LUND UNIVERSITY
Faculty of Science

The self-screening correction to RPA and *GW* – A model and *ab initio* study

Viktor Christiansson

Thesis submitted for the degree of Master of Science
Project duration: 8 months

Supervised by Prof. Ferdi Aryasetiawan

Department of Physics
Division of Mathematical Physics
June 2019

Abstract

The self-screening error in the GW approximation is studied by applying correction schemes beyond the random-phase approximation in model and *ab inito* calculations. Two model systems in the form of Hubbard dimers, with one and two orbitals per site respectively, are considered. Both the self-screening (GW-ss) and self-polarization (GW-sp) correction schemes are compared to ordinary GW calculations as well as to the exact results. It is found that GW-ss provides a significant improvement for small values of the interaction strength U_0, almost correctly describing the linear density response function and improving the spectral function over that obtained within GW. GW-sp instead correctly describes the HOMO-LUMO gap for the one-orbital model in the limit of large U_0, and provides a better agreement also for the spectral function in that regime, while still being approximately as accurate as GW-ss in describing the response and spectral function for lower interaction strengths. It is further found that GW-sp suffers from causality issues. The *ab initio* calculations were performed using a version of the *SPEX* code modified to include the self-screening correction. The correction is applied to four semiconductors: GaAs, ZnSe, Ge, and ZnO. It is found that the self-screening correction significantly improved the band gap for GaAs and ZnSe compared to ordinary GW calculations, while it marginally overestimates the width for Ge, and improves the ZnO one slightly. Further it is found to provide a small improvement also of the energy-placement of the semicore $3d$ states. A comparison is also made regarding the differences of the two correction schemes. It is concluded that the GW-ss scheme seems to provide the best correction for delocalized systems. GW-sp instead gives the better description for large interaction strengths, and hence more localized systems, although simultaneously being able to give a good agreement for small U_0 in the model calculations.

Acknowledgements

I would first and foremost like to thank my supervisor Professor Ferdi Aryasetiawan for all your help and guidance throughout this thesis. You have always provided insightful comments and valuable remarks, and I'm grateful for having gotten this opportunity to work with you. For all of this I owe you my deepest gratitude. Furthermore I would like to extend my thanks to Dr. Fredrik Nilsson for your assistance with helping me to familiarize myself with the codes, and in general assisting me with how to perform *ab initio* calculations. Finally I would also like to thank the other members in our group; especially Erik Linnér for our discussions and your critical proofreading, and Tor Sjöstrand for your additional help with the codes.

Contents

Acronyms

ARPES	–	Angle-resolved photoemission spectroscopy
BZ	–	Brillouin zone
cRPA	–	The constrained random-phase approximation
CBM	–	Conduction band minimum
DFT	–	Density functional theory
DMFT	–	Dynamical mean-field theory
DOS	–	Density of states
FLAPW	–	Full-potential linearized augmented-plane-wave
GWA	–	The GW approximation
$\mathbf{G_0W_0}$	–	One-shot GW
GW**-ss**	–	Self-screening corrected GW
GW**-sp**	–	Self-polarization corrected GW
HFA	–	The Hartree-Fock approximation
IBZ	–	Irreducible Brillouin zone
IPES	–	Inverse photoemission spectroscopy
IR	–	Interstitial region
LDA	–	Local density approximation
MPB	–	Mixed product basis
MT	–	Muffin-tin
RPA	–	The random-phase approximation
VBM	–	Valence band maximum

Section 1

Introduction

The calculation of the electronic structure of materials has long been a major area within condensed matter physics. Although the equation governing the behaviour of the electrons is known exactly, the Schrödinger equation, solving it efficiently and accurately has been at the center of attention in much work; the issue is the Coulomb term in the many-electron Hamiltonian, quickly rendering exact calculations of larger systems unfeasible. In order to simplify the problem, various approximations have been devised, such as the simple Hartree approximation, where the electrons are subjected to a potential originating from the average field of all the electrons, and the Hartree-Fock approximation (HFA), where the self-interaction in the Hartree approximation is removed. Another method is the local density approximation (LDA) [1] within density functional theory (DFT) [1, 2], where the potential beyond Hartree is approximated using the electron gas.

LDA has provided good results for various properties and is widely used today for a wide variety of applications. However, it is well-known that, for example, the description of band gaps in semiconductors is lacking [3], and it has also been found to place the energy of semicore states too high [4, 5]. In order to improve these, and other issues, theories beyond DFT are used. Green's function theory (see e.g. Refs. [6, 7]) is one such method useful to compute, e.g., the excitation energies of the system and expectation values of ground state properties. The self-energy Σ, which is a non-local and energy dependent potential, is used to incorporate interaction effects beyond the Hartree approximation. Although Σ in principle could be calculated exactly, for more complicated systems it is not feasible in practice: one must hence resort to approximations.

The GW approximation (GWA) [8], proposed by Hedin in 1965, is one of the main methods used today to accomplish this. Based on many-body perturbation theory using the Green's function formalism, Hedin showed that the self-energy and related quantities could be expanded in terms of the screened interaction, W, providing an exact set of self-consistency equations. Within GWA, the self-energy is schematically calculated using the lowest order term in the aforementioned expansion as $\Sigma = iGW$, where G is the Green's function, from which the name stems. GW is calculated within the random-phase approximation (RPA) [9, 10] in order to obtain the polarization and response functions of the system used in the computation of W.

GW is known to rectify the band gap problem of LDA very well (see e.g. Refs. [3,11]) and to in general further improve the band structure over LDA when comparing to experiments. As another example of its success, but also the limitations of GW, it is known to improve the energy of semicore $3d$ states in important semiconductors such as GaAs and Ge over the values calculated using LDA, yet it still places them too high in energy as compared to experiments [4,5]. Although in many respects improving the description of the electronic structure, GWA still suffers from a number of deficiencies, besides the remaining discrepancy in the semicore energies and band gap. In order to further improve it different vertex corrections can be included; however, this becomes difficult to do in practice. Firstly, including arbitrary higher

order terms for the self-energy in the expansion in W is not guaranteed to yield an improved result [12], and one therefore has to choose which to include carefully, and secondly these computations tend to be costly for higher order corrections, limiting their applicability for real systems. Nevertheless, fully summing up all the terms would provide the entirety of Hedin's equations, which are exact. For other types of problems, such as for for strongly correlated systems where GW is known to not provide an accurate description [12,13], different methods can be used instead to obtain better results. The poor description for strongly correlated systems which GW displays can be seen from its origin in a perturbation expansion in W. Strong onsite interactions, due to localized electrons, will therefore result in an incorrect picture, which is the case for these systems. For these materials instead dynamical mean-field theory (DMFT) (for a review see Ref. [14]) can be used, where the system under consideration is mapped to an impurity problem from which the Green's function can be calculated using a local self-energy. Another possibility would be to combine the two approaches, as done in GW+DMFT [15].

Another issue with GW, which has been the focus of this work, is self-screening. It is a type of self-interaction where an electron screens itself through the screened interaction W, due to it being calculated within RPA. The problem is readily seen for the hydrogen atom, where the correlation part of the self-energy becomes non-zero, causing the quasiparticle energy to be too high compared to the exact result [16]. Within the HFA, however, the correct result is obtained as it cancels the self-interaction introduced by the Hartree potential through the inclusion of exchange, while not including the correlation part of the self-energy from where the problem stems. The self-screening error has also been studied for models [17, 18], and it has furthermore been suggested that the self-screening could be responsible for the aforementioned error in the localized semicore states [19].

Two schemes to remedy the inherent self-screening error in RPA and GWA were proposed by Aryasetiawan et. al. in 2012 [19]. They were applied to a simple hydrogen dimer model, improving the agreement with the exact result when compared to GW for small interaction strengths. One of the schemes is based on that the Green's function can be divided into orbital components g_m, whereupon one can note that a given orbital is involved in the screening of itself in $\Sigma = \sum_m g_m W$ through its inclusion in the screened interaction. Physically this means that, as only one electron with a given spin can occupy an orbital, the electron will participate in screening itself while propagating, which clearly is unphysical. The second scheme is based on a similar principle, but instead from the point of view of the excitations in the system. Another approach to correct the self-screening error has recently been proposed by Wetherell et. al. [17], where a local density functional is used to derive a local potential which is added to the GW self-energy in order to remove the self-screening, and vertex corrections have also been devised to attempt to eliminate the error [18].

The purpose of the present work has been to further investigate the correction schemes for the self-screening error proposed in [19]. This was done in two ways: first by in more depth exploring model systems with variable parameters in the form of Hubbard dimers with one and two orbitals per site, respectively, and secondly by using *ab initio*, or first principles, calculations. The models were examined to see how well the schemes worked in different parameter ranges. These parameters were subsequently related to different physical regimes to see where the corrections provided improved (or worsened) the agreement with exact results in order to explore their usability. For example the effect on the density response function,

giving the response of the system to an external perturbation, and the renormalized Green's function was investigated for both models. The *ab initio* part was carried out by using a version of the *GW* code *SPEX* [11, 20, 26], which was modified to include the self-screening correction. The code was applied to four semiconductors, GaAs, ZnSe, Ge, and ZnO, and the results were compared to the band structures calculated within LDA and *GW*, and to experimental values, to ascertain the effect of the scheme. Specifically, the effects on the energy of the semicore states and the band gap were investigated.

The work is structured as follows. In Section 2 the necessary theory and theoretical background is presented. The schemes in [19] will be the basis for the work and are therefore explained in detail in Sections 2.7 - 2.8, preceded by a general introduction to the formalism and GWA. The model calculations using the proposed schemes will be carried out in a one-shot manner, that is only doing the first iteration of the self-consistency, in Section 3, and compared to the usual *GW* calculation within RPA as well as the exact solution of the systems. The investigation of real materials using the *ab initio* calculations of the electronic structure in section is detailed in Section 4, with the *SPEX* code and its modifications being described in Sections 4.1 - 4.2, and a description of the materials being investigated in Section 4.3. The results are presented and analyzed in Section 5 and the conclusions from the study are presented in Section 6. Finally, in Section 7 an outlook on potential future directions of work is explored.

Section 2

Theory

2.1 The many-electron Hamiltonian

In many-electron systems the term causing difficulties in solving the system exactly arises from the electron-electron interaction, $v(\mathbf{r}-\mathbf{r}') = 1/|\mathbf{r}-\mathbf{r}'|$, in the Hamiltonian

$$H = \sum_i -\frac{1}{2}\nabla_i^2 + V(\mathbf{r_i}) + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r_i}-\mathbf{r_j}|}. \tag{2.1}$$

H is here in Hartree atomic units, which will be used throughout this work, where $\hbar = m_e = e = 1/4\pi\epsilon_0 = 1$. In $V(\mathbf{r})$ the external potential and the electron-nucleon interaction are included. The electron-electron interaction becomes difficult to treat already for small systems, and increasingly so for calculations of the electronic structure of real materials. The issue is that the Schrödinger equation $H\Psi_n = E_n\Psi_n$ becomes ever more complicated to solve as the number of electrons in the system grows. As a consequence, many approximations of various types have been devised in order to obtain the energies and states in a simplified manner.

The aforementioned Hartree approximation is a simple mean-field approximation of the electron-electron interaction as a one-particle effective potential,

$$V_H(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{2.2}$$

using the electron density ρ. It originates from the idea that the electrons should feel the interaction from an average field from all of the electrons, instead of from each electron individually. A step beyond Hartree is the Hartree-Fock approximation (HFA) which removes the issue of self-interaction which is present in the Hartree potential, equation (2.2). It does this by including an exchange term besides the Hartree potential, resulting in the Hartree-Fock equation

$$\left(-\frac{1}{2}\nabla^2 + V(\mathbf{r}) + V_H(\mathbf{r})\right)\phi_{n\sigma}(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\Sigma_\sigma^x(\mathbf{r},\mathbf{r}')\phi_{n\sigma}(\mathbf{r}') = \varepsilon^{\mathrm{HF}}\phi_{n\sigma}(\mathbf{r}), \tag{2.3}$$

for a state $\phi_{n\sigma}$ with spin σ. The exchange potential is given by

$$\Sigma_\sigma^x(\mathbf{r},\mathbf{r}') = -v(\mathbf{r}-\mathbf{r}')\sum_m^{\mathrm{occ}}\phi_{m\sigma}(\mathbf{r})\phi_{m\sigma}^*(\mathbf{r}'). \tag{2.4}$$

Another, more rigorous, approach to approximate the problem is by going beyond the mean-field methods described above. One way is provided by the Green's function theory, described in Section 2.3, which will be the main framework utilized in this work.

2.2 The Hubbard Hamiltonian

The model calculations will be performed on a Hubbard dimer with one and two orbitals per site, respectively. That is, two sites with a Hamiltonian due to Hubbard [21] using the occupation number formalism:

$$H = \sum_{i,j,\sigma} h_{ij} c^\dagger_{i\sigma} c_{j\sigma} + \sum_{i,j} U_{ii,jj} n_{i\uparrow} n_{j\downarrow}. \tag{2.5}$$

Here $c_{j\sigma}$ is the annihilation operator, removing an electron of spin σ occupying orbital j, and $c^\dagger_{i\sigma}$ instead creates an electron, implying the first term to represent hopping in-between orbitals. The h_{ij} are the variable parameters governing the strength of these effects. Further the number operator $n_{i\sigma} = c^\dagger_{i\sigma} c_{i\sigma}$ counts the number of electrons present in orbital i, and the second term thus represents the interaction between the electrons, with $U_{ii,jj}$ providing its strength.

Two Hubbard models in the form of dimers (two sites) are used in this work: one with a single orbital per site and one with two orbitals per site. In the one-orbital model, both on-site and off-site Coulomb interaction is included. We thus set $U^{1\text{ orb}}_{11,11} = U^{1\text{ orb}}_{22,22} = U_0$ and $U^{1\text{ orb}}_{11,22} = U^{1\text{ orb}}_{22,11} = U_1$ here, as the two sites are interchangeable by symmetry. The Hamiltonian can then be written as

$$H^{1\text{ orb}} = \epsilon_0 \sum_{i\sigma} c^\dagger_{i\sigma} c_{i\sigma} + t \sum_{i\neq j,\sigma} c^\dagger_{i\sigma} c_{j\sigma} + U_0 \sum_i n_{i\uparrow} n_{i\downarrow} + U_1 \sum_{i\neq j} n_{i\uparrow} n_{j\downarrow}. \tag{2.6}$$

with $t = h_{ij}$ $(i \neq j)$ being the hopping parameter, and $\epsilon_0 = h_{ii}$ is the one-particle energy of the orbitals. For the two-orbital case, only the on-site interaction is taken into account in the two lower orbitals, $U_{1111} = U_{2222} = U_0$, yielding a Hamiltonian of the form

$$H^{2\text{ orb}} = \sum_{i\alpha} \epsilon_{i\alpha} c^\dagger_{i\alpha\sigma} c_{i\alpha\sigma} + \sum_{i\neq j} \sum_{\alpha\beta} \sum_\sigma t_{(i,\alpha)(j,\beta)} c^\dagger_{i\alpha\sigma} c_{j\beta\sigma} + U_0 \sum_{i\in\{1,2\}} n_{i(\alpha=1)\uparrow} n_{i(\alpha=1)\downarrow} \tag{2.7}$$

where $\epsilon_{i\alpha}$ is ϵ_0 for the two lower orbitals, $\alpha = 1$ on sites $i = 1,2$, and ϵ_1 for the two upper ones, $\alpha = 2$ on sites $i = 1,2$. The hopping terms included in this model are $t_{(1,1)(2,1)} = t_1$, $t_{(1,1)(2,2)} = t_{(1,2)(2,1)} = t_2$ and $t_{(1,2)(2,2)} = t_3$, where a reordering of (i,α) provides the same hopping strength due to symmetry. The two lower states can physically be interpreted as localized valence states, where the interaction occurs, and the two higher lying states as a delocalized conduction band, hence the lack of interaction. For a schematic view of the model systems see Figure 2.1.

2.3 The Green's function in many-body theory

In this section the theory for the Green's function (see e.g. Refs. [6, 7]) used in many-body theory at zero temperature will be described. The electron Green's function is defined as

$$G(1,2) = -i \left\langle \Psi_0 \right| T\hat{\psi}(1)\hat{\psi}^\dagger(2) \left| \Psi_0 \right\rangle \tag{2.8}$$

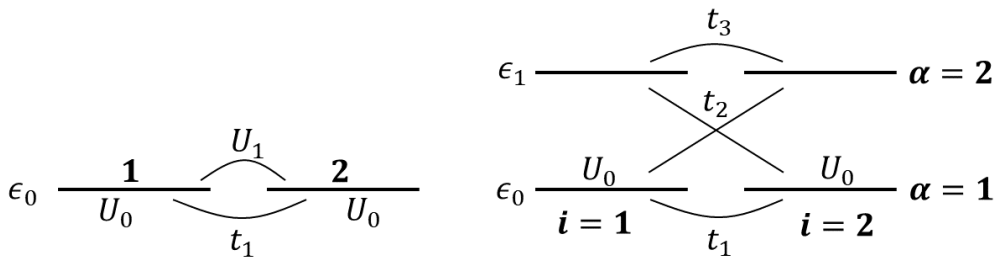

Figure 2.1: The two model Hubbard dimers used in this work, with the parameters corresponding to hopping (t_i), interaction (U_i) and orbital energies (ϵ_i) marked out for the different sites. Left: One orbital per site. Right: Two orbitals per site

where T is the time-ordering operator moving operators at later times to the left, with a possible sign change if two operators are interchanged due to their fermionic nature, and Ψ_0 is the ground state of the system. As $\psi^+(2)$ creates an electron at $(2) = (r_2,t_2) = (\mathbf{r}_2,\sigma_2,t_2)$ and $\psi(1)$ annihilates it at (r_1,t_1), the Green's function describes two processes; for $t_1 > t_2$ an electron is created at position r_2 at time t_2 and propagates to r_1 at time t_1 where it is annihilated, with the Green's function providing the probability amplitude for the propagation. Similarly for $t_1 < t_2$ an electron is annihilated at r_1 at t_1 and subsequently created again at the later time t_2 at r_2, or equivalently, a hole is created at (1) propagating to (2).

From the Green's function it is possible to obtain the total energy and the expectation value of any one-particle operator in the ground state as well as the one-particle excitation energies. We will limit the exposition of these properties to the excitation energies, as only these will be investigated in the current work. By using a time-independent Hamiltonian, the Green's function will only be dependent on the difference in time $t_1 - t_2$, and not their absolute values. $G(r_1,r_2;t_1 - t_2)$ can therefore be Fourier transformed with respect to the difference in time, and it can be written in a spectral representation as [6]:

$$\begin{aligned} G(r_1,r_2;\omega) = &\sum_{n(N-1)} \frac{\langle\Psi_0(N)|\,\hat{\psi}^\dagger(r_2)\,|\Psi_n(N-1)\rangle\,\langle\Psi_n(N-1)|\,\hat{\psi}(r_1)\,|\Psi_0(N)\rangle}{\omega + E_n(N-1) - E_0(N) - i\delta} \\ &+ \sum_{n(N+1)} \frac{\langle\Psi_0(N)|\,\hat{\psi}(r_1)\,|\Psi_n(N+1)\rangle\,\langle\Psi_n(N+1)|\,\hat{\psi}^\dagger(r_2)\,|\Psi_0(N)\rangle}{\omega - E_n(N+1) + E_0(N) + i\delta} \end{aligned} \tag{2.9}$$

$\Psi_n(N)$ is the n'th state of the N-particle system with energy $E_n(N)$. It is then clear that the denominators of G provide the exact excitation spectra, through its pole structure, for both of the $(N \pm 1)$-electron systems. It is in this form directly related to photoemission experiments, such as angle-resolved photoemission spectroscopy (ARPES) or inverse photoemission spectroscopy (IPES), where electrons are emitted from and absorbed in the material respectively. In ARPES the electrons are photoemitted by incident photons with energy ω, and the kinetic energy E_{kin} and momentum $\mathbf{k}$ of the outgoing electrons are measured, providing the excitation energies $E_{exc}(\mathbf{k})$ in the $(N-1)$-electron system [12]

$$E_{exc}(\mathbf{k}) = \omega - E_{kin}. \tag{2.10}$$

In IPES, instead electrons with known momentum and kinetic energy are absorbed by the material and an outgoing photon is measured providing the excitation energy of the $(N+1)$-electron system, similarly to equation (2.10).

