

## Excercise sheet 8

On this sheet, we will continue with theoretical methods for the simulation of excited states properties, but will go beyond the TDDFT formalism and instead use many-body method, namely the GW approximation and the Bethe-Salpeter Equation.

The modern implementations of the GW and BSE methods make these calculations relatively straightforward and mechanically not more difficult than DFT calculations with Quantum Espresso or similar codes. However, a significant part of the calculations is the identification of computational parameters for which the computed quasiparticle bandstructures or absorption spectra are as converged as possible. Often, it is not possible to fully converge the excitation spectra due to the computational cost of these methods and one has the unfavorable task of carefully finding a set of computational parameters that gives good results while keeping the calculations feasible.

There are a number of such parameters that have to be adjusted: similar to DFT calculations, the k-point grids used for the simulation need to be dense enough to sufficiently capture the influence of variation of the electron density and the electronic wavefunctions within the Brillouin zone. While the electrons with different k-vectors are essentially decoupled for local exchange-correlation functionals and one hence can relatively easily use very dense k-point grids in DFT calculations, GW and the BSE resemble hybrid functionals in the sense that all k-points in the Brillouin zone are coupled through non-local electron-electron interaction. This makes the scaling of the computational effort with the k-point grid density far less forgiving than in DFT.

Both GW and BSE methods require calculation of the non-interacting response function  $\chi_0$ , which in reciprocal space has the following form:

$$\chi_{0,GG'}(q, \omega) = 2 \sum_{n,m} \int_{BZ} \langle \varphi_{nk-q}^{KS} | e^{-i(q+G)\cdot r} | \varphi_{mk}^{KS} \rangle \langle \varphi_{mk}^{KS} | e^{i(q+G')\cdot r} | \varphi_{nk-q}^{KS} \rangle \frac{f_{nk-q}(1-f_{mk})}{\omega + \varepsilon_{nk-q} - \varepsilon_{mk} \pm i0^+} \frac{dq}{(2\pi)^3}$$

The cutoff energy then decides how many G and G' vectors we include in the Fourier transformation equation above, and with that also the matrix size of  $\chi_{0,GG'}$ . Also, the accuracy of  $\chi_{0,GG'}$  will be determined by the number of bands that we include in the summation over the band indices  $n$  and  $m$ .

Similarly, the self-energy in GW can be split into an exchange part and a correlation part, which in reciprocal space take the form

$$\begin{aligned} \Sigma_{nk}^x(\omega) &= \langle \varphi_{nk}^{KS} | \Sigma_{nk}^x(r, r') | \varphi_{nk}^{KS} \rangle = - \sum_m \int_{BZ} \sum_{G, G'} \frac{4\pi}{|G - G'|^2} \left| \langle \varphi_{nk}^{KS} | e^{i(q+G)\cdot r} | \varphi_{mk-q}^{KS} \rangle \right|^2 \frac{dq}{(2\pi)^3} \\ \Sigma_{nk}^c(\omega) &= \langle \varphi_{nk}^{KS} | \Sigma_{nk}^c(r, r', \omega) | \varphi_{nk}^{KS} \rangle \\ &= i \sum_m \int_{BZ} \sum_{G, G'} \frac{4\pi}{|G - G'|^2} \langle \varphi_{mk-q}^{KS} | e^{-i(q+G)\cdot r} | \varphi_{nk}^{KS} \rangle \langle \varphi_{nk}^{KS} | e^{i(q+G')\cdot r} | \varphi_{mk-q}^{KS} \rangle \times \\ &\quad \left( \int G_{mk-q}^0(\omega - \omega') \epsilon_{GG'}^{-1}(q, \omega) d\omega' \right) \frac{dq}{(2\pi)^3} \end{aligned}$$

Here as well, the number of to-be-included bands and the cutoff energies for the summations in  $\Sigma_{nk}^x$  and  $\Sigma_{nk}^c$  need to be carefully chosen through systematic variation. This will be a main part of this exercise sheet.

As most of the calculations are computationally quite heavy, it is strongly advised to run the simulations on the HPC. Like Quantum Espresso, YAMBO uses the math libraries provided by intel. For this reason, you need to load the intel64 module with

```
module load intel64
```

before you can run YAMBO on the HPC.

### Exercise 1: TDDFT for absorption spectra

In this exercise, you will use TDDFT to simulate the absorption spectrum of solid Argon, a wide bandgap insulator. Solid Argon has an FCC structure with a cubic lattice constant of 5.26 Å.

- (a) Take the structure from exercise 1 of the exercise sheet 7 and recalculate the electronic structure on a 12x12x12 k-point grid. Include 500 *unoccupied* bands in the nscf calculation, which we will use for a convergence test of the GW quasiparticle energies later.

