OCEAN

The UW-NIST BSE Project
Department of Physics
University of Washington

User's Guide, OCEAN version 2.0
updated May 15, 2017

Abstract

OCEAN is an *ab initio* Density Functional Theory (DFT) + Bethe-Salpeter Equation (BSE) code for calculations of core-level spectra. Currently the code allows for the calculations of x-ray absorption spectra (XAS), x-ray emission (XES), non-resonant x-ray inelastic x-ray spectra (NRIXS), and (direct) resonant inelastic x-ray scattering (RIXS) of periodic systems. The code is written in Fortran 90 with associated shell and Perl scripting.

Contents

Chapter 1

Overview

OCEAN provides a package to numerically solve the Bethe-Salpeter equation for core-level excitations. There are several steps in the process which use a variety of compiled programs and scripts. The entire processes is driven by the main script OCEAN.PL, though the experienced user can easily run subsections of the code individually. The OCEAN workflow is as follows:

1. Setup and parsing of the input file
2. Atomic calculation to construct PAW-style optimal projector functions (OPFs)
3. DFT calculation using one of several external DFT codes
4. Calculation to determine the core-hole screening
5. BSE calculation to get the final spectrum or spectra

The guide is broken up into sections to follow the path the code will take.

Table 1.1: Typographic conventions in this document

font	denotes
SMALL CAPS	names of programs
`typewriter font`	contents of files
‘`quoted typewriter font`’	file names
ROMAN CAPITALS	names of cards in the ‘`input`’ file
`slanted typewriter font`	commands executed at a command line

1.1 Acknowledgments and history

The OCEAN code is based historically off of the NIST BSE codes of Eric Shirley. The development of OCEAN grew out of collaboration between Eric Shirley and John Rehr at the University of Washington by way of the thesis project of John Vinson.

Select OCEAN and OCEAN-precursor papers (in reverse chronological order) :

- K. Gilmore, J. Vinson, E. L. Shirley, D. Prendergast, C. D. Pemmaraju, J. J. Kas, F. D. Vila, J. J. Rehr, Comp. Phys. Comm. **197**, 109 (2015)
- J. Vinson, J. J. Rehr, J. J. Kas, and E. L. Shirley, Phys. Rev. B **83**, 115106 (2011)
- E. L. Shirley, Ultramicroscopy **106**, 986 (2006)
- E. L. Shirley, Phys. Rev. Lett. **80**, 794 (1998)

A valence BSE code is now integrated into OCEAN for RIXS calculations:

- H. M. Lawler, J. J. Rehr, F. Vila, S. D. Dalosto, E. L. Shirley, and Z. H. Levine, Phys. Rev. B **78**, 205108 (2008)
- L. X. Benedict and E. L. Shirley, Phys. Rev. B **59**, 5441 (1999)
- L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. B **57**, 9385 (1998)
- L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. Lett. **80**, 4514 (1998)

The current list of contributors (in alphabetical order):

Keith Gilmore
Josh Kas
Yufeng Liang
Das Pemmaraju
David Prendergast
John Rehr
Eric Shirley
Fernando Vila
John Vinson

Chapter 2

Tutorials

This chapter will run through several tutorials showcasing some of the things OCEAN can do and highlighting some of the strategies once can use to make sure a result is "converged." The issue of error estimation in computational physics is a bit beyond this tutorial, but we will focus on what it takes to get a spectrum out of OCEAN that is a true representation of the underlying theories and approximations: DFT, pseudopotentials, BSE, &c. All of the initial inputs can be found in the '`EXAMPLES`' directory. Each section will use a single system to focus on some feature or settings.

2.1 Quick-start with $SrTiO_3$

Strontium titanate (STO) is a benchmark perovskite which is widely used experimentally as a substrate, interface, or doped to give interesting properties. Computationally it is a good testbed because it is undistorted and both the titanium and strontium are nominally d_0. Here we will just be interested in producing a rough calculation of the Ti $L_{2,3}$-edge x-ray absorption of STO.

- Go to the '`EXAMPLES/STO`' directory

All of the necessary files to run should be here. The main input is '`STO.in`' and will require some minor editing before we run.

When you run OCEAN you need to decide how many processors to run on. This can range from 1 on your laptop to thousands on a cluster. *Nb. When running locally I would recommend selecting fewer processors than your machine has.* You also need to know how your MPI (Message Passing Interface) is called. MPI is what allows the DFT codes to run across many, many nodes on a cluster. Two popular, open-source MPI distributions are MPICH and OpenMPI. To call MPICH you need to use "mpiexec -n N_P" whereas for OpenMPI the usage is "mpirun -n N_P" where N_P is the number of processors you want to run on. This information (the wrapper mpiexec/mpirun and N_P) is placed in the input file in PARA_PREFIX.

The second choice you will need to make before running is which DFT code you want to use. If only one is installed then your choice is easy. Currently OCEAN supports ABINIT and QuantumEspresso. Either will work fine, but for historical reasons the default is ABINIT. Selection between the two is accomplished via the DFT parameter which can be set to either "qe" or "abi."

- Edit '`STO.in`' to reflect your choices for DFT and parallel execution

Before looking at any of the other files or settings we want to get the calculation running.

- Run OCEAN: `/path/to/ocean/ocean.pl STO.in`

While that is running we can look at the other files that were included. First we have the pseudopotential files: '`*.fhi`' and '`*.fhi.UPF`'. Psuedopotentials are covered briefly in chapter 4. They are necessary for any OCEAN calculation. The files '`ti.opts`' and '`ti.fill`' provide a description of the titanium pseudopotential and tell OCEAN information about how to construct the matrix elements between the core and the valence electrons (see chapter 5). There is also the file '`photon1`' which describes the x-ray experiment being done. In this example we are looking at x-ray absorption, including both dipole and quadrupole terms in the electron-photon interaction. The photon file specifies the direction for the x-ray's polarization ϵ and momentum q as well as the approximate energy (which for the titanium $L_{2,3}$ edge is around 450 eV). A complete description of the photon file format is available in section 3.2.

As OCEAN runs it will move between several different directories. First it will parse the input file and put all of the settings for the run into the '`Common`' directory. Next it will carry out a series of atomic calculations to determine the localized basis and matrix elements for the core-level to valence excitations in the '`OPF`' directory. The ground-state DFT calculations will be carried out in '`DFT`', and then a translation into OCEAN formatted files will take place in '`PREP`'. The screening of the core-hole potential is done in '`SCREEN`', and finally the Bethe-Salpeter calculation is carried out in '`CNBSE`'.

When OCEAN has finished running there will be files called '`absspct*`' in the '`CNBSE`' directory. The naming scheme for these files is straight-forward, but long. Our run will have created the file '`absspct_Ti.0001_2p_01`'. This means that this '`absspct`' file is for titanium (Ti); it is the first titanium listed in the structure (0001); we ran for the $L_{2,3}$ edge or, equivalently, the $2p$ core levels (2p); and the corresponding photon file is '`photon1`' (01). The current OCEAN code can support up to 9999 different atomic sites and 99 different photon specifications. *If, for example, you were running on a very large, disordered cell like liquid water and comparing to NRIXS with a dozen different values for the momentum transfer you could run all of them and the outputs would be nicely labeled.*

We need to check the results, so we will plot the resulting '`absspct_Ti.0001_2p_01`' and compare to the included '`absspct`' file in the base '`STO`' directory.

- Open gnuplot (or your other favorite plotting tool) and compare your spectrum to the included one

The two should match very closely, but with the quality settings in the input there may well be differences between the DFT codes (see figure 2.1). The $3d$ transition metals highlight the

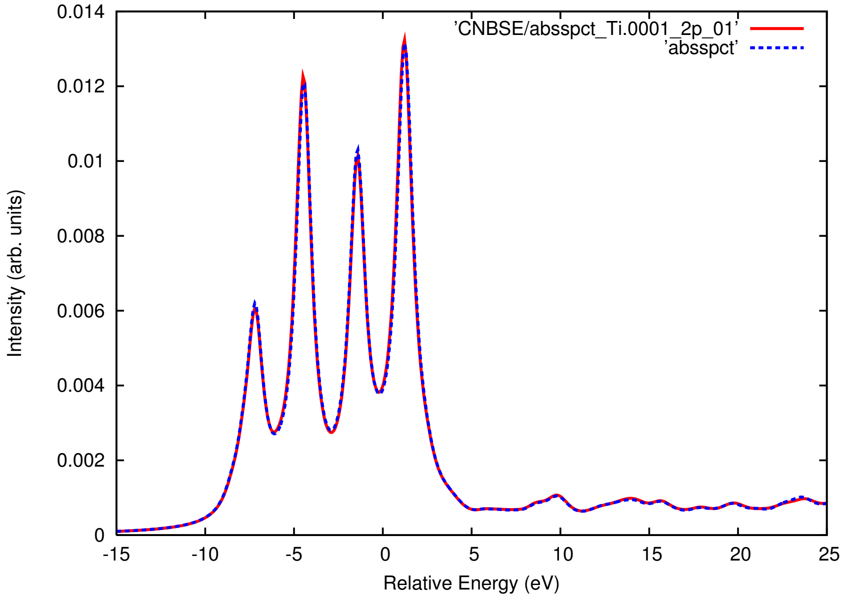

Figure 2.1: The titanium $L_{2,3}$ edge of strontium titanate as calculated using the STO example. Note the near, but not exact, match between the calculated and supplied spectra. A larger discrepancy might point to a bug or problem with your installation.

breakdown of a single-particle picture for x-ray absorption. A quick count shows that the intensity ratio between the L_3 and the L_2 edges should be 2:1 from the initial occupations of the $2p_{3/2}$ and $2p_{1/2}$. Instead for titanium we see a ratio of less than 1. The spin-orbit splitting of the $2p$ orbitals is not diagonal in the same basis as the conduction band d states, and so trying to move backward from the spectrum to the ground-state density of states (for either the $2p$ or $3d$ levels) can be difficult, but the BSE approach in OCEAN makes it easy.