The spectral function can be written as

$$A(r,r';\omega) = -\frac{1}{\pi}\,\mathrm{Im}\,G(r,r';\omega)\mathrm{sgn}(\omega-\mu) \tag{2.11}$$

where μ for an infinitely large system is the chemical potential. In calculations it can be obtained by integrating the trace of the spectral function as

$$N = \int_{-\infty}^{\mu} \mathrm{d}\omega Tr(A(\omega)). \tag{2.12}$$

with N being the number of particles in the system. The Green's function can subsequently be regained by a transformation [3]

$$G(r,r';\omega) = \int_{-\infty}^{\mu} \mathrm{d}\omega' \frac{A(r,r';\omega')}{\omega-\omega'-i\delta} + \int_{\mu}^{\infty} \mathrm{d}\omega' \frac{A(r,r';\omega')}{\omega-\omega'+i\delta}. \tag{2.13}$$

The peak structure of A, as for the poles of G, provides the excitation spectrum for both the $(N\pm1)$-electron systems.

Another equivalent way to rewrite G in equation (2.9) comes from finding the solution to the Fourier transform of the equation of motion for the Green's function.

$$(\omega - H_0)G(r,r';\omega) - \int \mathrm{d}r_1 \Sigma(r,r_1;\omega)G(r_1,r';\omega) = \delta(r-r'). \tag{2.14}$$

where H_0 is the one-particle Hamiltonian including the Hartree potential. The self-energy Σ, which is non-local and energy-dependent, thus incorporates the part of the interaction beyond the Hartree mean-field: both the exchange obtained in HFA and additional correlation effects. The solution is in general given by the Green's function [22]

$$G(r,r';\omega) = \sum_j \frac{\Psi_j(r,\omega)\Psi_j^\dagger(r',\omega)}{\omega - E_n(\omega)} \tag{2.15}$$

with the complex eigenvalues E_j and wavefunctions Ψ_j given by

$$H_0(r)\Psi_j(r,\omega) + \int \mathrm{d}r_1 \Sigma(r,r_1;\omega)\Psi_j(r_1,\omega) = E_j(\omega)\Psi_j(r,\omega). \tag{2.16}$$

As the definitions in equations (2.9) and (2.15) must be the same, the $\mathrm{Re}(E_j)$ provide the excitation energies E of the system, as it gives a pole when $\mathrm{Re}(E_j(\omega_j)) = \omega_j$. For a large system they are also referred to as quasiparticle energies, where the wavefunction of the quasiparticle is defined as $\Psi_j(r, \mathrm{Re}(E_j(\omega_j)))$, and the imaginary part provides the inverse lifetime [12]. Furthermore, equation (2.16) is known as the quasiparticle equation for $\omega = \omega_j$, as it then provides the quasiparticles in the system. By solving the equation for $\Sigma = 0$ the result is a non-interacting Green's function

$$G^0(r,r';\omega) = \sum_n \frac{\varphi_n(r)\varphi_n^*(r')}{\omega - \varepsilon_n} \tag{2.17}$$

with φ_n and ε_n being the non-interacting wavefunctions and energies respectively. G^0 will be used extensively in the following sections, as the full G becomes impractical for realistic calculations.

2.4 The linear density response function

How a system responds to an external perturbation is an important property to know in order to describe its electronic structure. The response function, R, [6,7] gives the change in the density ρ due to the application of an external field φ according to

$$R(1,2) = \frac{\delta\rho(1)}{\delta\varphi(2)}. \tag{2.18}$$

The spectral representation of the time-ordered linear density response function can be shown to take the form

$$R(r,r';\omega) = \sum_{n\neq 0}^{exc} \left[\frac{\langle\Psi|\,\hat{\rho}(r)\,|n\rangle\,\langle n|\,\hat{\rho}(r')\,|\Psi\rangle}{\omega - E_n + E_0 + i\delta} - \frac{\langle\Psi|\,\hat{\rho}(r')\,|n\rangle\,\langle n|\,\hat{\rho}(r)\,|\Psi\rangle}{\omega + E_n - E_0 - i\delta} \right]. \tag{2.19}$$

with the density operator being $\hat{\rho}(r) = \hat{\psi}^\dagger(r)\hat{\psi}(r) = \sum_{ij} c_i^\dagger c_j \phi_i^*(r)\phi_j(r)$. $|\Psi\rangle$ is the ground state with energy E_0 and the sum is limited to run only over the excited states $|n\rangle$. As such, it can readily be seen that the denominator of (2.19) gives the exact excitation energies of the system. The spectral function of the response, S, given by

$$S(r,r';\omega) = -\frac{1}{\pi}\,\mathrm{Im}\,R(r,r';\omega)\mathrm{sgn}(\omega) \tag{2.20}$$

therefore exhibits a peak structure providing the excitation spectrum.

2.4.1 The Random-Phase Approximation

Within the random-phase approximation (RPA) [9, 10], the response of the system to an external perturbation φ is considered to be the response of a non-interacting system to the perturbation as well as the Hartree potential induced by it. Schematically this can be written using the non-interacting response, P^0, to the total field $V = \varphi + V_H$ which will be further explained in Section 2.6, as:

$$\delta\rho = R^{\mathrm{RPA}}\delta\varphi = P^0(\delta\varphi + \delta V_H) \tag{2.21}$$

Furthermore, the Hartree potential changes due to the induced charge density according to $\delta V_H = v\delta\rho$. Equation (2.21) can then be rewritten as

$$R^{\mathrm{RPA}}\delta\varphi = (P^0 + P^0 v R^{\mathrm{RPA}})\delta\varphi \tag{2.22}$$

As stated before the response is a system dependent property, and since the external field φ was chosen arbitrarily it can be dropped and the RPA equation takes the form

$$R^{\mathrm{RPA}}(r,r';\omega) = P^0(r,r';\omega) + \int \mathrm{d}r_1\,\mathrm{d}r_2 P^0(r,r_1;\omega)v(r_1 - r_2)R^{\mathrm{RPA}}(r_2,r';\omega). \tag{2.23}$$

How this is calculated in practice is left for Section 2.6.

2.5 Hedin's equations

In 1965 Lars Hedin derived an expansion of the self-energy within many-body perturbation theory where Σ was expanded in a screened interaction W, instead of in the bare interaction v which was known to produce incorrect results in metals [8]. The resulting equations, known as Hedin's equations, are an exact set of self-consistency equations which since their inception have been used extensively in electronic structure calculations. The equations are derived from the equation of motion of the Green's function, however they will only be presented here and the derivations can be found in Ref. [8].

In the first of Hedin's equations the self-energy is expanded as

$$\Sigma(1,2) = i \int \mathrm{d}3\,\mathrm{d}4 W(1,3) G(1,4) \Lambda(4,2,3) \tag{2.24}$$

where

$$\Lambda(1,2,3) = -\frac{\delta G^{-1}(1,2)}{\delta V(3)} = \delta(1-2)\delta(1-3) + \frac{\delta \Sigma(1,2)}{\delta V(3)} \tag{2.25}$$

is an introduced vertex function, with the functional derivative being with respect to the total field V. Λ can in turn also be written as an integral equation, involving the Green's function and the self-energy, as the second of Hedin's equations

$$\Lambda(1,2,3) = \delta(1-2)\delta(1-3) + \int \mathrm{d}4\,\mathrm{d}5\,\mathrm{d}6\,\mathrm{d}7 \frac{\delta \Sigma(1,2)}{\delta G(4,5)} G(4,6) G(7,5) \Lambda(6,7,3). \tag{2.26}$$

Schematically, the screened interaction is related to the bare interaction v through the dielectric function, ϵ, as $W = \epsilon^{-1} v$. Formally it can be shown to be expressible in terms of the polarization function P in the third of Hedin's equations as

$$W(1,2) = v(1,2) + \int \mathrm{d}3\,\mathrm{d}4 v(1-3) P(3,4) W(4,2). \tag{2.27}$$

The polarization function P, introduced for the non-interacting case in Section 2.4, gives the response of the density to the total field and is related to the Green's function through the fourth of Hedin's equations:

$$P(1,2) = -i \int \mathrm{d}3\,\mathrm{d}4 G(1,3) \Lambda(3,4,2) G(4,1^+). \tag{2.28}$$

Here the 1^+ is used to denote that the time is evaluated infinitesimally later than t_1. Another equivalent way to compute the screened interaction is in terms of the response function in Section 2.4. Equation (2.27) can be rewritten as

$$W(1,2) = v(1,2) + \int \mathrm{d}3\,\mathrm{d}4 v(1-3) R(4,3) v(4,2). \tag{2.29}$$

with the help of which the aforementioned inverse of the dielectric function, effectively screening the bare interaction in $W = \epsilon^{-1} v$, can be written schematically as

$$\epsilon^{-1} = 1 + vR. \tag{2.30}$$

R can in turn be obtained through an integral equation involving the polarization

$$R = P(1,2) + \int \mathrm{d}3\,\mathrm{d}4 P(1,3)v(3-4)R(4,2), \tag{2.31}$$

which can be seen to take the same form as the RPA equation in equation (2.23), only differing in the polarization function used. Finally, the fifth of Hedin's equations (also known as the Dyson equation) provides the Green's function itself and can be related to the non-interacting G^0 by

$$G(1,2) = G_0(1,2) + \int \mathrm{d}3\,\mathrm{d}4 G_0(1,3)\Sigma(3,4)G(4,2). \tag{2.32}$$

Further, the spectral function A can now be rewritten as [22]

$$A(\omega) = \frac{1}{\pi}\sum_n \frac{|Im\Sigma_n(\omega')|}{|\omega' + \Delta - \varepsilon - \mathrm{Re}(\Sigma_n(\omega'))|^2 + |\,\mathrm{Im}(\Sigma_n(\omega'))|^2} \tag{2.33}$$

to take proper care of the shift in Fermi level $\Delta = E_f - \epsilon_f$ when going from the non-interacting G^0, with ϵ_f and eigenenergies ε_n, to the system described by G, with fermi level E_f. Here the n index means the matrix element of the object with respect to the non-interacting wavefunctions in the G^0 system.

Albeit that the five integral equations equations (2.24), (2.26) – (2.28) and (2.32), collectively known as Hedin's equations, are supposed to be solved self-consistently, this becomes impractical for realistic systems due to the large numerical cost associated. Instead approximations have to be employed to simplify the problem and the GW approximation, the one being used in this work, will be described in Section 2.6.

2.6 The GW approximation

The GW approximation (GWA) [8,12], used to simplify the calculation of the self-energy, was proposed in 1965 by Hedin and has since then become a widely used method to calculate, e.g., the electronic structure of materials. The approximation is based on neglecting the correction $\delta\Sigma/\delta V$ to the vertex function when calculating the self-energy in equation (2.24), which yields a vertex function of the form $\Lambda(1,2,3) = \delta(1-2)\delta(2-3)$. The resulting self-energy expression then takes the form,

$$\Sigma(1,2) = iG(1,2)W(1,2). \tag{2.34}$$

It is diagrammatically shown in Figure 2.2 and describes the propagation of the particle through the propagator G subjected to the screened interaction W. After a Fourier trans-

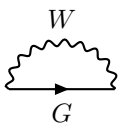

Figure 2.2: Diagram for the self-energy calculated within GW.

formation of the time coordinates we obtain the energy-dependent and non-local self-energy as

$$\Sigma(r,r';\omega) = \frac{i}{2\pi}\int \mathrm{d}\omega' G(r,r';\omega+\omega')W(r,r';\omega') \tag{2.35}$$

which can be used in the quasiparticle equation from Green's function theory, equation (2.16), in order to obtain the quasiparticle energies, E_i, of the system.

It should be noted that if the screened interaction W in equation (2.34) is replaced by the bare interaction v the HFA is recovered. Further, noting that the screened interaction can be written as $W = v + W^c$, where W^c is the correlation part of the screened interaction, the total self-energy can be divided as

$$\Sigma(r,r';\omega) = \Sigma^x(r,r') + \Sigma^c(r,r';\omega). \tag{2.36}$$

This implies that the GWA includes the exchange part from the HFA, Σ^x_σ, and additionally the effects of screening are taken into account in Σ^c, the correlation part of the self-energy coming from W^c:

$$\Sigma^c_\sigma(r,r';\omega) = \frac{i}{2\pi}\int \mathrm{d}\omega' G(r,r';\omega+\omega')W^c(r,r';\omega'). \tag{2.37}$$

GWA is calculated within RPA, described in Section 2.4.1, and the response function R is hence obtained from the RPA equation, (2.23). The screened interaction can then be obtained from the response function as $W = v + vRv$, formally using the Hedin's equations (2.29), where the response has been replaced by the RPA one. The polarization propagator from equation (2.28) is obtained within RPA by also neglecting the vertex corrections here, identically as for Σ, and by using a non-interacting G_0, yielding the expression

$$P^0(r,r';\omega) = -i\int \frac{\mathrm{d}\omega'}{2\pi}\sum_\sigma G^0_\sigma(r,r';\omega+\omega')G^0_\sigma(r',r;\omega'). \tag{2.38}$$

where the G^0_σ is now explicitly defined for each spin. The self-energy is similarly obtained by replacing the Green's function with the non-interacting one, where the exchange part can be written as

$$\Sigma^x_\sigma(r,r') = -\sum_m^{\mathrm{occ}} \varphi_m(r)\varphi^*_m(r')v(r-r'), \tag{2.39}$$

and the Dyson equation, (2.32), can subsequently be used to calculate the renormalized Green's function as $G = G^0 + G^0\Sigma G$.

Although Hedin's equations should be solved self-consistently, by updating G and calculating a new self-energy until the result is converged, this is seldom done in GW calculations. Instead, a one-shot GW (G_0W_0) approach is usually used, where the screened interaction is calculated only once from the non-interacting G^0. In practical applications the input into a GW calculation is commonly obtained from a preceding mean-field or DFT calculation, and it is well known that, unless fully self-consistent, the result will depend on the initial values. However, one reason that the calculations commonly are carried out in this way is that the full self-consistent scheme is very numerically heavy, and thus too demanding to be fully employed. Another reason for justifying that G_0W_0 is used instead of a fully self-consistent

approach is that self-consistent GW is known to worsen the result for a number of properties, such as a worsening bandwidth, over those computed using a one-shot approach [12, 23].

As mentioned in the introduction, GWA has been successful at describing many properties of various systems, including almost correctly predicting the correct band gap and improving the description of high lying core states, known as semicore states, in semiconductors. Despite the wide applicability and use, there are still several shortcomings which GWA suffers from, as was previously discussed. The approach to correct them is by including corrections to the simple vertex function used within GW, and, as previously noted, if the entire $\delta\Sigma/\delta V$ is utilized the equations become exact. However these corrections, in the form of including additional diagrams, are difficult objects, and further it is not certain that including arbitrary diagrams will necessarily improve the results, implying care has to be taken when deciding which diagrams to include [12].

One of the problems GW faces is the inherent issue originating from self-screening. This can directly be seen in the case of the hydrogen atom, as was done by Nelson et. al. in [16]. The eigenenergy for the $1s$ state using the HFA becomes the correct result, as it cancels the self-interaction introduced by the Hartree term, while GW gives an erroneous quasiparticle energy. Even when the exact wavefunctions and energies are used as input the error persists, and the error is thus not from the initial data, but instead due to the approximation itself. The error is found to originate from the correlation part of the self-energy Σ^c, which schematically is calculated as

$$\Sigma^c = iGW^c = iG(W - v). \tag{2.40}$$

Looking at the hydrogen atom there is only one electron present, and therefore no other electron able to screen it. As such the interaction it experiences should be the bare interaction, that is $W = v$, and hence $W^c = W - v = 0$, implying Σ^c to be zero. However, when calculating within GW the correlation part of the self-energy was found to be $\Sigma^c \neq 0$, W^c must therefore also be non-zero. The only possibility for this to be the case is if the only electron has screened itself, clearly illustrating the issue.

2.7 Self-screening correction for the GW approximation

The self-screening error can be seen to arise directly in the equations for the calculation of the self-energy within GW. The non-interacting G^0 can be written as [12, 19]

$$G^0_\sigma(r,r';\omega) = \sum_m g_{m\sigma}(r,r';\omega) = \sum_m \frac{\varphi_{m\sigma}(r)\varphi^*_{m\sigma}(r')}{\omega - \varepsilon_{m\sigma}} \tag{2.41}$$

where $\varepsilon_{m\sigma}$ is the eigenenergy corresponding to the state $\varphi_{m\sigma}$ of the non-interacting Hamiltonian H^0, and a finite broadening δ is used as $\varepsilon_{n\sigma} \to \varepsilon_{n\sigma} \pm i\delta$ with $(+)$ for an occupied and $(-)$ for an unoccupied state. Using this non-interacting G^0 the self-energy in equation (2.34) can be rewritten as

$$\Sigma_\sigma(1,2) = i\sum_m g_{m\sigma}(1,2)W(1,2). \tag{2.42}$$

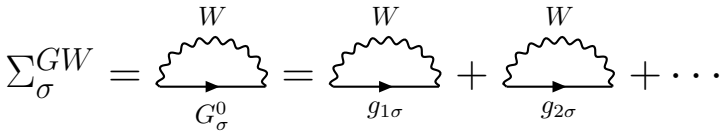

Figure 2.3: Diagram for the self-energy calculated within ordinary GW, with the individual terms shown explicitly.

for which the diagram is shown in Figure 2.3 similarly to the one in Figure 2.2, but now with the individual orbital Green's functions $g_{m\sigma}$ being used as the propagators.

The self-screening correction for GW proposed in [19] consists of identifying that a given $g_{m\sigma}$, describing an electron propagating as illustrated in Figure 2.3, should not be involved in the screening process of itself, through its inclusion in W. This is the case because another electron with identical spin cannot occupy that same orbital. However, in equation (2.42) this is indeed the case as W is calculated within RPA from the polarization propagator $P^0 = -i\sum_\sigma G_\sigma^0 G_\sigma^0$ which involves each of the $g_{m\sigma}$ lines in the calculation. The Σ is hence calculated in GW with electrons screening themselves.

To remove the self-screening, a unique screened interaction can be calculated for each of the $m\sigma$ in equation (2.42) by going beyond RPA in the calculation of W. The correction consists of calculating unique $W_{m\sigma}$ for each $m\sigma$ where the given $g_{m\sigma}$ has been omitted in the calculations. This is done by setting

$$G_{m\sigma}^0 = G_\sigma^0 - g_{m\sigma} \tag{2.43}$$

from which the polarization function can be calculated as

$$P_{m\sigma}^0(r,r';\omega) = -\frac{i}{2\pi}\int \mathrm{d}\omega' \left[G_{m\sigma}^0(r,r';\omega+\omega')G_{m\sigma}^0(r',r;\omega') + G_{-\sigma}^0(r,r';\omega+\omega')G_{-\sigma}^0(r',r;\omega')\right]. \tag{2.44}$$

The full Green's function of the opposite spin is used here as both spins are allowed in the same orbital, and thus $g_{m-\sigma}$ should be included in the screening process. Each of the polarization functions $P_{m\sigma}$ can then be used to calculate different response functions $R_{m\sigma}$ within the RPA as in usual GW calculations,

$$R_{m\sigma}(r,r';\omega) = P_{m\sigma}^0(r,r';\omega) + \int \mathrm{d}r_1\,\mathrm{d}r_2 P_{m\sigma}^0(r,r_1;\omega)v(r_1-r_2)R_{m\sigma}(r_2,r';\omega) \tag{2.45}$$

and finally a unique screened interaction for each $m\sigma$ can be found as

$$W_{m\sigma}(r,r';\omega) = v(r-r') + \int \mathrm{d}r_1\,\mathrm{d}r_2 v(r-r_2)R_{m\sigma}(r_2,r_1;\omega)v(r_1-r') \tag{2.46}$$

with the given $g_{m\sigma}$ line removed. The correction scheme then alters equation (2.42) to explicitly become [19]

$$\Sigma_\sigma^{GW\text{-ss}}(1,2) = i\sum_m g_{m\sigma}(1,2)W_{m\sigma}(1,2). \tag{2.47}$$

The diagram for the self-screening corrected GW (GW-ss) can be seen in Figure 2.4 and be directly compared to the ordinary GW diagram in Figure 2.3. It can be noted that the

form of the diagrams are the same, however, the modification causes the propagation of $g_{m\sigma}$ to occur without the presence of itself in the screening process of v to W; the self-screening has hence been removed from the equations.

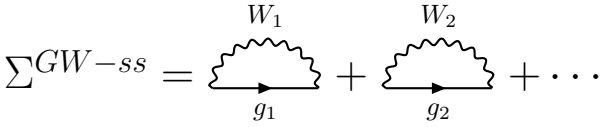

Figure 2.4: Diagram for the self-energy calculated with the self-screening corrected GW.

