excitations involved are different by those involved in GW (see Fig. 3)). Based on (5), we can now directly compare the different approximations. Both GW and T-matrix allow one to describe physical processes involving three particles: the particle which is added to the system and the electron-hole pair that it creates. Ideally, one would propagate the three particles together, which can be numerically heavy. Therefore one chooses to propagate a pair and to treat the third particle in a kind of mean-field of the other two. This is illustrated in Fig. (3): in GW one propagates together the electron-hole pair created by the additional particle, whereas in the T-matrix one propagates together the additional electron (additional hole) and the excited electron (hole left behind) (pp T-matrix) or the additional electron (additional hole) and the hole left behind in the electron excitation (excited electron) (eh T-matrix).

As shown in Fig. 3 the additional particle with a given spin couples with a second, excited, particle (electron or hole) that can have spin up or down. In the GWA, electron and hole have necessarily the same spin. The T-matrix approximation contains instead also spin-flips.

In general none of the three possibilities (GWA, particle-particle or electron-hole T-matrix) will give an exhaustive description. This reflects the dilemma of how to decide which two-particle

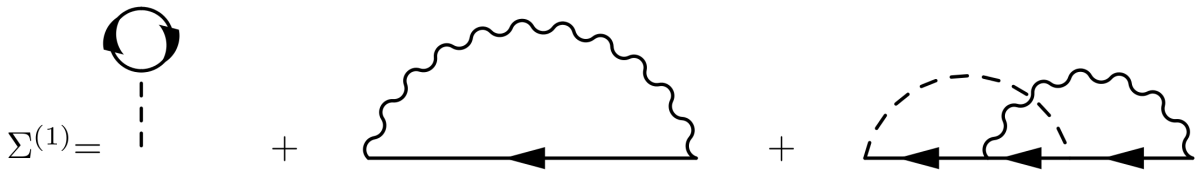

Fig. 4: *Diagrams corresponding to the self-energy obtained with the first iteration of the screened T-matrix. The diagrams, from left to right, represent the Hartree, GW, and second-order screened exchange (SOSEX) terms, respectively.*

correlation to privilege in the description of a (at least) three-particle problem. It suggests to work with combinations.

The success of the GWA is due to the screening of the Coulomb interaction with respect to Hartree-Fock. In the T-matrix one uses a rough approximation for the screening, i.e. $L \approx L_0$. One can go beyond this approximation and include the Hartree potential or even a local part of the xc self-energy, $\Sigma_{xc}^{loc}(1)\,\delta(1,1)$, in the variation $\delta G^{-1}/\delta U_{ext}$ contained in $L = \delta G/\delta U_{ext} = -G\,(\delta G^{-1}/\delta U_{ext})\,G$. This yields $L(4,2;5,2) \simeq G(4,7)\,\varepsilon^{-1}(7,2)\,G(7,5)$ where ε^{-1} is a test-charge test-charge screening function $\varepsilon^{-1} = 1 + v_c\chi$ if only the Hartree part is included, otherwise at least a partially test-charge test-electron screening $\varepsilon^{-1} = 1 + \left(v_c + \frac{\delta\Sigma_{xc}^{loc}}{\delta\rho}\right)\chi$ [6]. This makes the expression much more, though not fully, consistent: now ε^{-1} contains a large part of the derivative of the self-energy, that is also considered in the effective interaction, and L contains the screening of the formerly independent propagators GG, that is itself based on the two-particle correlation function. One arrives then at a screened version of the T-matrix (referred to as T_s in the following), where the bare Coulomb potential in the correlation part of the (unscreened) T-matrix is replaced by the screened one. The pp screened T-matrix reads

$$\begin{aligned} T_s(1,2;1',4) = &- v_c(1,2)\,\delta(1,1')\,\delta(4,2) + v_c(1,2)\,\delta(2,1')\,\delta(4,1) \\ &+ iW(1,2)\,G(1,3)\,G(2,5)\,T_s(3,5;1',4). \end{aligned} \tag{10}$$

One can write a similar equation also for the eh screened T-matrix. The first iteration of the pp and eh screened T-matrix equations produces the same self-energy, which contains the Hartree and GW contributions, and a term corresponding to the second-order screened exchange (SOSEX) (see Fig. 4). We hence can conclude that GW is contained in this screened T-matrix approach, which moreover contains higher-order terms, which can be identified as vertex corrections. For short-range interactions, where screening is not important, the screened T-matrix reduces to the T-matrix, which is suitable for treating short-range correlation. For long-range interaction, where, instead, screening is important, we find that the screened T-matrix behaves more like GW (in its first iteration, indeed, it gives GW and SOSEX, which is actually already used to improve GW), which is capable of taking into account long-range correlation. Therefore the screened T-matrix should be able to capture the physics of systems with effective short-range interactions as well as of systems with effective long-range interactions.

3 The Hubbard dimer

The Hubbard Hamiltonian can be obtained from the many-body Hamiltonian in (1) in which the field operators are expressed in the atomic basis (we will consider only one basis function per atom). Retaining only the diagonal elements of the Coulomb interaction (indicated as U) and indicating as $-t$ and ε_0 the off-diagonal and diagonal elements, respectively, of the kinetic energy operator, one arrives at

$$\hat{H} = -t \sum_{\langle i,j \rangle} \sum_{\sigma} \hat{c}^{\dagger}_{i\sigma} \hat{c}_{j\sigma} + U \sum_{i} \hat{c}^{\dagger}_{i\uparrow} \hat{c}_{i\uparrow} \hat{c}^{\dagger}_{i\downarrow} \hat{c}_{i\downarrow} + \varepsilon_0 \sum_{i} \sum_{\sigma} \hat{c}^{\dagger}_{i\sigma} \hat{c}_{i\sigma} + \hat{V}_0 \,. \tag{11}$$

Here $c^{\dagger}_{i\sigma}$ and $c_{i\sigma}$ are the creation and annihilation operators for an electron at site (atomic position) i with spin σ, while U is the on-site (spin-independent) interaction, $-t$ is the hopping energy, and ε_0 the orbital energy. The summation $\sum_{\langle i,j \rangle}$ is restricted to the nearest-neighbor sites. The Hamiltonian further contains a potential V_0 that can be chosen to fix the zero-energy scale. The model does not contain a long-range interaction, hence everything linked to non-locality is not treated, but we can suppose that these kind of issues are well treated by Hartree-Fock, GW, etc. Moreover, to keep the discussion simple, we will consider only sites that have the same energy, thus neglecting any inhomogeneity which can mimic the role of the crystal potential in a solid [24]. The physics of the Hubbard model arises from the competition between the hopping term, which prefers to delocalize electrons, and the on-site interaction, which favors localization. The ratio U/t is a measure for the relative contribution of both terms and is the intrinsic, dimensionless coupling constant of the Hubbard model. Depending on this ratio the system is a metal or an insulator.
This model can be solved exactly for small clusters. In the following we will consider a "molecule" made up of two equivalent sites, each with one orbital. We will look at the one- and two-electron cases (quarter and half filling, respectively) in order to cover a wide physics.

3.1 Exact solution

In order to calculate the exact 1-GF for the Hubbard dimer we start from its Lehmann representation (2) and we project it onto the (orthonormal) one-electron site basis $\{\varphi_i(x)\}$, such that[4] $G(x_1, x_2; \omega) = \sum_{ij} G_{ij}(\omega) \varphi_i(x_1) \varphi^*_j(x_2)$ and

$$G_{ij\sigma}(\omega) = \lim_{\eta \to 0^+} \left(\sum_m \frac{\langle \Psi_0^N | \hat{c}_i | \Psi_m^{N+1} \rangle \langle \Psi_m^{N+1} | \hat{c}_j^{\dagger} | \Psi_0^N \rangle}{\omega - (E_m^{N+1} - E_0^N) + \mathrm{i}\eta} + \sum_n \frac{\langle \Psi_0^N | \hat{c}_j^{\dagger} | \Psi_n^{N-1} \rangle \langle \Psi_n^{N-1} | \hat{c}_i | \Psi_0^N \rangle}{\omega - (E_0^N - E_n^{N-1}) - \mathrm{i}\eta} \right), \tag{12}$$

with the indices i, j running over the sites. The ground-state wavefunction and energy $|\Psi_0^N\rangle$ and E_0^N of the N-electron system as well as the wavefunctions and energies $|\Psi_m^{N\pm1}\rangle$ and $E_m^{N\pm1}$ of

[4] This expression for the 1-GF comes naturally by using the following basis transformation for the field operators: $\hat{\psi}(\mathbf{x}) = \sum_i \hat{c}_i \varphi_i(\mathbf{x})$ and $\hat{\psi}^{\dagger}(\mathbf{x}) = \sum_i \hat{c}_i^{\dagger} \varphi_i^*(\mathbf{x})$, where $\hat{c}_i$ and $\hat{c}_i^{\dagger}$ are annihilation and creation operators, respectively.

the ($N\pm1$)-electron system can be obtained by diagonalizing the Hamiltonian for $N-1$, N, and $N+1$ electrons, separately (as shown in Appendix A).