2.2 Core-level shifts with NH_4NO_3

Ammonium nitrate is a wide band gap crystalline solid. Nitrogen atoms occupy central positions in the two constituent ions NH_4^+ and NO_3^-. Given the difference in chemical environment between hydrogen and oxygen nearest neighbors it is of no surprise that there is a significant shift in the energy of the $1s$ electrons between the two nitrogen sites. Therefore NH_4NO_3 makes a good example to look at how we can account for these shifts within OCEAN. As always with our tutorials, feel free to start the calculation and then read while it is running.

2.2.1 Overview of core-level shifts

There are several methods used for calculating energy levels in core-level spectroscopy. For completeness we will give an overview of them here. The most straightforward, popular in quantum chemistry techniques, is to directly compare the energies of initial and final states.

This requires that you calculate explicit initial and final states as well as the core level wave function. Pseudopotential-based techniques need a bit more creativity due to their lack of explicit core states. Methods that carry out a self-consistent calculation in the presence of the core hole (full core hole, half core hole, or eXcited core hole) can make use of the energy differences between DFT calculations with and without the core hole. This gives a relative energy shift that can be calibrated to give meaningful information about the core level shifts both within a cell and between different cells.[**?**]

The OCEAN code has neither explicit core levels nor self-consistent calculations with a core hole, and so we need a different approach. In the BSE approach the only elements we need are the bare energies of the electron and the hole, and then the interaction terms.

$$\begin{aligned} H_{\mathrm{BSE}} &= H_0 + V_x - W \\ H_0 &= \epsilon_{n,\mathbf{k}} - \epsilon_h \end{aligned} \tag{2.1}$$

where the non-interacting part is the quasiparticle energies of the electron $\epsilon_{n,\mathbf{k}}$ and core hole ϵ_h. The correct determination of the excitation energies, the onset of the edge or position of an exciton, relies on an accurate determination of all three components of the BSE Hamiltonian (as well as the validity of neglecting higher-order interaction terms). This tutorial is focused on how H_0 is determined in OCEAN and the effects of core-level shifts on the spectra.

The conduction band energies in H_0 are, as a first approximation, the DFT energies. Were we to have the exact density functional, then the highest-occupied state would have a meaningful energy. A better approach is the GW method which approximates the one-electron Green's function in an interacting many-electron system instead of the non-interacting DFT. For many systems (especially *s*-*p* bonded materials) the DFT energies and wavefunctions, even within the local-density approximation (LDA), are remarkably similar to what would be produced by a GW calculation. The famous DFT underestimation of band gaps is an incredibly small error in comparison to the magnitude of the approximation being made to get there, a reflection of what Walter Kohn referred to as the "short-sightedness" of electrons in condensed matter. The typical result of correcting the DFT energies with a GW calculation is a moderate increase in the band gap and a 3-5 % (5-10 %) stretch of the conduction (valence) bands, pushing the energies away from the Fermi level. For our NH_4NO_3 example any correction to the band gap is present for both nitrogen species, and we will use the DFT energies for the calculation.

A GW calculation for the core level is neither possible nor feasible, and it is this removal energy we want to try and approximate. Because we are not relaxing the system in the presence of the hole (see the various core-hole approaches mentioned previously) this requires two calculations 1) the energy of the core-level electron and 2) the total energy from the relaxation of the rest of the system once that electron is removed.

1. For calculating the energy of the core level we assume that the core level wave function is the same between chemical environments. This means that the kinetic energy of the core level is fixed, and therefore changes in the core-level energy can only come from changes in the local potential integrated over the core-state's density. This comes directly from the DFT calculation as the total Kohn-Sham potential.

2. The response of the rest of the system to the removal of the core-level electron is given (within linear response) by the screening of the core-hole potential that we are already calculating. An additional factor of 1/2 is picked up from adiabatically moving the charge of this electron.

To evaluate both 1 and 2 we approximate the core as a delta function at its location τ. The energy of the core hole (as it would then appear in the BSE Hamiltonian H^0) is given by

$$\epsilon_h = V_{\mathrm{KS}}(\tau) - 1/2W(\tau) + X \tag{2.2}$$

where X is an offset which will not vary from site to site and encapsulates the core's kinetic energy term and also the screening response of the other core electrons – neither of which change from site to site by assumption.

Now that we are able to generate both $\epsilon_{n,\mathbf{k}}$ and ϵ_h we can accurately determine the excitation spectra energies — so long as the interacting parts of the BSE are also carefully calculated. The absolute alignment is not *ab initio*, but requires only a single, element-specific offset X.

2.2.2 Ammonium Nitrate CLS

Measurements of the nitrogen 1*s* x-ray emission (Kα) from ammonium nitrate show two well separated peaks. Calculations of using several different computational tools concur that these two peaks belong separately to the ammonium and nitrate ions.[**?**] In contrast we can look at the local, angular-momentum-projected density of states around the nitrogen atoms in the ground state. (This is a standard feature of DFT codes, and here we are using the output from QE). In figure 2.2 we show the *p*-type states around both kinds of nitrogen in ammonium nitrate, and clearly the main contributions to the nitrogen Kα emission will be at the same energy for both the NH_4^+ and NO_3^-. This discrepancy is because the ground-state calculation has no core-level shift.

Having hopefully given some introduction to the idea of core-level shifts we can now turn to calculating spectra.

- Go to the '`EXAMPLES/AN`' directory
- Edit '`nh4no3.in`' to set your DFT and MPI (remember tutorial 1?)
- Run OCEAN: */path/to/ocean/ocean.pl nh4no3.in*

Core-level shifts are calculated in the screening stage. The parameter CORE_OFFSET controls the calculation. If it is set to "false" no calculations are done. If you inspect the '`nh4no3.in`' input file you will see that this is the case for the example. Within OCEAN there are three options for this parameter:

1. false – don't run any shifts
2. true – only interested in a single cell or testing

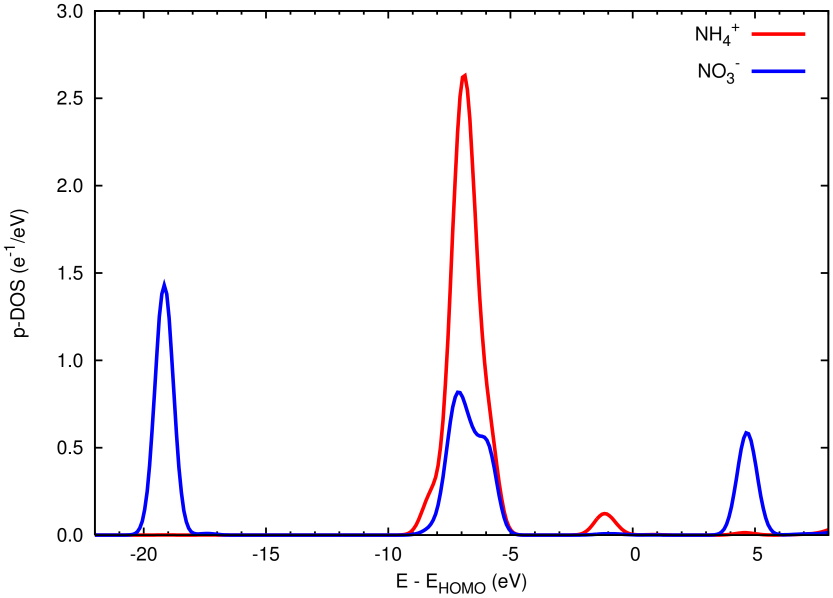

Figure 2.2: The projected $l = 1$ (p-type) DOS around the two nitrogen sites in ammonium nitrate. In the ground state both the ammonium and nitrate nitrogen have significant p character at the same energy (5-10 eV below the HOMO). In contrast XES measurements and calculations show the ammonium peak 4-5 eV below the nitrate peak.