The self-energy can as usual be split into the exchange and correlation parts $\Sigma = \Sigma^x + \Sigma^c$ where now

$$\Sigma^x_\sigma(r,r') = -\sum_m^{\text{occ}} \varphi_m(r)\varphi^*_m(r')v(r-r') \tag{2.48}$$

and

$$\Sigma^c_\sigma(r,r';\omega) = \frac{i}{2\pi}\sum_m \int \mathrm{d}\omega' g_{m\sigma}(r,r';\omega+\omega')W^c_{m\sigma}(r,r';\omega) \tag{2.49}$$

respectively, with $W^c_{m\sigma} = W_{m\sigma} - v = vR_{m\sigma}v$. The Fock part clearly remains the same as in GW, as it does not contain any screening, and the change enters only through the correlation part of the self-energy where the screening is included. Moreover, the quantities $G_{m\sigma}$, P_m, R_m, and W_m introduced above are only auxiliary objects meant to be used for the calculation of the self-energy. This fact, and the consequences attached to attempting to compare them with their physical counterparts used in GW, will be illustrated and shown in Sections 3.1 and 3.2

Despite of the more phenomenological origin of the modifications, the scheme can be shown to be partly equivalent to adding vertex corrections [19]. By including higher order exchange diagrams, it is found that parts of these cancel the self-screening terms in the direct diagrams. It should furthermore be noted that the current scheme is not applicable to a full self-consistent calculation, as it requires the Green's function to be written in the form of equation (2.41), which is not necessarily possible for the renormalized G using the Dyson equation. Nevertheless, in principle it should be a possibility to extend the scheme, although no work has been carried out in that direction at this point.

2.8 Self-polarization correction for RPA

Another approach also proposed in [19], with the same intention of removing the self-screening error, is to remove the self-polarization caused by an electron-hole excitation instead of removing the self-screening caused by an electron in a given orbital. Instead of directly correcting Σ, it provides a correction of the response within RPA. The polarization propagator, P, can within RPA be found from the individual polarizations p_α corresponding to the excitations of the system, denoted by α. It is calculated as

$$P^0(r,r';\omega) = \sum_\alpha p_\alpha(r,r';\omega), \tag{2.50}$$

where the sum is over each excitation $\alpha = (n,m,\sigma)$ in the system between the occupied state φ_n and the unoccupied state φ_m for a given spin. We have [12, 19]

$$p_\alpha = \frac{\varphi_{m\sigma}(r)\varphi^*_{n\sigma}(r)\varphi^*_{m\sigma}(r')\varphi_{n\sigma}(r')}{\omega - \varepsilon_m + \varepsilon_n + i\delta} - \frac{\varphi_{m\sigma}(r')\varphi^*_{n\sigma}(r')\varphi^*_{m\sigma}(r)\varphi_{n\sigma}(r)}{\omega + \varepsilon_m - \varepsilon_n - i\delta}. \tag{2.51}$$

where the denominator gives the energy difference between the non-interacting states, $\varepsilon_m - \varepsilon_n$, with a small broadening $i\delta$.

The RPA equation (2.23) can schematically be expanded in the p_α as

$$R^{RPA} = P^0 + P^0 v R^{RPA} = [1 - P^0 v]^{-1} P^0 = \sum_\alpha [1 - \sum_\beta p_\beta v]^{-1} p_\alpha, \tag{2.52}$$

where the non-interacting system's response, P^0, is screened by $[1 - \sum_\beta p_\beta v]$ in order to obtain the response of the system within RPA. Similarly to how $g_{m\sigma}$ screened itself through W when calculating Σ in GWA, p_α screens itself here when calculating R, resulting in what is termed self-polarization. To remedy this, a new polarization is introduced,

$$P_\alpha = P^0 - p_\alpha, \tag{2.53}$$

with a given excitation removed, and thereupon the self-polarization corrected response function is introduced [19]

$$R^{\text{sp}} = \sum_\alpha [1 - P_\alpha]^{-1} p_\alpha. \tag{2.54}$$

The usual GW approximation can thereafter be used to calculate the self-energy as in equation (2.35), but now using R^{sp} to obtain the screened interaction, $W = v + vR^{sp}v$, providing the self-polarization corrected GW (GW-sp) scheme. Similarly to the case of the self-screening correction, also this approach can be shown to be partly equivalent to including additional diagrams.

2.9 The *GW* approximation in matrix representation

In order to simplify the treatment of GWA numerically for the models it is possible to rewrite the equations in matrix form. It can be noted that the eigenstates φ_m can be expanded in the basis orbitals ϕ_i as $\varphi_m = \sum_i c_i^m \phi_i$, and as a result the polarization function can be written as

$$P(r,r';\omega) = \sum_n^{\text{occ}} \sum_m^{\text{unocc}} \sum_{\alpha\beta\gamma\kappa} c^n_\alpha \phi_\alpha(r) c^m_\beta \phi_\beta(r) c^n_\gamma \phi_\gamma(r') c^m_\kappa \phi_\kappa(r') P_{nm}(\omega), \tag{2.55}$$

assuming the orbital wavefunctions to be real. c_i^m is the i'th expansion coefficient for eigenstate m and $P_{nm}(\omega)$ are the terms in equation (2.51). It can as such be written as a matrix in the $\phi_\alpha(r)\phi_\beta(r)$ product basis: $P_{(\alpha\beta)(\gamma\kappa)}$. Similarly R in the RPA equation, (2.23), can be seen to take on the same form as P above, and by including the expansion coefficients and the summations in the ω dependent part for simplicity we can rewrite it as

$$\sum_{\alpha\beta\gamma\kappa} \phi_\alpha(r)\phi_\beta(r)\phi_\gamma(r')\phi_\kappa(r')R_{(\alpha\beta)(\gamma\kappa)}(\omega) = \sum_{\alpha\beta\gamma\kappa} \phi_\alpha(r)\phi_\beta(r)\phi_\gamma(r')\phi_\kappa(r')P_{(\alpha\beta)(\gamma\kappa)}(\omega)$$
$$+ \sum_{\alpha'\beta'\gamma\kappa} \sum_{\alpha\beta\gamma'\kappa'} \phi_\alpha(r)\phi_\beta(r)\phi_\gamma(r')\phi_\kappa(r')P_{(\alpha\beta)(\gamma'\kappa')}(\omega)R_{(\alpha'\beta')(\gamma\kappa)}(\omega)$$
$$\times \int \mathrm{d}r_1\, \mathrm{d}r_2 \phi_{\gamma'}(r_1)\phi_{\kappa'}(r_1)v(r_1 - r_2)\phi_{\alpha'}(r_2)\phi_{\beta'}(r_2)$$

By identifying the last integral as $V_{(\gamma'\kappa')(\alpha'\beta')}$, the matrix element of the interaction in the $\phi_\alpha\phi_\beta$ basis, we can for a given $(\alpha\beta)(\gamma\kappa)$ write

$$R_{(\alpha\beta)(\gamma\kappa)}(\omega) = P_{(\alpha\beta)(\gamma\kappa)}(\omega) + \sum_{\alpha'\beta'\gamma'\kappa'} P_{(\alpha\beta)(\gamma'\kappa')}(\omega)V_{(\gamma'\kappa')(\alpha'\beta')}R_{(\alpha'\beta')(\gamma\kappa)}(\omega). \tag{2.56}$$

This can be seen as a matrix equation

$$R = P + PVR \tag{2.57}$$

in the product basis $\phi_\alpha(r)\phi_\beta(r)$. The same argument also holds for the self-screening and self-polarization corrected response functions and as such they can be treated numerically as matrices in the same way.

The next step to use the equations numerically is to find W and Σ in matrix form in the orbital wave function basis. The exchange part of the self-energy can be found in matrix form by taking the matrix elements of Σ^x in equation (2.39), assuming real wave functions:

$$\langle\phi_k|\Sigma^x|\phi_p\rangle = -\int \mathrm{d}r\, \mathrm{d}r' \phi_k(r) \sum_m^{\text{occ}} \varphi_m(r)\varphi_m(r')v(r - r')\phi_p(r')$$
$$= -\sum_{i,j} \sum_m^{\text{occ}} c_i^m c_j^m V_{(ki,jp)}$$

where the eigenstates φ_m have been expanded in the orbital basis. To find the correlation part of the self-energy we first find the imaginary part, whereupon the real part can be acquired through a Hilbert transform. The correlation part of the self-energy in equation (2.37) can in turn be rewritten using the spectral representation of W^c [12]

$$D(r,r';\omega) = -\frac{1}{\pi}\,\mathrm{Im}\, W^c(r,r';\omega)\mathrm{sgn}(\omega) \tag{2.58}$$

and by splitting the sum into occupied and unoccuped states. For the occupied ones the expression can be shown to take the form

$$\mathrm{Im}\,\Sigma^c_{occ}(r,r';\omega) = \pi \sum_m^{occ} \varphi_m(r)\varphi_m(r')D(r,r';\varepsilon_m - \omega)\theta(\varepsilon_m - \omega) \tag{2.59}$$

and for the unoccupied states it becomes

$$\mathrm{Im}\,\Sigma^c_{unocc}(r,r';\omega) = -\pi \sum_m^{unocc} \varphi_m(r)\varphi_m(r')D(r,r';\omega - \varepsilon_m)\theta(\omega - \varepsilon_m) \tag{2.60}$$

The spectral function of the correlation self-energy [12]

$$\Gamma(r,r';\omega) = -\frac{1}{\pi}\,\mathrm{Im}\,\Sigma^c(r,r';\omega)\mathrm{sgn}(\omega-\mu) \tag{2.61}$$

can therefore be written as

$$\Gamma_{occ}(r,r';\omega) = \sum_{m}^{occ} \varphi_m(r)\varphi_m(r')D(r,r';\varepsilon_m-\omega)\theta(\varepsilon-\omega) \tag{2.62}$$

and

$$\Gamma_{unocc}(r,r';\omega) = \sum_{m}^{unocc} \varphi_m(r)\varphi_m(r')D(r,r';\omega-\varepsilon_m)\theta(\omega-\varepsilon) \tag{2.63}$$

when summing over the the occupied and unoccupied states respectively. As the terms in the sum over all states are independent, the full spectral representation of the self-energy will only be $\Gamma = \Gamma_{occ} + \Gamma_{unocc}$. We rewrite this into matrix form for the occupied part as, remembering that $W^c = vRv$,

$$\langle\phi_k|\Gamma_{occ}|\phi_p\rangle = -\frac{1}{\pi}\sum_{m}^{occ}\sum_{ij} c_i^m c_j^m\,\mathrm{Im}\left[VR(\varepsilon_m-\omega)V\right]_{(ki)(pj)}\theta(\varepsilon_m-\omega) \tag{2.64}$$

where the φ's have been expanded in the ϕ-basis functions and $[VRV]$ denotes a matrix multiplication. The unoccupied part is similarly

$$\langle\phi_k|\Gamma_{unocc}|\phi_p\rangle = -\frac{1}{\pi}\sum_{m}^{unocc}\sum_{ij} c_i^m c_j^m\,\mathrm{Im}\left[VR(\omega-\varepsilon_m)V\right]_{(ki)(pj)}\theta(\omega-\varepsilon_m) \tag{2.65}$$

The real part can subsequently be obtained by a Hilbert transform

$$\mathrm{Re}\,\Sigma^c_{kp}(r,r';\omega) = \int_{-\infty}^{\infty}\frac{\Gamma_{kp}(r,r';\omega')}{\omega-\omega'}, \tag{2.66}$$

where in the matrix representation the transform is carried out for each individual element. The full self-energy in matrix form then simply becomes,

$$\langle\phi_k|\Sigma(\omega)|\phi_p\rangle = \Sigma_{kp} = \Sigma^x_{kp} + \mathrm{Re}(\Sigma^c_{kp}(\omega)) + i\,\mathrm{Im}(\Sigma^c_{kp}(\omega)) \tag{2.67}$$

with the imaginary parts being obtained from Γ_{kp} using equation (2.61). In the case of GW-sp the equations become identical, except for using the self-polarization corrected response in W, and for GW-ss several different $R_{m\sigma}$ and $W_{m\sigma}$ are calculated, though each identically as described here, and thereafter Σ^c is calculated by using the W^c_m as described in Section 2.7. To obtain the Green's function after a GW calculation, the Dyson equation, equation (2.32), is used. In matrix form it can be written as

$$G = \left[\mathbb{1} - G^0\Sigma\right]^{-1} G^0. \tag{2.68}$$

with the matrix elements $\langle\phi_k|G^0|\phi_p\rangle$ straightforwardly obtained by expanding equation (2.17) in the ϕ_k basis.

Section 3

Model calculations

To ascertain the accuracy of the correction schemes presented in Sections 2.7 and 2.8 the two model dimers will be investigated in the following sections. The one-orbital system is calculated analytically, while the calculations for the two-orbital one are carried out numerically. Since we will only be interested in G_0W_0 calculations, GW will henceforth refer to the one-shot approach unless otherwise specified.

3.1 One-orbital system

The one-orbital model is simple enough to be solved analytically since it consists of only two electrons, with opposite spin, occupying a dimer with two sites and one orbital on each. The Hamiltonian of the system is given by equation (2.6) with the orbital set at zero energy, i.e. $\epsilon_0 = 0$, as this only provides a constant shift. It takes the form

$$H^{1\text{ orb}} = \begin{pmatrix} U_1 & 0 & t & t \\ 0 & U_1 & -t & -t \\ t & -t & U_0 & 0 \\ t & -t & 0 & U_0 \end{pmatrix}$$

using the basis states in the occupation representation formalism

$$|1\rangle = c^+_{1\uparrow}c^+_{2\downarrow}|0\rangle = |\uparrow\cdot\downarrow\rangle \quad |3\rangle = c^+_{1\uparrow}c^+_{1\downarrow}|0\rangle = |\uparrow\downarrow\cdot\rangle$$
$$|2\rangle = c^+_{1\downarrow}c^+_{2\uparrow}|0\rangle = |\downarrow\cdot\uparrow\rangle \quad |4\rangle = c^+_{2\uparrow}c^+_{2\downarrow}|0\rangle = |\cdot\uparrow\downarrow\rangle$$

with the dot separating the two sites as $|\text{site } 1 \cdot \text{site } 2\rangle$. The diagonalization of the Hamiltonian and the calculation of the exact response function in equation (2.19) has been done in a previous work [24], however with $U_1 = 0$. The response was in matrix form found to be

$$R_{11,11}(r,r';\omega) = R_{22,22}(r,r';\omega) = 2x^2\left[\frac{1}{\omega-\varepsilon_3+\varepsilon_1+i\eta} - \frac{1}{\omega+\varepsilon_3-\varepsilon_1-i\eta}\right]$$
$$R_{11,22}(r,r';\omega) = R_{22,11}(r,r';\omega) = -R_{11,11}(r,r';\omega)$$
$$R_{12,12}(r,r';\omega) = R_{21,21}(r,r';\omega) = 16(x^2-y^2)^2\left[\frac{1}{\omega-\varepsilon_4+\varepsilon_1+i\eta} - \frac{1}{\omega+\varepsilon_4-\varepsilon_1-i\eta}\right]$$

using the product basis $\phi_i\phi_j$ as outlined in Section 2.9. The ε_i are the energies of states ϕ_i and are together with the x and y variables obtained from the diagonalization to be

$$2x^2 + 2y^2 = 1, \qquad x = 2yt/(\varepsilon_4 - U_1)$$
$$\varepsilon_1 = \frac{1}{2}(U_0 + U_1 - \sqrt{(U_0-U_1)^2+16t^2})$$
$$\varepsilon_2 = U_1, \qquad \varepsilon_3 = U_0$$
$$\varepsilon_4 = \frac{1}{2}(U_0 + U_1 + \sqrt{(U_0-U_1)^2+16t^2}).$$

The exact Green's function is straightforwardly found by also diagonalizing the $N\pm1$ Hamiltonians and finding the matrix elements with the N-particle states, according to equation (2.9). The matrix elements of the resulting Green's function are

$$\begin{aligned}G_{11} = G_{22} = &\frac{\frac{1}{2}(x+y)^2}{\omega + E_1(N-1) - \varepsilon_1 - i\delta} + \frac{\frac{1}{2}(x+y)^2}{\omega - E_1(N+1) + \varepsilon_1 + i\delta} \\ &+ \frac{\frac{1}{2}(x-y)^2}{\omega + E_2(N-1) - \varepsilon_1 - i\delta} + \frac{\frac{1}{2}(x-y)^2}{\omega - E_2(N+1) + \varepsilon_1 + i\delta}\end{aligned}$$

$$\begin{aligned}G_{12} = G_{21} = &- \frac{\frac{1}{2}(x+y)^2}{\omega + E_1(N-1) - \varepsilon_1 - i\delta} + \frac{\frac{1}{2}(x+y)^2}{\omega - E_1(N+1) + \varepsilon_1 + i\delta} \\ &- \frac{\frac{1}{2}(x-y)^2}{\omega + E_2(N-1) - \varepsilon_1 - i\delta} + \frac{\frac{1}{2}(x-y)^2}{\omega - E_2(N+1) + \varepsilon_1 + i\delta}\end{aligned}$$

where $E_1(N-1) = -t$, $E_2(N-1) = t$, $E_1(N+1) = U_0+2U_1-t$ and $E_2(N+1) = U_0+2U_1+t$.

When calculating the mean-field states used as input into GWA the only states which need to be considered are the one-particle ones as both electrons will feel the same mean-field on both sites due to symmetry, and it can as such be neglected. The Hamiltonian therefore takes the simple form

$$H^{\text{MF 1 orb}} = \begin{pmatrix} 0 & t \\ t & 0 \end{pmatrix}$$

Also the response function within RPA was in [24] found to be

$$R^{\text{RPA}}(r,r';\omega) = \varphi_A(r)\varphi_B(r)\left[\frac{a_{no}}{\omega - \Delta_{\text{no}} + i\delta} - \frac{a_{no}}{\omega + \Delta_{\text{no}} - i\delta}\right]\varphi_A(r')\varphi_B(r') \tag{3.1}$$

where $\varphi_A(r) = \frac{1}{\sqrt{2}}(\phi_1(r)-\phi_2(r))$ is the unoccupied antibonding state with energy $\varepsilon_A = t$, and $\varphi_B(r) = \frac{1}{\sqrt{2}}(\phi_1(r)+\phi_2(r))$ is the occupied bonding state with eigenenergy $\varepsilon_B = -t$. Thus the two states are split by $\Delta\varepsilon = 2t$, and $\Delta_{\text{no}} = \sqrt{\Delta\varepsilon^2 + 4U_{AB,AB}\Delta\varepsilon}$, the excitation energy for normal GW, and $a_{no} = 2\Delta\varepsilon/\Delta_{\text{no}}$ will depend on the chosen parameters. t will be used as the unit here throughout the section and is hence set to 1. Furthermore, the $U_{ij,kl}$ matrix elements in the orbital basis are the ones used as parameters and $U_{AB,AB}$ must therefore be found in this basis. This is done by expanding the φ's:

$$\begin{aligned}U_{AB,AB} &= \int \mathrm{d}r\,\mathrm{d}r'\varphi_A(r)\varphi_B(r)v(r-r')\varphi_A(r')\varphi_B(r') \\ &= \frac{1}{4}(U_{11,11} + U_{22,22} - U_{11,22} - U_{11,22}) \\ &= \frac{1}{2}(U_0 - U_1).\end{aligned}$$

With knowledge of the RPA response function, the self-energy $\Sigma = \Sigma^x + \Sigma^c$ can be calculated using GW. The following derivation of Σ^c and Σ^c_{ss} has previously been done in [19], but is redone here as it serves to illustrate the scheme using an exactly solvable system. In the final part of the section further investigation of the model beyond what was previously done will

also be presented, including the self-polarization correction and a comparison between the schemes.