The Davidson algorithm used for the iterative diagonalization of the Kohn-Sham-Hamiltonian at each k-point in the nscf calculation becomes unstable, if many unoccupied bands are to be calculated. For this reason, switch to the more stable (but also slower) conjugate gradient algorithm implemented in Quantum Espresso by adding the line

```
diagonalization='cg'
```

to the &electrons block.

Also, like on exercise sheet 7, make sure again that Quantum Espresso does not consider non-symorphic symmetry operations, which YAMBO does not like.

- (b) We will now use YAMBO to calculate the GW corrected band gap of solid Argon based on the DFT electronic structure from (a). The preparation of the GW calculation follows the same steps as for the TDDFT calculations of exercise sheet 7: Go to TMP\_DIR/Ar.save and run the command

```
~/bin/yambo/p2y -M (adjust the path if you put the yambo files somewhere else)
```

Then move the created SAVE folder into a separate folder yambo (or whatever you prefer) and do a setup run. An input file for the setup run can be generated by the command

```
~/bin/yambo/yambo -M -i -Q -F yambo-in_setup
```

This will write a yambo input file yambo.in\_setup, which you run with

```
~/bin/yambo/yambo -M -F yambo-in_setup
```

As on the previous sheet, the option “-M” tells yambo not to use the parallelization routines, which cannot be done on the HPC unless the job is executed using a job submission script (which does not make sense here, because the p2y and setup calculations are quite fast). Usually, YAMBO opens the generated input file in the text editor vi, which is part of linux. The option “-Q” tells YAMBO to only generate the file and not to open it.

After the setup run is finished, we now need to generate an input file for the GW calculation. Run the command

```
~/bin/yambo/yambo -M -Q -d -p p -g n -V par -F yambo.in_gw
```

Let’s go through the options: “-d” tells YAMBO to calculate the dynamical, i.e. frequency dependent, dielectric matrix that will be used to calculate the dielectric Coulomb interaction  $W$  entering the GW self-energy. In order to make the calculations easier, we will directly calculate the inverse dielectric function used for the dielectric screening using the plasmon pole approximation (PPA), where

$$\epsilon_{GG'}^{-1}(q, \omega) = \delta_{GG'} + R_{GG'}(q) \left( \frac{1}{\omega - \Omega_{GG'}(q) + i0^+} - \frac{1}{\omega + \Omega_{GG'}(q) + i0^+} \right)$$

This is invoked by the option “-p p”. YAMBO uses the Godby-Needs flavor of the PPA (Phys. Rev. Lett. 62, 1169 (1989)) and fits the residuals  $R_{GG'}$  and  $\Omega_{GG'}(q)$  so that the PPA reproduces the real inverse dielectric function for the frequencies  $\omega = 0$  and  $\omega = i\omega_{PPA}$ . This circumvents the costly necessity of the dielectric function  $\epsilon_{GG'}(q, \omega)$  for each frequency. This means that YAMBO explicitly calculates the dielectric function for two frequencies and then proceeds to fit the PPA.

The option “-g n” instructs YAMBO to calculate the GW quasiparticle energies from the GW self-energy as a one-shot correction to the DFT band energies, using a Newton extrapolation scheme. “-V par” lets YAMBO write the necessary keywords for parallelization of the calculations into the input file.

Have a look at the generated input file. The block at the end of the file defines for which k-points and bands we actually want to calculate the quasiparticle energies. By default, YAMBO calculates the GW corrections for all k-points in the irreducible grid (in our case all 72 points) and all bands included in the DFT electronic structure (504 in our case). This will take forever and is unnecessary, as we are only interested in corrections for a few bands around the Fermi energy. Set this block to

```
%QPkrange
1|1|3|6|
%
```

which will cause YAMBO to calculate quasiparticle corrections for the third, fourth, fifth and sixth band at the first k-point in the list, which is the Brillouin zone center. This will give us a corrected value for the fundamental band gap and will also include the second valence and conduction bands.

We also need to decide, how many bands we want YAMBO to sum over for the calculation of the polarization function and the correlation part of the GW self-energy. This is defined by the blocks `GbndRnge` and `BndsRnXp`. Set the second number in each block to 50, which will instruct YAMBO to sum over 50 occupied and unoccupied bands for now.