3. [any number] – running many cells, e.g., molecular dynamics snapshots.

Choosing the number to set CORE_OFFSET can be tricky because we have no idea what values the Kohn-Sham potential or screening calculation will give ahead of time. Instead, we can use the option "true" and the code will figure out what value to set CORE_OFFSET to such that the average offset is zero. Go ahead and change the input so that CORE_OFFSET is "true" and rerun the calculation. If you'd like you can use the script "re_run.sh" by setting the directory where OCEAN was installed.

This time the code has performed additional calculations to get the Kohn-Sham potential at each nitrogen site. The difference is remarkable – the main parts of the ammonium and nitrate XES are no longer overlapping, but instead separated by several eVs. The core-level shifts are unique to each site and each screening radius. As such they are stored within the '`SCREEN`' directory by site and radius, i.e., '`SCREEN/zN_0001/`'. If you look at these files you will see that the two nitrate and two ammonium nitrogens are almost the same (small differences are inevitable) as we expect since they are equivalent sites.

2.2.3 Further details

The calculation of the core-level shifts are carried out in the '`SCREEN`' directory because the core-hole screening enters in (Eq. 2.2). The auxiliary script CORE_SHIFT.PL takes care of the calculation, running the necessary post-processing for either ABINT or QUANTUMESPRESSO.

2.3 RIXS of diamond

Resonant inelastic x-ray scattering (RIXS) is a photon-in–photon-out spectroscopy where the response of the system can depend strongly on both the incoming and outgoing x-ray photons. Several effects can be probed with RIXS, including the phonons and spin-density waves, but OCEAN is a BSE code, and, therefore, it is only capable of exploring the excitonic, band-structure contributions to RIXS spectra. After a brief overview of the theory we will walk through an example looking at the carbon K edge in diamond. In addition to the regular input file the diamond example has a shell script '`diamond.sh`'. You will need to edit this script to include the path of the OCEAN executables.

2.3.1 RIXS theory

Before embarking on RIXS calculations we implore you to read the excellent reviews by Kotani and Shin (Rev. Mod. Phys. **73**, 203, 2001) and Ament *et al.* (Rev. Mod. Phys. **83**, 705, 2011). With OCEAN we are able to calculate what Ament *et al.* refer to as direct RIXS. To be filled in later.

2.3.2 Method

The way we go about RIXS calculations with OCEAN is a two step process: a set of core-level calculations followed by a valence one. First, we carry out a core-level BSE calculation to determine the intermediate states. From Eq. **??** above we have

$$\sigma(\mathbf{k},\hat{\epsilon},\omega;\mathbf{k}',\hat{\epsilon}',\omega') \propto \sum_f \left| \sum_M \frac{\langle f|\hat{d}^\dagger_{\mathbf{k}',\hat{\epsilon}'}|M\rangle\langle M|\hat{d}_{\mathbf{k},\hat{\epsilon}}|i\rangle}{\omega-(\varepsilon_M-\varepsilon_i)+i\Gamma_M} \right|^2 \delta(\varepsilon_f-\varepsilon_i+\omega'-\omega) \tag{2.3}$$

but now we denote the intermediate states as $M = m, \alpha$ with both the conduction-band electronic states m and the specific core site α. If a system has more than one atomic site of the same element (or with edges with overlapping energy ranges, e.g., $\mathrm{V}_x\mathrm{O}_y$) then each site must be included coherently in the sum over intermediate states. We can rewrite Eq. 2.3

$$\sigma(\mathbf{k},\hat{\epsilon},\omega;\mathbf{k}',\hat{\epsilon}',\omega') \propto \sum_f \left| \langle f|\hat{d}^\dagger_{\mathbf{k}',\hat{\epsilon}'} \sum_{M=m,\alpha} \frac{|m\rangle\langle m|\hat{d}_{\mathbf{k},\hat{\epsilon};\alpha}|i\rangle}{\omega-(\varepsilon_m+\Delta_\alpha)+i\Gamma_\alpha} \right|^2 \delta(\varepsilon_f+\omega'-\omega) \tag{2.4}$$

by setting ε_i to zero and breaking out the explicit reference to which site the photon operator $\hat{d}$ is acting on. We are able to separate out the **absorption** by site because the core orbital is localized and has no overlap with the core orbitals at another site.

We can now define the conduction-band–core-hole "exciton" as

$$|y(\omega,\mathbf{k},\hat{\epsilon};\alpha)\rangle \equiv \sum_m \frac{|m\rangle\langle m|\hat{d}_{\mathbf{k},\hat{\epsilon};\alpha}|i\rangle}{\omega-(\varepsilon_m+\Delta_\alpha)+i\Gamma_\alpha} \tag{2.5}$$

where the scare-quotes around exciton remind us that this is not an eigenstates of the two-particle BSE Hamiltonian. Instead we are interested in what we could call the experimentally-measured exciton which is the weighted, coherent sum over all possible intermediate states that can be created by our transition operator. The weighting is from the resonant energy denominator, broadened by the core-hole lifetime. (Additional broadening could come from the lifetime of the conduction band states which make up m and possibly from phonon coupling, but these are neglected here.) In OCEAN the exciton is solved for using a two-particle, electron-hole basis. The electron orbitals are in a Bloch representation with band index n and crystal momentum $\mathbf{q}$ to represent the electron part of the exciton, while the holes are indexed by α to represent the atomic site and principle, angular, and azimuthal quantum numbers and also form Bloch states with momentum $\mathbf{q}-\mathbf{k}$. By construction we are conserving momentum, hence the hole and electron are offset by the momentum of the absorbed photon.

Inserting our basis into Eq. 2.5 we get

$$\begin{aligned}
|y(\omega,\mathbf{k},\hat{\epsilon};\alpha)\rangle &= \sum_{\substack{c,\mathbf{q}\\ \alpha,\mathbf{q}-\mathbf{k}}} \frac{|c\mathbf{k};\alpha,\mathbf{q}-\mathbf{k}\rangle\langle c\mathbf{k};\alpha,\mathbf{q}-\mathbf{k}|\hat{d}_{\mathbf{k},\hat{\epsilon};\alpha}|i\rangle}{\omega - H_{\text{BSE}} + i\Gamma_\alpha} \\
&= \sum_{\substack{c,\mathbf{q}\\ \alpha,\mathbf{q}-\mathbf{k}}} \frac{|c\mathbf{k};\alpha,\mathbf{q}-\mathbf{k}\rangle\langle c\mathbf{k}|\hat{d}_{\mathbf{k},\hat{\epsilon};\alpha}|\alpha,\mathbf{q}-\mathbf{k}\rangle}{\omega - H_{\text{BSE}} + i\Gamma_\alpha}
\end{aligned} \tag{2.6}$$

where, because we have assumed the excited state is a single electron-hole pair, the only connection between the initial state i and the intermediate state is reduced to the transition matrix element between the core level α and conduction band c.

We have been a bit sloppy, but our problem is now reduced to

$$\begin{aligned}
|x(\omega)\rangle &= \frac{1}{\omega - H + i\eta}|b\rangle \\
|x(\omega)\rangle &= \frac{1}{A(\omega) + i\eta}|b\rangle \\
Ax &= b
\end{aligned}$$

where we renamed $y \to x$ and $\Gamma \to \eta$ to better match standard linear algebra conventions. This can be approximated using iterative techniques, such as the generalized minimal residual method used in OCEAN (Saad and Schultz, 1987). See section **??** for a summary and reference to the parameters that control this approximation within OCEAN. Iterative linear algebra techniques rely on the finite imaginary term η to avoid dividing by 0 and instabilities from finite-precision arithmetic. Here we have an advantage as the imaginary term physically corresponds to the inverse lifetime of the intermediate state Γ which can be quite large compared to the other energy scales in the Hamiltonian, making convergence easier than for an infinitesimal η.