The exchange part from equation (2.39) can directly be written

$$\Sigma^x(r,r') = -v(r-r')\varphi_B(r)\varphi_B(r') \tag{3.2}$$

for which the diagonal matrix elements in the $\varphi_{A,B}$ basis are

$$\begin{aligned}\langle\varphi_A|\Sigma^x|\varphi_A\rangle &= -\int \mathrm{d}r\,\mathrm{d}r'\varphi_A(r)\varphi_B(r)v(r-r')\varphi_A(r')\varphi_B(r') = -\frac{1}{2}(U_0-U_1)\\ \langle\varphi_B|\Sigma^x|\varphi_B\rangle &= -\int \mathrm{d}r\,\mathrm{d}r'\varphi_B(r)\varphi_B(r)v(r-r')\varphi_B(r')\varphi_B(r') = -\frac{1}{2}(U_0+U_1)\end{aligned}$$

The correlation part can be calculated using equation (2.37) where the non-interacting Green's function in equation (2.17) is used,

$$G^0_\sigma = \frac{\varphi_A(r)\varphi_A(r')}{\omega-\varepsilon_A+i\delta} + \frac{\varphi_B(r)\varphi_B(r')}{\omega-\varepsilon_B-i\delta}. \tag{3.3}$$

Since $W^c = vRv$, Σ^c is obtained as

$$\Sigma^c(r,r';\omega) = \frac{\lambda_1(r,r')}{\omega+\Delta_{\mathrm{no}}-\varepsilon_B-i\delta} + \frac{\lambda_2(r,r')}{\omega-\Delta_{\mathrm{no}}-\varepsilon_A+i\delta} \tag{3.4}$$

where contour integration has been used to carry out the frequency integral. All the space dependent parts have been hidden inside $\lambda(r,r')$ for simplicity in accordance with [19],

$$\begin{aligned}\lambda_1(r,r') &= a_{no}\varphi_B(r)\varphi_B(r')\int \mathrm{d}r_1\,\mathrm{d}r_2 v(r-r_2)v(r'-r_1)\varphi_A(r_1)\varphi_B(r_1)\varphi_A(r_2)\varphi_B(r_2)\\ \lambda_2(r,r') &= a_{no}\varphi_A(r)\varphi_A(r')\int \mathrm{d}r_1\,\mathrm{d}r_2 v(r-r_2)v(r'-r_1)\varphi_A(r_1)\varphi_B(r_1)\varphi_A(r_2)\varphi_B(r_2)\end{aligned} \tag{3.5}$$

The diagonal matrix elements of the correlation part of the self-energy can finally be found to be

$$\begin{aligned}\langle\varphi_A|\Sigma^c|\varphi_A\rangle &= \frac{a_{no}}{\omega+\Delta_{\mathrm{no}}-\varepsilon_B-i\delta}U_{AB,AB}U_{AB,AB} + \frac{a_{no}}{\omega-\Delta_{\mathrm{no}}-\varepsilon_A+i\delta}U_{AA,AB}U_{AA,AB}\\ &= \frac{1}{4}\frac{a_{no}(U_0-U_1)^2}{\omega+\Delta_{\mathrm{no}}-\varepsilon_B-i\delta} + 0\end{aligned} \tag{3.6}$$

where $U_{AA,AB} = 0$ is found by again expanding $\varphi_{A,B}$ in the orbital basis. Similarly the matrix element for the bonding state becomes

$$\langle\varphi_B|\Sigma^c|\varphi_B\rangle = \frac{1}{4}\frac{a_{no}(U_0-U_1)^2}{\omega-\Delta_{\mathrm{no}}-\varepsilon_A+i\delta} \tag{3.7}$$

Next the self-screening correction described in Section 2.7 is applied to the system. Rewriting the Green's function of the system in equation (3.3) in terms of g_m according to equation (2.41), the non-interacting Green function of the system is

$$G^0_\sigma = \frac{\varphi_A(r)\varphi_A(r')}{\omega-\varepsilon_A+i\delta} + \frac{\varphi_B(r)\varphi_B(r')}{\omega-\varepsilon_B-i\delta} = g_A + g_B. \tag{3.8}$$

Following the prescription we define $G_{m\sigma} = G_\sigma - g_{m\sigma}$ and obtain

$$G_A = G^0_\sigma - g_A = g_B \tag{3.9}$$
$$G_B = G^0_\sigma - g_B = g_A \tag{3.10}$$

for each of the spins, and the polarization function can straightforwardly be found by carrying out the integral in

$$P_{A\sigma}(r,r';\omega) = -\frac{i}{2\pi}\int \mathrm{d}\omega' \left[G_A(r,r';\omega+\omega')G_A(r',r;\omega') + G^0_{-\sigma}(r,r';\omega+\omega')G^0_{-\sigma}(r',r;\omega')\right]. \tag{3.11}$$

Since G_AG_A has poles only in the upper half-plane the contour can be closed in the lower one giving no contribution to the polarization function. Hence $P_{A\sigma} = -iG^0_\sigma G^0_\sigma$, which is half the one found for the uncorrected case, namely $P = -2iG^0G^0$. In the same way it follows that $P_{B\sigma} = P_{A\sigma}$. Using the RPA equation (2.23), it can easily be verified that the response function will be of the same form as for the uncorrected case

$$R^{ss}_{A\sigma}(r,r';\omega) = \varphi_A(r)\varphi_B(r)\left[\frac{a_{ss}}{\omega - \Delta_{\mathrm{ss}} + i\delta} - \frac{a_{ss}}{\omega + \Delta_{\mathrm{ss}} - i\delta}\right]\varphi_A(r')\varphi_B(r'). \tag{3.12}$$

where only $\Delta_{ss} = \sqrt{\Delta\varepsilon^2 + 2U_{AB,AB}\Delta\varepsilon}$ and $a_{ss} = \Delta\varepsilon/\Delta_{\mathrm{ss}}$ have changed for the self-screening correction compared to the usual RPA. Furthermore, $R^{ss}_{B\sigma} = R^{ss}_{A\sigma}$. A comparison between the approximate results and the exact results will be presented later on in Section 5.1.1. Proceeding identically to the calculation of Σ above for the uncorrected case, the exchange part is noted to be the same as in equation (3.2), since only the screening in the correlation part will be affected. Next the diagonal matrix elements for the correlation part becomes

$$\begin{aligned}\langle\varphi_A|\Sigma^c_{ss}|\varphi_A\rangle &= \frac{1}{4}\frac{a_{ss}(U_0-U_1)^2}{\omega + \Delta_{\mathrm{ss}} - \varepsilon_B - i\delta}\\ \langle\varphi_B|\Sigma^c_{ss}|\varphi_B\rangle &= \frac{1}{4}\frac{a_{ss}(U_0-U_1)^2}{\omega - \Delta_{\mathrm{ss}} - \varepsilon_A + i\delta}.\end{aligned} \tag{3.13}$$

To check how the correction affects the system the HOMO-LUMO gap between the two states can also be calculated. The gap caused by the one-particle and exchange energies are the same for both cases

$$\varepsilon_A - \varepsilon_B + \langle\varphi_A|\Sigma^x|\varphi_A\rangle - \langle\varphi_B|\Sigma^x|\varphi_B\rangle = 2t + U_1,$$

and the total gaps can thus be calculated to be using a perturbation example of the quasi-particle equation, which will be further explained in Section 4.1,

$$\begin{aligned}\Delta E^{GW} &= 2t + U_1 + \langle\varphi_A|\Sigma^c|\varphi_A\rangle - \langle\varphi_B|\Sigma^c|\varphi_B\rangle = 2t + U_1 + \frac{\Delta\varepsilon}{\Delta_{\mathrm{no}}}\frac{(U_0-U_1)^2}{\Delta\varepsilon + \Delta_{\mathrm{no}}}\\ \Delta E^{GW-ss} &= 2t + U_1 + \langle\varphi_A|\Sigma^c_{ss}|\varphi_A\rangle - \langle\varphi_B|\Sigma^c_{ss}|\varphi_B\rangle = 2t + U_1 + \frac{\Delta\varepsilon}{2\Delta_{\mathrm{ss}}}\frac{(U_0-U_1)^2}{\Delta\varepsilon + \Delta_{\mathrm{ss}}}.\end{aligned}$$

Additionally, when using the self-polarization correction outlined in section 2.8 instead of the self-screening correction, which was not done in [19], we find the response function to

only differ from the self-screening ones by a factor of two. This also causes the correlation part of the self-energy to be twice as large:

$$R^{sp} = 2R^{ss}_{A\sigma} \tag{3.14}$$
$$\Sigma^c_{sp} = 2\Sigma^c_{ss}, \tag{3.15}$$

while the exchange part remains unaffected as usual. That the results would become identical for the two approaches, except for the factor of two, could also be seen directly from a simple argument. The expression for P^0 in equation is found by doing the contour integration in equation (2.38) using $G^0_\sigma G^0_\sigma + G^0_\sigma G^0_\sigma$. This is what gives us the expression for the p_α found in equation (2.51), where a given excitation α is a combined index of an occupied $n\sigma$ and an unoccupied $m\sigma$. This occurs because the terms in the contour integration of $(\sum_n g_n)(\sum_m g_m)$ provides zero contribution if the poles are in the same half-plane, and a non-zero one for products which have poles in both half-planes, the latter being the case for a term with a product of an occupied and an unoccupied state. When the self-screening correction is used, a given $g_{n\sigma}$ is removed from the calculation of P and hence all combinations $\alpha = (n\sigma,m\sigma)$ with this n are removed. For this simple model there is only one other state for a given n to couple to, and therefore only one excitation to remove. The self-polarization correction corresponds to removing one excitation from the polarization and as there only exists one per spin the corrected polarization functions are the same.

The difference arises when calculating the response function, where, using self-screening, only the excitation involving spin σ is a part of $R_{m\sigma}$, while for self-polarization both spins are summed over to give R^{sp}, because all excitations in the system are involved. Since both spins provide equal polarization and response functions, due to not using a spin-polarized system, we get a factor of two more with the self-polarization correction. Besides this, the rest of the calculation is the same and the excitation energies therefore remain identical. It should however be noted that the methods are not directly related in this way for larger systems with several unoccupied or occupied states, as will become clear already for the two-orbital system to be treated in the next section. The result found here also gives a first indication to the auxiliary nature of the self-screening $R^{ss}_{m\sigma}$, as only one of the spins is involved in calculating it. In the other two cases both spins are used. This will cause the weight of the excitation peak to be about halved as compared to the exact case, as will be further discussed in Section 5.1.1.

For the exact case the HOMO-LUMO gap can be found to be [19]

$$\Delta E^{\text{exact}} = -2t + U_1 + \sqrt{(U_0 - U_1)^2 + 16t^2}. \tag{3.16}$$

By expanding the above expressions with respect to $\frac{U_0 - U_1}{2t} < 1$, we see that the exact result and the self-screening corrected one both yield

$$\Delta E \approx 2t + U_1 + \frac{t}{2}\frac{(U_0 - U_1)^2}{4t^2} \tag{3.17}$$

as was previously shown in [19], while GW and GW-sp have a factor of two larger in the last term. Thus GW-ss agrees with the exact HOMO-LUMO gap for weak to medium interaction strengths, which also will be analyzed further in Section 5.1.1.

By using the Dyson equation, (2.32), a new renormalized G can be calculated using the self-energy. It is readily seen that the off-diagonal elements Σ_{AB} and Σ_{BA} are zero since they will involve $U_{AA,AB} = U_{AB,BB} = 0$. The Dyson equation can thus be written as the matrix equation

$$\begin{pmatrix} G_{BB} & 0 \\ 0 & G_{AA} \end{pmatrix} = \begin{pmatrix} G^0_{BB} & 0 \\ 0 & G^0_{AA} \end{pmatrix} + \begin{pmatrix} G^0_{BB} & 0 \\ 0 & G^0_{AA} \end{pmatrix} \begin{pmatrix} \Sigma_{BB} & 0 \\ 0 & \Sigma_{AA} \end{pmatrix} \begin{pmatrix} G_{BB} & 0 \\ 0 & G_{AA} \end{pmatrix} \tag{3.18}$$

which can be inverted as $G = [\mathbb{1} - G^0\Sigma]^{-1}G^0$ and is found to take the form

$$\begin{pmatrix} G_{BB} & 0 \\ 0 & G_{AA} \end{pmatrix} = \begin{pmatrix} \frac{G^0_{BB}}{1-G^0_{BB}\Sigma_{BB}} & 0 \\ 0 & \frac{G^0_{AA}}{1-G^0_{AA}\Sigma_{AA}} \end{pmatrix} \tag{3.19}$$

This expression holds for all three cases with Σ, Σ_{ss} and Σ_{sp} being used respectively. These renormalized Green's functions will be compared to the exact one in Section 5.1.1.

3.2 Two-orbital system

The second model considered is a Hubbard dimer with two orbitals per site. Although not being an overly complicated system it is cumbersome to solve analytically, and therefore a numerical approach was taken. Also for this model only two electrons with total spin 0 are considered, giving a total of 16 states in the exact case. The code produced to solve the problem is based on a previous work [24]. The diagonalization of the Hamiltonian and calculation of the response function of the system in question within RPA and for the exact case are maintained as before, while the calculation of the self-energy and Green's function, as well as the self-screening and self-polarization schemes, have been added. In the model the parameters U_0, t_2, t_3, and ϵ_1 will be varied to see their effects on the results, while $t_1 = 1$ will be kept fixed as a reference point as in the one-orbital system. ϵ_0 is set to 0 as only the separation between the orbital energies is important since an overall constant can always be added.

The exact Hamiltonian is obtained in matrix form by calculating the matrix elements $\langle i| c^+_{k\sigma}c_{l\sigma} |j\rangle$, for all states i,$j = 1, \ldots, 16$, and the orbitals k,$l = 1, \ldots, 4$ are labeled according to Figure 2.1. The non-zero matrix elements are assigned values according to the two-orbital Hamiltonian in equation (2.7). After a subsequent diagonalization, the response function is calculated using equation (2.19) and the previously obtained matrix elements. The exact Green's function is directly obtained using equation (2.9) after also the $N \pm 1$ Hamiltonians have been diagonalized and the matrix elements with the $N = 2$ ground state have been calculated. The self-energy is in turn simply acquired by inverting the Dyson equation as $\Sigma = G_0^{-1} - G^{-1}$.

To calculate the quantities within GWA, instead a mean-field Hamiltonian for each of the

spins is solved self-consistentlyt:

$$H^{\text{MF 2 orb}} = \begin{pmatrix} b_1^2 U_0 & t_1 & 0 & t_2 \\ t_1 & b_2^2 U_0 & t_2 & 0 \\ 0 & t_2 & \epsilon_1 & t_3 \\ t_2 & 0 & t_3 & \epsilon_1 \end{pmatrix} \tag{3.20}$$

with the matrix elements being in the basis states where the electron under consideration occupies each of the four orbitals in Figure 2.1. Here the $n_{i\uparrow}n_{i\downarrow}$ term has been approximated as $n_{i\uparrow}\langle n_{i\downarrow}\rangle$, and similarly for the down-spin electron, such that the given electron only feels the average presence of the other one. To ensure convergence a mixing parameter $x = 10^{-5}$ is introduced, with the expansion coefficients b_i of the ground state in the orbital ϕ_i being modified as $b_i = \sqrt{(b_i^{\text{old}})^2(1-x) + (b_i^{\text{new}})^2 x}$ to mix the old and new densities after each iteration. As the system is not spin-polarized, the mean-occupation will be the same for both spins. The non-interacting Greens function can subsequently be found using equation (2.17), and the GW, GW-ss and GW-sp calculations are performed as outlined in sections 2.6 – 2.8, using the matrix formalism described in Section 2.9. The results are presented and discussed in Section 5.1.2.

It was noted for the one-orbital model that the direct correspondence between the self-screening and self-polarization correction schemes found there does not necessarily apply to larger systems. Indeed it is seen that they differ for the two-orbital model, even though they are based on the same principle of removing the issue of an electron screening itself. To see this we consider the four eigenstates arising from diagonalizing the mean-field Hamiltonian, schematically drawn in Figure 3.1a.

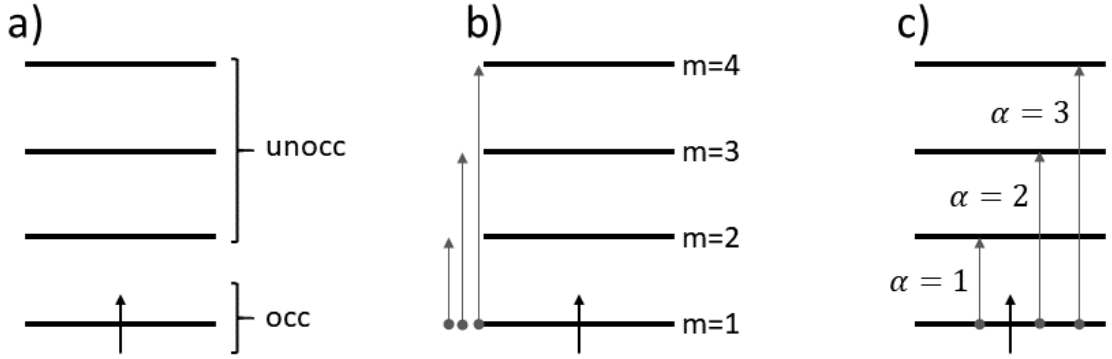

Figure 3.1: a) Schematic drawing of the eigenstates of the non-interacting system with only the lowest orbital occupied. b) The system as viewed from the self-screening perspective, where the excitations involving $g_{1\uparrow}$ have been marked out. c) The system from the self-polarization perspective, with all possible excitations for the up-spin electron being shown.

When observing the system from the point of view of the self-screening correction, the important aspect is how removing a given state will affect the polarization. If we take $(m\sigma) = 1\uparrow$, which is the only occupied state, it is clear that upon removal of this $g_{1\uparrow}$ all the excitations will be removed for the spin, shown in Figure 3.1b. This occurs when doing the

contour integration in equation (2.38) since occupied states will be coupled to unoccupied states to give the polarization in the form of equations (2.50) and (2.51), as also discussed for the one-orbital model. This means that the only channels still contributing to the polarization are for the other spin, i.e. effectively we could see it as $P_{1\uparrow} = -i(0 + G_\downarrow G_\downarrow)$. If instead the self-polarization correction is used, the contour integration has already been carried out and the system can be viewed as having three excitations for each spin, drawn in Figure 3.1c. The new polarization propagators calculated in equation (2.53) remove one of these excitations, $\alpha = 1,2,3$, each.

Comparing the two approaches it becomes clear that they are similar but not the same, as neither of the P_α, with only one excitation removed, is the same as $P_{1\sigma}$ where all excitations, three in this case, have been removed for the given spin. Nevertheless, if instead $m = 2$ (or $m = 3,4$) was considered we would see that $P_{m\sigma=2\uparrow} = P_{\alpha=1}$, as can be seen directly from that the only excitation involving $m = 2$ is the one designated $\alpha = 1$ in Figure 3.1c. Thus the same excitation is removed in both schemes, creating identical polarization functions. A general conclusion can then be drawn: a system with only one possible excitation, for any given state, will have the same modified polarization propagators, P_α and $P_{m\sigma}$, for both schemes. However, whenever any state is involved in two or more possible excitations there will be polarization propagators which will be different for the two approaches; the self-screening scheme will remove all of the excitations involving this state, while the self-polarization at most removes one excitation.

Section 4

Ab initio calculations for real materials

Calculating material properties from first principles is an important application for electronic structure methods. As was described in section 2.6, *GW* has proven to provide good agreement with experiments in many cases and is today a widely used method. To carry out the *ab inito* calculations in this work an older version of the *GW* code *SPEX* [11,20,26], previously modified by Rei Sakuma [25] was used. It was further modified in this work to also include the self-screening correction for *GW*. As input to *SPEX* the DFT code *FLEUR* [20] is used to provide the initial wavefunctions and eigenenergies used in the *GW* calculations.

4.1 The SPEX code

SPEX is an all-electron code, treating both valence and core electrons, using the full-potential linearized augmented-plane-wave (FLAPW) method. The polarization matrix is expanded in a mixed-product basis (MPB), denoted $M_I^{\mathbf{k}}$, consisting of products of the basis functions within the FLAPW method, obtained from *FLEUR*. The basis consists of interstitial plane waves in a so-called interstitial region surrounding muffin-tin spheres, which are non-overlapping, wherein there are numerical wavefunctions [26]. The DFT wavefunction for a given band n takes the following form

$$\varphi^{\sigma}_{n\mathbf{k}}(\mathbf{r}) = \begin{cases} \frac{1}{\sqrt{N}} \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} \sum_{p=0}^{1} A^{n\mathbf{k}\sigma}_{almp} u^{\sigma}_{almp}(\mathbf{r} - \mathbf{R}_a) \quad , \quad \mathbf{r} \in MT(a) \\ \frac{1}{\sqrt{V}} \sum_{|\mathbf{k}+\mathbf{G}| \leq G_{max}} c^{n\mathbf{k}\sigma}_{\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}} \quad , \quad \mathbf{r} \in IR \end{cases} \tag{4.1}$$

with the muffin-tins (MT) being centered at an atom a at position $\mathbf{R_a}$ in the unit cell, and $c^{n\mathbf{k}\sigma}_{\mathbf{G}}$ and $A^{n\mathbf{k}\sigma}_{almp}$ being expansion coefficients with the condition to have continuity in its value and first derivative at the boundary between the muffin-tin and interstitial regions (IR). The u^{σ}_{almp} are numerical functions, N the number of particles, and V the volume of a crystal. For more details on the FLAPW basis and the mixed product basis employed in *SPEX* see e.g. Refs. [26, 27]. The energy corresponding to a given band n and k-point $\mathbf{k}$ is $\varepsilon^{\sigma}_{n\mathbf{k}}$, providing the DFT band structure.