3.1.1 $N = 1$

- **One site**
  In the limit $t \to 0$ the two-site Hubbard model should represent two isolated atoms [25, 26]. In order to compare this limiting case with the solution of an isolated atom, we calculate the exact 1-GF for the case of a one-site Hubbard model with one electron. We choose as ground state the spin-up configuration $|\uparrow\rangle$ which has energy ε_0 (equivalently, the spin-down configuration could be chosen). There is only one state with $N-1 = 0$ electrons, i.e., the vacuum $|\ \rangle$ with energy 0, and only one state with $N+1 = 2$ electrons, i.e., $|\uparrow\downarrow\rangle$ with energy $\varepsilon_0 + U$. Moreover, note that, since we consider a spin-independent Hamiltonian, the 1-GF in (12) is diagonal in spin. Therefore the exact one-particle Green function reads
  $$G^{\uparrow} = \frac{1}{\omega - \varepsilon_0 - \mathrm{i}\eta}\,, \qquad G^{\downarrow} = \frac{1}{\omega - U - \varepsilon_0 + \mathrm{i}\eta}\,. \tag{13}$$
  There is only a removal energy, $\omega = \varepsilon_0$, and an addition energy, $\omega = \varepsilon_0 + U$, for this system. Note that by taking $U = 0$ in (13) one gets the non-interacting G_0.
  The exact self-energy can be obtained from the Dyson equation $\Sigma(\omega)=G_0^{-1}(\omega)-G^{-1}(\omega)$
  $$\Sigma(\omega) \;=\; \begin{pmatrix} 0 & 0 \\ 0 & U \end{pmatrix}, \tag{14}$$
  which reflects the fact that the electron in the ground-state can interact only with an additional spin-down electron.

- **Two sites**
  Using the information in Tables 1-2 of the Appendix A we can build the 1-GF at quarter filling. The 1-GF has the following structure
  $$G(\omega) = \begin{pmatrix} G^{\uparrow}_{11} & G^{\uparrow}_{12} & 0 & 0 \\ G^{\uparrow}_{21\uparrow} & G^{\uparrow}_{22} & 0 & 0 \\ 0 & 0 & G^{\downarrow}_{11} & G^{\downarrow}_{12} \\ 0 & 0 & G^{\downarrow}_{21} & G^{\downarrow}_{22} \end{pmatrix}. \tag{15}$$
  with
  $$G^{\uparrow}_{ij}(\omega) = \frac{(-1)^{i-j}}{2}\left(\frac{1}{\omega - (\varepsilon_0{+}t) + \mathrm{i}\eta} + \frac{(-1)^{i-j}}{\omega - (\varepsilon_0{-}t) - \mathrm{i}\eta}\right), \tag{16}$$
  and
  $$\begin{aligned} G^{\downarrow}_{ij}(\omega) =& \frac{(-1)^{i-j}}{4}\left(\frac{1}{\omega - (\varepsilon_0{+}t) + \mathrm{i}\eta} + \frac{1}{\omega - (\varepsilon_0{+}t{+}U) + \mathrm{i}\eta}\right) \\ &+ \frac{1}{2}\left(\frac{\frac{1}{a^2}\left(1 + \frac{4t}{(c-U)}\right)^2}{\omega - \left(\varepsilon_0{+}t{-}\frac{c-U}{2}\right) + \mathrm{i}\eta} + \frac{\frac{1}{b^2}\left(1 - \frac{4t}{(c+U)}\right)^2}{\omega - \left(\varepsilon_0{+}t{+}\frac{c+U}{2}\right) + \mathrm{i}\eta}\right), \end{aligned} \tag{17}$$

where $c = \sqrt{16t^2+U^2}$, $a = \sqrt{2\left(\frac{16t^2}{(c-U)^2}+1\right)}$, and $b = \sqrt{2\left(\frac{16t^2}{(c+U)^2}+1\right)}$. The one-particle Green function is spin-diagonal; the spin-down block has only the electron part, whereas the spin-up block has both electron and hole parts. This is consistent with the fact that with a spin-up electron in the ground state one can have only spin-up holes. Note also that $G_{11\uparrow} = G^{U=0}_{11\uparrow}$ and $G_{12\uparrow} = G^{U=0}_{12\uparrow}$, whereas

$$G^{U=0}_{ij\downarrow}(\omega) = \frac{(-1)^{i-j}}{2}\left(\frac{1}{\omega-(\varepsilon_0+t)+\mathrm{i}\eta}+\frac{(-1)^{i-j}}{\omega-(\varepsilon_0-t)+\mathrm{i}\eta}\right). \tag{18}$$

One can interpret the poles of the Green function using a simple molecular picture with the energies of the bonding and antibonding orbitals at $\varepsilon_0 - t$ and $\varepsilon_0 + t$, respectively, and with the spin-up electron in the bonding orbital. If $U = 0$ one can remove (add) a spin-up (spin-down) electron from (to) the bonding orbital ($\omega = \varepsilon_0 - t$), and one can add a spin-up/spin-down electron to the antibonding orbital ($\omega = \varepsilon_0 + t$). When the on-site electron-electron interaction U is switched on, the energy for the addition of a spin-down electron to the bonding state evolves to $\omega = \varepsilon_0 + t + (U-c)/2$, and the addition energy to the antibonding state is split into $\omega = \varepsilon_0 + t$ and $\omega = \varepsilon_0 + t + U$. The addition of a spin-down electron, furthermore, can excite the system, giving rise to the satellite $\omega = \varepsilon_0 + t + (U+c)/2$.

We now discuss two limiting cases, which are directly related to the importance of correlation: the noninteracting limit and the atomic limit.

- **noninteracting limit: $U \to 0$**
  Of course in this limit one retrieves the noninteracting Green function. The spectral weight of the pole $\omega = \varepsilon_0+t+(c+U)/2$, which becomes $\omega = \varepsilon_0+3t$, is suppressed. Note that $\omega = \varepsilon_0 + 3t$ deviates from the antibonding peak by an energy $\Delta = 2t$, which is a pole of the polarizability P (as shown later in Eq. (22)): this pole of the Green function involves hence an excitation of the system, which justifies the identification of the related peak as a satellite.
- **atomic limit: $t \to 0$**
  We first notice that for $t = 0$ the ground state energy goes to $E_0 = \varepsilon_0$ so that it approaches degeneracy with the other doublet state in the $N = 1$ spin-up subspace (see Table 1 in Appendix A). However for any small but finite t this degeneracy is lifted. In this limit the poles of the Green function reduce to the addition and removal energies of two isolated atoms, one with one electron and the other one empty [25, 26]. Indeed the spin-up poles retain the same equal weight ($\pm 1/2$) and go towards $\omega = \varepsilon_0$, which can be interpreted as the removal energy of an atom with a spin-up electron and the addition energy of an empty atom, respectively. The spin-down poles $\omega = \varepsilon_0 + t$ and $\omega = \varepsilon_0 + t - (c - U)/2$, and $\omega = \varepsilon_0 + t + U$ and $\omega = \varepsilon_0 + t + (c + U)/2$ (satellite) merge at $\omega = \varepsilon_0$ and $\omega = \varepsilon_0 + U$, respectively, all with equal weight $(\pm 1/4)$ (see Appendix A). These two energies can be interpreted

as the energy for the addition of a spin-down electron to an empty atom and to an atom with a spin-up electron, respectively.

The self-energy $\Sigma(\omega) = G_0^{-1}(\omega) - G^{-1}(\omega)$, has the following structure

$$\Sigma(\omega) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \Sigma_{11}^{\downarrow} & \Sigma_{12}^{\downarrow} \\ 0 & 0 & \Sigma_{12}^{\downarrow} & \Sigma_{11}^{\downarrow} \end{pmatrix}. \tag{19}$$

The self-energy has only a spin-down part, i.e., the electron in the system interacts only with spin-down electrons. This is in line with the fact that our system consists of one electron with spin up. When we add a second electron to the system, it can have spin up or spin down. If the second electron has spin up, the two electrons are locked in a configuration where two different sites are occupied, and therefore there is no interaction. If the second electron has spin down, the two electrons can interact.

3.1.2 $N = 2$

Using the information in Tables 1-3 of the Appendix A we can build the 1-GF at half filling. It has the structure given in Eq. (15), with the components gives by

$$\begin{aligned} G_{ij}^{\sigma}(\omega) =& \frac{(-1)^{i-j}}{2a^2} \left(\frac{\big(1 + 4t/(c{-}U)\big)^2}{\omega - \big(\varepsilon_0 - t + (c{+}U)/2\big) + \mathrm{i}\eta} + \frac{(-1)^{i-j}\big(1 - 4t/(c{-}U)\big)^2}{\omega - \big(\varepsilon_0 + t + (c{+}U)/2\big) + \mathrm{i}\eta} \right) \\ &+ \frac{1}{2a^2} \left(\frac{\big(1 + 4t/(c{-}U)\big)^2}{\omega - \big(\varepsilon_0 + t - (c{-}U)/2\big) - \mathrm{i}\eta} + \frac{(-1)^{i-j}\big(1 - 4t/(c{-}U)\big)^2}{\omega - \big(\varepsilon_0 - t - (c{-}U)/2\big) - \mathrm{i}\eta} \right), \end{aligned} \tag{20}$$

which enjoys a nice spin symmetry $G_{ij}^{\uparrow} = G_{ij}^{\downarrow}$. We can now study the limits $U \to 0$ and $t \to 0$.

- **noninteracting limit: $U \to 0$**
  In this limit the ground state becomes $|\Psi_0\rangle = \frac{1}{2}\big(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle + |\,\updownarrow, \cdot\rangle + |\cdot, \updownarrow\rangle\big)$ (where $|1,\, 2\rangle$, with occupations of the sites 1, 2 given by $\cdot$, $\uparrow$, $\downarrow$, $\updownarrow$ indicate Slater determinants) with energy $E_0 = 2(\varepsilon_0 - t)$, and the one-particle Green function reduces to the noninteracting one

$$G_{ij\sigma}^{U=0}(\omega) = \frac{(-1)^{i-j}}{2} \left(\frac{1}{\omega - (\varepsilon_0{+}t) + \mathrm{i}\eta} + \frac{(-1)^{i-j}}{\omega - (\varepsilon_0{-}t) - \mathrm{i}\eta} \right). \tag{21}$$

  Note that this is the same as the one for the spin-up block of the one-electron case.