We can substitute our exciton from Eq. 2.5 into Eq. 2.4 and then expand to alleviate the sum over final states

$$\sigma(\mathbf{k},\hat{\epsilon},\omega;\mathbf{k}',\hat{\epsilon}',\omega') \propto \sum_f \left| \langle f|\hat{d}^\dagger_{\mathbf{k}',\hat{\epsilon}'} \sum_\alpha |y(\omega,\mathbf{k},\hat{\epsilon};\alpha)\rangle \right|^2 \delta(\varepsilon_f+\omega'-\omega) \tag{2.7}$$

$$\propto \sum_f \left| \langle f|\hat{d}^\dagger_{\mathbf{k}',\hat{\epsilon}'}|\bar{y}(\omega,\mathbf{k},\hat{\epsilon})\rangle \right|^2 \delta(\varepsilon_f+\omega'-\omega)$$

$$\propto -\mathfrak{Im}\left[\langle \bar{y}(\omega,\mathbf{k},\hat{\epsilon})|\hat{d}_{\mathbf{k}',\hat{\epsilon}'}\frac{1}{\omega'-\omega-H}\hat{d}^\dagger_{\mathbf{k}',\hat{\epsilon}'}|\bar{y}(\omega,\mathbf{k},\hat{\epsilon})\rangle\right] \tag{2.8}$$

where $\bar{y}=\sum_\alpha y(\alpha)$. The operator $\hat{d}^\dagger_{\mathbf{k}',\hat{\epsilon}'}$ acting on the excitonic state y fills the core hole α leaving a valence hole (and the conduction electron). The Hamiltonian in Eq. 2.8 is acting on a space containing a valence hole and a conduction band electron, and is approximated in OCEAN using the valence BSE solver, originally released as AI2NBSE. This is done by inserting a set of valence and conduction band states (eigenstates of the ground-state DFT Hamiltonian)

$$\sigma(\mathbf{k},\hat{\epsilon},\omega;\mathbf{k}',\hat{\epsilon}',\omega') \propto -\mathfrak{Im}\left[\langle \bar{y}(\omega,\mathbf{k},\hat{\epsilon})|\hat{d}_{\mathbf{k}',\hat{\epsilon}'}|vc\rangle\langle vc|\frac{1}{\Omega-H}|v'c'\rangle\langle v'c'|\hat{d}^\dagger_{\mathbf{k}',\hat{\epsilon}'}|\bar{y}(\omega,\mathbf{k},\hat{\epsilon})\rangle\right] \tag{2.9}$$

and then using the Haydock method to generate the spectra with respect to the energy loss $\Omega=\omega'-\omega$. The states v and c refer to valence and conduction band states, respectively, and the momentum dependence has been dropped for legibility.

2.3.3 RIXS script: diamond.sh

In the example directory is the script '`diamond.sh`' which will run all the steps necessary for the RIXS calculation of diamond. This script can either be run directly (on a workstation or interactive session) or pasted into a batch script, but you'll need to edit the variable "OCEAN_BIN" to point to your OCEAN installation directory. As remarked in the previous section the calculation is broken up into several steps. The provided '`diamond.sh`' script can be used as a template for your own RIXS runs.

2.3.4 First, calculate the x-ray absorption

For a variety of reasons, as a first step in calculating RIXS we almost always want to run an absorption calculation. If we are comparing to experiment we will need to establish an alignment between the energies measured in the experiment versus the calculated energies which have no absolute core-level alignment. On the other hand if we are calculating ahead of an experiment we will probably want to look at the absorption to try and pick interesting features to investigate. In any case, the core-level BSE is a much cheaper calculation (usually fewer $1s$ or $2p$ states than valence bands, and the core-levels are nicely localized in space), and it will be easier and faster to check the convergence parameters by looking at the absorption than it would be by running the full RIXS.

As with any tutorial feel free to start it before reading the section. In this case run it using the script '`diamond.sh`' which may require editing to specify the path to your OCEAN

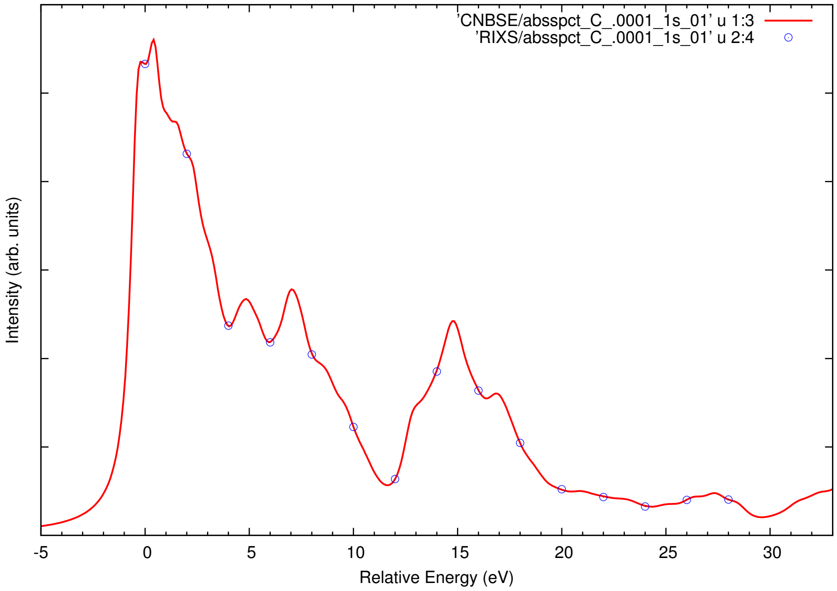

Figure 2.3: The calculated XAS at the carbon K edge for diamond (red line) and the same calculation done using GMRES for a select set of 15 points (blue circles).

installation. There are a couple of differences in this input file '`diamond.in`' from some of the other tutorials that should be noted. The first is the presence of 'photon_q' (the momentum transferred to the electrons by the photons) which must be present and non-zero for all RIXS (or valence) calculations. This is both physically true, the scattered or emitted photons are going to be measured at a non-zero angle from the incoming beam and therefore represent a finite momentum transfer, and because parts of the calculation diverge in the limit of zero momentum transfer. Some BSE codes go to great length to work around the divergence — we just use a finite momentum transfer (which can be very, very small). The second is that we use a negative number for 'edges' which tells the code we want to calculate every element with a matching Z. In our case we are asking for every carbon atom ($Z = 6$). This is a good habit for RIXS calculations as physically you must include every matching site in the intermediate sum, but the code has no checks to enforce this.

For your own, future calculations you'll want to spend some time with the absorption calculations. You can see how many bands are needed to cover the energy range you are interested in, and check the convergence of both the k-points and xmesh. Valence calculations (including the final part of RIXS) scale with the size of xmesh to the second power. Unlike the core-level calculation, the valence-band hole can be anywhere in the unit cell. These calculations take a bit longer than the core-level BSE. But, for the example we won't bother with second-guessing or checking and move right along.

2.3.5 Second, run the core-level code again

The second (and third) steps are both run by the helper code RIXS.PL. We now much determine the excitonic states at each incident photon energy (and polarization) which requires switching the algorithm in the BSE code. This is done via the line in DIAMOND.SH that says "echo gmres > Common/cnbse.solver", which switches us to using the GMRES method. There are two ways to determine which energies to use for the incoming photons. In the diamond example we make use of 'cnbse.gmres.erange' which allows us to set the energy range and step size. The units are eVs and the range is given with respect to the same energy axis of the '`absspct`' files from the previous absorption calculation (this is why we do that first). In DIAMOND.SH the line 'echo 0 30 2 > Common/cnbse.gmres.erange' tells the code to start at 0 and calculate every 2 eV until it hits 30. The last input needed to determine the excitons is the description of the incoming photon. By default the code will use every available photon file, but by using "photon_in" we are able to tell it to only use the first one '`photon1`'.

After the core-level BSE has run a new set of '`absspct`' files has been created, but these have a slightly different format. You should plot '`CNBSE/absspct_C_.0001_1s_01`' using columns 1 and 3 and '`RIXS/absspct_C_.0001_1s_01`' using columns 2 and 4 (see Fig. 2.3). The RIXS version has only a few points, running from 0 to 30 and spaced every 2 eV, but these points should lie on top of the curve from CNBSE. More importantly, the '`RIXS`' directory is now full of '`echamp_C_`' files. These contain the information about the excitons $y_{\hat{\epsilon},\alpha}(\omega)$. For the diamond example we should have 30 fileechamp_C_ files: 15 energy steps times 2 carbon atoms per cell times 1 incoming photon files. The code also has generated '`cksv`' files which contain the transition information for a valence-band electron to fill the core hole $\langle v|\hat{d}^{\dagger}_{\mathbf{k},\hat{\epsilon}}|\alpha\rangle$. Together these are used to construct the seed vectors for the valence-level BSE calculations.

2.3.6 Third, run valence-level calculations

The last step in the RIXS calculation is a set of valence BSE runs to solve for the final conduction-electron–valence-hole spectra. The number of valence runs in our example is 15: 15 energy steps times 1 photon-in–photon-out pair. In general, if you do not specify 'photon_in' and 'photon_out' like we have in this example the code will run for every available pair $N_{\text{photon}}(N_{\text{photon}} - 1)$. We exclude scattering into the same photon polarization and direction. Often experiments are set up such that the detector is parallel to the incoming photon's polarization, requiring that $\hat{\epsilon} \cdot \hat{\epsilon}' = 0$ and suppressing the A^2 term in the electron-photon interaction.