In order to calculate the wavefunctions and energies, *FLEUR* takes as input material data in the form of the unit cell, as will be outlined in Section 4.3. Related to the actual numerical calculations several adjustable cut-off parameters are introduced in *FLEUR*, such as G_{max} in equation (4.1), which gives the plane wave cut-off of the wavefunction, and l_{max}, the angular momentum cut-off inside the muffin-tins. Also the muffin-tin radius for the atoms can be varied, and the number of bands and k-points in the discretized BZ, for which the states are calculated, have to be chosen. For a given set of parameters the code is then iterated until the charge density is converged. Thereafter the cut-offs are improved in new runs until there

is convergence in properties of interest. When doing the DFT computations in this work the convergence of the band structure and total energy was checked. Furthermore, various approximations of the exchange-correlation functional are available; in this work only LDA has been emplyed however. A short overview of DFT and LDA, which is used in this work, are presented in Appendix A.

In short, the flow of a *SPEX* calculation is as follows. The polarization P_{IJ} and Coulomb matrices v_{IJ} are first obtained in the MPB. The polarization takes the form [26]

$$P_{IJ}(\mathbf{k},\omega) = \sum_{\sigma} \sum_{\mathbf{q}}^{\mathrm{BZ}} \sum_{n}^{\mathrm{occ}} \sum_{n'}^{\mathrm{unocc}} M_{Inn'}^{\sigma\mathbf{kq}} M_{Jnn'}^{\sigma\mathbf{kq}*} \left(\frac{1}{\omega + \varepsilon_{n\mathbf{q}}^{\sigma} - \varepsilon_{n'\mathbf{q}+\mathbf{k}}^{\sigma} + i\delta} - \frac{1}{\omega - \varepsilon_{n\mathbf{q}}^{\sigma} + \varepsilon_{n'\mathbf{q}+\mathbf{k}}^{\sigma} - i\delta} \right) \tag{4.2}$$

where the summation over the Brillouin zone (BZ) is carried out using the tetrahedron method. For notational simplicity the quantity $M_{Inn'}^{\sigma\mathbf{kq}}$ is introduced according to

$$M_{Inn'}^{\sigma\mathbf{kq}} = \left\langle M_I^{\mathbf{k}} \varphi_{n\mathbf{q}}^{\sigma} \middle| \varphi_{n'\mathbf{q}+\mathbf{k}}^{\sigma} \right\rangle = \int M_I^{\mathbf{k}*}(\mathbf{r}) \varphi_{n\mathbf{q}}^{\sigma*}(\mathbf{r}) \varphi_{n'\mathbf{q}+\mathbf{k}}^{\sigma}(\mathbf{r})\, \mathrm{d}\mathbf{r}, \tag{4.3}$$

where $M_I^{\mathbf{k}}(\mathbf{r})$ is the mixed product basis used to expand products of the $\varphi_{n\mathbf{k}}^{\sigma}(\mathbf{r})$ wavefunctions in the separate regions. Inside the MT the basis describes products between two numerical functions $u_{almp}^{\sigma*}$ and $u_{al'm'p'}^{\sigma}$, and I is a combined index of the a,l,l',m,m',p and p' indices. The resulting products are then artificially made into Bloch functions, providing the **k**-dependence. The MPB inside the IR is instead a new set of interstitial plane waves, which are products of each other, and the index I represents a new **G** with at most twice the cutoff, $\leq 2G_{max}$, for the new basis. For more details on the mixed product basis see Refs. [26,28]. To simplify the calculation of the dielectric function ϵ, used to calculated the screened interaction W, the polarization matrix is transformed into the Coulomb eigenbasis, $E_{\mu}^{\mathbf{k}}$, with eigenvalues $v_{\mu}(\mathbf{k})$. The equations used in GWA to obtain W can now simply be rewritten as [26]:

$$\epsilon_{\mu\nu}(\mathbf{k},\omega) = \delta_{\mu\nu} - \sqrt{\mathbf{v}_{\mu}(\mathbf{k})} \mathbf{P}_{\mu\nu}(\mathbf{k},\omega) \sqrt{\mathbf{v}_{\nu}(\mathbf{k})} \tag{4.4}$$

$$W_{\mu\nu}(\mathbf{k},\omega) = \sqrt{v_{\mu}(\mathbf{k})} \epsilon_{\mu\nu}^{-1}(\mathbf{k},\omega) \sqrt{v_{\nu}(\mathbf{k})} \tag{4.5}$$

The self-energy is subsequently found by carrying out the contour integration similarly to as in equation (2.35). As usual the self-energy is split up into the exchange and correlation parts

$$\left\langle \varphi_{n\mathbf{q}}^{\sigma} \middle| \Sigma_x^{\sigma} \middle| \varphi_{n\mathbf{q}}^{\sigma} \right\rangle = -\sum_{\mathbf{k}}^{\mathrm{BZ}} \sum_{n'}^{\mathrm{occ}} \sum_{I,J} v_{IJ}(\mathbf{k}) M_{Inn'}^{\sigma\mathbf{kq}*} M_{Jnn'}^{\sigma\mathbf{kq}} \tag{4.6}$$

$$\left\langle \varphi_{n\mathbf{q}}^{\sigma} \middle| \Sigma_c^{\sigma}(\omega) \middle| \varphi_{n\mathbf{q}}^{\sigma} \right\rangle = \frac{i}{2\pi} \sum_{\mathbf{k}}^{\mathrm{BZ}} \sum_{n'}^{\mathrm{all}} \sum_{\mu,\nu} E_{\mu nn'}^{\sigma\mathbf{kq}*} E_{\nu nn'}^{\sigma\mathbf{kq}} \int \mathrm{d}\omega' \frac{W_{\mu\nu}^{c}(k,\omega')}{\omega + \omega' - \varepsilon_{n'\mathbf{q}+\mathbf{k}}^{\sigma} + i\delta \mathrm{sgn}(\varepsilon_{n'\mathbf{q}+\mathbf{k}}^{\sigma})} \tag{4.7}$$

where the overlap matrices $E_{\nu nn'}^{\sigma\mathbf{kq}}$ are defined similarly to equation (4.3), but instead of the MPB the Coulomb eigenbasis is used. The complete frequency dependent self-energy is then obtained and can further be analyzed outside of *SPEX*. It is used to find the quasi-particle energy from equation (2.16), where first order perturbation theory is employed to simplify

the equation. Herein the self-energy is taken as the one in the DFT basis, as usually within first order perturbation theory, instead of $\Psi^\sigma_{n\mathbf{q}}$. The quasiparticle energies providing the bandstructure are then to first order given by [3]

$$E_{n\mathbf{q}} = \varepsilon_{n\mathbf{q}} + \langle \Sigma^x \rangle + \langle \Sigma^c(E_{n\mathbf{q}}) \rangle - \langle v^{xc} \rangle \tag{4.8}$$

with $-v^{xc}$ being present to remove the double counting from the DFT energies.

As the Coulomb matrix v and the screened interaction W diverge at the Γ-point ($\mathbf{k} = \mathbf{0}$), this is treated as a special case due to the important contribution it yields. Reference [26] provides the details on how the Γ-point is treated in the *SPEX* code by splitting up the matrix into body, wing, and head elements and treating these separately.

To minimize the computational effort needed, *SPEX* reduces the set of k-points to the irreducible BZ (IBZ), from which the full BZ can be obtained using the possible space and time-reversal symmetries allowed in the system. In this way the quantities under consideration only have to be calculated for these IBZ points, significantly reducing the number of k-points required. In equations (4.2), (4.6), and (4.7) the integrations must be carried out for a larger number of points due to the additional dependence on $\mathbf{k} + \mathbf{q}$. However, it can efficiently be done by instead using an extended IBZ. In addition, the this version of the code is parallelized to calculate P_{IJ}, $\varepsilon_{\mu\nu}$, and $W_{\mu\nu}$ for each of the k-points in the IBZ individually, since they do not depend on one another. As a consequence the self-energy is also split up into components $\Sigma^{\mathbf{k}}(\mathbf{q})$ for each $\mathbf{k}$ and $\mathbf{q}$ in equation (4.7) [25], with the full self-energy finally being re-summed as $\Sigma(\mathbf{q}) = \sum_{\mathbf{k}} \Sigma^{\mathbf{k}}(\mathbf{q})$ at the end of the calculation for each state $\varphi^\sigma_{n\mathbf{q}}$.

4.2 The self-screening correction in *SPEX*

The modification of the *SPEX* code to incorporate the self-screening correction in section 2.7 is in essence to calculate a new $P^m_{IJ}(\mathbf{k},\omega)$, for a given band in equation (4.2), with the state labelled by m removed. Then for each of these new polarization matrices, $\epsilon^m_{\mu\nu}$ and $W^m_{\mu\nu}$ are calculated and subsequently used in a modified version of equation (4.7) to obtain the self-energy. The expression is altered according to

$$\left\langle \varphi^\sigma_{n\mathbf{q}} \right| \Sigma^\sigma_c(\omega) \left| \varphi^\sigma_{n\mathbf{q}} \right\rangle = \frac{i}{2\pi} \sum_{\mathbf{k}}^{\mathrm{BZ}} \sum_{m}^{\mathrm{all}} \int \mathrm{d}\omega' \frac{W^c_m(k,\omega')}{\omega + \omega' - \varepsilon^\sigma_{m\mathbf{q}+\mathbf{k}} + i\delta \mathrm{sgn}(\varepsilon^\sigma_{m\mathbf{q}+\mathbf{k}})} \tag{4.9}$$

where the W^c_m has been connected with the corresponding g_m, as in equation (2.49); that is the state m for which W^c_m is calculated is the same as in the denominator. The summation over the μ and ν basis as well as the products with $E^{\sigma\mathbf{kq}}_{\mu nm}$ have been omitted for simplicity. Once more it can be noted that only the correlation part is affected by the self-screening correction and the exchange part remains the same as in the original version. This is because all screening effects are taken into account in W^c, while $v = W - W^c$ is the bare interaction for which the screening does not come into effect.

In practice, the bracket in equation (4.2) can be considered as a particle-hole factor, originating from pairs of occupied and unoccupied bands. When calculating the polarization these factors are multiplied by the transformation matrices, and the products are summed over all $\mathbf{q}$-points, bands, and spins in order to give the polarization at $\mathbf{k}$. In order to remove

a given band to obtain P^m, it suffices to remove all of the particle-hole factors involving that given m as either an occupied or unoccupied state. For a spin-polarized calculation this is done by directly setting the given (ph-factor) = 0, while for the non-polarized case the factor is instead set to 1/2, as the contribution from the other spin is equivalent to the one removed. This approach eliminates the need to explicitly do the sum over spins in the latter case, as would otherwise have been necessary, and instead retains that the sum can be substituted for a factor of 2, i.e. $\sum_\sigma = 2$.

The dielectric matrix and screened interaction are thereafter calculated for each of the P^m according to equations (4.4) and (4.5). Extra care has to be taken when calculating the contribution from the Γ-point, as stated in the previous section. As a consequence of the scheme, the aforementioned head and wing elements also would need to be treated for each of the bands, however currently this has not been implemented and only the body elements are treated properly; the non-corrected elements are presently used instead for the wing and head elements.

Due to that the P^m, ϵ_m, ϵ_m^{-1}, and W_m arrays have to be saved for each of the bands, the memory and time requirements quickly become huge as many unoccupied states are necessary for a correct description. A way to reduce the computational effort significantly is therefore to reduce the correction to only encompass a few bands, and the modifications were implemented with this in mind. When doing calculations one begins by first identifying the n_{ssc} states to be treated with the self-screening correction from the DFT results. Then the self-screening corrected version of *SPEX* is set to create $n_{ssc} + 1$ different P_m, with $m = 0$ denoting the uncorrected case. When spin-polarization is used, the number of n_{ssc} which have to be calculated are twice the number of bands removed, as $P^{m\uparrow} \neq P^{m\downarrow}$ in this case, although through the rest of this work we will limit ourselves to non-polarized systems. Calculating ϵ_m, ϵ_m^{-1} and W_m are done by looping over the routines using the $n_{ssc} + 1$ different quantities. Finally when doing the summation for the self-energy the different W_m are connected with their corresponding g_m, as usual in GW-ss. However, now the bands n which are not treated with the self-screening correction will all use the normal screened interaction $W_n = W_0$ from GWA. The expression schematically takes the form

$$\Sigma = \Sigma^x + \sum_{n'} g_{n'} W^c_{n'} \tag{4.10}$$

$$W^c_{n'} = \begin{cases} W_m & \text{if } n' = m \text{ subject to self-screening correction} \\ W_0 & \text{if } n' \text{ not subject to self-screening correction} \end{cases} \tag{4.11}$$

Since the self-screening is expected to be largest for localized states, a choice can be made to only calculate the self-screening corrected W_m for these states, while all other retain the usual W. For instance, GaAs is a good candidate to employ this approximate scheme too. As can be seen in Figure 4.1 there are narrow semicore bands located around -15 eV, relative to the highest occupied state, in the DFT bandstructure obtained from *FLEUR* (details on the input used for these calculation are given in Section 4.3). It is known that narrow bands correspond to localized states, and by analyzing the corresponding density of states (DOS) in Figure 4.1 it is is indeed seen that the band originates almost completely from the $3d$ orbitals. Employing the approximate self-screening correction can then be done by calculating P^m for

the $m = 1, \ldots, 5$ bands from the localized $3d$ orbitals, with the rest of the bands using the common P calculated for $m = 0$.

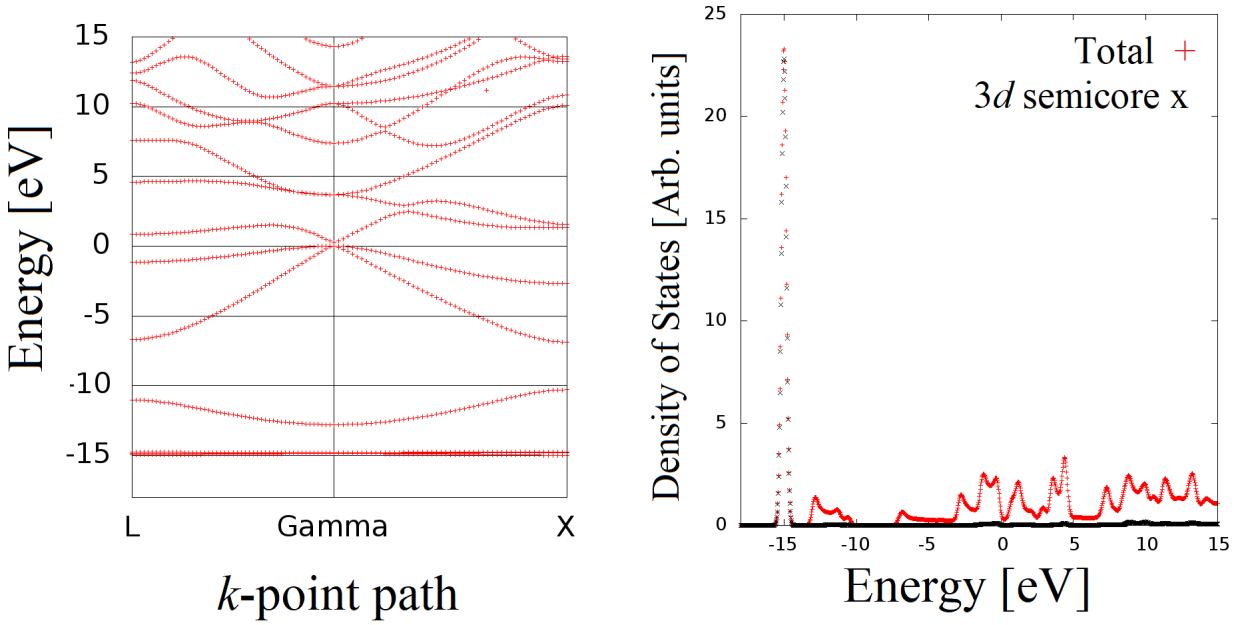

Figure 4.1: Left: GaAs Bandstructure using the DFT code *FLEUR*. Right: GaAs density of states using *FLEUR*, with the contribution from the $3d$ orbitals marked out.

It is of course possible to also include more of the higher lying bands using the correction, as was also done in additional computations for bands to just above the Fermi level for comparison. The elements the scheme was applied to are given in the next section. For the semiconductors considered the properties investigated are the positions of semi-core states, by calculating their removal energies relative the lowest unoccupied states, and the width of the band gap. The results obtained using the modified code are presented in Section 5.1.3 where they are compared to experimental values, and those calculated within GWA.

4.3 Material data

The only input which is required for an *ab initio* calculation is information about the unit cell of the material in question, in the form of the lattice constant a_0 (in atomic units), the primitive lattice vectors $\vec{a}_i$ for the given crystal structure, and the basis. The basis details the atomic positions and their respective atomic numbers, Z, also providing the number of electrons. In the current work calculations using the self-screening correction scheme, as outlined in the previous section, have been performed for four bulk semiconductors; GaAs, ZnSe, Ge, and ZnO. The input data used is given in Table 4.1 and was taken from References [29, 30].

The diamond structure is a face-centered cubic lattice with a two atom basis. It has

Material	Structure	a_0	c
GaAs [a]	Zincblende	10.68	
ZnSe [a]	Zincblende	10.71	
Ge [a]	Diamond	10.70	
ZnO [b]	Wurtzite	6.14	9.828

Table 4.1: Crystal structure and lattice constant a_0 in atomic units for the materials used in this work. The c parameter is also given for wurtzite ZnO. The structure is further detailed in the text. [a] Ref. [29], [b] Ref. [30]

primitive lattice vectors

$$\begin{aligned} \vec{a}_1 &= (0.0, 0.5, 0.5)a_0 \\ \vec{a}_2 &= (0.5, 0.0, 0.5)a_0 \\ \vec{a}_3 &= (0.5, 0.5, 0.0)a_0 \end{aligned} \tag{4.12}$$

and the basis atoms are separated with a displacement of $a_0/4$ along the diagonal, placed at the origin. For Ge the basis was chosen as:

$$\begin{aligned} Z_1 = 32,\ (0.125, 0.125, 0.125)a_0 \\ Z_2 = 32,\ (-0.125, -0.125, -0.125)a_0. \end{aligned} \tag{4.13}$$

The zincblende structure is the same as the diamond structure except that the basis consists of two different elements, causing each type to be surrounded by four nearest neighbours of the other element. For both GaAs and ZnSe the basis was chosen as:

$$\begin{aligned} Z_1,\ (0\ ,0, 0) \\ Z_2,\ (0.25, 0.25, 0.25)a_0. \end{aligned} \tag{4.14}$$

where $Z_1 = 31$, $Z_2 = 33$ for GaAs and $Z_1 = 30$, $Z_2 = 34$ for ZnSe. For the ZnO calculations the example input file provided by the *FLEUR*-project found in Ref. [30] was used. For a detailed description of the wurtzite structure see e.g. Ref. [31].

Section 5

Results and discussion

In the following, the results obtained during this work will be presented and discussed. First the model calculations will be analyzed, where the validity of the GW-ss and GW-sp schemes will be compared to the common GW approximation in different parameter ranges. In this part the parameters are given in units of $t = 1$. Thereafter the *Ab initio* results for the band structure of bulk semiconductors, for which the unit cells are given in Section 4.3, will be presented and compared to experimental values obtained from the literature.

5.1 Model calculations

5.1.1 One-orbital system

To analyze the effect of the two schemes the excitation energies of the system can be found from the spectral function of the response, S, which is straightforwardly obtained from the response functions R, R^{RPA}, $R^{ss}_{m\sigma}$, and R^{sp}. As previously argued, the self-screening corrected response functions are only auxiliary quantities and are merely $R^{ss}_{A\sigma} = R^{sp}/2$ for this model. Therefore only the self-polarization corrected one will be compared to the exact and RPA

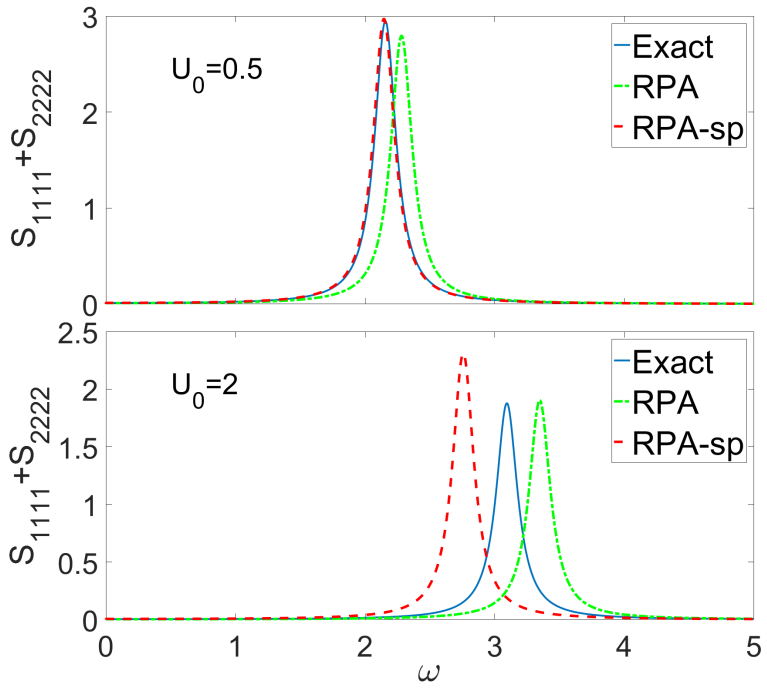

Figure 5.1: The trace of the spectral part of the response function for the exact (blue, full line), RPA (green, dash-dotted line), and RPA with self-polarization correction (red, dashed line). Upper: $t = 1$, $U_0 = 0.5$, $U_1 = 0.2$, lower: $t = 1$, $U_0 = 2$, $U_1 = 0.2$.

response functions. In Figure 5.1 the traces of the spectral function of the response, $S_{1\,\mathrm{orb}} = S_{AA,AA} + S_{BB,BB} = S_{11,11} + S_{22,22}$ which is independent of basis, are plotted for $U_0 = 0.5$ and 2, and $U_1 = 0.2$.