- **atomic limit: $t \to 0$**
  In the atomic limit there are no double occupancies (Heitler-London limit), therefore the two electrons, one with spin up and the other with spin down, are localized one on one site and the other on the other site with equal probability, i.e. the ground state is the singlet $|\Psi_0\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle\big)$. The spectral function thus shows, for each spin, two peaks

with the same spectral weight $(1/2)$, one for the removal of an electron (peak at ε_0), and one for the addition of a second electron (peak at $\varepsilon_0 + U$), which represents the removal and addition energies, respectively, of an isolate atom with one electron.

As far as the self-energy is concerned, unlike for the case at half filling, it has a spin-up and spin-down block, with the two blocks being the same. This again reflects the fact that the system is symmetric with respect to spin.

3.2 GW

The GW equations, as well as all the other approximations to the self-energy, should, in principle, be solved self-consistently, since the self-energy is a functional of the one-body Green function. In this chapter, however, we will not use self-consistency. In particular for GW we will use G_0 to build W and Σ: $\Sigma = v_H + iG_0W_0$, with $v_H = -iv_cG_0$ and $W_0 = [1 + iv_cG_0G_0]^{-1}v_c$. This approach is called G_0W_0 or one-shot GW, and it is often used in practice (although one uses a G_0 that already contains part of the interaction through a suitable single-particle Hamiltonian, like KS, HF, or quasiparticle self-consistent GW), since a fully self-consistent procedure is computationally demanding, especially for large systems. Besides this, it is important to stress that the failures of GW we discuss in this chapter are not solved by self-consistency since they are rooted in the fundamental structure of the GW approximation.

3.2.1 $N = 1$

- **One site**
  The RPA polarizability $P(\omega) = -i \int \frac{d\omega'}{2\pi} G(\omega+\omega') G(\omega')$ is zero, consistent with the fact that the system has only one state (the site orbital) in which it could be excited, which would however require a spin-flip, which is not allowed. Therefore there is no response of the system, and thus the screened potential equals the bare Coulomb potential, $W = U$. The GW self-energy, therefore, is equal to the exact one, and, consequently, also the Green function. Therefore, together with the fact that GW is also exact for the empty atom, for two separate one-site Hubbard models the GWA yields the exact solution, contrary to the $t \to 0$ limit of the two sites, as we will show in the following. This shows again the relation of the GWA to the classical description of the system charge: for the one-site Hubbard model, where the electron is well confined, the classical description of the system works well (i.e. one knows where the electron is). For the two-site Hubbard model, where, instead, the electron shows its quantum nature, this classical picture fails. This is in line with the size-consistency problem GW suffers from [8, 9] and it is analogous to what is observed in DFT [27].

- **Two sites**
  The RPA polarizability $P = -iGG$ is given by

$$P^{\uparrow}_{ij}(\omega) = \frac{(-1)^{i-j}}{4} \left(\frac{1}{\omega - 2t + \mathrm{i}\eta} - \frac{1}{\omega + 2t - \mathrm{i}\eta} \right). \tag{22}$$

Note that, since we have only one electron with spin up in the ground state the polarizability is not zero only for the spin-up block. Moreover the same result is obtained whether the noninteracting or exact G is used, since they are equal for the spin-up block. The RPA screened interaction $W(\omega) = \big(1 - UP(\omega)\big)^{-1}U$ becomes

$$W_{ij}(\omega) = U\delta_{ij} + (-1)^{i-j}\frac{U^2 t}{\omega^2 - h^2}\,, \tag{23}$$

with $h^2 = 4t^2 + 2Ut$. The self-energy $\Sigma(\omega) = v_H + \frac{i}{2\pi}\int d\omega' G(\omega+\omega')\,W(\omega')e^{i\omega'\eta}$, has the following structure,

$$\Sigma(\omega) = \begin{pmatrix} \Sigma_{11\uparrow} & \Sigma_{12\uparrow} & 0 & 0 \\ \Sigma_{12\uparrow} & \Sigma_{11\uparrow} & 0 & 0 \\ 0 & 0 & \Sigma_{11\downarrow} & \Sigma_{12\downarrow} \\ 0 & 0 & \Sigma_{12\downarrow} & \Sigma_{22\downarrow} \end{pmatrix}. \tag{24}$$

A striking difference with respect to the exact self-energy in Eq. (19) is the nonzero elements of the spin-up block. Moreover, since $G_\uparrow = G_{0\uparrow}$, G_0W_0, G_0W, GW_0, and GW (with G the exact Green function) give the same expression for the spin-up self-energy. As we will discuss in the following this is a direct consequence of the self-screening error in the GWA.

The 1-GF can be obtained from

$$G^{GW}(\omega) = \Big(G_0^{-1}(\omega) - \Sigma(\omega)\Big)^{-1}. \tag{25}$$

It reads

$$G^{GW}_{ij\uparrow}(\omega) = (-1)^{i-j}\left(\frac{\frac{1}{4}+\frac{2t+h}{4A}}{\omega-\omega_1^+ + i\eta} + \frac{\frac{1}{4}-\frac{2t+h}{4A}}{\omega-\omega_1^- - i\eta}\right) + \frac{\frac{1}{4}-\frac{2t+h}{4A}}{\omega-\omega_2^+ + i\eta} + \frac{\frac{1}{4}+\frac{2t+h}{4A}}{\omega-\omega_2^- - i\eta}\,, \tag{26}$$

$$G^{GW}_{ij\downarrow}(\omega) = (-1)^{i-j}\left(\frac{\frac{1}{4}+\frac{2t-h+U/2}{4B}}{\omega-\omega_3^+ + i\eta} + \frac{\frac{1}{4}-\frac{2t-h+U/2}{4B}}{\omega-\omega_3^- + i\eta}\right) + \frac{\frac{1}{4}-\frac{2t+h-U/2}{4C}}{\omega-\omega_4^+ + i\eta} + \frac{\frac{1}{4}+\frac{2t+h-U/2}{4C}}{\omega-\omega_4^- + i\eta}\,, \tag{27}$$

where

$$A = \sqrt{\big(h+2t\big)^2 + \frac{2U^2t}{h}}\,,\; B = \sqrt{\left(h-2t-\frac{U}{2}\right)^2 + \frac{2U^2t}{h}}\,,\; C = \sqrt{\left(h+2t-\frac{U}{2}\right)^2 + \frac{2U^2t}{h}}$$

and the poles are at

$$\omega_1^{+/-} = \big(2\varepsilon_0 - h \pm A\big)/2\,, \qquad \omega_2^{+/-} = \big(2\varepsilon_0 + h \pm A\big)/2\,,$$
$$\omega_3^{+/-} = \big(2\varepsilon_0 + h + U/2 \pm B/2\big)\,, \qquad \omega_4^{+/-} = \big(2\varepsilon_0 + h + U/2 \pm C\big)/2\,.$$

We can now study the noninteracting limit and the atomic limit.

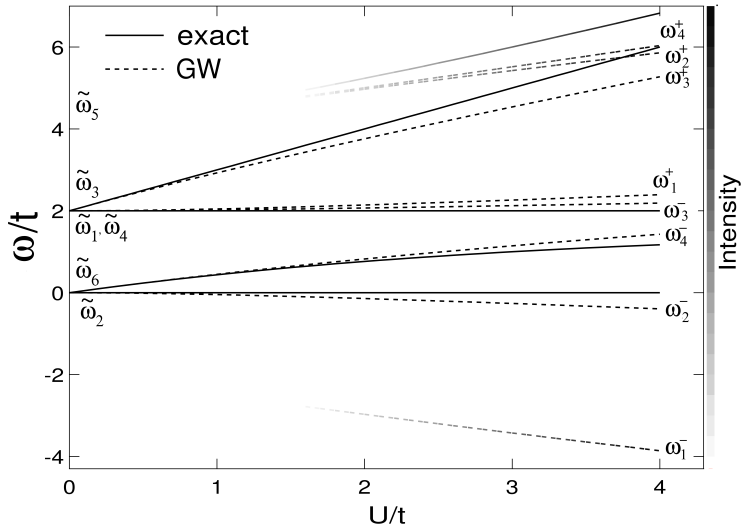

Fig. 5: *Two-site Hubbard model at quarter filling from Ref. [10]: comparison between exact (solid lines) and GW (dashed lines) renormalized excitation energies ω/t as function of U/t, with $\varepsilon_0 = t = 1$ eV. The thin lines represent weak satellites, which appear with increasing interaction U. The labels on the left of the figure refer to the exact energies, with $\tilde{\omega}_1 = \tilde{\omega}_4 = \varepsilon_0 + t$, $\tilde{\omega}_2 = \varepsilon_0 - t$, $\tilde{\omega}_3 = \varepsilon_0 + t + U$, $\tilde{\omega}_5 = \varepsilon_0 + t + U/2 + c/2$, and $\tilde{\omega}_6 = \varepsilon_0 + t + U/2 - c/2$, whereas the labels on the right refer to the GW energies.*

– **noninteracting limit: $U \to 0$**
  In this limit the interacting GW Green function reduces of course to the noninteracting one, with the poles ω_1^+, ω_2^-, ω_3^+, ω_3^-, and ω_4^- collapsing to the bonding and antibonding energies $\omega = \varepsilon_0 - t$ and $\omega = \varepsilon_0 + t$, respectively. The poles ω_1^-, ω_2^+, and ω_4^+, instead, collapse to $\omega = \varepsilon_0 - 3t$ and $\omega = \varepsilon_0 + 3t$ with zero intensity. Note also that the energies $\omega = \varepsilon_0 - 3t$ and $\omega = \varepsilon_0 + 3t$ deviate from the bonding and antibonding peaks, respectively, by $\pm 2t$, which are the poles of the RPA polarizability P. This means that these energies arise from excitations of the system. The poles ω_1^-, ω_2^+, and ω_4^+ can be, therefore, identified as satellites.