Other important parameters are a whole number of cutoff energies. `NGsBlkXp` defines the number of G vectors to be included in the calculation of the interacting and non-interacting response/polarization functions  $\chi/\chi_0$ . This is an important convergence parameter, which can have significant effects on the accuracy of the results but also can quickly make the calculations expensive. It should thus be carefully chosen. Set it to 100 eV for now, which is a reasonable value for many materials. Similarly, `EXXRLvcs` and `VXCRLvcs` define the number for G vectors for the calculation of the exchange part of the GW self-energy and for the exchange-correlation potential used in the one-shot correction. As these are not screened, the calculations typically require a significantly higher energy cutoff to achieve convergence, but the calculations are also much faster than those for the correlation part. Set both values to 500 eV. Also notice the definition of another energy in the input file. `PPAPntXp` is the second frequency that is used for the fitting procedure of the plasmon pole approximation. The default value of 1 Ha typically works well.

Finally, we need to define how YAMBO should parallelize the calculations. As for the TDDFT calculation, we need to manually define the number of q-point, G-vector, conduction band groups etc. For the parallelization of the calculation of the polarization function X, use the same values as in exercise 2 of exercise sheet 7. For the parallelization of the self-energy, we have three options: parallelization over the k-points, for which we want to calculate the quasiparticle energies, parallelization over quasiparticle bands at each k-point of interest, and parallelization over the bands that are summed over in the calculation of the self-energy. Set

```
X_all_q_CPU= "1 4 5"
```

```
X_all_q_ROLEs= "q qp b".
```

With this, everything is prepared. Start the calculation with the provided submission script. As for the TDDFT calculations from exercise sheet 7, YAMBO first calculates the dipole matrix elements for transitions between occupied and unoccupied bands and then the non-interacting response function  $\chi_0$  at each q-point in the Brillouin zone sampling (here 72) and for each of the two frequencies. This is the most costly part of the calculation, if we are only interested in a few bands at one single k-point. It then calculates, rather quickly, the PPA for the inverse dielectric function for each of the q-points.

After this (run level EXS), it calculates the exchange part of the self-energy and the exchange-correlation potential. This takes only a few seconds in this case.

Finally, YAMBO calculates the correlation part of the self-energy and the quasiparticle corrections. This step only takes a minute in this case, but becomes the most time-consuming step if we are interested in corrections at many k-point and/or for many bands.

YAMBO writes a file `o.qp`, which contains the results. For each k-point and band that we were interested in, it reports the DFT energy ( $E_0$ ) and the quasiparticle correction ( $E-E_0$ ). The desired quasiparticle band energies are then just the sum of the third and fourth columns reported in the file. Clearly,  $G_0W_0$  corrects the value of the fundamental gap (between the fourth and the fifth band) predicted by DFT by about 5 eV.

- (c) The results from (b) were obtained using 50 bands in the summation for the calculation of the non-interacting response function and for the non-local correlation. While this made the calculations relatively fast, 50 bands are probably not enough to yield converged results.

Repeat the calculations for 100, 200, 300 and 500 bands. Here, it makes sense to run all simulations simultaneously by creating a separate folder for each of the four calculations and copying the input file `yambo.in_gw`, the submission script and the SAVE folder into each of them. Before doing that, go into the SAVE folder and delete all files that the previous YAMBO run wrote into the folder: the file `ndb.QP`, which contains the quasiparticle energies, `ndb.HF_and_locXC`, and all files of the formats `ndb.pp*` and `ndb.dip_*`. Adjust the input files in each folder and submit the calculations to the HPC. Finally, plot the fundamental band gap as a function of the number of included bands. You will see that you need to include about 300 bands to converge the fundamental band gap to within 0.01 eV<sup>1</sup>. Note that the convergence of energy differences (e.g. band gaps) is usually faster than the convergence of the absolute band energies.

The somewhat slow convergence of the quasiparticle corrections with the number of included bands is a factor that makes GW computationally quite costly. As mentioned in the lecture, a recently proposed remedy is using 'effective energy techniques' (EET) to include the contribution of higher-energy bands in an approximate way. YAMBO uses the approach proposed by Bruneval and Gonze (Phys. Rev. B 78, 085125 (2008)), which indeed uses a single effective energy for this purpose.

Repeat the GW calculation with 50 bands and an additional EET correction. To do this, add the following lines to the input file:

```
XTermKind= "BG"
XTermEn= 40.00000 eV
GTermKind= "BG"
GTermEn= 40.81708 eV
```

This instructs YAMBO to use the Bruneval-Gonze EET correction for both the calculation of the non-interacting response function (X) and for the self-energy (G). The effective energies are defined through the keywords `XTermEn` and `GTermEn`, here set to the default values.

Notice that the correction works quite well even for the small number of bands and brings the predicted band gap close to the converged value.

Before the next step, delete all the SAVE folders except for one, as each of them should be about 1-1.5 Gb in size.