For a small difference, $(U_0 - U_1) = 0.3$, it is clear that the self-polarization correction is almost identical to the exact result, while the RPA response provides an incorrectly, too large, excitation energy. When $(U_0 - U_1)$ is increased the error within RPA increases while the self-polarization correction worsens and gives a too small excitation energy. As can be seen in Figure 5.1, around $(U_0 - U_1) = 1.8$ RPA is just slightly better at describing the peak position than when using the self-polarization correction, although the errors are of opposite sign.

As the peak structure of S provides the excitation energies, so does the poles of R. We have obtained the position of these analytically in Section 3.1 as Δ_{no} for RPA, Δ_{ss} for self-screening and self-polarization, and $\Delta_{\mathrm{exact}} = \varepsilon_3 - \varepsilon_1$ for the exact case. These are plotted for a fixed $U_1 = 0.2$ in Figure 5.2 as a function of U_0. It is clear that for small $(U_0 - U_1)$ the excitation energy calculated using the correction schemes is almost identical to the exact one. Although RPA is in better agreement for $(U_0 - U_1) > 1.6$ and in decent agreement also for smaller values, this is more an artifact of that it first underestimates the energy for negative $(U_0 - U_1)$ and thereafter overestimates it for a range of U_0, whereas the corrected ones follow closely the exact curve for small values, whereafter it immediately underestimates the energy.

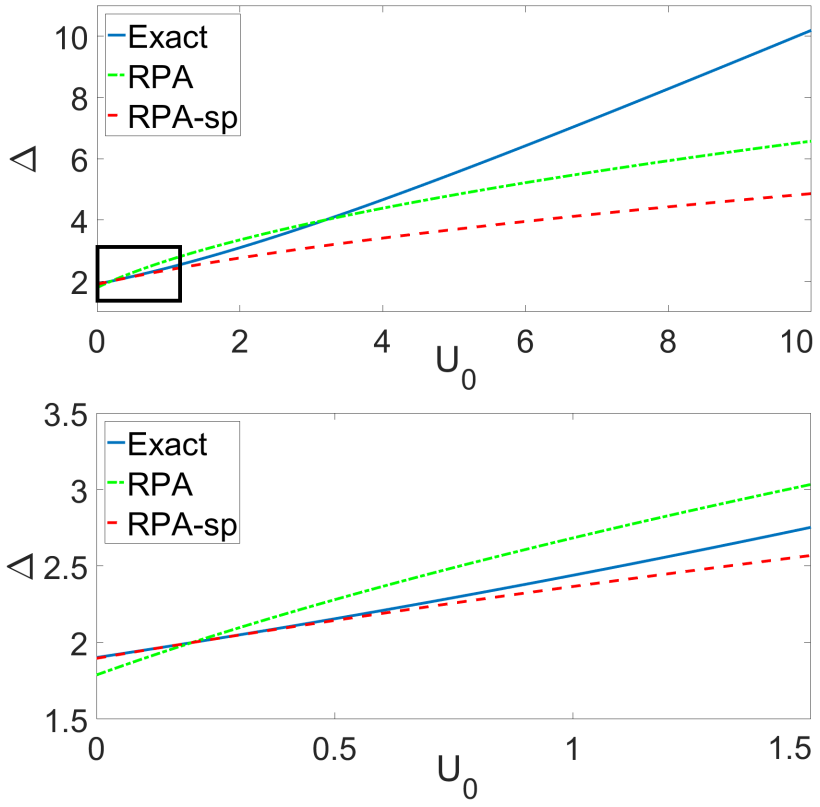

Figure 5.2: Excitation energy Δ for the exact (blue, full line), RPA (green, dash-dotted line), and self-polarization corrected (red, dashed line) cases plotted as a function of U_0 for a fixed $U_1 = 0.2$. The lower figure shows a zoom in of the box in the upper one.

Also the HOMO-LUMO gap, ΔE, for the full range of interaction strengths can provide insight into how well the corrected schemes work. By plotting the gap in Figure 5.3, for a fixed $U_1 = 0.2$, it is seen that the factor of two by which GW-ss and GW-sp differ comes into effect. GW-ss proves to be significantly better than the rest for small $(U_0 - U_1)$, as was indicated by expanding the gap in Section 3.1 whereupon the exact result was obtained, and almost perfectly describes the behaviour of it. Up to $U_0 \approx 1$ the self-screening correction provides the best agreement, and in-between $U_0 \approx 1$ and $U_0 \approx 2.6$ GWA is in best agreement in the absolute value of the gap. Even though GW seems to be better, it is not because it describes the trend of the gap for increasing U_0 well, but rather due to first overestimating and then subsequently underestimating it, providing an accidental improvement in the region. After $U_0 \approx 2.6$ GW-sp clearly gives the best agreement with the exact result. In this case it is also evident that GW-sp overestimates the gap for a longer duration than both GW and GW-ss and thereafter underestimates it. However, a distinction compared to the accidental improvement seen for GW becomes clear by further increasing U_0 to exaggerated values, as can be seen in Figure 5.4. It is here visible that GW remains slightly better than GW-ss, in agreement with the previous results for larger $(U_0 - U_1)$ obtained for S above. It nevertheless

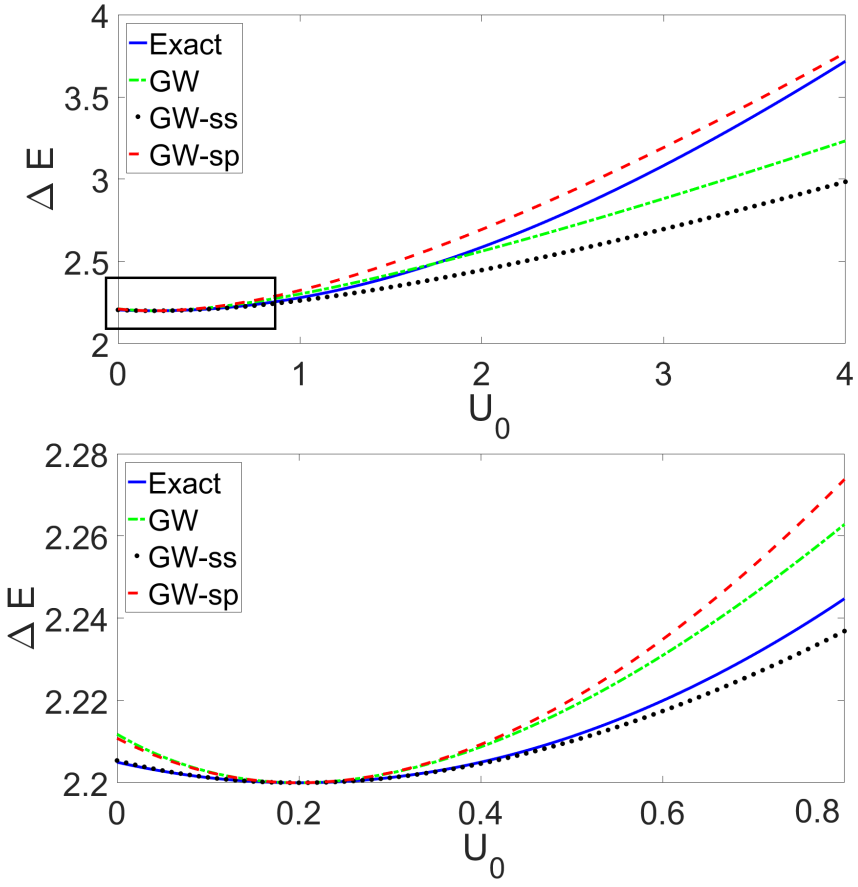

Figure 5.3: The HUMO-LUMO gap plotted as a function of U_0 for the exact (blue, full line), GW (green, dash-dotted line), self-screening corrected GW (black, dotted line) and self-polarization corrected (red, dashed line) cases for fixed $U_1 = 0.2$. The lower figure shows a zoom-in for small U_0 marked out by the box in the upper figure.

becomes clear that GW-sp gives a significantly better description of the trend of the gap for large values of U_0, while for small $(U_0 - U_1)$ being approximately of the same accuracy as that of GW. The improvement for GW-sp is therefore probably not only due to the accidental crossing which GW was attributed with, but also from the fact that it seems to give a better picture of the behaviour of the gap at very large U_0 in general. This is verified by expanding ΔE for $U_0 \gg 1$, where it is found that both the exact and GW-sp provide $\Delta E^{\text{exact}} = \Delta E^{GW-\text{sp}} \approx U_0$ while GW and GW-ss goes as $U_0/2$. It is thus clear that GW-sp provides a qualitatively correct description of the gap for large U_0, something the other schemes fail to do. This could potentially be a very important feature if it also extends to more complicated systems, however, it has not been investigated for the more complex two-orbital model in this work.

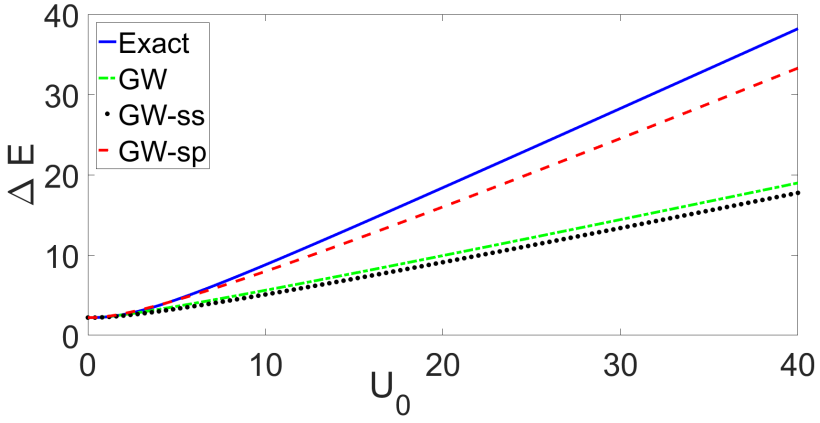

Figure 5.4: The HUMO-LUMO gap plotted for exaggerated values of U_0 for the exact (blue, full line), GW (green, dash-dotted line), GW-ss (black, dotted line) and GW-sp (red, dashed line) for fixed $U_1 = 0.2$.

Finally, also the spectral functions A using the renormalized G, calculated within GW, GW-ss and GW-sp, can be compared to the exactly calculated one. The effect is limited for the main peak in the regime of small U_0, yet GW-ss and GW-sp seem to improve the placement of the satellite features by about half of the error in GW, which can be seen in Figure 5.5. Increasing U_0 maintains the improved description of the satellite features for both correction schemes, while the placement of the GW-ss main peak worsens slightly compared to GW. However, similarly to as for the HOMO-LUMO gap the GW-sp scheme seems to also provide the best description of the $(N \pm 1)$ excitation energies in this regime. Note that the heights of the peaks are not directly comparable as the $i\delta$ broadening has been chosen differently for the exact and approximate schemes, where A for the latter are calculated using equation (2.33) to properly account for the shift. Their integrated weights on the other hand are the same, as they should due to it being related to the number of particles in the system according to equation (2.12).

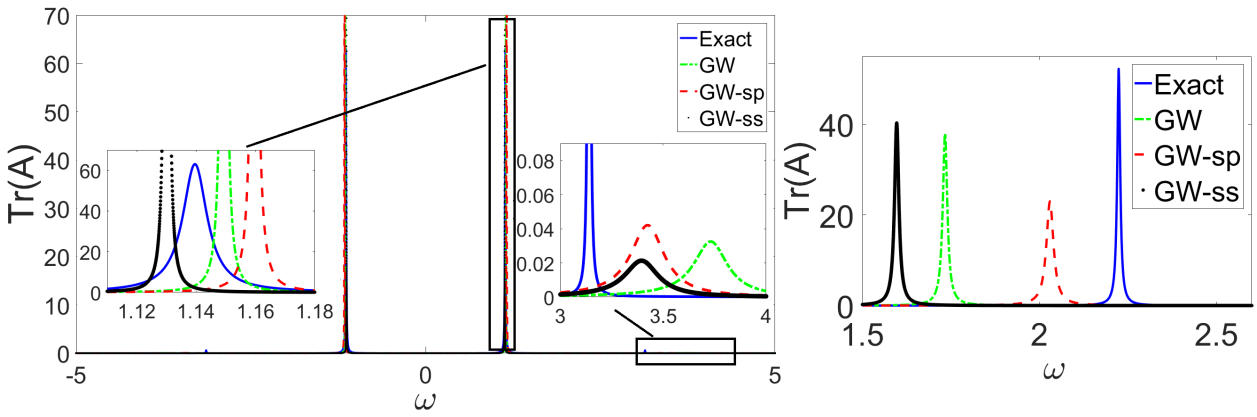

Figure 5.5: The trace of the spectral function A for the exact (blue, full line), GW (green, dash-dotted line), GW-ss (red, dashed line) and GW-ss (black, dotted line) cases. Left: $U_0 = 1$, $U_1 = 0.2$. The left inset is a zoom-in of the marked main peaks, and the right inset is of the satellite features. Right: The main peaks for $U_0 = 5$, $U_1 = 0.2$.

5.1.2 Two-orbital system

The two-orbital system is also analyzed for all four cases: the exact one, normal RPA and GW, self-screening corrected RPA and GW, and self-polarization corrected RPA and GW. The effects of the schemes will be compared by observing how they affect the matrix elements of S, which gives the excitation energies in the system, as well as A, giving the excitation energies for the $N \pm 1$ electron systems. Due to being a more complicated system with additional variable parameters (U_0, t_2, t_3, ϵ), all parameters except one will be kept fixed at a time to investigate their respective effects.

In Figure 5.6 U_0 is varied between 0.5,1 and 2 with $\epsilon_1 = 2$, $t_2 = 0.2$ and $t_3 = 0.5$ being kept fixed for the $S_{11,11} + S_{22,22}$ elements. It is here seen that, similarly to the one-orbital model, the self-polarization correction and self-screening correction give very good agreement for small interaction strengths U_0. Both the main peak and the satellite are well-described for these schemes, while RPA only captures the satellite in a satisfying manner. In addition, also the weight of the main excitation peak is improved when using the self-polarization correction. Furthermore we once again see the auxiliary status of the $R_{m\sigma}$ calculated using the self-screening correction scheme. $R_{1\sigma}$, which is plotted here, is approximately halved as compared to both the exact and the self-polarization cases. This occurs, as explained in Section 3.2, because all excitations for a given spin are removed from the polarization function when the ground state, $m = 1$, is removed within the scheme. The only contribution therefore comes from the other spin and as the system is not spin-polarized the two spin channels provide an equal weight, leading to the weight becoming halved. If instead another $g_{m\sigma}$ was chosen, the effect would be seen on the peak corresponding to the given excitation, which then is approximately halved, while the rest retain their weights from both spins.

As U_0 is increased, the agreement of the position of the main excitation peak worsens for the self-screening and self-polarization schemes. For $U_0 = 1$ the agreement is still slightly better than the overestimaton found for RPA, but similarly to the one-orbital model the excitation energy is instead underestimated. For both schemes, however, the weights for

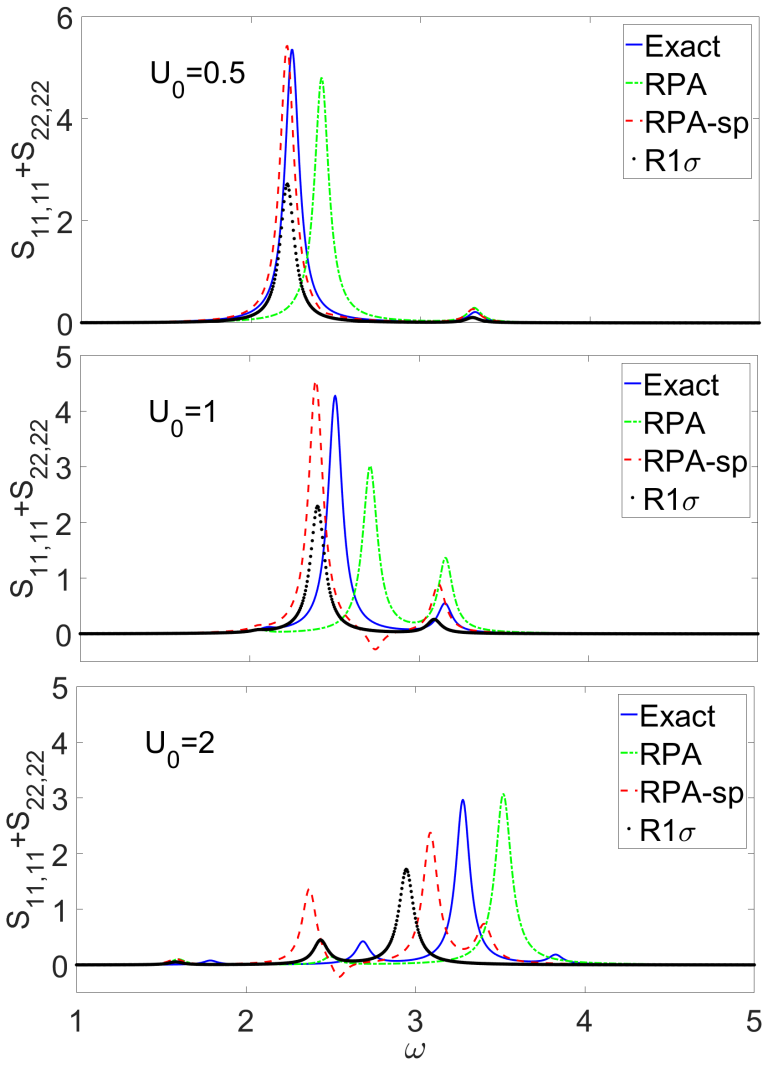

Figure 5.6: $S_{11,11}+S_{22,22}$ plotted as a function of ω for the exact (blue, full line), RPA (green, dash-dotted line), $R_{1\sigma}$ from the self-screening corrected RPA (black, dotted line), and self-polarization corrected RPA (red, dashed line) cases for varied U_0 and fixed $\epsilon_1 = 2$, $t_2 = 0.2$ and $t_3 = 0.5$. Upper: $U_0 = 0.5$. Middle: $U_0 = 1$. Lower $U_0 = 2$.

both peaks present are significantly better. Increasing U_0 further the peak structure becomes more complicated, and already for $U_0 = 2$ it is difficult to draw any conclusion as to which describes the structure the best. The placement of the main excitation peak seems to be as good within RPA as it is for the self-polarization scheme, while the weight seems to favour RPA. The self-screening scheme gives a too low excitation energy, but a better weight as it still is only roughly halved compared to the exact response. Increasing U_0 even more than what is shown here, it is found that the correction schemes give worse excitation energies than RPA, although all at poor agreement with the exact result. The case is similar to the one-orbital case though, where RPA now underestimates the excitation energy instead of overestimating it, as for smaller U_0. While RPA and the self-screening scheme only give rise to three visible peaks, it is found that the self-polarization one also predicts the higher energy

excitation found in the exact case, although at a lower energy. As the weight is significantly larger than the exact peak it is likely that too much of the weight from the main peak has been transferred here, as well as to the too large peak at lower energy. An important issue with the self-polarization scheme can however be noted both for $U_0 = 1$ and 2: S, which should be strictly positive for $\omega > 0$, is found to be negative in a short range. The scheme therefore has problems with causality, which need to be addressed.