The final RIXS spectra are named a bit differently than the XAS or XES spectra files. They are a function of photon in, incoming energy, and photon out, but not dependent on which atomic site (all sites much be summed over coherently). The naming scheme is then something like '`rxsspct_C_.0000_1s_01.08.02`', where we are looking at the carbon 1s edges using the 1st photon file as our incoming beam, at the 8th energy step, and the 2nd photon file as the outgoing x-ray. (*N.b.* we are planning on getting rid of the unused '0000' which for '`absspct`' or '`xesspct`' files denote the atomic site.) The format of the '`rxsspct`' files follows that of the valence BSE output. For RIXS only the energy in column 1 and ϵ_2 in column 3 are

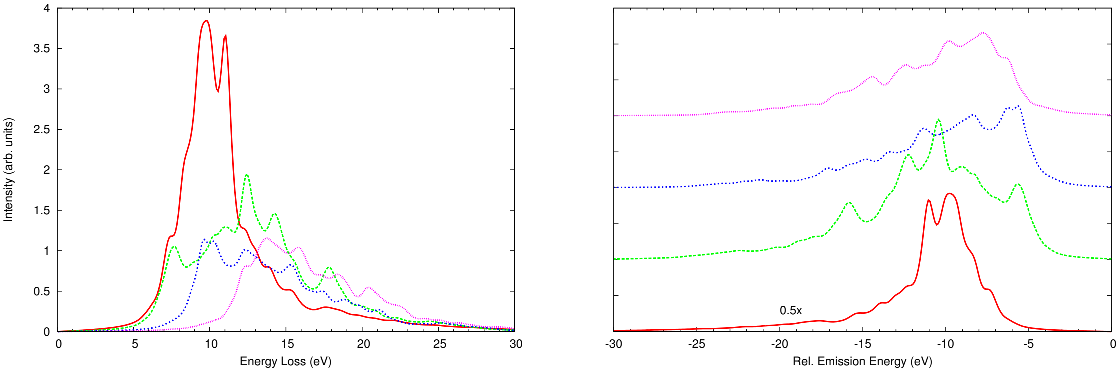

Figure 2.4: The first four calculated energies of the diamond K-edge RIXS plotted either as a function of (a) energy loss or (b) emission energy. In (b) the first energy has been scaled down by 0.5 and the spectra are offset vertically for clarity.

meaningful. The energy scale is in eV, but it is measuring the loss $\Omega = \omega - \omega'$, see Fig. 2.4(a) and (b).

Chapter 3

Inputs

The OCEAN package is run using the PERL script OCEAN.PL which takes care of reading in the input file and then running the various executables. This chapter will introduce the various controls and explain their formatting and use.

3.1 Running OCEAN

In addition to the main input file there are several auxiliary files. The '`*.opts`' and '`*.fill`' files are described in sections 5.0.3 and 5.0.2 respectively, and the pseudopotentials are touched upon briefly in chapter 4. The next section, 3.2, covers the '`photon?`' files which specify the necessary information about the x-ray probe.

The input parser for OCEAN is relatively basic. Any text following a comment character, '#,' is ignored up to the line return. If a file has been switching around operating systems it may have non-unix line returns which can cause problems. If the parser returns odd errors try running DOS2UNIX. The parser is case ignoring and converts all inputs to lower case, requiring any filenames in the input file to be lower case. Each card must be on its own line followed by the value or values to be assigned it. All card values can be surrounded by braces ({}) and multi-value cards must be surrounded by braces. For example, the NKPT card must be a list of 3 integers and so must have the following form: "NKPT { 2 2 2 }" whereas the OCCOPT card takes only one integer and may therefore be listed: "OCCOPT 1" or "OCCOPT { 1 }". Line returns may be present within an input.

After parsing the cards are all stored as small, fortran formatted files in the '`Common`' directory. The package then goes through the following steps:

1. An atomic calculation is done to create the optimal projector functions (OPFs).
2. Slater F and G integrals are computed at the atomic level
3. ABINIT or QuantumEspresso input files are prepared
4. ABINIT (or QuantumEspresso) is run to get DFT valence density and wave functions.

This requires usually 3 runs. One self-consistent run to get the converged density and two additional runs to calculate wave functions for the screening calculation and for the final states.

5. The wave functions and density are converted from ABINIT (or QE) output to the form expected by OCEAN.
6. The core electrons are screened self-consistently in the presence of the core-hole in the atomic system
7. The spherical part of the direct interaction is screened using RPA. χ^0 is calculated using real-space integration around the site of interest. The RPA is coupled with a form of the Levine-Louie-Hybersten screening model.
8. The overlap is calculated between the core wave function and the conduction band using the specified photon operator, XAS or RXIS.
9. The Bethe-Salpeter equation (BSE) is iteratively solved using a form of the Haydock recursion algorithm.

3.2 Photon files

The '`photon`' files determine the specifics of the x-ray operator. A single OCEAN calculation may involve a single '`photon`' file or many, depending on the experiment in question. There is currently a hard limit of 99 (which may be excessive), and each should be numbered.

As an example, if one is simulating non-resonant inelastic x-ray scattering NRIXS (also called x-ray Raman or XRS) the experiment being compared to may include a dozen different values for the momentum transfer. This is easily done with OCEAN and the user needs simply to create 12 files named '`photon1`' through '`photon12`'.

3.2.1 Photon operators

In OCEAN we currently support 3 photon operators:

- dipole: $\hat{\epsilon} \cdot \mathbf{r}$
- quad: $\hat{\epsilon} \cdot \mathbf{r} + i/2\,(\hat{\epsilon} \cdot \mathbf{r})\,(\mathbf{q} \cdot \mathbf{r})$
- NRIXS: $e^{i\mathbf{q}\cdot\mathbf{r}}$

For most light elements the combination of a small core wave-function (small $\mathbf{r}$) and small photon energy (therefore small $\mathbf{q}$) conspire to make dipole and quad effectively the same, i.e., they are in the dipole limit. If you are ever unsure just make two '`photon`' files and check!

3.2.2 Photon format

Below is a sample '`photon`' file. Note that for the photon file, depending on the calculation, only some of the information in it will affect the calculation. Look back to the operators in the previous section. They don't use all the same variables.

dipole	operator
cartesian 0 0 1 end	$\hat{\epsilon}$
cartesian 0 1 0 length 8.55 inverseangstrom end	$\hat{q}$ $\|\mathbf{q}\|$
9000	X-ray energy (eV)

The file has four sections. In the first we select our operator. In the second we set the direction for the x-ray polarization vector ϵ (currently no circularly polarized light). The next section's meaning depends on the operator in play. For quad it sets the q-vector direction for the incoming light, while for NRIXS it sets the direction and magnitude of the momentum **transfer** (the length specifier is optional unless NRIXS is selected). For dipole the q-vector doesn't enter in. The fourth section is only important for quad and it sets the magnitude of q by way of the photon's energy in eV.

3.3 Complete list of ocean inputs

The input of OCEAN can be divided into several sections

- run info
- structural information of the material
- non-structural DFT parameters, e.g., convergence
- atomic and screening info
- parameters that control the BSE section and its convergence
- parameters that only affect the output spectrum

Run Control ....

Purpose: Control aspects of the run

Optional Cards: PARA_PREFIX, SER_PREFIX, DFT, and SCRATCH

Structural information ....

Purpose: Specify the structure

Required Cards: ACELL, RPRIM, NTYPAT, ZNUCL, NATOM, TYPAT, XRED, and DIEMAC

Optional Cards: NSPIN, SMAG, LDAU, and ZYMB

Non-Structural DFT parameters

Purpose: Control convergence of the DFT calculation

Required Cards: PP_LIST and ECUT

Recommended Cards: NKPT, NGKPT, NBANDS, SCREEN.NBAND, OCCOPT, and FBAND

Optional Cards: TOLDFE, TOLWFR, NSTEP, SCREEN.NKPT, ABPAD, DFT_ENERGY_RANGE, SCREEN_ENERGY_RANGE, K0, and DEGAUSS

Atomic and Screening Info

Purpose: Controls the OPF and screening sections of the calculation

Required Cards: OPF.FILL, OPF.OPTS, and EDGES

Recommended Cards: SCREEN.SHELLS, CNBSE.RAD, SCFAC, and CORE_OFFSET

Optional Cards: OPF.HFKGRID

BSE Parameters

Purpose: Parameters used in the BSE calculation

Recommended Cards: CNBSE.XMESH, CNBSE.MODE, CNBSE.NITER, METAL, SPIN-ORBIT, PHOTON_Q, CNBSE.STRENGTH, CNBSE.SOLVER, CNBSE.GMRES.ELIST, CNBSE.GMRES.ERANGE, CNBSE.GMRES.NLOOP, CNBSE.GMRES.GPRC, and CNBSE.GMRES.FFFF

Spectrum Information

Purpose: Controls the spectrum that is written out, not any aspect of the actual calculation

Recommended Cards: CNBSE.BROADEN

Optional Cards: CNBSE.SPECT_RANGE

These CARDS are listed below in (almost) the same order as in the table above. Each CARD description is of this form:

♦ **CARD** `required arguments` *type*
or
♦ **CARD** `required argument list[ length of list ]` *type*

The type is one of *Required*, *Recommended*, or *Optional*. The argument list is a brief statement of the valid arguments to the card. Arguments in square brackets are optional. The text description explains the arguments and their uses more fully. Example uses of the card look like this:

```
* brief description of the example
CARD  arguments
```

For arguments that are 2-D arrays the ordering is in normal math / C ordering.