- (d) Besides the number of included bands, the cutoff energies also need to be tested for convergence, especially the cutoff energy for the response function and dielectric matrix. A real convergence is often difficult to achieve, as the computational cost increases quite steeply with the number of included G vectors (the size of the dielectric matrix increases quadratically with the number of G vectors). Another complication is that the cutoff energy of the response function is not independent of the number of included bands, so that in principle the response function cutoff energy and the number of included bands need to be converged simultaneously. This is, however, often not done in the literature, as it is quite a time-

---

<sup>1</sup> This is actually a quite reasonable number, materials such as zinc oxide require more than 3000 bands for a similar convergence.

consuming procedure. For the purpose of this exercise, it is enough to converge both parameters independently.

Do another set of GW calculations for cutoff energies of 150 and 200 eV (you can also go higher, if you want) and summation over 300 bands (without EET, for comparison with the case of 100 eV cutoff). Does the fundamental band gap significantly change? Which cutoff energy should one choose to obtain a good prediction of the fundamental band gap?

- (e) With the parameter choice finished, we are now ready to recalculate the GW quasiparticle states for all 72 k-points, which we will need both for the BSE calculations in the next step and for a plot of the full GW quasiparticle bandstructure. Prepare a new folder for this calculation and change the input file to reflect the following requirements: We want the quasiparticle energies at all 72 k-points and for the first 8 bands. Use a cutoff obtained from (d), which will give us a good prediction of the GW electronic structure. Use 50 bands and the EET correction.

Warning: These calculations will take between 2 and 3 hours, because we now calculate the corrections for 72 points instead of just one point. As an alternative, you can just take the predicted fundamental band gap value and use the scissor correction technique from exercise 1 of exercise sheet 7 for the calculations in (f).

Before you calculate the absorption spectrum from solution of the Bethe-Salpeter equation, use the YAMBO postprocessing program ypp to plot the corrected bandstructure. For this, we first need to remove time-reversal symmetry from the saved data, because ypp cannot do the interpolation while it is still included. Go to the folder that contains SAVE and run the command

```
~/bin/yambo ypp -Q -M -y
```

This writes the file ypp.in. Open it and uncomment the keyword RmTimeRev by removing the # in front of it. Now run

```
~/bin/yambo ypp -Q -M
```

This should create a folder FixSymm, which in turn contains another SAVE folder, which contains the YAMBO data with removed time-reversal symmetry. Put the provided ypp.in\_bandinterp into the FixSymm folder and enter the folder. BANDS\_bands allows the user to choose what bands he/she wants to include into the bandstructure plot, BANDS\_kpts defines the k-point path (the same path as used for the phonon dispersion of silicon in exercise 2 of exercise sheet 5) and BANDS\_steps tells ypp how many k-point on the path to interpolate. All k-point coordinates are given in units of reciprocal lattice vectors (rlu). Crucially, we also want to include the calculated quasiparticle corrections for the interpolated bandstructure. This is done with the line GfnQPdb= "E < ./SAVE/ndb.QP", which tells ypp to read the quasiparticle energies (E) from the file ndb.QP, which is located in the folder SAVE.

You can generate this input file yourself with the command ~/bin/yambo/yp -Q -M -s b -V qp and adapting it. "-V qp" instructs ypp to add the keywords related to quasiparticle or scissor corrections to the input file, which are quite a few. The unneeded keywords were deleted from ypp.in\_bandinterp compared to the raw generated version.

Now run ypp with

```
~/bin/yambo/yp -M -F ypp.in_bandinterp
```

This will interpolate the bandstructure and write the results in the file o.bands\_interpolated, which can be plotted with gnuplot or similar programs. Here, you want to make sure that ypp really includes the quasiparticle corrections. Look out for a line “<---> [06.02] External/Internal QP corrections”, which appears if ypp understood that you want to include these corrections and successfully found and analysed the quasiparticle data.

- (f) We will now finally calculate the optical spectrum of Ar from solution of the Bethe-Salpeter Equation. Stay in the folder, where you calculated the GW quasiparticle energies, but create a new input file yambo.in\_bse. To not confuse the input file generator, it is probably best to delete all files that were generated by the GW calculation (i.e. ndb.dip\_\*, ndb.pp\*, ndb.HF\_and\_locXC). However, be careful that you do not delete ndb.QP, because it contains the quasiparticle corrections that we need for the BSE calculation. After this is done, return back into the calculation folder and run the command

```
~/bin/yambo/yambo -M -Q -b -o b -k sex -y h -V par -F yambo.in_bse
```