Varying t_2 instead in Figure 5.7, it is seen that an increase in the parameter causes both RPA and the correction schemes to provide better estimates of the excitation energy of the main peak, as well as the peak weights. The improvement, however, is more pronounced

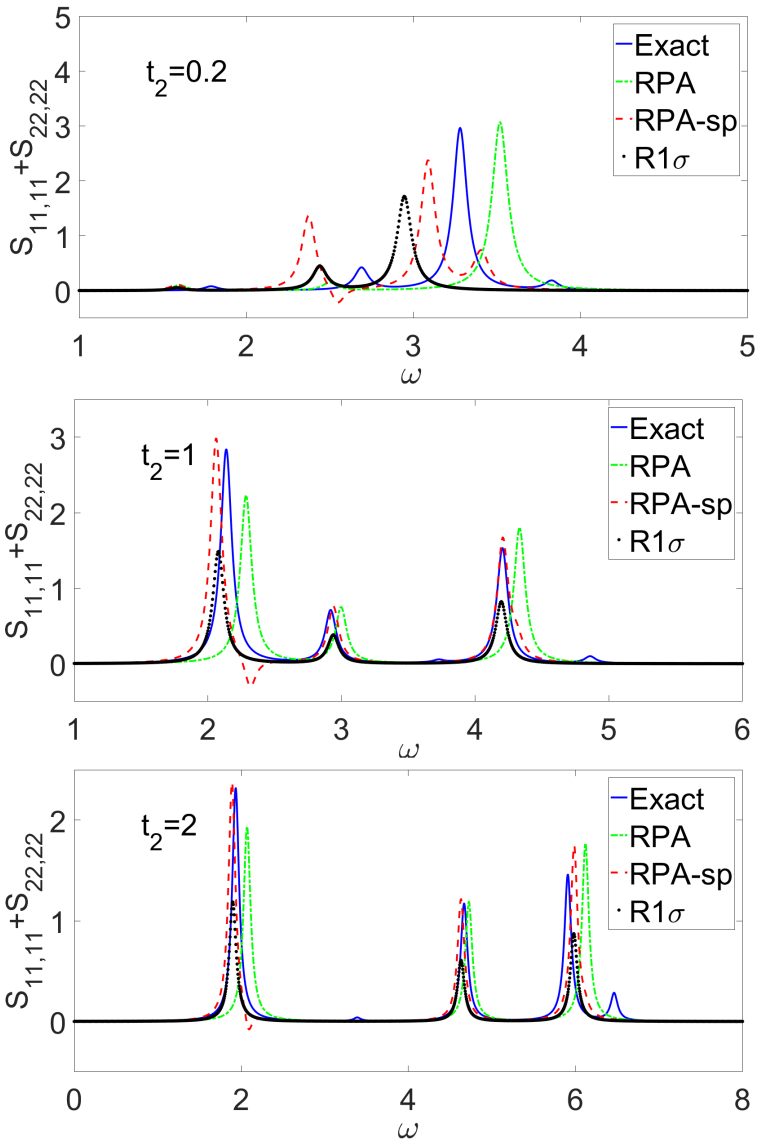

Figure 5.7: $S_{11,11}+S_{22,22}$ plotted as a function of ω for the exact (blue, full line), RPA (green, dash-dotted line), $R_{1\sigma}$ from the self-screening corrected RPA (black, dotted line), and self-polarization corrected RPA (red, dashed line) cases for varied t_2 and fixed $U_0 = 1$, $\epsilon_1 = 2$, and $t_3 = 0.5$. Upper: $t_2 = 0.2$. Middle: $t_2 = 1$. Lower $t_2 = 2$.

for the correction schemes, indicating that the increase in hopping strength between the lower and higher lying orbitals effectively mitigates the effect of the interaction. This is reasonable physically, since allowing the electrons to occupy the higher lying orbitals to a larger extent would on average decrease the interaction experienced by an electron, as there only is interaction in the lower ones. Furthermore, for large t_2 the negative spectral part of the response diminishes and will for $t_2 = 4$ be completely removed. By instead decreasing t_2 the negative weight increases. The correction to the three large peaks have almost shifted them to the exact result, showing an improvement over RPA in the structure, albeit still lacking the two smaller peaks present in the exact response.

Likewise when instead t_3 is increased, as shown in Figure 5.8, the effect is an improvement for the correction schemes. For $t_3 = 2$ the main excitation peak and the lower satellite feature are almost correctly described, a considerable improvement over RPA. The remaining weight

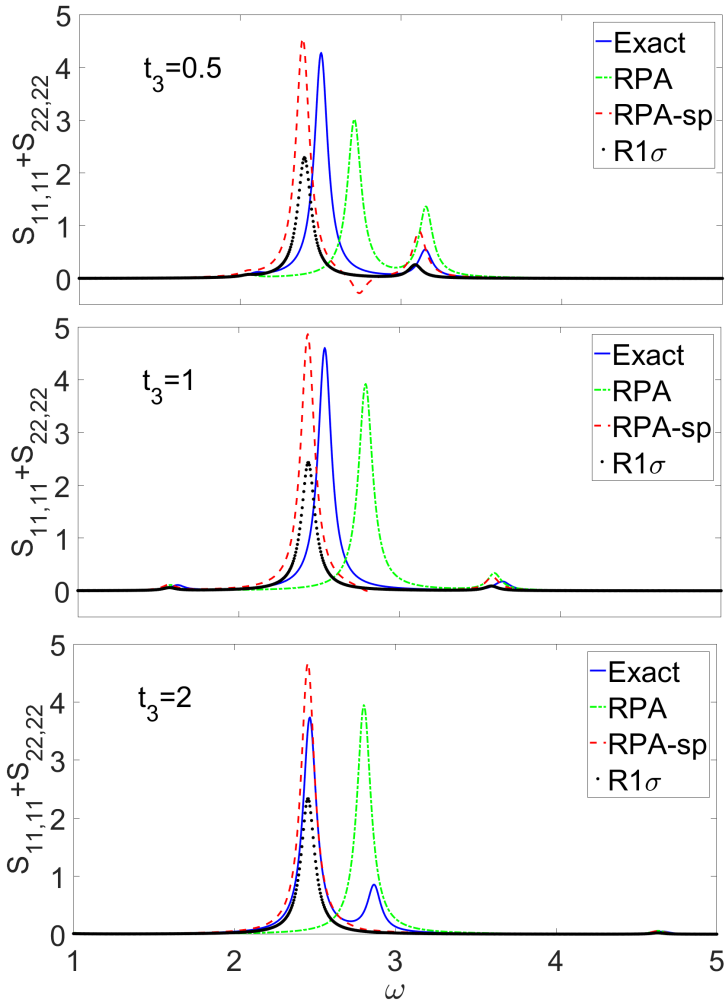

Figure 5.8: $S_{11,11}+S_{22,22}$ plotted as a function of ω for the exact (blue, full line), RPA (green, dash-dotted line), $R_{1\sigma}$ from the self-screening corrected RPA (black, dotted line), and self-polarization corrected RPA (red, dashed line) cases for varied t_3 and fixed $U_0 = 1$, $\epsilon_1 = 2$, $t_2 = 0.2$. Upper: $t_3 = 0.5$. Middle: $t_3 = 1$. Lower $t_3 = 2$.

difference for the main peak could possibly originate in that the excitation just above the main one, which is not captured with the approximate methods, has a part of this weight. In addition the causality issues from the negative spectral part of the response is cured, while for $t_3 = 1$ it is present but almost negligible. The improvement could as before be seen physically by that the increased hopping in-between the higher lying orbitals makes it energetically beneficial to move away from the lower orbitals with interaction strength U_0, thereby effectively decreasing the mean interaction.

The final parameter, ϵ_1, is varied in Figure 5.9 and does not seem to improve or worsen the agreement to any larger extent. The largest effect is instead to remove other excitations than the main one, and limiting the effect of varying other parameters due to the increased gap. The increased gap between the orbitals causes an equivalent gap in the energy of the states between the lowest two and the highest two, when observing from the mean-field point of view, decreasing the probability of occupying the higher states as this is less favourable energetically. Despite the apparent lack of improvement with increasing ϵ_1, the aforementioned limited effect of varying other parameters such as U_0 should be taken into account. By more carefully studying the excitation energy of the main peak, it is found that the increase in ϵ_1 marginally worsens the excitation energy within RPA. This is expected as the effect of localizing the electrons more in the lower orbitals effectively would make the electrons feel a larger interaction, which is the regime RPA is known to work worse for. When

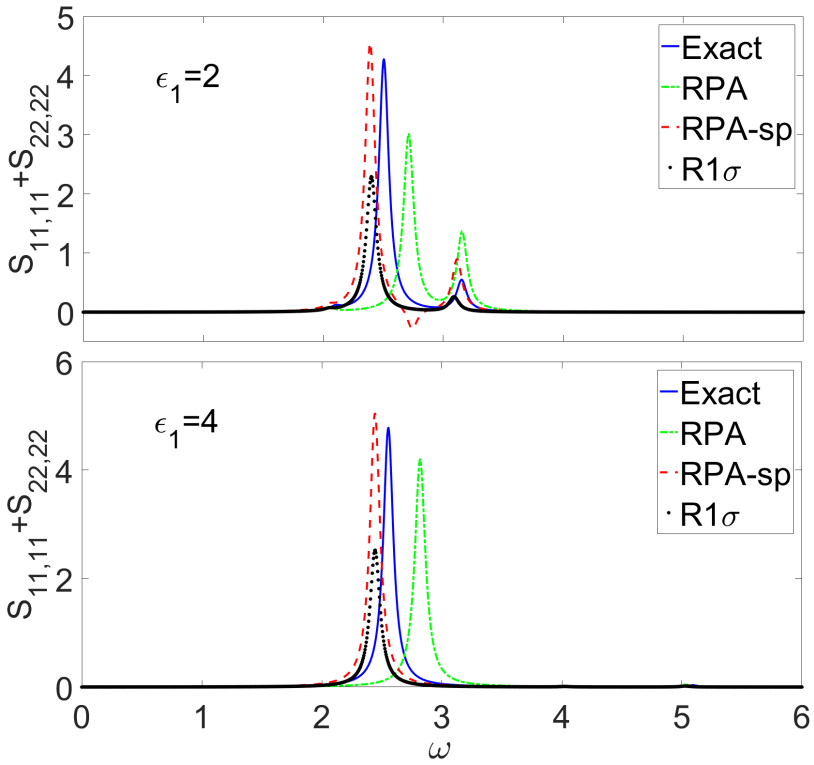

Figure 5.9: $S_{11,11}+S_{22,22}$ plotted as a function of ω for the exact (blue, full line), RPA (green, dash-dotted line), $R_{1\sigma}$ from the self-screening corrected RPA (black, dotted line), and self-polarization corrected RPA (red, dashed line) cases for varied ϵ_1 and fixed $U_0 = 1$, $t_2 = 0.2$ and $t_3 = 0.5$. Upper: $\epsilon_1 = 2$. Lower $\epsilon_1 = 4$.

observing the position of the excitation within the self-polarization scheme it is found that the increase does not worsen the result, albeit not improving it either. This could be taken as evidence for that the self-screening still is a non-negligible part of the error for relatively localized states, although not the major contribution. In the regime of large U_0 the RPA positions are still better, yet both underestimate the peak position once more and it could thus potentially be explained with the accidental improvement observed in the one-orbital model.

Also the renormalized Green's function is compared in Figure 5.10 where $\mathrm{Tr}(A)$ has been plotted, note however that equation (2.11) is used, implying that the shift is not properly taken into account. However, the general trends for both forms of A were found to agree for the one-orbital model and the qualitative picture are therefore believed to be the same here as well, although the absolute peak positions are incorrect. The effect of the correction schemes are less visible here than for the response but a few features can still be noted. Firstly, the main peaks are very slightly improved using the self-polarization correction over the other other two approaches for small U_0, with the self-screening correction seemingly being the worst. The effect of the $[\mathbb{1} - G^0\Sigma]^{-1}$ factor in calculating $G = [\mathbb{1} - G^0\Sigma]^{-1}G^0$ is twofold; it shifts the main peaks and gives rise to additional satellite features. Since G^0 remains

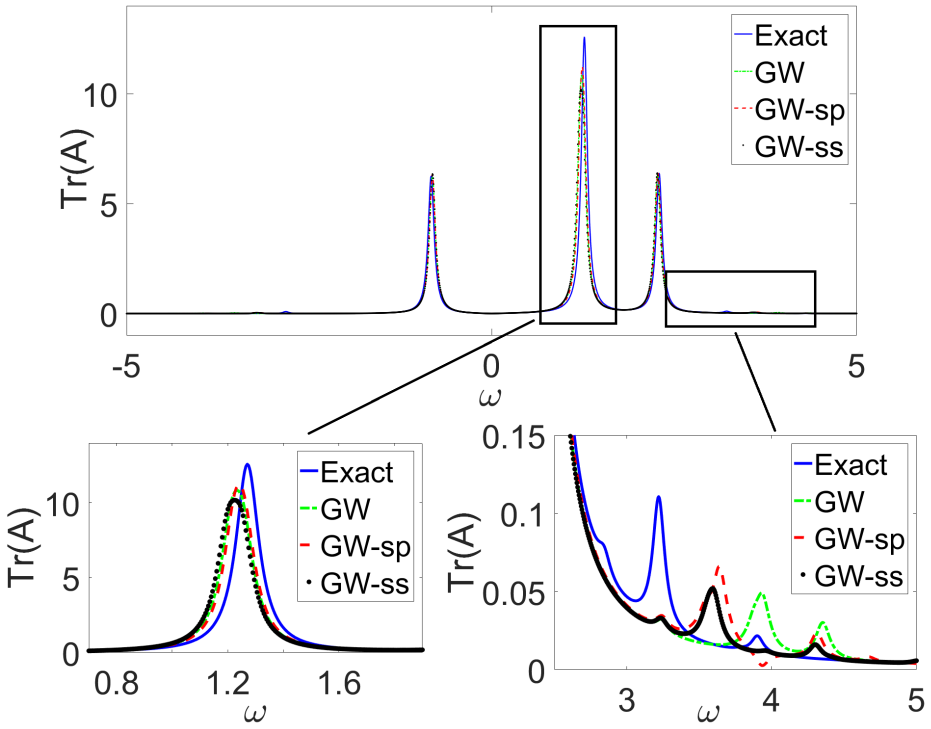

Figure 5.10: $\mathrm{Tr}(A)$ plotted as a function of ω using the exact (blue, full line), GW (green, dash-dotted line), GW-sp (red, dashed line), and GW-ss (black, dotted line) results for $U_0 = 1$, $\epsilon_1 = 2$, $t_2 = 0.2$ and $t_3 = 0.5$. Lower left: Zoom-in of the marked peak. Lower right: Zoom-in of the marked satellite features.

unaffected by the schemes, it is reasonable that the main peaks remain largely unaffected in the regime of small U_0. The satellites are on the other hand more interesting to compare

and provide the second major feature to be compared for the schemes. They are all shifted to higher energies compared to the exact case, or lowered in the $\omega < 0$ part of the spectrum, and it is readily seen that both of the correction schemes give significant improvements to their positions compared to GW. Finally, the issue of causality once again appears for the self-polarization case, where A has a slightly negative part around the lower satellite features not shown in the figure. For larger interaction strengths the main peaks are shifted more compared to the exact result as would be expected, while the satellite features become increasingly difficult to analyze. Smaller interaction strengths instead give a negligible shift of the main peaks, but retain the improved positions of the satellite features described above.

5.1.3 *Ab initio* results for GaAs, ZnSe, Ge, and ZnO

To evaluate the self-screening correction scheme for real materials its effects on the band structure was investigated. The results of these computations for the semiconductors GaAs, ZnSe,Ge, and ZnO are presented in this section. The properties which were investigated were the band gap and the position of the semicore states, and a comparison is made between LDA, GW, GW-ss with only the $3d$ semicore states treated, and GW-ss with additional bands treated to above the Fermi energy. The semicore positions were measured in two ways to distinguish different effects; relative to the valence band maximum (VBM) and relative to the conduction band minimum (CBM) at the Γ-point. The values stated as the experimental energy positions relative to the CBM are obtained by adding the VBM experimental value with the band gap. Although the $3d$ bands are mostly dispersionless, there is in fact a splitting between some of them and the values stated here as the semicore energies are the average for the bands at the Γ-point.

In Table 5.1 the results from the calculations for GaAs are presented. The computations for the correction scheme were carried out for the five Ga $3d$ semicore states, as well as for the $3d$ and the 15 higher lying bands (20 in total), where the valence and conduction bands are included. As expected, the band gap is significantly improved within all of the GW schemes compared to LDA, and is in agreement with other reported values [26]. However, the band gap was improved even further by also including the valence and conduction bands in the correction scheme, almost correctly predicting the experimental value. This result

GaAs	LDA	GW	GW-ss (5: $3d$)	GW-ss (20)	Experiment
Band gap	0.30	1.38	1.38	1.49	1.52 [1]
Semicore (VBM)	14.84	16.53	16.64	16.56	18.8 [2]
Semicore (CBM)	15.14	17.92	18.02	18.05	20.3 *

Table 5.1: Band gap and energy levels of the semicore Ga $3d$ states (in eV) in GaAs for LDA, GW, GW-ss with only the $3d$ bands removed, and GW-ss with 20 bands removed. [1] Ref. [32], [2] Ref. [33]. * The experimental CBM value is obtained by adding the VBM with the band gap.

was initially unexpected as these are more extended states, and were as such believed to not be as affected by the self-screening. On the other hand, the desired improvement in the semicore states was not observed and the energies were only improved by about 0.1 eV for

$n_{\text{ssc}} = 5$. The inclusion of additional states in the correction scheme ($n_{\text{ssc}} = 20$) seems to remove this improvement when the VBM is used as the reference point, however, if instead the CBM is used as reference the improvement is slightly larger. This indicates that the inclusion of these bands in the scheme lowers energy of the valence band and increases the energy of the conduction band, in total leading to the improved band gap. This would seem to be in-line with the known fact that band gaps within HFA are too wide, since removing the self-screening brings it closer to the unscreened picture provided by Hartree-Fock. The obtained results for the 3*d* states within *GW* are about 2 eV lower than those found in [5]. This could, at least in part, be attested to using different exchange-correlation functionals providing a slightly better LDA starting point of 15.4 eV relative the VBM, but it is likely also due to the difference in the methods used for the *GW* calculations. The results are more in agreement with the 16.9 eV reported in [26], where another version of the *SPEX* code was used, indicating that the value obtained here is indeed accurately calculated.

Similarly the band gap for ZnSe, found in Table 5.2, is seen to significantly improve over the LDA value using the *GW* schemes. The gap is further improved by the inclusion of the additional higher lying states, as for GaAs, and almost halves the error in this case. Also here the effect on the semicore states is on the order of 0.1 eV, far from the 2 eV discrepancy which *GW* displays compared to experiment. The same tendency is once more found when comparing the results using the VBM or CBM as reference points: the $n_{\text{ssc}} = 20$ calculation is worsened and improved, respectively, relative the *GW*-ss calculation where only the Zn 3*d* bands are removed. Also for ZnSe the *GW* result for the semicore states in [5] is better than those found here by approximately 1.2 eV.

ZnSe	LDA	*GW*	*GW*-ss (5: 3*d*)	*GW*-ss (20)	Experiment
Band gap	1.04	2.45	2.45	2.60	2.8 [3]
Semicore (VBM)	6.55	7.39	7.47	7.39	9.2 [2]
Semicore (CBM)	7.59	9.83	9.92	9.99	12.0 *

Table 5.2: Band gap and energy levels of the semicore Zn 3*d* states (in eV) in ZnSe for LDA, *GW*, *GW*-ss with only the 3*d* bands removed, and *GW*-ss with 20 bands removed. [3] Ref. [34], [2] Ref. [33]. * The experimental CBM value is obtained by adding the VBM with the band gap.

Germanium is known to have an indirect band gap [32], and the semicore states were therefore only calculated with the VBM as reference at the Γ-point. Furthermore, the unit cell contains two Ge atoms, giving in total 10 Ge 3*d* bands to be treated with the correction. As before, there is a slight improvement in the energy of the semicore levels, and the worsened result for the $n_{\text{ssc}} = 20$ case is once more attributed to the lowering of the valence band. LDA does not predict any band gap and finds it to be a metal, although only with a small overlap of the bands around the Fermi level. It is, however, correctly described as a semiconductor using *GW*, with an indirect band gap where the CBM is at the L point and the VBM at the Γ-point. The band gap obtained is already in very good agreement with the experimental value for *GW*, and it is furthermore in agreement with other reported values [26]. The application of the correction to the states around the Fermi level is seen to produce a slightly worsened value for Ge on the other hand. This is attributed to that the correction removes

the self-screening error, which is still present although cancelled by other effects in GW, causing the band gap to widen regardless of the already good description.

Ge	LDA	GW	GW-ss (10: $3d$)	GW-ss (20)	Experiment
Band gap	–	0.72	0.72	0.80	0.74[1]
Semicore (VBM)	24.75	26.94	26.99	26.93	29.6[4]

Table 5.3: The indirect band gap and energy levels of the semicore $3d$ states (in eV) in Ge for LDA, GW, GW-ss with only the $3d$ bands removed, and GW-ss with 20 bands removed. [1] Ref. [32], [4] Ref. [35].