3.4 Run Control

♦ **PARA_PREFIX** `para_prefix` *Optional*
This is necessary for parallel runs. Depending on your system you may need so helper to launch parallel jobs such as MPIRUN, MPIEXEC, SRUN, APRUN, &c. The OCEAN scripts are designed to try and parse this to figure out how many processors you are trying to run across, e.g., 'mpirun -n 12' will be parsed to tell OCEAN to prepare DFT runs across 12 cores. If you need to include additional flags, e.g., 'mpirun -n 16 --map-by ppr:8:socket' just make sure that the first number is the number of processors.

♦ **SER_PREFIX** `ser_prefix` *Optional*
On some systems CUT3D will hang when launched on multiple processors, but must be launched with the parallel job launcher from above, i.e., 'mpirun -n 1'.

♦ **DFT** `dft` *Optional*
This controls which DFT code to use by specifying either 'qe' for QuantumEspresso or 'abi' for ABINIT. By default ABINIT is used.

♦ **SCRATCH** `scratch` *Optional*
Recommended for compute clusters with local scratch space so that ABINIT or Quantum Espresso will write scratch files somewhere reasonable. The path '`scratch`' should include an absolute path. Otherwise the default behavior is that scratch files will be written to the '`DFT`' directory.

3.5 Structural Information

♦ **ACELL** `acell[ 3 ]` *Required*
ACELL sets the scaling (in Bohr) for the primitive vectors of the unit cell which are in turn set by RPRIM.

```
* For Li metal: BCC structure, a = b = c = 6.595 Bohrs
```

```
ACELL { 6.595  6.595  6.595 }
```

♦ **RPRIM** `rprim[ 3 ][ 3 ]` *Required*

RPRIM sets the primitive vectors for defining the unit cell.

```
* For a BCC
RPRIM{ 1.0  0.0  0.0
       0.0  1.0  0.0
       0.5  0.5  0.5  }

* or the more symmetric
RPRIM{ -0.5  0.5  0.5
        0.5 -0.5  0.5
        0.5  0.5 -0.5  }
```

♦ **NTYPAT** `ntypat` *Required*

The card NTYPAT gives the number of different types of atoms in the cell.

♦ **ZNUCL** `znucl[ntypat]` *Required*

`znucl`[] is a list of length `ntypat` that gives the atomic numbers of the types of atoms that are in the cell.

♦ **NATOM** `natom` *Required*

The card NATOM gives the number of total atoms in the cell.

♦ **TYPAT** `typat[natom]` *Required*

The card TYPAT lists each atom by the rank it has in the ZNUCL list. The order they are listed in will correspond to XRED. (This may be fixed soon to be less idiosyncratic.)

♦ **XRED** `xred[natom,3]` *Required*

The XRED card lists the reduced coordinates, (x, y, z), of the all `natom` atoms in the cell.

```
* For a diamond-like structure with 2 atoms per unit cell
XRED{ 0.0  0.0  0.0    # atom 1
      0.25 0.25 0.25 } # atom 2
```

◆ **DIEMAC** `diemac` *Required*

The MACropscopic DIElectric constant `diemac` is required for both the DFT stage and also as a parameter in the screening calculations. It is included here as a physical property of the system of interest.

◆ **NSPIN** `nspin` *Optional*

Can choose to have spin=1 (paramagnetic) or spin=2. Currently, only collinear spin is allowed.

◆ **SMAG** `smag` *Optional*

Controls the initial magnetism for the DFT calculations. See the hematite example and the ABINIT or QUANTUMESPRESSO documentation for more information.

◆ **LDAU** `ldau` *Optional*

Controls the $+U$ parameters for the DFT run. Only works with programQuantumESPRESSO currently. See the hematite example.

◆ **ZSYMB** `zsymb` *Optional*

Each element listed in ZNUCL also needs a symbol when running using QUANTUMESPRESSO. By convention this will be the atomic symbol: Fe for iron, and the code can automatically populate this for you. However, sometimes the same element needs to be marked as inequivalent, either because of local magnetic order or charge-state ($+U$). For example, in hematite one needs two different iron species in order to set one site as spin up and one as spin down. When this is the case the user needs to list iron twice ZNUCL and provide zsymb.

```
ntypat 3
znucl{ 26 26 8 }
zsymb{ Fe1 Fe2 O }
```

3.6 DFT Parameters

◆ **PP_LIST** `files[ntypat]` *Required*

The names of the pseudo-potential files need to be listed here in the same order as znucl Currently all these files need to be all lower case.

◆ **ECUT** `ecut` *Required*

The plane-wave basis is truncated according to `ecut` measured in Rydberg. The plane-wave cut-off is determined mostly from the pseudo-potentials being used and convergence should be checked with respect to ground state energy.

◆ **NKPT** `nkpt[3]` *Recommended*

NKPT sets the number of k-points used to sample the cell for calculation of the final states. The required number of k-points will vary based on system size (inversely with unit cell volume), and in general metals will require larger grids than insulators. Convergence should always be checked against the k-point sampling. If this is not specified than OCEAN will attempt to guess at a reasonable value.

◆ **NGKPT** `ngkpt[3]` *Recommended*

DFT calculations require a self-consistently determined ground state density for the Kohn-Sham Hamiltonian. As such, in OCEAN, first a ground state calculation is done using only the occupied states and a few unoccupied states for convergence. This density is then used as input to the other DFT calculations to get wave functions. NGKPT determines the k-point sampling for this ground state calculation. Care should be taken to ensure enough k-points are used to give a converged density.

◆ **NBANDS** `nbands` *Recommended*

The number of total bands for the final-state wave-functions is set by NBAND. This includes all valence states but not core states.

```
* In diamond there are two C atoms each with 4 valence electrons.
* To include only the bottom 4 conduction bands
nband 8
```

◆ **SCREEN.NBAND** `screen.nband` *Recommended*

SCREEN.NBAND sets the number of bands to be calculated for the screening wave-functions. The screening calculation requires a large number of bands as the screened potential only converges in the `screen.nbands` → ∞ limit. In practice including bands around 100 eV above the fermi level should be sufficient, but convergence should be checked.

◆ **OCCOPT** `occopt` *Recommended*

OCCOPT controls how ABINIT determines occupation. The allowed values of `occopt` are listed in the ABINIT documentation, but the two most important are 1 and 3 as spin-dependent DFT is not yet compatible with OCEAN.

`occopt` = **1**

States are all doubly degenerate and either occupied or empty depending on band. Suitable for insulators. For Quantum Espresso this is equivalent to OCCUPATIONS = 'fixed'.

`occopt` = **3**

States are all doubly degenerate but can have fractional occupations depending on fermi level. Suitable for metals. For Quantum Espresso this is equivalent to OCCUPATIONS = 'smearing' and smearing = 'fermi-dirac'.

◆ **FBAND** `fband` *Recommended*

FBAND determines the number of unoccupied bands to be included in the SCF calculation of the density. For insulators the default (`fband` = 0.125) is fine, but for metals care must be made to ensure that the highest band in the density calculation has no occupation weight. The number of extra bands is determined by the formula n = `natom` * `fband`. Current this is only enabled for the ABINIT version of the code. The QUANTUM ESPRESSO version will only use the defaults.

◆ **TOLDFE** `toldfe` *Optional*

TOLDFE sets the convergence parameter for the density or SCF run. For ABINIT when the total energy changes by less than `toldfe` in two consecutive SCF runs the density is assumed to be converged. The default (`toldfe` = 10^{-6}) might be a little too permissive, but is adequate for many systems. For QUANTUM ESPRESSO this sets the CONV_THR parameter for the SCF calculation.

◆ **TOLWFR** `tolwfr` *Optional*

The convergence criterium for the (NSCF) wave-function calculations is set by TOLWFR. The default (`tolwfr` = 10^{-16}) should be sufficient. For the QUANTUM ESPRESSO version, this parameter sets the CONV_THR for the NSCF calculations.

◆ **NSTEP** `nstep` *Optional*

The maximum number of iterations the DFT code attempt. If convergence is not reached after `nstep` iterations then ABINIT will continue with a silent warning and QUANTUMESPRESSO will fail. The default is 50.