Let us again quickly go through the options: “-b” instructs yambo to calculate the non-interacting response function at each q-point and from that the static (i.e. not frequency-dependent) dielectric matrix without electron-hole interaction (run level em1s in the generated input file). We then want to do an optics calculation using the Bethe-Salpeter Equation (-o b) that uses a screened Coulomb interaction (or “exchange”) In the Bethe-Salpeter kernel (-k sex). Here, we could also use other approximations in the kernel, for example a Hartree potential, ALDA or HF. The calculations would then correspond to TDDFT calculations using the Casida Equation, i.e. using a transition basis instead of solving the Dyson Equation of the response function. Finally, we want to obtain the eigenvalue spectrum of the effective excitonic Hamiltonian built during the BSE calculation using a Haydock solver. The Haydock solver has the significant advantage that it does not explicitly diagonalize the excitonic Hamiltonian but uses mathematical minimization techniques to directly calculate the macroscopic dielectric function. This method is much faster and needs less memory than direct diagonalization. At the same time, it has the disadvantage that we do not have access to the eigenvectors of the excitonic Hamiltonian and hence only get the spectrum but no information about the excitations that give rise to the spectrum.

Open the generated input file. The use of a transition basis means that we describe the eigenvectors/wavefunctions of the excitations in the system as a superposition of band-to-band transitions, each transition being a product of a Kohn-Sham wavefunction corresponding to a hole with a given k-vector and a Kohn-Sham wavefunction corresponding to an electron at a given wavevector  $k'$  (for an optical excitation  $k \approx k'$ ). Here, the Bethe-Salpeter kernel will couple transitions between different bands and at different k-points with each other, which will lead to wavefunctions of the excited states that (usually) significantly differ from the underlying Kohn-Sham wavefunctions. By default, YAMBO solves the Bethe-Salpeter Equation within the Tamm-Dancoff approximation (TDA), i.e. it only considers transitions from occupied valence bands into unoccupied conduction bands, but not in the opposite direction. The TDA can be deactivated by setting BSEmod=”coupling”. The transition basis implies that the

results should become more accurate if we include more valence and conduction bands. However, as the BSE kernel couples two transitions together, doubling the number of included valence and of included conduction bands should nominally increase the computational time by a factor 16. One thus has to carefully choose how many valence and conduction bands one needs to (or can) include for sufficiently accurate results.

A look at the quasiparticle bandstructure gives a few hints how we should choose the valence and conduction bands: the first three valence bands are degenerate (or almost degenerate in the GW bandstructure) at the Brillouin zone center and other high-symmetry points. This originates in the cubic symmetry of solid argon, similar to the bandstructure of silicon. Due to this degeneracy, it does not make much sense to include less than three valence bands in the calculation, otherwise we would neglect some transitions that energetically match (and thus might couple to) transitions that we do include. The fourth and lowest valence band, on the other hand, is energetically well separated from the other valence band and lies about 26-27 eV below the conduction band minimum. Due to this separation, it is less likely that transitions from this band have a strong impact onto the dielectric function near the electronic band gap, at energies below 20 eV. As for the conduction bands, the minimum of second-lowest conduction band is located about 18-19 eV above the valence band and it might have an impact.

First solve the BSE for a basis of three valence bands and one conduction band. For this, change the BSEBands block to

```
% BSEBands
  2 |  5 |
%
```

The BSE calculation automatically includes all 72 k-points of the irreducible grid. We also have to decide how many G vectors we want to include for the calculation of the bare Coulomb potential (BSENGexx), for the screened Coulomb interaction in the BSE kernel (BSENGBlk) and for the non-interacting response function (NGsBlkXs). Set BSENGexx to 500 eV and NGsBlkXs and BSENGBlk to 200 eV. As for the GW calculation, all cutoffs technically need to be adjusted to ensure sufficient convergence, but we will skip this step here. For reasons of implementation (and, in a way, common sense) BSENGBlk always needs to be equal to or smaller than NGsBlkXs. If one can afford it, it makes sense to use the same value for both parameters. Use again 50 bands for the calculation of the polarization function/non-interacting response function and EET by manually adding the line XTermKind= "BG". We also still are missing the inclusion of quasiparticle corrections. You can ask YAMBO to include the necessary keywords by running the above generation command again but replacing "-V par" by "-V qp". Or, simpler, you just add the line KfnQPdb= "E < ./SAVE/ndb.QP" to the input file. Now, only some cosmetic changes are necessary. We need to tell YAMBO the energy range for which we want it to calculate and print the macroscopic dielectric function. This is done with the keyword BenRange. Tell YAMBO to plot the dielectric function from 0 eV to 20 eV by changing the second number. You also should increase BenSteps so that enough energy steps are included in the plot of the dielectric function to make the plot smooth. An energy step of 0.01 eV, i.e. BenSteps=2000, is probably a good choice. The peak broadening that is applied to the excitation energies for the calculation of the dielectric function is defined in the BDmRange block. The standard value of 0.1 eV might blur out the features in the plot a bit much, but this is almost purely a matter of taste. The first and second numbers in the block



allow definition of a broadening value for the low energy range of the dielectric function, which then linearly increases up to the second value towards the high-frequency end of the plot.