– **atomic limit: $t \to 0$**
  In this limit all the spin-up poles acquire an equal weight $(\pm 1/4)$ and go towards $\omega = \varepsilon_0$, in agreement with the exact solution. The unphysical poles, and hence the self-screening problem, cannot be detected in the atomic limit because the excitation energy $2t \to 0$. The spin-down poles ω_3^+ and ω_4^+ (satellite) go towards $\omega = \varepsilon_0 + \frac{U}{2}$, each with weight $(\pm 1/2)$, whereas ω_3^- and ω_4^- merge at $\omega = \varepsilon_0$ with zero weight. This is in contrast with the exact solution where the four poles go towards $\omega = \varepsilon_0$ and $\omega = \varepsilon_0 + U$ with equal weight $(\pm 1/4)$. The error stems from the fact that the GW self-energy is static in this limit (it consists of the Hartree potential $U/2$ only for the spin-down block). The exact atomic-limit self-energy, instead, has an extra term that is frequency-dependent, namely

$$\Sigma_{ij}^{\downarrow}(\omega) = \delta_{ij} \frac{U}{2} \left(1 + \frac{U}{2(\omega - \varepsilon_0) - U + \mathrm{i}\eta} \right).$$

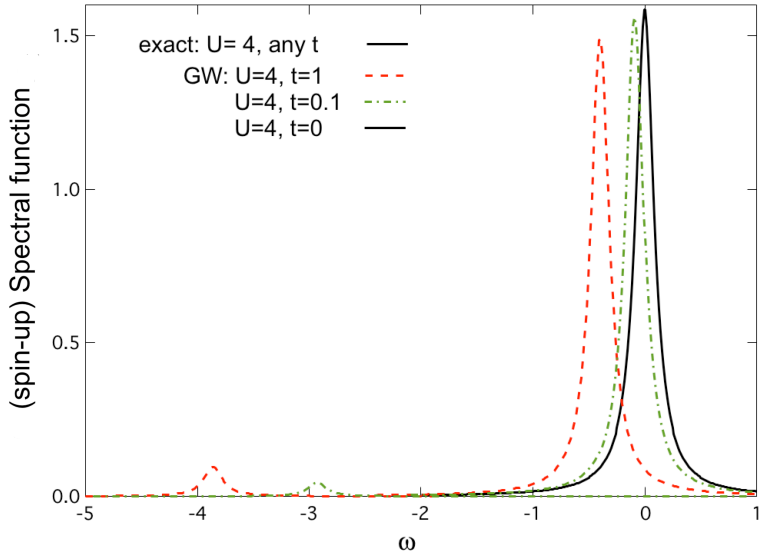

Fig. 6: *Two-site Hubbard model at quarter filling: comparison between exact (black continuous line) and GW (red dashed, green dash-dotted, and black continuous lines for* $t = 1$, 0.1, *and* 0, *respectively) spectral function (removal part) at* $U = 4$ *and* $\varepsilon_0 = t$.

This self-energy shows, in addition to the Hartree potential, a frequency-dependent term that creates the extra pole $\omega = \varepsilon_0 + U$ besides the only pole $\omega = \varepsilon_0$ of the noninteracting Green function (for $t \to 0$). The appearance of the peak $\omega = \varepsilon_0 + U$ is an effect of "strong correlation". One can understand that the problems of GW in the atomic limit arise from the interpretation of the charge density: in GW, it is treated as a classical charge distribution, namely half an electron on each atom in the limit $t \to 0$, that responds to the additional electron or hole. Instead, it should rather express the probability for an electron to be on one or the other atom, so that the additional electron can meet an empty or occupied atom with equal probability, which leads to the peak splitting.

In Fig. 5 we compare the GW and exact addition/removal energies. The GW approximation yields two satellites more than the exact solution, namely, ω_1^- and ω_2^+. These satellites come from the poles of the spin-up Green function and are produced by the frequency-dependent spin-up self-energy. This is a consequence of the self-screening problem the GWA suffers from. The direct consequences of the self-screening error on photoemission spectra are illustrated in Fig. 6, where we reported the spectral function $A(\omega)$ for the spin-up channel. We see that the exact result shows only one peak, which corresponds to the site energy ε_0, whereas the GW approximation gives in general spurious peaks at any value of the hopping, except at $t = 0$. This is due to the fact that one and the same electron can be used twice; for example, once to be removed, and at the same time to screen this perturbation. At $t = 0$, instead, the electron does not hop between the two sites, therefore there is no screening in the system, $W = v_c$ and we retrieve the HF exact result. The spectral function for the spin-down channel at quarter filling is reported

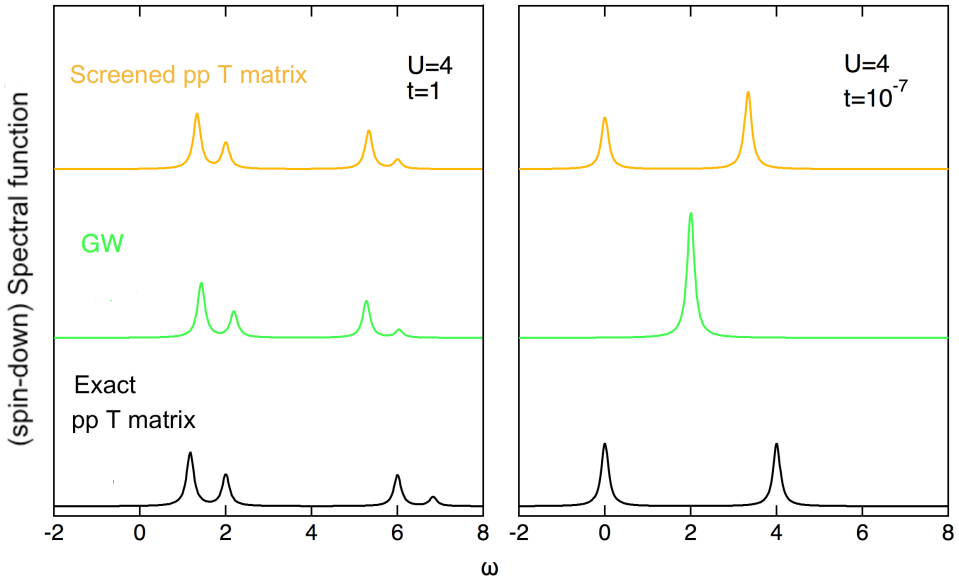

Fig. 7: *Spectral function at quarter filling: exact (lower panel) vs. GW (middle panel), pp T-matrix, and screened pp T-matrix (upper panel).*

in Fig. 7. The GWA produces a spectral function profile in good agreement with the exact one at moderately strong interaction U/t (middle left panel). In the atomic limit (middle right panel), instead, GW produces only one peak.

The inclusion of an explicit vertex in the self-energy, namely a two-point $\Gamma = \delta + f_{xc}P$ for valence bands and, e.g., $\Gamma = \delta$ for the well separated conduction bands, removes by construction the self-screening error for the one-electron case. Indeed, using this vertex it is clear that the self-energy remains unchanged for the spin-down block, whereas for the spin-up block $\Sigma_{xc} = -v_H$. Therefore, we arrive at

$$\Sigma(\omega) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \Sigma^{\downarrow}_{11} & \Sigma^{\downarrow}_{12} \\ 0 & 0 & \Sigma^{\downarrow}_{12} & \Sigma^{\downarrow}_{22} \end{pmatrix}, \tag{28}$$

with $\Sigma_{11\downarrow}$ and $\Sigma_{12\downarrow}$ within the GW approximation. We now get a self-energy with the same structure as the correct one. It is clear that now the spin-up block of the one-particle Green function equals the noninteracting Green function, like the exact one. For the spin-down block we get four poles as in the exact case, although they still differ from the exact values by the GW error.

We point out that it is essential to use a three-point vertex in order to get the correct number of poles. Indeed, by using the two-point vertex $\Gamma = \delta + f_{xc}P$ both for valence and conduction bands, one would get a zero self-energy for the spin-down block. Therefore the spin-down Green function would equal the noninteracting one, which has only two poles instead of four as the exact interacting Green function: there would be no satellites since there would be no screening.

One can readily understand that these vertex corrections will not correct the wrong atomic limit of GW, since the two problems have a different nature; therefore a more complex vertex, able to introduce an additional frequency-dependence in Σ, is needed to fix both.

3.2.2 $N = 2$

In order to calculate the one-body Green function within the GWA (G_0W_0), we need the following ingredients,

$$P_{ij}^{\sigma}(\omega) = \frac{(-1)^{i-j}}{4}\left(\frac{1}{\omega - 2t + \mathrm{i}\eta} - \frac{1}{\omega + 2t - \mathrm{i}\eta}\right), \tag{29}$$

$$W_{ij}(\omega) = U\delta_{ij} + (-1)^{i-j}\frac{2U^2t}{\omega^2 - h'^2}, \tag{30}$$

$$\Sigma_{ij}^{\sigma}(\omega) = \frac{U}{2}\delta_{ij} + \frac{U^2t}{2h'}\left(\frac{1}{\omega - (\varepsilon_0 + t + h') + \mathrm{i}\eta} + \frac{(-1)^{i-j}}{\omega - (\varepsilon_0 - t - h') - \mathrm{i}\eta}\right), \tag{31}$$

with $h'^2 = 4t^2 + 4tU$. The one-body Green function hence reads

$$G_{ij\sigma}^{GW} = (-1)^{i-j}\left(\frac{\frac{1}{4} + \frac{h'+2t+\frac{U}{2}}{4A'}}{\omega - \omega_1^+ + \mathrm{i}\eta} + \frac{\frac{1}{4} - \frac{h'+2t+\frac{U}{2}}{4A'}}{\omega - \omega_1^- - \mathrm{i}\eta}\right) + \frac{\frac{1}{4} + \frac{-h'-2t+\frac{U}{2}}{4B'}}{\omega - \omega_2^+ + \mathrm{i}\eta} + \frac{\frac{1}{4} - \frac{-h'-2t+\frac{U}{2}}{4B'}}{\omega - \omega_2^- - \mathrm{i}\eta} \tag{32}$$

with

$$A' = \sqrt{\left(2t + h' + \frac{U}{2}\right)^2 + \frac{4U^2t}{h'}}, \quad B' = \sqrt{\left(2t + h' - \frac{U}{2}\right)^2 + \frac{4U^2t}{h'}},$$

and

$$\omega_1^{+/-} = \frac{1}{2}\left(2\varepsilon_0 - h' + \frac{U}{2} \pm A'\right), \quad \omega_2^{+/-} = \frac{1}{2}\left(2\varepsilon_0 + h' + \frac{U}{2} \pm B'\right).$$

We can now study the two limits $U \to 0$ and $t \to 0$.