◆ **SCREEN.NKPT** `screen.nkpt[3]` *Optional*

The screening calculation is less sensitive to k-point sampling than the final states and is therefore calculated on a smaller grid which is set by SCREEN.NKPT. The default setting (SCREEN.NKPT = 2 2 2) is sufficient for a wide variety of systems though very small or very large unit cells may require more or only the gamma point respectively.

◆ **ABPAD** `abpad` *Optional*

During the non-self-consistent calculation of the wavefunctions, the convergence of the highest bands can be very slow. The calculation speed can be increased by adding some throw-away bands to the top. These bands are not considered when checking for convergence and should also not be used in the subsequent steps of the calculation. The parameter abpad adds bands to the calculation and tells ABINIT to ignore them using ABINIT's nbdbuf setting. This has no effect for the QUANTUM ESPRESSO version.

◆ **DFT_ENERGY_RANGE** `dft_energy_range` *Optional*

Instead of setting the number of bands (`nbands`) the user may request an energy range in eV for the final-state wave-functions. This is only an estimate using the volume of the unit cell and may be unreliable. This parameter is only used if no value for `nbands` is given. The default value is 25 eV.

♦ **SCREEN_ENERGY_RANGE** `screen_energy_range` *Optional*

Instead of setting the number of bands for the screening (`screen.nband`) the user may request an energy range in eV. This is only an estimate using the volume of the unit cell and may be unreliable. This parameter is only used if no value for `screen.nband` is given. The default value is 100 eV.

♦ **K0** `k0[3]` *Optional*

The DFT states are calculated using a shifted k-point grid. The first k-point is given by

$$\frac{1}{N_k(1)\cdot k_0(1)}, \frac{1}{N_k(2)\cdot k_0(2)}, \frac{1}{N_k(3)\cdot k_0(3)}$$

where N_k is given by nkpt The default shift is $1/8, 2/8, 3/8$.

♦ **DEGAUSS** `degauss` *Optional*

Sets the broadening of the smearing function (in Ryd.) for metallic occupations in the DFT

3.7 Atomic and Screening Info

♦ **OPF.FILL** `z z.fill` *Required*

Each type of atom you want spectra of requires a '`*.fill`' file which is explained in 5.0.2. Eventually the code will handle doing an arbitrary set of edges for a given system so this input could have as many (`z z.fill`) pairs as you have types of atoms, `ntypat`. Right now you MUST only put one type here due to bugs. The file '`z.fill`' can in principle have any name, but it is recommended a reasonable naming scheme is kept that associates it with a specific pseudo-potential file. The number `z` is the atomic number of the site of interest.

```
* Looking at Oxygen K-edge
8  o.fill
```

♦ **OPF.OPTS** `z z.opts` *Required*

Each type of atom you want spectra of requires a '`*.opts`' file which is explained in 5.0.2. Eventually the code will handle doing an arbitrary set of edges for a given system so this input could have as many (`z z.opts`) pairs as you have types of atoms, `ntypat`. Right now you MUST only put one type here due to bugs. The file '`z.opts`' can in principle have any name, but it is recommended a reasonable naming scheme is kept that associates it with a specific pseudo-potential file. The number `z` is the atomic number of the site of interest.

```
* Looking at Oxygen K-edge
8  o.opts
```

◆ **EDGES** `edges[3,*]` *Required*

Each edge entry consists of 3 integers. When the first is greater than 0 it denotes the **index** of the atom (in the order set by XRED), and when it is less than zero it sets the (negative) Z, allowing you to specify **all** of a given element. The second and third numbers are the principle and angular quantum number. For instance ‘ `1 2 1` ’ would run the $L_{2,3}$ edge of the first atom in the input, while ‘ `-22 2 1` ’ would run the $L_{2,3}$ edges of every titanium atom in the system.

◆ **SCREEN.SHELLS** `shells[]` *Recommended*

The screening calculation is a hybrid of RPA and a model based off of Levine-Louie-Hybersten see section 6. The cross-over between these two regimes is set by `shells` in Bohr and several different radii can be chosen to look at the convergence. In the `shells` $\rightarrow \infty$ limit the RPA is being used for the entire calculation, but convergence is usually reached around 3 - 4 Bohr. If a large radius is being used than the supercell defined by PAW.NKPT must have dimensions larger than the radius chosen or the math could be odd. The default `shells` $= 3.5$ Bohr is ok.

This parameter should not exceed 6 Bohr in the current version of the code.

◆ **CNBSE.RAD** `cnbse.rad` *Recommended*

One screening radius as defined by SCREEN.SHELLS is used in the BSE calculation and is specified by CNBSE.RAD.

◆ **SCFAC** `scfac` *Recommended*

The Slater integrals are calculated in an atomic program and are generally scaled by some factor before use. For 3d transition metals a value of 0.8 is often used, while for f-electron atoms a value of 0.6 may be more appropriate. The value is more or less taken to be independent of the chemical environment and only a function of the element in question.

◆ **CORE_OFFSET** `core_offset` *Recommended*

This card determines wether or not core-level shifts are run, and, if so, sets the offset. Acceptable values are: true, false, or any number. If set to true the code will determine a value of offset so that the average shift among all sites being run is zero.

◆ **OPF.HFKGRID** `opf.hfkgrid` *Optional*

This sets two parameters for the program HFK.X. The first is a grid and should be left at the default value of 2000. The second determines the maximum number of projectors per angular momentum channel that can be created. In the future this should dynamically resize, but in the present the default of 20 will work for many calculations. If the program HFK.X fails it may be solved by increasing this number. It will say something about ‘ne’ being too small.

3.8 BSE Parameters

♦ **CNBSE.XMESH** `xmesh[3]` *Recommended*

When the wave-functions are converted into the NIST BSE format they are condensed onto a grid controlled by CNBSE.XMESH. The states are then projected onto a localized basis set, using the PAW formalism. By default OCEAN will attempt to guess a reasonable setting for this mesh.

♦ **CNBSE.MODE** `mode` *Recommended*

The mode can be either "XAS" or "XES" (NRIXS or XRS are treated the same as XAS). By default the code uses "XAS."

♦ **CNBSE.NITER** `niter` *Optional*

This sets the number of Haydock iterations. The default is 100.

♦ **METAL** `metal` *Optional*

The METAL card determines wether the code expects to determine occupation by band or by energy with respect to the fermi energy. The parameter `metal` can be set to either *.true.* or *.false.* (default). If `metal` is set to *.true.* the code will complain unless OCCOPT is 3 and similarly for `metal` = *.false.* and `occopt` = 1. The use of OCEAN on metals requires some further testing.

♦ **SPIN-ORBIT** `spin-orbit` *Optional*

The spin-orbit splitting between, e.g., the L_2 and L_3 edges, is automatically calculated within OCEAN, but can be overridden by setting SPIN-ORBIT ≥ 0. This could be useful for heavy elements where the $L_{2,3}$ splitting is many 100s of eVs and the two can be considered separately (set it to 0), or for investigating the effect of spin-orbit strength on the fine structure.

♦ **PHOTON_Q** `photon_q[3]` *Optional*

For UV/Vis and RIXS calculations the valence and conduction band states must be offset from each other (by the momentum that is absorbed or transferred). This sets that momentum offset in units of the reciprocal lattice vectors of the system. By default it is 0 0 0.

♦ **CNBSE.STRENGTH** `cnbse.strength` *Optional*

This card sets the strength for the two interaction terms in the BSE Hamiltonian. By default it should be set to 1, and for emission the code automatically sets it to 0. It can be changed for curiosity's sake.

♦ **CNBSE.SOLVER** `cnbse.solver` *Optional*

Choose between using the Haydock algorithm (full spectrum, no exciton) or the GMRES algorithm (single energy, exciton).

♦ **CNBSE.GMRES.ELIST** `cnbse.gmres.elist[*]` *Optional*
A list of energies (of arbitrary length) for the code to run the GMRES algorithm. The energies are given in eV, and have the same scaling/shifting as the standard Haydock spectra, e.g., by default the lowest unoccupied state is at zero without the core hole. Note, if both this and CNBSE.GMRES.ERANGE are specified then CNBSE.GMRES.ERANGE will be used.

♦ **CNBSE.GMRES.ERANGE** `cnbse.gmres.erange[3]` *Optional*
Run the GMRES algorithm on a set of evenly spaced energies specified by starting energy, ending energy, and step size (all in eV). This is especially useful for reproducing RIXS planes, in contrast to CNBSE.GMRES.ELIST where the energy points can be arbitrarily spaced. If both CNBSE.GMRES.ERANGE and CNBSE.GMRES.ELIST are specified this will be used.

♦ **CNBSE.GMRES.NLOOP** `cnbse.gmres.nloop` *Optional*
The version of the GMRES algorithm implemented in OCEAN will grow to be a subspace of no larger than NLOOP before it is restarted. By default this is set to 80, and should probably not be changed without careful timing tests.