Finally we need to define the parallelization. Use the same numbers as for the TDDFT calculations on exercise sheet 7. For the calculation of the actual BSE, YAMBO offers three parallelization levels: parallelization by k-points (k), by electron-hole pairs (eh), and by transitions (t). Usually, it seems advisable to choose the number of k groups as large as permitted by the memory of the computer, using only one eh group and distributing the CPUs of each k group over transition groups. Use 10 k groups and 2 transitions groups here, ie. 20 CPUs in general:

```
BS_CPU= "10 1 2"  
BS_ROLEs= "k eh t"
```

Now change `yambo.in_gw` in `j.submit.sh_gw` to `yambo.in_bse` and submit the job to the HPC to run the simulation. You will see in the LOG files that there is an additional "Kernel" step now, which calculates the actual BSE kernel. This is followed by a BSE solver step, which in this case consists of an iterative calculation of the eigenvalue spectrum with the Haydock technique. This will write the same files as for the TDDFT optics calculations on exercise sheet 7. With one important difference: the results now contain the stronger electron-hole coupling effects. Plot the imaginary part of the dielectric function from `o.eps_q1_haydock_bse` and compare the results with the TDDFT spectra from exercise sheet 7 and with the experimentally measured spectra. The BSE peak position is in quite good agreement with experiment.

- (g) Repeat the calculation from (f), but include the second-lowest conduction band as well and see how the results change.

Clearly, the position of the lowest excitonic peak is slightly downshifted by about 30 meV. Apart from this, the inclusion of the sixth band had a noticeable effect on the peak intensities and, crucially, added additional features above 17 eV that were absent before. One hence always has to keep in mind to include enough bands in the calculation so that the dielectric function is sufficiently converged within the energy range over which one is plotting.

As mentioned earlier, we should also check convergence with the cutoff energies. Also, and more crucially, the Bethe-Salpeter spectra depend heavily on the density of the underlying k-point grid. The k-point sampling imposes a periodicity on the results, because we do not include Kohn-Sham states that have a longer periodicity as that given by the smallest non-vanishing k-vector. For that reason, an  $N \times N$  k-point sampling in reciprocal space corresponds to a supercell of  $N \times N$  unit cells in real space. This can be problematic for excitons, which can have wavefunctions that extend over 10-100 nm and thus span many unit cells of the system. If the k-point sampling is too small the exciton starts interacting with excitons in neighboring supercells, which will change the excitation peak energies in the calculated spectra compared to the case of a sufficiently isolated exciton. However, to test this here, we would have to completely recalculate the GW quasiparticle energies, which goes beyond the scope of this exercise, but is the subject of exercise 2

## Exercise 2: Excitons in 2D BN

Here, we will calculate the optical spectrum of monolayer boron nitride (BN) as an example for a 2D material.

- (a) As always, first optimize the atomic positions and lattice constants of monolayer BN or take the geometry obtained on exercise sheet 2. Then do a non-selfconsistent calculation of the electronic structure of monolayer BN on a grid of  $12 \times 12 \times 1$  k-points and with 400 unoccupied bands. As a preparation for (e) it is best if you place the BN layer in the middle of the unit cell, i.e. with z-coordinates of 0.5 for both boron and nitrogen atoms (see BN.nscf.in\_ref).
- (b) To make the process of checking the convergence of the optical spectra easier, we will now calculate the fundamental band gap of BN and will then use a scissor correction to shift the DFT conduction bands for the BSE calculations. Interestingly, the nature and exact band gap size of monolayer BN is still not clearly experimentally established, but it appears to be in the region of 7 eV.

Let's use the results from (a) to calculate the band gap at the K point of the hexagonal Brillouin zone using the GW approximation. Follow the steps from Exercise 1 (b) to prepare and run the GW calculation. Calculate quasiparticle corrections for the third to sixth band at the K point (k-point 19 in the list of irreducible k-points). Use 400 bands for the calculation of the polarization function and the self-energy. You can also add an additional EET correction for both parts. Use a cutoff energy of 100 eV for the polarization function and of 600 eV for the exchange part of the self-energy and the exchange-correlation potential and use the same parallelization parameters as in Exercise 1 (b). Also use a truncated Coulomb potential like in Exercise 2 of Exercise sheet 7. Make sure that the z-value given in the block (in Bohr) is equal to or slightly smaller than the lattice constant in z-direction in your calculation. An input file is provided as reference.