- **noninteracting limit: $U \to 0$**
  In this limit $h' \approx 2t$, therefore the poles ω_1^+ and ω_2^- collapse to the poles of G_0, $\omega = \varepsilon_0 + 2t$ and $\omega = \varepsilon_0 - 2t$, whereas ω_1^- and ω_2^+ collapse to $\omega = \varepsilon_0 - 3t$ and $\omega = \varepsilon_0 + 3t$. These last two poles have zero intensity at $U = 0$ and deviate from the bonding/antibonding energies by $\pm 2t$ (poles of P), which justify their identification as satellites.

- **atomic limit: $t \to 0$**
  In the limit $t \to 0$ we have $h' \approx 0$, from where it follows that the poles ω_1^+ and ω_2^+ merge at $\omega = \varepsilon_0 + \frac{U}{2}$, whereas the poles ω_1^- and ω_2^- go towards $\omega = \varepsilon_0$. This is in contrast with the exact solution where the poles go towards the poles of the isolated atom ($\omega = \varepsilon_0$ and $\omega = \varepsilon_0 + U$). This can again be understood considering that GW treats the charge/spin density as a classical charge distribution, namely an half electron with half spin up and half an electron with half spin down on each atom that respond to the additional electron or hole in the atomic limit.

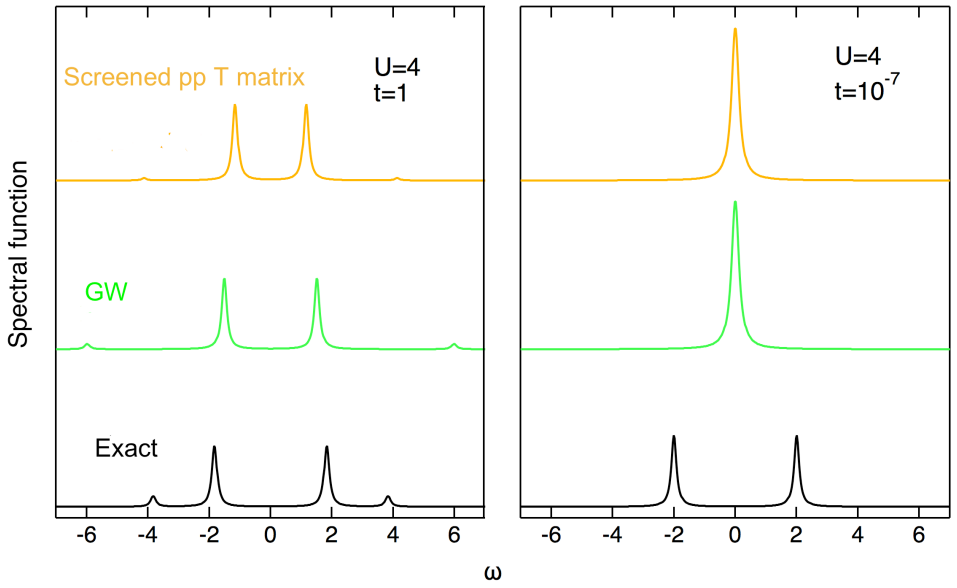

Fig. 8: *Spectral function at half filling: exact (lower panel) vs. GW (middle panel) and screened pp T-matrix (upper panel).*

The spectral function at half filling is reported in Fig. 8: for moderately strong interaction (middle left panel), GW performs quite well (although the position of the satellites is overestimated), whereas at strong correlation (middle right panel), no gap is observed.[5]

Finally, we compare the LUMO at quarter filling with the HOMO at half filling. The LUMO in the N-electron system should be equal to the HOMO in the $N+1$-electron system. This occurs for the exact case, and it is the pole at $\omega = \varepsilon_0 + t - (c-U)/2$. Within the GW approximation the two energies, namely ω_4^- at quarter filling and ω_2^- at half filling, are different. The difference stems from the different polarizability P, which is used to build the screening for the N and $N+1$ cases. This simply shows again that the problem arises from the use of a TC-TC screening, which depends only on the charge density of the system but not on the charge that is to be screened. Indeed, the removal of an electron from the $N+1$-electron system should be screened by N electrons only—which can be expressed through vertex corrections. The source of this error comes from the fact that direct and exchange interactions are not treated on equal footing in GW.

3.3 T-matrix

Also for the T-matrix approximation we adopt the same protocol adopted for the GWA, namely we use G_0 to build T and Σ. Projected in the (orthonormal) site-basis of the Hubbard model the

[5] We note that the Hubbard dimer at half filling enjoys an interesting symmetry (particle-hole symmetry) [28] by setting $\varepsilon_0 = -U/2$ and $V_0 = U$ in the Hubbard Hamiltonian. With this choice the exact 1-GF has the same expression as in (20) but with $\varepsilon_0 = -U/2$. For the GW and T-matrix 1-GF one has to set $\varepsilon_0 = 0$. For $U \neq 0$ the particle-hole symmetry is lost in the GW and T-matrix approximations due to the use of G_0. We enforce this symmetry by dropping the terms U/2 in the poles and corresponding weights of the GW and T-matrix 1-GF.

(unscreened) T-matrix becomes

$$T^{\sigma\sigma'}_{ilkj}(\omega) = -\mathrm{i}\delta_{il}\delta_{jk}\left(\bar{T}^{\sigma\sigma'}_{1,ij}(\omega) - \delta_{\sigma\sigma'}\bar{T}^{\sigma\sigma}_{1,ij}(\omega)\right), \tag{33}$$

with $\bar{T}^{\sigma\sigma'}_1(\omega) = \left(1 + UL^{\sigma\sigma'}_0(\omega)\right)^{-1}U$, from which

$$\Sigma^{\sigma}_{ij}(\omega) = -\mathrm{i}\int \frac{d\nu}{2\pi}\, G^{\bar{\sigma}}_{ji}(\nu)\, \bar{T}^{\sigma\bar{\sigma}}_{ij}(\omega \pm \nu). \tag{34}$$

Here $\bar{\sigma}$ indicates a spin opposite to σ, the sign '$+$' refers to the particle-particle contribution for which $L^{\sigma\sigma',pp}_{0,ij}(\omega) = -\mathrm{i}\int \frac{d\omega'}{2\pi}\, G^{\sigma}_{ij}(\omega')\, G^{\sigma'}_{ij}(\omega-\omega')e^{-\mathrm{i}\omega'\eta}$, and the sign '$-$' refers to the electron-hole contribution for which $L^{\sigma\sigma',eh}_{0,ij}(\omega) = -\mathrm{i}\int \frac{d\omega'}{2\pi}\, G^{\sigma}_{ij}(\omega')G^{\sigma'}_{ij}(\omega'-\omega)e^{\mathrm{i}\omega'\eta}$. More details on the spin and time structure of the T-matrix can be found in Refs. [3, 4].

The screened T-matrix approximation is more difficult to handle, because the screened interaction is frequency dependent. However, in the atomic limit $t \to 0$ there is no screening in the system and the screened T-matrix reduces to the unscreened one. For $t > 0$ we assume that for the model used here the major contribution to the T-matrix arises from the on-site screened interaction. This is dominated by the bare interaction U, which justifies to take the screened interaction in its static ($\omega = 0$) limit. In this case the structure of the screened T-matrix is the same as for the unscreened T-matrix with the onsite screened Coulomb interaction $W = U - (1+\delta_{N,2})U^2t/h^2$, with $h^2 = 4t^2 + 2Ut(1+\delta_{N,2})$ and N the total number of electrons in the system, replacing U. We notice that with this approximation, in particular assuming a static W, the screened T-matrix will not reduce anymore to the unscreened T-matrix in the atomic limit.