For ZnO only the LDA, GW, and GW-ss with $n_{\text{ssc}} = 20$ computations where carried out due to time constraints. However, as the results follow the previously described trend of a worsened semicore energy with the VBM as a reference and an improved one if the CBM instead is used, it is very likely that there would be an improvement in the Zn $3d$ semicore states of ZnO as well if the correction was limited to them. It can be noted that the Zn $3d$ states are not well separated from an O 2p state in LDA, which is a known issue [4]. The average was therefore taken to include this O 2p state as well, although it should not make any difference as they are close in energy. The position of the semicore states calculated within GW is in good agreement with that found in [4] at 6.4 eV relative the VBM, and once more the effect of the correction scheme on the localized states is almost negligible. It can further be noted that the O 2p state separates in the GW calculations, enabling them to be averaged separately. The band gap is once again widened, although the discrepancy for GW is too large for the self-screening correction to correct it.

ZnO	LDA	GW	GW-ss (20)	Experiment
Band gap	0.76	2.63	2.76	3.44 [1]
Semicore (VBM)	5.27	6.31	6.30	7.5, 8.8 [2]
Semicore (CBM)	6.48	8.94	9.07	10.9, 12.2 *

Table 5.4: Band gap and energy levels of the semicore $3d$ states (in eV) in ZnO for LDA, GW, and GW-ss with 20 bands removed. [1] Ref. [32], [2] Ref. [33]. * The experimental CBM value is obtained by adding the VBM with the band gap.

To finalize the discussion on the self-screening correction implemented in *SPEX*, a comparison of the approximate increase in time and memory required to carry out the computations is needed. As usual in GW calculations the main time duration is spent on the self-energy summation, however this remains largely unaffected with the addition of the correction. This comes from that the sum over the bands n' is carried out irrespectively of whether or not the correction scheme is used; the only increase in time is from a short loop deciding which W_m to use. The preceding calculation of P_m, ϵ_m and W_m are largely affected though, as they must be carried out $(n_{\text{scc}} + 1)$ times. Already for 10-20 bands, the time requirements are on the same order, or larger than, the duration for the computation of Σ. Treating ~ 100 bands would thus be very time consuming in the current implementation. It should be noted though that the calculations for each m are independent until the summation in Σ, and could

as such be parallelized. Thus the treatment of a large number of bands within the scheme with approximately the same time-requirement as for a usual GW calculation would be possible, although with an increase in the number of cores required. The main issue would then instead be the memory requirements, as the arrays have to be stored as $(n_{\mathrm{ssc}}+1)\cdot(\text{size of P})$, et cetera. A solution would be to have the arrays calculated in turn and saved to the disk, with the extra time requirement this would entail, until they are used in the summation in Σ.

Section 6

Conclusions

For the model calculations the main conclusion which can be drawn is regarding how the correction schemes from [19] compare to GWA with a change in interaction strength U_0. For small U_0 they are found to significantly improve the agreement with the exact results, and for some properties they are seen to almost eliminate all errors incurred by RPA and GWA. Specifically for the one-level model the schemes work very well for describing the excitation energy of the system and improving the satellite features of A in the regime of small interaction strengths. With an increase in the interaction, the excitation energy worsens, but the improved description of the satellites is maintained. The differences between the self-screening and self-polarization schemes is furthermore evident here, with a halved weight of $R^{ss}_{m\sigma}$ compared to R^{sp} due to the treatment of the spin. Also the behaviour of the HOMO-LUMO gap shows the contrast clearly; the self-screening correction provides the exact result for small $(U_0 - U_1)$, while the self-polarization correction does so in the limit of large interaction strengths.

The analysis is notably more complicated for the two-level system, due to the greater complexity of the system and its larger number of variable parameters. The vastly improved description at smaller interactions is a recurring result of the schemes though, with an almost entirely cured error in the excitation energy and considerably improved satellite features once more. When U_0 is increased there are additional features which can be compared, and it is increasingly difficult to make an all-encompassing assessment. However, the excitation energies seem to worsen by using the corrections while the satellite placements in A improve for a large U_0. When also adjusting the t_2 and t_3 parameters to increase the hopping strength in-between the orbitals, it becomes clear that the effective decrease of the interaction, from the delocalization of the electrons away from the lower orbitals, causes the corrections to improve the excitation energies again.

It is known that the delocalization improves the description within RPA as well, however, evidently the self-screening error seems to be a large part of the remaining error of these more delocalized electrons. This is seen from the substantial improvement of the correction schemes in these regimes, and to connect these results to the calculations of real materials, the found improvement in the more delocalized valence and conduction bands. On the other hand, an increasing U_0 localizes, and hence correlates, the electrons, and GWA and RPA are known to work poorly in this regime. As there is little improvement for larger U_0, and even worsened results for the excitation energy, from the self-screening correction, a conclusion could be drawn that the issue of GWA and RPA from localization, is not mainly from self-screening. Instead it originates in other correlation effects. However, since the satellite features are notably affected and improved for this regime as well, it is likely that self-screening is still a non-negligible part of the error. Furthermore, the large improvement observed for the self-polarization correction in both the HOMO-LUMO gap and the spectral function supports that the self-screening is indeed an important contributor to the error, and the two schemes complement each other in the different regimes. Drawing far reaching

conclusions to their applicability solely based on model calculations should, nevertheless, be done with caution. That RPA is better at describing some features for larger U_0 could possibly be due to cancellation effects. As the error in the excitation energy is of a different sign for RPA and the correction schemes, it is a possibility that too much has been removed in the corrections compared to the real case. That is, a part of the self-screening error in RPA is already effectively cancelled by neglecting higher order corrections. The correction schemes then include this effective cancellation, from its RPA starting point, while at the same time being cancellations of the self-screening error themselves. As such, for higher interaction strengths too much is removed, and the corrected schemes are worsened compared to RPA.

The problem with causality arising from the self-polarization method is a known issue within many-body perturbation theory. It does, however, not appear in the self-screening correction scheme. As such it should originate in their inherent differences: the way in which the corrections are performed. Going back to the difference in the polarization propagators, P_α and $P_{m\sigma}$, the disparity begins with the additional $P_{1\sigma}$, where all the excitations for the given spin are removed, which is not present for the self-polarization. Viewing this from the point of view of a $g_{m\sigma}$ it means that, when calculating W_m within the self-screening correction, all excitations involving itself are removed, for both the occupied and unoccupied states. On the other hand, in the two-orbital system the self-polarization correction only fully removes the unoccupied $g_{m\sigma}$ from interacting with themselves in $[1 - P_\alpha v]^{-1} p_\alpha$, while an occupied state is present in P_α in the form of the other allowed excitations. This could conceivably offer an explanation to the causality issue as the self-polarization correction removes a part, causing a negative S, while the self-screening neglects the given $g_{m\sigma}$ entirely. As the one-level system is not plagued by the issue it seems to be a reasonable origin, although additional investigation would be required to fully identify it.

From the *ab inito* calculations two conclusion can be drawn related to the accuracy of the self-screening correction scheme; the semicore states seems to be largely unaffected by the corrections, while the improvement in the band gap is substantial, compared to the GW error for GaAs and ZnSe. A reason for the discouraging results found for the $3d$ states might be explained by that the calculations were performed on bulk materials, which have many possible screening channels. Removing a single one for a given state might therefore produce only a limited effect. It is possible that atomic and molecular systems or surface calculations therefore would provide a larger improvement. The improvement in the band gap is in-line with that the removal of a part of the screening should shift the gap closer to the HFA value, where there is no screening. As Hartree-Fock is known to overestimate the value compared to the experimental gap, and GW on the other hand underestimates it, the improvement is reasonable from a numerical point of view.

It could further be speculated that it potentially would be sufficient to apply the correction scheme only to the bands closest to the gap in order to obtain the observed improvement, although that would require further calculations to verify. Additional calculations with improved parameters should also be carried out to verify the results, although it seems unlikely that it would alter the shift relative GW to any larger extent. Especially the number of unoccupied bands included in the GW calculations, which in the present calculations might have been insufficiently converged due to time limitations, should be investigated. In a preliminary GaAs calculation with almost twice as many bands, however, no notable difference in the GW values for neither the semicore energy nor the band gap was observed, and it is

therefore probably not a large source of error. It is furthermore also a possibility that the approximate treatment of the head and wing elements of the matrices in *SPEX*, as those from the non-corrected case, could affect the outcome, and a proper treatment of them would be required to further assess the correction.

Finally, to summarize the results found in this study it was found that the self-screening correction was able to correctly describe both models for small interaction strengths with a very high accuracy, while the self-polarization correction only was able to give an improved picture over RPA and GW for the response and spectral functions. Instead, it correctly describes the HOMO-LUMO gap in the one-orbital model and gives a substantial improvement of the spectral function in the regime of large U_0. As such it would seem as if the self-screening correction is a good approach to correct the self-screening error for less localized systems, while the self-polarization correction on the other hand would be the choice for localized ones, while still providing an as accurate description of the response function and spectral function as GW-ss for small U_0. It is with this in mind not too unexpected that the largest improvement for the real materials was for the states close to the band gap, as these originate in more delocalized orbitals, even though the self-screening error intuitively would be the largest for the localized states. That the semicore states, which come from the localized $3d$ orbitals, are not corrected much relative to the GW results could, at least partly, be related to the observed inadequacy of the self-screening correction in the localized regime. By instead applying the self-polarization correction, which from the model calculations is known to work better for the localized states, the energy placement could possibly be improved even further than what was observed here. To test this prediction further investigations are required.

Section 7

Outlook

There exists several aspects of the self-screening and self-polarization schemes to where further effort could be directed. One possibility would be to develop the scheme to properly incorporate self-consistency, although the usual issues with self-consistent GW and the further increase in computational cost might make *ab initio* calculations unfeasible. It would nevertheless be be of interest to see how well the schemes perform, although the theory would have to be reworked to allow for self-consistency. Higher order corrections beyond RPA could potentially also be included at the G_0W_0 level, however how these would be affected by the corrections is not directly evident and would require an extension of the theory.

Another direction would be to thoroughly investigate the causality error occurring in the self-polarization scheme. As the correction schemes are seen to provide somewhat different results for larger U_0 with for example the excitation energy being better described by the self-polarization correction (though keeping in mind that $R_{m\sigma}$ in fact is an auxiliary quantity) it would be an interesting prospect to see if resolving the causality problem would affect the results even further. The origin of the problem would first have to be identified, whereupon an enhancement of the scheme could be developed and applied to models to test its effect, and eventually also to real systems. Finding a way of combining the beneficial effects of the self-screening and self-polarization corrections could further be a profitable, albeit more complicated, venue.

On the *ab initio* side more effort could be put into calculating the properties of additional materials to ensure that the obtained results applies also to other systems. First the correct treatment of the Γ-point must be included though, to ensure that the head and wing elements are treated correctly within the method. This is currently being implemented. Another material which at the moment is under investigation is Gd, with bands originating in 4f orbitals which are incorrectly described within both LDA and GWA. Although the effect on the localized $3d$ semicore states was smaller than expected, it is of interest to see the result of the scheme on different types of materials. As a further verification of the results obtained in this work, an application of the scheme to another GW code, independently, would be beneficial for comparison. Furthermore, as was noted in the discussion, atomic or molecular systems could potentially be impacted to a greater extent by the correction and are as such interesting candidates to for future work.

Finally, the application of the self-screening correction scheme to the constrained random-phase approximation (cRPA) [36] is another avenue, currently being at an early state of development. cRPA is known to overscreen in the downfolding procedure [37], and it is hoped that the reduction of the screening due to the correction could potentially improve this issue. A potential future use of an improved cRPA scheme could be in the downfolding procedure in multitier GW+EDMFT (GW+extended DMFT) [38].

The results obtained within this study are intended to be published in the future.

Appendix A

Density Functional Theory

In density functional theory [1, 2] it is shown that the the Hamiltonian, and with it all properties of the system, are determined by the ground state density of the system. The scheme used in practical calculations is due to Kohn and Sham [1], and is based on replacing the interacting system with a non-interacting one. The auxiliary non-interacting system is further chosen to have the same ground state density as that of the full interacting system. The Kohn-Sham equation,

$$\left(-\frac{1}{2}\nabla^2 + V_{\mathrm{KS}}\right)\varphi_n(r) = \varepsilon_n\varphi_n(r) \tag{A.1}$$

gives the non-interacting wavefunctions and energies, however, only the ε_n of the highest occupied value is a real physical quantity. Nevertheless, the other energies are often interpreted as excitation energies or used as input into other methods [39], with both roles being used in this work for the calculation of the bandstructure of real materials. Furthermore, the Kohn-Sham equation must be solved self-consistently with respect to the density,

$$n(\mathbf{r}) = \sum_{\sigma}\sum_{n}^{\mathrm{occ}} |\varphi_n(\mathbf{r},\sigma)|^2. \tag{A.2}$$

The Kohn-Sham potential is given by

$$V_{\mathrm{KS}} = V_H(r) + v_{ext}(r) + V_{xc}(r) \tag{A.3}$$

where V_H is the usual Hartree potential, v_{ext} is some external potential, and V_{xc} is known as the exchange-correlation potential. The exchange-correlation potential incorporates both interaction effects beyond the Hartree mean-field, as well as the rest of the interacting kinetic energy not treated in the non-interacting system in equation (A.1). It is formally obtained as a functional derivative of the exchange-correlation functional E_{xc} as $V_{xc} = \delta E_{xc}[n]/\delta n$. If the functional was known exactly the problem would have been straightforward to solve, however this is not the case. Instead various approximate forms are treated in the literature, and one of the most common is the local density approximation [1] which is used in this work. It is based on assuming the functional to be given by that of the electron gas ϵ_{xc}, with the density given by the non-interacting system, according to

$$E_{xc}[n] = \int \mathrm{d}\mathbf{r} n(\mathbf{r})\epsilon_{xc}(n(\mathbf{r})). \tag{A.4}$$

Using this assumed form of the functional it is then possible to self-consistently solve equation (A.1).

Bibliography

[1] W. Kohn, L. J. Sham, *Self-consistent Equations Including Exchange and Correlation Effects*. Phys. Rev., 140, A1133 (1965).

[2] P. Hohenberg, W. Kohn, *Inhomogeneous Electron Gas*. Phys. Rev., **136**, B864 (1964).

[3] G. A. Wilfried, L. Jönsson, J. W. Wilkins, *Quasiparticle Calculations in Solids*, in *Solid State Physics, Volume 54*. Eds.: H. Ehrenreich, F. Spaepen. Academic Press (2000).

[4] S. Massidda, R. Resta, M. Posternak, A. Baldareschi, *Polarization and dynamical charge of ZnO within different one-particle schemes*. Phys. Rev. B **52**, R16977 (1995).

[5] F. Aryasetiawan, O. Gunnarsson, *3d semicore states in ZnSe, GaAs, and Ge*. Phys. Rev. B **54**, 17564 (1996).

[6] A. L. Fetter, J. D. Walecka, *Quantum Theory of Many-Particle Systems*. Dover Publications Inc. (2003).

[7] L. Hedin, S. Lundqvist, *Effects of Electron-Electron and Electron-Phonon Interactions on the One-Electron States of Solids* in *Solid State Physics Volume 23*. Eds.: F. Seitz, D. Turnbull, H. Ehrenreich. Academic Press (1969).

[8] L. Hedin, *New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem*. Phys. Rev. **139**, A796 (1965).

[9] D. Bohm, D. Pines, *A Collective Description of Electron Interactions: III. Coulomb Interactions in a Degenerate Electron Gas*. Phys. Rev. **92**, 609 (1953).

[10] D. Pines, *Elementary excitations in solids*. W.A. Benjamin, Inc. (1963).

[11] A. Schindlmayr, C. Friedrich, E. Sasioglu, S. Blügel, *First-Principles Calculation of Electronic Excitations in Solids with SPEX*. Z. Phys. Chem. 224, 357 (2010). (Accessed from `arXiv:1110.1596`)

[12] F. Aryasetiawan, O. Gunnarsson, *The GW method*. Rep. Prog. Phys. **61**, 237 (1998).

[13] E. L. Shirley, R. M. Martin, *GW quasiparticle calculations in atoms*. Phys. Rev. B **47**, 15404 (1993).

[14] A. Georges, G. Kotliar, W. Krauth, M. J. Rozenberg, *Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions*. Rev. Mod. Phys. **68**, 13 (1996).

[15] S. Biermann, F. Aryasetiawan, A. Georges, *First-Principles Approach to the Electronic Structure of Strongly Correlated Systems: Combining the GW Approximation and Dynamical Mean-Field Theory*. Phys. Rev. Lett. **90**, 086402 (2003).

[16] W. Nelson, P. Bokes, P. Rinke, R. W. Godby, *Self-interaction in Green's-function theory of the hydrogen atom*. Phys. Rev. A **75**, 032505 (2007).

[17] J. Wetherell, M. J. P. Hodgson, R. W. Godby, *The GW self-screening error and its correction using a local density functional*. Phys. Rev. B **97**, 121102(R) (2018).

[18] P. Romaniello, S. Guyot, L. Reining, *The self-energy beyond GW: Local and nonlocal vertex corrections*. J. Chem. Phys. **131**, 154111 (2009).

[19] F. Aryasetiawan, R. Sakuma, K. Karlsson, *GW approximation with self-screening correction*. Phys. Rev. B **85**, 035106 (2012).

[20] *The FLEUR project*, `https://www.flapw.de/site/`

[21] J. Hubbard, *Electron correlations in narrow energy bands*. Proc. R. Soc. Lond. A, **276** 238 (1963).

[22] F. Aryasetiawan, *Self-energy of ferromagnetic nickel in the GW approximation*. Phys. Rev. B **46**, 13051 (1992).

[23] U. von Barth, B. Holm, *Self-consistent GW_0 results for the electron gas: Fixed screened potential W_0 within the random-phase approximation*. Phys. Rev. B **54**, 8411 (1996). Erratum: Phys. Rev. B **55**, 10120 (1997).

[24] V. Christiansson, *Model study of the linear density response function - A comparison between the exact result and RPA*. Bachelor thesis, Lund university. Available from `https://lup.lub.lu.se/student-papers/search/publication/8915173`

[25] R. Sakuma, *Customized Spex Documentation*. Available from

`http://www.matfys.lth.se/staff/Rei.Sakuma/spex/cspex_doc.pdf` Accessed on 20190404.

[26] C. Friedrich, S. Blügel, A. Schindlmayr, *Efficient implementation of the GW approximation within the all-electron FLAPW method*. Phys. Rev. B **81**, 125102 (2010).

[27] C. Friedrich, A. Schindlmayr, S. Blügel, *Efficient calculation of the Coulomb matrix and its expansion around k=0 within the FLAPW method*. Computer Physics Communications Vol. 180, Issue 3, 347-359 (2009).

[28] F. Aryasetiawan, O. Gunnarsson, *Product-basis method for calculating dielectric matrices*. Phys. Rev. B **49** 16214 (1994).

[29] N. W. Ashcroft, N. D. Mermin, *Solid State Physics*. Brooks/Cole, Cengage learning (1976).

[30] *ZnO example input file for FLEUR*. Available at
`https://www.flapw.de/site/examples/elements/#zinc` Accessed on 20190502

[31] H. Morkoç, Ü. Özgür, *Zinc Oxide: Fundamentals, Materials and Device Technology*. WILEY-VCH Verlag (2009).

[32] C. Kittel, *Introduction to Solid State Physics* 7th ed. John Wiley & Sons, Inc. (1996).

[33] L. Ley, R. A. Pollak, F. R. McFeely, S. P. Kowalczyk, D. A. Shirley, *Total valence-band densities of states of III-V and II-VI compounds from x-ray photoemission spectroscopy.* Phys. Rev. B **9**, 600 (1974).

[34] Ed.: H. E. Ruda. *Widegap II–VI Compounds for Opto-Electronic Applications.* Chapman & Hall (1992).

[35] W. B. Jackson, J. W. Allen, *Experimental self-energy corrections for various semiconductors determined by electron spectroscopy.* Phys. Rev. B **37**, 4618 (1988).

[36] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, I. Lichtenstein, *Frequency-dependent local interactions and low-energy effective models from electronic structure calculations.* Phys. Rev. B **70**, 195104 (2004).

[37] H. Shinaoka, M. Troyer, P. Werner, *Accuracy of downfolding based on the constrained random-phase approximation.* Phys. Rev. B **91**, 245156 (2015).

[38] F. Nilsson, L. Boehnke, P. Werner, F. Aryasetiawan, *Multitier self-consistent GW+EDMFT.* Phys. Rev. Materials **1**, 043803 (2017).

[39] R. M. Martin, *Electronic Structure Basic Theory and Practical Methods.* Cambridge University Press (2010).