Run the calculation and use the output file o.qp to derive the difference in band gaps between the DFT and GW calculations. This should give you a GW band gap of about 7.5 eV.

If you want, you can also do an additional Quantum Espresso calculation using a hybrid functional, for comparison. An input file is provided that uses the hybrid functional HSE06. This, however, strongly underestimates the band gap, even though it still yields much better results than PBE or LDA calculations.

- (c) Use the steps from Exercise 1 (f) to set up a new calculation of the optical spectrum. Include all four valence bands and the lowest four conduction bands in the calculation of the BSE. Use 250 bands for the calculation of the non-interacting response function, which appears to be sufficiently converge the optical spectrum. In combination with the EET approach, one probably could also get away with 50 bands (which you can try, if you want) as well. In general, the BSE optical spectra are far less sensitive to the number of included bands in the calculation of the response function than the GW approximation, but convergence should still be checked.

Use a cutoff of 100 eV for the response function and the screened Coulomb interaction in the BSE and 600 eV for the Coulomb/exchange contribution to the BSE kernel. Do not forget to also include a truncated Coulomb potential, as this significantly improves the quality of the predicted peak positions (and it does so for the correct reasons). Use the necessary keywords from Exercise 1 of exercise Sheet 7 to include a correction of the band energies through the scissor shift technique. Also, change the diagonalization method from Haydock ("h") to explicit diagonalization ("d") as a preparation for (e) and remove the # in front of the keyword WRbsWF (this will write the eigenvalues and eigenvectors to the disk, which we need for the analysis in (e)).

Run the BSE calculation and plot the resulting imaginary part of the macroscopic dielectric function.

Compare this to a second calculation, where you increase the cutoff energy to  $NGsBlkXs=200$  eV but keep everything else the same. Was a cutoff energy of 100 eV good choice <sup>2</sup>?

- (d) As mentioned in the previous exercise, the main convergence parameter for BSE optical spectra is the density of the k-point grid, which is used for the calculations. This is especially important for 2D materials: in a 2D material, the dielectric screening is inhomogeneous, as many electric field lines will go through the medium surrounding the 2D material instead of through the 2D material itself. This has the effect that the q-resolved dielectric function features a strong variation close to the  $q=0$  point (i.e. the point of vanishing momentum transfer), see Phys. Rev. B 88, 245309 (2013). Due to our discrete sampling, however, we can only capture the dielectric function at specific discrete q-points and assume that the dielectric function only changes weakly in the vicinity of these discrete points. This can introduce strong errors if the sampling is too coarse. This is especially the case near  $q=0$ , where the inverse dielectric function has a value of exactly 1 (if a truncated Coulomb potential is used). For coarse grids, this causes an artificial *underscreening* in the vicinity of the Brillouin zone center and hence an overestimation of exciton binding energies. This underscreening is reduced if one increases the density of the k-point grid, so that the area that is represented by the  $q=0$  point shrinks. A number of methods to correct for this have been proposed in recent times <sup>3</sup>, but unfortunately, they are not yet implemented in YAMBO.

To test this, recalculate the DFT electronic structure on a  $24 \times 24 \times 1$  k-point grid and redo the BSE calculation. You can reuse the yambo input file from the calculations in (c), with the one difference that you have to change the `QpntsRXs` block to reflect that you now have 61 irreducible points in the grid instead of 19.

Due to the increased k-point sampling density, the calculations will also take quite bit longer, about one hour.

Compare the optical spectrum obtained with a  $24 \times 24 \times 1$  sampling with those from (c). The increased sampling density had a significant effect on the peak positions in the spectrum. Calculate the exciton binding energy (compared to the electronic band gap at the K point) in both cases and compare them to the recent results reported in <https://arxiv.org/pdf/1807.06413.pdf>.

It is quite possible that even the  $24 \times 24 \times 1$  k-point grid does not yet allow for fully converged peak positions and exciton binding energies. If you are interested, do another calculation for a  $36 \times 36 \times 1$  grid. For this, increase the walltime in the job submission script to 12:00:00. You can also ask for 2 computational nodes to further speed up the calculation, but don't forget to then also adjust the number of CPUs in the mpirun command on the bottom of the file from 20 to 40 and to change parallelization options in the BSE input file accordingly.

- (e) Let's now further analyse the excitation spectrum and have a look at the actual excitations that give rise to the peaks in the macroscopic dielectric function.

---

<sup>2</sup> This cutoff energy corresponds to the inclusion of local field effects in the optical spectra, which have a noticeable quantitative effect, but can be largely neglected if one is just interested in a purely qualitative picture.