3.3.1 $N = 1$

- **One site**
  In this case the particle-particle correlator $L^{\sigma\sigma',pp}_0(\omega) = -\mathrm{i}\int \frac{d\omega'}{2\pi}\, G^{\sigma}(\omega')\, G^{\sigma'}(\omega-\omega')e^{-\mathrm{i}\omega'\eta}$ is zero and $T = U$. The self-energy, therefore, is equal to the exact one, and, consequently, also the Green function

- **Two sites**
  In order to calculate the one-body Green function within the pp T-matrix approximation ($\mathrm{G_0T^{pp}_0}$), we need the following ingredients,

$$L^{\sigma\bar{\sigma},pp}_{ij}(\omega) = -\frac{1}{4}\left(\frac{1}{\omega - 2(\varepsilon_0+t) + \mathrm{i}\eta} + \frac{(-1)^{i-j}}{\omega - 2\varepsilon_0 + \mathrm{i}\eta}\right), \tag{35}$$

$$\bar{T}^{\sigma\bar{\sigma},pp}_{ij}(\omega) = U\delta_{ij} + \frac{U^2}{4}\left(\frac{1}{\omega - 2\varepsilon_0 - \frac{U}{2} - 2t + \mathrm{i}\eta} + \frac{1}{\omega - 2\varepsilon_0 - \frac{U}{2} + \mathrm{i}\eta}\right), \tag{36}$$

  from which the self-energy reads

$$\Sigma^{\uparrow}_{ij} = 0, \tag{37}$$

$$\Sigma^{\downarrow}_{ij} = \frac{U}{2}\delta_{ij} + \frac{U^2}{8}\left(\frac{1}{\omega - \varepsilon_0 - \frac{U}{2} - 3t + \mathrm{i}\eta} + \frac{(-1)^{i-j}}{\omega - \varepsilon_0 - \frac{U}{2} - t + \mathrm{i}\eta}\right), \tag{38}$$

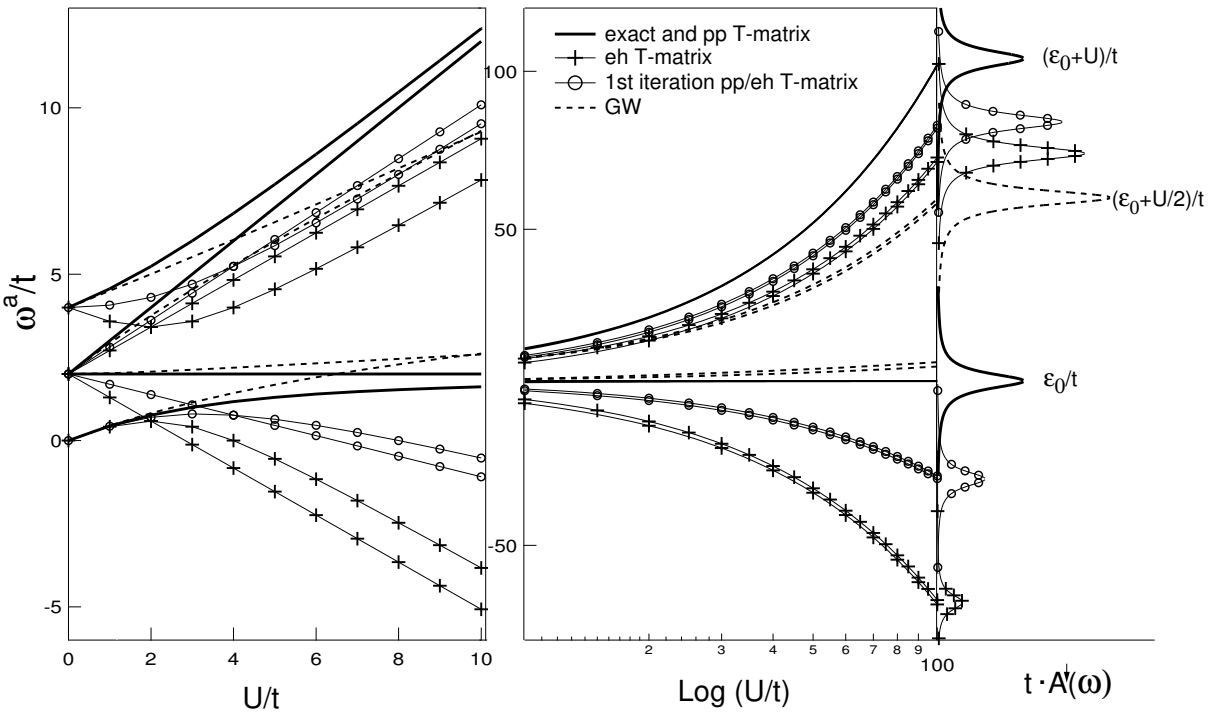

Fig. 9: *Two-site Hubbard model at quarter filling from Ref. [3]: comparison between the exact spin-down renormalized addition energies ω^a/t (solid lines) as function of U/t (left panel) and Log(U/t) (right panel) and the results obtained from GW (dashes), particle-particle (solid lines, equal to the exact result), electron-hole (crosses), and 1st iteration unscreened T-matrix (circles). In the atomic limit the spectral function, i.e., the peak positions and weights, is illustrated on the right-hand side, upon multiplying by t and taking the $t \to 0$ limit.*

which equals the exact self-energy. Of course this self-energy produces the exact 1-GF. Note that the unscreened T-matrix approximation does not suffer from the self-screening problem of GW, as one can expect since no screening is involved.

The eh T-matrix gives a self-energy with the same structure as the pp T-matrix (hence no self-screening), but with poles that are shifted by U from the exact ones, so that one cannot retrieve the exact result.

Moreover, since the pp T-matrix gives the exact results, it becomes clear that its screened version worsens the results, except in the atomic limit, where $W \to v_c$ and one has to get back the unscreened T-matrix results. As already discussed, this is not the case within the static approximation to W we are considering here. However, even with an approximate W the pp screened T-matrix is better than GW as shown in Fig. 9, where the renormalized addition energies ω^a/t for the spin-down channel are reported vs. U/t. Particularly interesting is the spectral function at $U/t \to \infty$: unlike the GWA, the T-matrix approximation "sees" where the electron is, although only the pp T-matrix "sees" well. Indeed, the eh T-matrix yields two peaks with the correct spectral weight, but at the wrong position, namely $\varepsilon_0 + U/\sqrt{2}$ and $\varepsilon_0 - U/\sqrt{2}$. Therefore, it is clear that in the case

of the Hubbard dimer with one electron, the particle-particle contribution to the T-matrix describes the essential physics.

Also note from Fig. 9 that the eh T-matrix and the pp T-matrix give the same results at the first iteration, since $\Sigma^{eh,(1)} = \Sigma^{pp,(1)}$. In particular, in the atomic limit their first iteration shows two peaks in the spin-down spectral function, although they are not correctly located and they do not have the same spectral weight. It hence contains the right physics, although the results are still poor. Finally the spectral function calculated using the pp screened T-matrix is shown in Fig. 7: although introducing screening in the pp T-matrix corrupts the exact result, the spectral function profile is in good agreement with the exact one and of similar quality as the GWA at moderately strong interaction U/t (upper left panel). In the atomic limit (middle right panel), instead, the screened pp T-matrix is superior to GW.

3.3.2 $N = 2$

In this case neither the pp nor the eh unscreened T-matrix reproduce the exact result. In particular, in the case of the pp T-matrix approximation we get

$$L_{ij}^{\sigma\bar{\sigma},pp} = -\frac{1}{4}\left(\frac{1}{\omega - 2(\varepsilon_0+t) + \mathrm{i}\eta} - \frac{1}{\omega - 2(\varepsilon_0-t) - \mathrm{i}\eta}\right), \tag{39}$$

$$T_{ij}^{\sigma\bar{\sigma},pp}(\omega) = U\delta_{ij} + \frac{U^2 t}{2\bar{h}}\left(\frac{1}{\omega - 2\varepsilon_0 - \bar{h} + \mathrm{i}\eta} - \frac{1}{\omega - 2\varepsilon_0 + \bar{h} - \mathrm{i}\eta}\right), \tag{40}$$

$$\Sigma_{ij}^{\sigma}(\omega) = \frac{U}{2}\delta_{ij} + \frac{U^2 t}{4\bar{h}}\left(\frac{1}{\omega - \varepsilon_0 - t - \bar{h} + \mathrm{i}\eta} + \frac{(-1)^{i-j}}{\omega - \varepsilon_0 + t + \bar{h} - \mathrm{i}\eta}\right), \tag{41}$$

with $\bar{h}^2 = 4t^2 + 2tU$. The one-body Green function hence reads

$$G_{ij\sigma}^{T^{pp}} = (-1)^{i-j}\left(\frac{\frac{1}{4} + \frac{\bar{h}+2t+\frac{U}{2}}{4A''}}{\omega - \omega_1^+ + \mathrm{i}\eta} + \frac{\frac{1}{4} - \frac{\bar{h}+2t+\frac{U}{2}}{4A''}}{\omega - \omega_1^- - \mathrm{i}\eta}\right) + \frac{\frac{1}{4} + \frac{-\bar{h}-2t+\frac{U}{2}}{4B''}}{\omega - \omega_2^+ + \mathrm{i}\eta} + \frac{\frac{1}{4} - \frac{-\bar{h}-2t+\frac{U}{2}}{4B''}}{\omega - \omega_2^- - \mathrm{i}\eta} \tag{42}$$

with

$$A'' = \sqrt{\left(2t + \bar{h} + \frac{U}{2}\right)^2 + \frac{2U^2 t}{\bar{h}}}\,, \quad B'' = \sqrt{\left(2t + \bar{h} - \frac{U}{2}\right)^2 + \frac{2U^2 t}{h'}}\,,$$

and

$$\omega_1^{+/-} = \frac{1}{2}\left(2\varepsilon_0 - \bar{h} + \frac{U}{2} \pm A''\right)\,, \quad \omega_2^{+/-} = \frac{1}{2}\left(2\varepsilon_0 + \bar{h} + \frac{U}{2} \pm B''\right).$$

Similar equations hold for the eh T-matrix approximation and the pp and eh (statically) screened T-matrix approximation. The pp T-matrix performs rather well over a wide U/t range as one can see in Fig. 10. In particular the satellite energies are better described than in GW. The energies calculated using the electron-hole unscreened T-matrix (not shown in the figure), instead, show divergencies. Finally, the screened pp T-matrix is overall superior to the GW and unscreened T-matrix in the selected U/t range in the left panel of Fig. 10. In the $U/t \to \infty$ limit, all approximations studied are rather poor. Interestingly if one looks at the first iteration, both

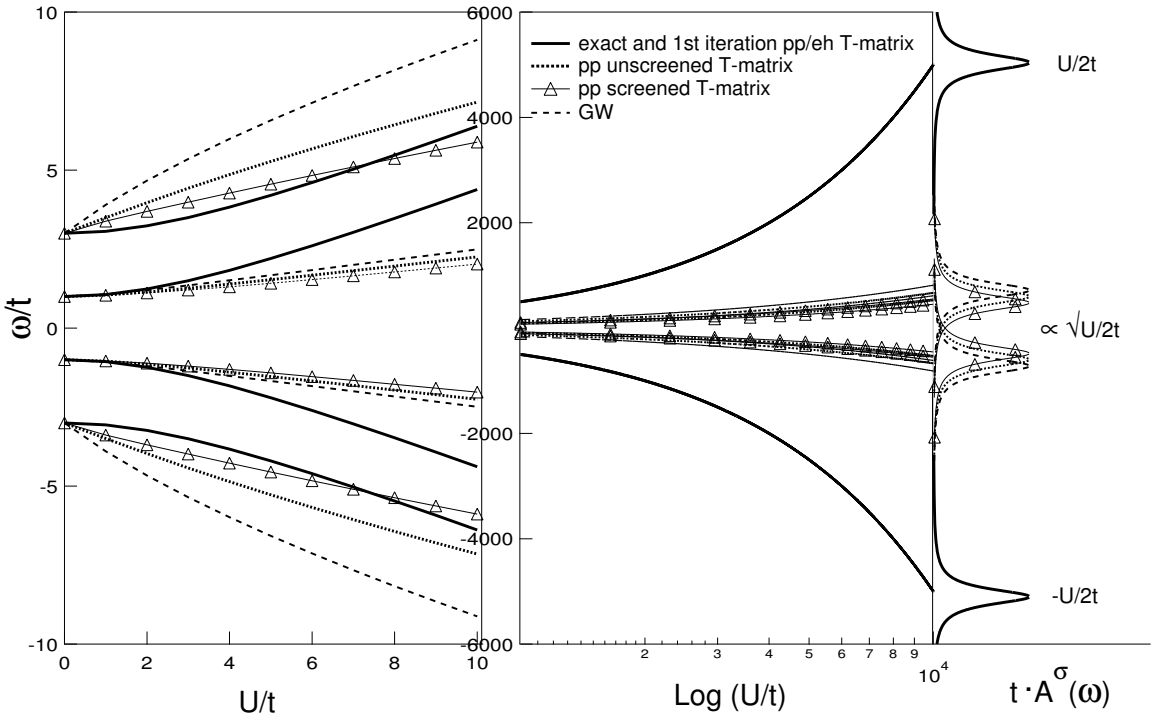

Fig. 10: *Two-site Hubbard model at half filling from Ref. [3]: comparison between the exact renormalized addition/removal energies ω/t (solid lines) as function of U/t (left panel) and Log(U/t) (right panel) and the results obtained from GW (dashes) and particle-particle (dots) and screened T-matrix (triangles). In the atomic limit the spectral function, i.e., the peak positions and weights, is illustrated on the right-hand side, upon multiplying by t and taking the $t \to 0$ limit.*

the particle-particle and electron-hole contributions to the T-matrix give the exact results for all t. Finally the spectral function calculated using the pp screened T-matrix is shown in Fig. 8 where one can better appreciate the superiority of the pp screened T-matrix approximation (at least for the description of satellite energies) with respect to the GWA at moderately strong interaction U/t (upper left panel). In the atomic limit (middle right panel), instead, the screened pp T-matrix fails just as GW.

Why is the pp T-matrix approximation exact for one electron in the atomic limit, and not for two electrons? To derive the unscreened T-matrix we used the approximation $\frac{\delta G}{\delta U_{ext}} \approx GG$. In the case of one electron this is not an approximation, but it is the exact time-ordered response, and therefore the (pp) unscreened T-matrix yields the exact result for one electron. This is not the case for two electrons for which $\frac{\delta G}{\delta U_{ext}} \approx GG$ is a rough approximation, and one needs to include some screening. The screened matrix indeed improves over the unscreened T-matrix for two electrons even with an approximate RPA screening; such approximate screening is instead dramatic for one electron and a more accurate screening is needed (as pointed out above the exact screening would yield the unscreened T-matrix and hence an exact result for one electron). *One should hence use a screened T-matrix with a screened interaction adapted to the system.*

4 Conclusions and outlook

In this chapter we discussed the importance of vertex corrections in relation to two major shortcomings of the GW approximation to the self-energy: the self-screening error and the incorrect atomic limit. Because of the self-screening, the GWA produces extra unphysical removal and addition energies in the spectrum of the Hubbard dimer with one electron. We showed that this error can be corrected by a two-point vertex $\Gamma = \delta + f_{xc}P$, derived from time-dependent density functional theory, but it should be used only for the valence state, for which it produces exact results. For the conduction state, instead, this two-point vertex produces fewer poles than the exact solution. In fact, it seems to be more reasonable to stick to the TC-TC description of the screening for the conduction state, with $\Gamma = \delta$, which yields the correct number of energies, although the values are still different from the exact ones. One could extrapolate these findings for the case of more bands in the situation where the valence bands are similar and the conduction bands are localized elsewhere or with different spin. The comparison with the exact solution for the two-site Hubbard model sheds light on another feature of the GW approximation: in the atomic limit ($t \to 0$) the GW solution for the two-site model does not reduce to the solution for two isolated sites; this is caused by the description of the electrons as an average charge distribution, instead of a probability. The approximate three-point vertex, which cures the self-screening problem and yields the correct number of poles in the one-particle Green function, is not sufficient to correct the atomic limit. To correct the atomic limit we have derived a vertex function from a screened version of the T-matrix approximation. The screened T-matrix, in which Hartree and exchange terms are treated on an equal footing, reduces to the T-matrix in the atomic limit when a dynamically screened interaction is used, which gives the exact result for the Hubbard dimer at quarter-filling; even with an approximate W it is better than GW at quarter filling, whereas at half filling it is overall superior to both the pp T-matrix and GW over a wide U/t range. This means that the vertex corrections derived from this version of the pp screened T-matrix can improve over GW. In the atomic limit at half-filling, however, the screened T-matrix performs as badly as GW. It remains hence to understand "*Why, in the atomic limit, is the pp screened T-matrix exact for one electron but not for two electrons? Which approximations we use in the derivation of the T-matrix self-energy is harmless for the one-electron case, but dramatic for the two-electron case? And why? Will realistic systems be more forgiving than the Hubbard dimer when the screened pp T-matrix approximation is used?*"

Acknowledgements

I would like to acknowledge the important contribution of my former PhD student Stefano Di Sabatino to the study of the Hubbard model using various methods and approximations. I also want to thank Lucia Reining with whom I have enjoyed working on the topics presented in this chapter as well as my collaborators Friedhelm Bechstedt and Arjan Berger for fruitful discussions and joint work.

A Solutions for 2±1 electrons

To treat the case at 1/4 and 1/2 filling we need to diagonalize the two-site Hamiltonian for one, two, and three electrons. Since the Hamiltonian conserves particle number we can diagonalize it separately for the various N sectors. The eigenstates of the system will be linear combinations of Slater determinants, which are denoted by the kets $|1\ ,\ 2\rangle$, with occupations of the sites 1 and 2 given by $\cdot$, $\uparrow$, $\downarrow$, $\updownarrow$. We choose $V_0 = 0$, so that the vacuum state $\zeta^{N=0}\rangle = |\ \cdot,\ \cdot\ \rangle$ has zero energy for the dimer. Moreover, the following convention as to fermionic ordering is used

$$|\ \updownarrow, \updownarrow\rangle = c^\dagger_{2\downarrow} c^\dagger_{2\uparrow} c^\dagger_{1\downarrow} c^\dagger_{1\uparrow} |0_1\, 0_2\rangle. \tag{43}$$

A.1 One electron

In the case of one electron, the basis vectors are $|\uparrow, \cdot\rangle$, $|\downarrow, \cdot\rangle$, $|\cdot, \uparrow\rangle$, and $|\cdot, \downarrow\rangle$. In this basis the Hamiltonian matrix reads

$$H = \begin{pmatrix} \varepsilon_0 & 0 & -t & 0 \\ 0 & \varepsilon_0 & 0 & -t \\ -t & 0 & \varepsilon_0 & 0 \\ 0 & -t & 0 & \varepsilon_0 \end{pmatrix}, \tag{44}$$

and we get the eigenvalues E_i and corresponding eigenvectors in Table 1.