<sup>3</sup> Phys. Rev. B 73, 233103 (2006), Physical Review B **94**, 155406 (2016)

Go to the folder where you did the BSE calculations for the 24x24 grid and run the command

```
~/bin/yambo/ypp -M -e s
```

This will call ypp and tell it to write the excitation spectrum into two files, one ordered by energy and one ordered by intensity. Look at the file `o.exc_E_sorted`. Clearly, the lowest energy peak consists of two contributions, are very strong one and a much weaker one, which are at about the same energy. The second peak has a similar structure, but weaker intensities.

YAMBO can also tell us about the composition of the excitations in terms of Kohn-Sham orbitals. Run the command

```
~/bin/yambo/ypp -M -e a -Q -F ypp.in_amp
```

and then open the created file `ypp.in_amp`. For each of the states in the range defined by `States`, ypp will write a file that contains the Kohn-Sham orbital contributions. Change `States` to "1-4" to request output for the first 4 states, i.e. the first two exciton peaks. Here, ypp will combine all excitations with an energy separation smaller than `Degen_Step` to single excitation. This is supposed to correct numerical errors in the calculation which might lead to a small energy separation of excitations which should actually be degenerate. Run ypp with this input file. You will see that ypp merges states 1 and 2, as well as 3 and 4. This will create four files `o.exc_weights_*`, one for each excitation. Open `o.exc_weights_at_1`. Ypp lists a number of transitions that contribute to the exciton, each in terms of the band index and k-point of the hole (`Band_V` and `K_V` ibz) and the electron (`Band_C` and `K_C` ibz) in the symmetry-reduced k-point grid. The energies in the final column are given without the quasiparticle corrections.

We can see here, that the excitation is mainly composed of transitions between the valence band maximum (band 4) and conduction band minimum (band 5) at the K points of the hexagonal Brillouin zone (k-point index 19 for the 12x12x1 grid and 61 for the 24x24x1 grid).

The same is true for the third excitation, but its composition slightly differs from that of excitation 1. It is thus likely that excitation 1 is the excitonic ground state and excitation 3 is the first excited state (like in a hydrogen atom).

Finally, let's also have a look at the excitonic wavefunction of excitation 1. Run

```
~/bin/yambo/ypp -M -e w -Q -F ypp.in_wfn
```

which creates an input file for the calculation of the exciton wavefunction. Change the format to "x", which will create an output file that can be plotted using XCrysDen or VESTA, and the `Direction` to "123" (creating the data for a 3D plot of the wavefunction, as opposed to averaging it). Now, we need to decide how many unit cells we want to include in the supercell that will contain the plot of the exciton wavefunction. Change this to `11 | 11 | 1 |`, i.e. we want a supercell of 121 unit cells (ypp requires the supercell dimension to be odd integer multiple of the unit cell lattice vectors in each direction). Also, we need to decide where to place the hole for the plot. The exciton wavefunction depends on both the three coordinates of the holes and the three coordinates of the electron, which we, obviously, cannot really plot. For this reason, it is customary to fix the hole at a point and only plot the electronic part of the wavefunction. The resulting plot will depend on the position of the hole, so it should be chosen reasonably. We will here assume that the hole is located in a  $p_z$  orbital of nitrogen, which strongly contributes to the highest energy valence band. We will thus place the hole slightly above the nitrogen atom in the unit cell (see the reference `ypp.in_wfn`, but make sure to that

the coordinates in fit to your coordinates from step (a)! ). The x- and y- coordinates can be obtained from the position of the nitrogen atom in terms of the lattice vectors ( $1/3, 1/3$  for  $60^\circ$  angle between the lattice vectors and  $1/3, -1/3$  for a  $120^\circ$  angle) in combination with the lattice vectors reported by YAMBO at the beginning of the BSE calculation. Also change FFTGvecs to 800 eV to save some computational time.

After the preparations are done, use the provided job script to submit and run the ypp calculation on the HPC. The calculation will take about 10 minutes. Visualize the data in XCrysDen or VESTA. The exciton wavefunction has a three-fold symmetry axis, which mirrors the symmetry of the underlying crystal lattice. The exciton wavefunction is quite localized around the position of the hole (the grey ball in VESTA), with a radius of about  $15 \text{ \AA}$ , i.e. about 5 lattice constants. This strong localization was to be expected from the large exciton binding energy of this exciton of about 2 eV, which implies that the bound electrons and holes need to be rather close together.

In a similar way, you can calculate and plot the exciton wavefunction of the second absorption strong peak. In that case, the exciton wavefunction is more delocalized but does not appear to show any nodes that would point towards it being a higher order excitation.