E_i	$\|\uparrow, \cdot\rangle$	$\|\downarrow, \cdot\rangle$	$\|\cdot, \uparrow\rangle$	$\|\cdot, \downarrow\rangle$
$\varepsilon_0 - t$	0	$1/\sqrt{2}$	0	$1/\sqrt{2}$
$\varepsilon_0 - t$	$1/\sqrt{2}$	0	$1/\sqrt{2}$	0
$\varepsilon_0 + t$	0	$1/\sqrt{2}$	0	$-1/\sqrt{2}$
$\varepsilon_0 + t$	$1/\sqrt{2}$	0	$-1/\sqrt{2}$	0

Table 1: *Eigenvalues and corresponding eigenvectors for the one-electron sector.*

A.2 Two electrons

In the case of two electrons, the basis vectors are $|\uparrow, \downarrow\rangle$, $|\downarrow, \uparrow\rangle$, $|\uparrow, \uparrow\rangle$, $|\downarrow, \downarrow\rangle$, $|\updownarrow, \cdot\rangle$, and $|\cdot, \updownarrow\rangle$. In this basis the Hamiltonian matrix reads

$$H = \begin{pmatrix} 2\varepsilon_0 & 0 & 0 & 0 & -t & -t \\ 0 & 2\varepsilon_0 & 0 & 0 & t & t \\ 0 & 0 & 2\varepsilon_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\varepsilon_0 & 0 & 0 \\ -t & t & 0 & 0 & 2\varepsilon_0{+}U & 0 \\ -t & t & 0 & 0 & 0 & 2\varepsilon_0{+}U \end{pmatrix}, \tag{45}$$

which yields the eigenvalues E_i and corresponding eigenvectors in Table 2, where

$$c = \sqrt{16t^2{+}U^2} \quad a = \frac{\sqrt{2}}{(c{-}U)}\sqrt{16t^2 + (c{-}U)^2}\,, \quad b = \frac{\sqrt{2}}{(c{+}U)}\sqrt{16t^2 + (c{+}U)^2}\,.$$

E_i	$\vert\uparrow,\downarrow\rangle$	$\vert\downarrow,\uparrow\rangle$	$\vert\uparrow,\uparrow\rangle$	$\vert\downarrow,\downarrow\rangle$	$\vert\uparrow\downarrow,\cdot\rangle$	$\vert\cdot,\uparrow\downarrow\rangle$
$2\varepsilon_0+(U-c)/2$	$\frac{4t}{a(c-U)}$	$-\frac{4t}{a(c-U)}$	0	0	$1/a$	$1/a$
$2\varepsilon_0+(U+c)/2$	$-\frac{4t}{b(c+U)}$	$\frac{4t}{b(c+U)}$	0	0	$1/b$	$1/b$
$2\varepsilon_0+U$	0	0	0	0	$-1/\sqrt{2}$	$1/\sqrt{2}$
$2\varepsilon_0$	0	0	0	1	0	0
$2\varepsilon_0$	0	0	1	0	0	0
$2\varepsilon_0$	$1/\sqrt{2}$	$1/\sqrt{2}$	0	0	0	0

Table 2: *Eigenvalues and corresponding eigenvectors for the two-electron sector.*

The first three solutions are singlets, the last three triplets. The first singlet is the ground state. It is interesting to consider the limits $U \to 0$ and $U \to \infty$ or $t \to 0$. The expansion of the coefficients for these limits are:
$U \to 0$:

$$c \approx 4t\left(1+\frac{U^2}{32t^2}\right) \tag{46}$$

so $\frac{1}{c-U} \approx \frac{1}{4t(1+U^2/32t^2-U/4t)} \approx \frac{1}{4t}\left(1+\frac{U}{4t}-\frac{U^2}{32t^2}+\frac{U^2}{16t^2}\right) \approx \frac{1}{4t}\left(1+\frac{U}{4t}+\frac{U^2}{32t^2}\right)$ and

$$\begin{aligned} a &= \sqrt{2}\sqrt{16t^2/(c-U)^2+1} \approx \sqrt{2}\sqrt{(1+\frac{U}{4t}+\frac{U^2}{32t^2})^2+1} \approx 2\sqrt{1+U/4t+U^2/16t^2} \\ &\approx 2\left(1+U/8t+U^2/32t^2-U^2/128t^2\right) \end{aligned} \tag{47}$$

$$\begin{aligned} b &= \sqrt{2}\sqrt{16t^2/(c+U)^2+1} \approx \sqrt{2}\sqrt{(1-\frac{U}{4t}+\frac{U^2}{32t^2})^2+1} \\ &\approx 2\sqrt{1-U/4t+U^2/16t^2} \approx 2\left(1-U/8t+3U^2/128t^2\right) \end{aligned} \tag{48}$$

In the non-interacting ($U=0$) case, we have hence a weight of $1/2$ for double occupation.
For $U \to \infty$

$$c \approx U\left(1+\frac{8t^2}{U^2}\right) \tag{49}$$

$$a \approx \sqrt{2}\sqrt{U^2/4t^2+1} \approx \frac{U}{\sqrt{2}t}\left(1+2t^2/U^2\right) \tag{50}$$

$$b \approx -a \tag{51}$$

In particular we get $a(c-U) \approx \frac{U}{\sqrt{2}t}\left(1+2t^2/U^2\right)8t^2/U = \frac{8t}{\sqrt{2}}(1+2t^2/U^2)$ so that the coefficients on the single occupations are $\pm 1/\sqrt{2}$, whereas double occupancies are impossible (Heitler-London limit).
For $t \to 0$ we have the same expansion as for $U \to \infty$ since the ratio t/U is what matters; the only peculiarity is the fact that the ground state energy goes to zero, so that it approaches degeneracy with the threefold degenerate triplet states (case $t=0$). (However for any small but finite t the ground state is the singlet $S=0$).

A.3 Three electrons

The basis vectors for three electrons are $|\uparrow, \updownarrow\rangle$, $|\downarrow, \updownarrow\rangle$, $|\updownarrow, \uparrow\rangle$, and $|\updownarrow, \downarrow\rangle$, and the Hamiltonian matrix, the same as for one electron but with negative (hole) hopping

$$H = \begin{pmatrix} 3\varepsilon_0{+}U & 0 & t & 0 \\ 0 & \varepsilon_0{+}U & 0 & t \\ t & 0 & \varepsilon_0{+}U & 0 \\ 0 & t & 0 & \varepsilon_0{+}U \end{pmatrix}. \tag{52}$$

This yields eigenvalues E_i and corresponding eigenvectors in Table 3.

E_i	$\|\uparrow, \updownarrow\rangle$	$\|\downarrow, \updownarrow\rangle$	$\|\updownarrow, \uparrow\rangle$	$\|\updownarrow, \downarrow\rangle$
$3\varepsilon_0{-}t+U$	0	$-1/\sqrt{2}$	0	$1/\sqrt{2}$
$3\varepsilon_0{-}t+U$	$-1/\sqrt{2}$	0	$1/\sqrt{2}$	0
$3\varepsilon_0{+}t+U$	0	$1/\sqrt{2}$	0	$1/\sqrt{2}$
$3\varepsilon_0{+}t+U$	$1/\sqrt{2}$	0	$1/\sqrt{2}$	0

Table 3: *Eigenvalues and corresponding eigenvectors for the three-electron sector.*

References

[1] A.L. Fetter and J.D. Walecka: *Quantum Theory of Many-Particle Systems* (Benjamin, New York, 1964)

[2] L. Hedin, Phys. Rev. **139**, A796 (1965)

[3] P. Romaniello, F. Bechstedt, and L. Reining, Phys. Rev. B **85**, 155131 (2012)

[4] R.M. Martin, L. Reining, and D.M. Ceperley: *Interacting Electrons: Theory and Computational Approaches* (Cambridge University Press, 2016)

[5] G. Strinati, Riv. Nuovo Cimento **11**, 1 (1988)

[6] F. Bruneval, F. Sottile, V. Olevano, R. Del Sole, and L. Reining, Phys. Rev. Lett. **94**, 186402 (2005)

[7] M. Springer, F. Aryasetiawan, and K. Karlsson, Phys. Rev. Lett. **80**, 2389 (1998)

[8] N.E. Dahlen, R. van Leeuwen, and U. von Barth, Phys. Rev. A **73**, 012511 (2006)

[9] A. Stan, N.E. Dahlen, and R. van Leeuwen, Europhys. Lett. **76**, 298 (2006)

[10] P. Romaniello, S. Guyot, and L. Reining, J. Chem. Phys. **131**, 154111 (2009)

[11] M. Puig von Friesen, C. Verdozzi, and C.-O. Almbladh, Phys. Rev. Lett. **103**, 176404 (2009)

[12] M. Puig von Friesen, C. Verdozzi, and C.-O. Almbladh, Phys. Rev. B **82** 155108 (2010)

[13] S. Di Sabatino, J. A. Berger, L. Reining, and P. Romaniello, J. Chem. Phys. **143**, 024108 (2015)

[14] S. Di Sabatino, J. A. Berger, L. Reining, and P. Romaniello, Phys. Rev. B **94**, 155141 (2016)

[15] S. Di Sabatino, J. Koskelo, J. A. Berger, and P. Romaniello, Phys. Rev. Research **3**, 013172 (2021)

[16] W. Nelson, P. Bokes, P. Rinke, and R.W. Godby, Phys. Rev. A **75**, 032505 (2007)

[17] F. Aryasetiawan, R. Sakuma, and K. Karlsson, Phys. Rev. B **85**, 035106 (2012)

[18] M. Guzzo, PhD thesis, École Polytechnique (2013)

[19] W. Hanke and L.J. Sham, Phys. Rev. B **21**, 4656 (1980)

[20] G. Strinati, Phys. Rev. B **29**, 5718 (1984)

[21] G. Strinati, Phys. Rev. Lett. **49**, 1519 (1982)

[22] V.M. Galitskii, Zh. Eksp. Teor. Fiz. **34**, 251 (1958) [Sov. Phys. JETP **7**, 104 (1958)]

[23] L.P. Kadanoff and G. Baym: *Quantum Statistical Mechanics*
(W.A. Benjamin, New York, 1964)

[24] M. Vanzini, L. Reining, and M. Gatti, Eur. Phys. J. B **91**, 192 (2018)

[25] J. Hubbard, Proc. R. Soc. London, Ser. **A 276**, 238 (1963)

[26] J. Hubbard, Proc. R. Soc. London, Ser. **A 277**, 237 (1964)

[27] A.J. Cohen, P. Mori-Sanchez, and W.T. Yang, Science **321**, 792 (2008)

[28] G. Stefanucci and R. van Leeuwen:
Nonequilibrium Many-Body Theory of Quantum Systems: A Modern Introduction
(Cambridge University Press, 2013)