♦ **CNBSE.GMRES.GPRC** `cnbse.gmres.gprc` *Optional*
The version of the GMRES algorithm implemented in OCEAN uses the one-electron energies as a pre-conditioner with a Lorentzian broadening set by GPRC (in Ha.). By default this is 0.5 Ha. which should be reasonable for K edges.

♦ **CNBSE.GMRES.FFFF** `cnbse.gmres.ffff` *Optional*
Sets the convergence criterion for the GMRES.

3.9 Spectrum Information

♦ **CNBSE.BROADEN** `broadening` *Recommended*
The amount of Lorentzian broadening included in the spectra is set by CNBSE.BROADEN. This is a required convergence parameter for the continued fraction that is our Haydock approximation to the BSE, but setting it to the core-hole intrinsic broadening is recommended since features sharper than that won't be observable. Given in eV the default is `broadening` = 0.1 eV.

♦ **CNBSE.SPECT_RANGE** `esteps emin emax` *Optional*
This controls the output spectrum (or spectra) from a run. The energy of the spectrum is set by setting the conduction band minimum and core-hole binding energy to zero. All systems will include a small amount of excitonic binding (depending on l selection rules) and spin-orbit

splitting can lower the first peak significantly, eg. $L_{2,3}$ splitting in 4d transition metals. These parameters do not change the BSE calculation, only the output of the spectra at the end. The energy range is given in eV and by default `emin` $= -40.817$ and `emax` $= 68.028$ (-1.5 to 2.5 Ha.). The parameter `esteps` gives the number of energy points used for the plot, and by default it is 1200.

Chapter 4

Ground-State Wave Functions

In principle any pseudo-potential DFT code can be used. In our implementation we have an interfaces for using ABINIT [1, 7, 6, 5] and QuantumEspresso [4, 8].

There are two important differences between what is required for a good core-level calculation with OCEAN and more traditional ground-state DFT calculations. The first is that spectroscopy calculations often require denser k-point sampling than a calculation of only the density. Secondly, a psuedopotential that is suitable for ground-state calculations may have poor scattering characteristics at higher energies that will make up the spectrum being calculated. With these brief warnings the rest of this chapter will be spent going over various aspects of the calculation and mention briefly strategies for ensuring accuracy and convergence.

You need a structure and a set of pseudo-potentials to begin a calculation. From there, you must check convergence with respect to all of the adjustable parameters of the calculation. Not to worry, there are not really that many. In this section we will focus on just a few.

4.1 Pseudo-potentials

Pseudo-potentials must be carefully constructed to ensure that they adequately mimic the all-electron system being investigated. For general purpose pseudo-potentials this means having good transferability so they can be used in a variety of chemical environments. Some literature exists on the topic, and the reader is invited to read up on it[**?**] .

As mentioned above for excited state calculations states as high as 50-100 eV above the fermi level must be adequately represented by the pseudo-potentials used. This often requires a harder pseudo-potential than might otherwise be necessary for other calculations. Additionally the way in which chemical bonding affects the valence and low-lying conduction bands is often influenced by the inclusion of semi-core states in the calculation. Better agreement is often seen with the inclusion of semi-core states, eg. 2s and 2p states in a 3d transition metal. This has been seen to also improve the high-energy scattering.[]

In the future the authors hope to collect a database of pseduopotentials to add to the ease of using OCEAN, and all users will be invited to contribute their carefully tested ones as well.

Currently the format of the pseudopotentials must match that of .fhi. Users can download compatible pseudopotentials from the ABINIT website [1] or build their own using the OPIUM code.

4.2 Self-consistent Density Calculation

A DFT calculation requires, as one might expect, a density. In the self-consistent loop, wavefunctions are converged in an inner loop, and the density is updated from there to form an outer loop. Once the density has been converged, in this usage it changes less than the parameter TOLDFE over the course of two successive updates, this calculation is complete. The density, $n(\mathbf{r})$, depends on two (2) major convergence factors; ECUT and NGKPT. What follows is a brief discussion of these and other parts of the density calculation.

In plane-wave based codes the basis size is truncated according to the ECUT parameter, specifying the largest energy plane-wave. This cut-off energy is largely dominated by the pseudo-potentials, the harder the pseudo-potential the higher the cut-off energy needs to be. It should be noted that the scaling of the DFT section with respect to the number of plane waves is no better than N*LOG(N), and the number of plane waves goes as the cut-off to the 3/2 power. When the total energy no longer changes as a function of an increasing energy cut-off the calculation is converged. A reasonable value would be 1meV per atom or less.

The second important factor is the k-space sampling used for the self-consistent calculation. As in all numerical work, integrations are replaced by finite sums. The required k-space sampling is highly material dependent. A material with a small unit cell or a metal will both tend to require a larger number of k-points. This can be guided by experience, but should always be checked by increasing the sampling and looking at the total energy or perhaps the pressure.

The last consideration is the treatment of electron occupancies. For insulators this is straightforward, fill all the lowest bands, but for metal the situation requires more care. Choosing a value of OCCOPT that is greater than 1 will cause ABINIT to smear the occupation levels near the fermi energy. This helps lead to faster convergence by better sampling the fermi surface. Along with metallic occupations comes the need to include enough extra bands in the density calculation to hold the smeared out electrons. The parameter FBAND determines how many extra bands are added. These two parameters follow directly from their ABINIT definitions and more can be found in ABINIT's help. When using QUANTUMESPRESSO smearing is always enabled, but this may change in the future.

4.3 K-point sampling

A sum over a few wave functions ψ_k is substituted for a integral over the entire Brillouin zone. In principle knowledge of the crystal symmetry of the unit cell could allow for intelligent choices of points to sample at with non-equal weight, but that is not the approach here. Instead a regular mesh of points in reciprocal space is used and the mesh must be dense enough to ensure that the sampling is sufficient.

In practice the required k-point grid scales inversely with unit cell size, but testing runs at different sizes and checking that the spectrum is unchanged should be done for each calculation. Unit cells that are oblong may benefit from k-point meshes that have a larger number of k-points along the dimensions with longer **reciprocal** lattice vectors. The only way to ensure convergence w.r.t. k-point sampling is to increase or decrease the number of k-points and make sure that the results are not changing. For metals a large number of k-points may be required in order to get a good sampling of the Fermi surface.

4.4 Number of Bands

The excited state (excitonic) wave-function at any given energy may include contributions from much higher energy DFT wave-functions. Additionally the attractive core-hole potential will pull some states down in energy. If one is interested in spectra up to 20 eV above the edge, the number of bands needed in the calculation will be larger than if one is only interested in the DOS up to 20 eV. This convergence should be tested explicitly.

4.5 Beyond DFT

In principle these wave-functions should include self-consistent self-energy corrections, eg. GW, but in practice it has been found that for many systems self-energy corrections only affect the energies and not the wave-functions. Even single-shot self-energy calculations can be computationally expensive and we provide several methods to approximate the GW correction; ad-hoc band stretching and the many-pole self-energy which are both presented later.

LDA+U work is underway and will be a feature of OCEAN 1.1. Better integration with GW is a planned feature for later in the 1.x series.

Chapter 5

OPFs

In OCEAN we build on top of one-electron orbitals calculated using pseudopotentials. This significantly reduces the computational cost of the DFT calculations, but near the nucleus the electron wavefunctions diverge significantly from an all-electron calculation. For clarity we will refer to electronic states from a pseudopotential calculation as “pesudized” in contrast to “all-electron” wavefunctions. The matrix elements between a core orbital and a pseudized valence or conduction band state are not well defined. The pseudized valence states are not orthogonal to the core levels because core states explicitly don’t exist in a pesudopotential calculation. To overcome this problem we make use of the optimal projector functions (OPFs).

The OPF are solutions to an atomic DFT calculation that depends on the element being investigated not the material. The framework was developed by E. L. Shirley [**?**] and is similar to the projector augmented wave (PAW) formalism developed by Blöchl [2]. We would encourage the reader to peruse Blöchl’s paper and will only summarize the method here without much justification. The pseudopotentials are constructed such that the pseudized orbitals are equal to the all-electron ones outside of the cut-off radius r_c. Further, In principle this section can be run once for a given pseudopotential and reused. The important outputs of this section are a collection of all-electron and pseudo projector augmented waves. The all-electron core wave function is expressed as coefficients of the all-electron OPFs. The psuedo OPFs will be used later to express the valence/conduction electron wave functions. In OCEAN the construction of all-electron and pseudized atomic states is taken care of by the program HFK.X.

The central idea is that a small number of local orbitals can adequately describe the region of space around the core-hole excitation site. These atom-centered orbitals are $\phi_{\nu,l,m}$, where l and m are the expected angular and azimuthal quantum numbers, and ν allows several projectors per angular momentum channel. The number of projects ν is determined automatically by the code and ranges around 4 for many systems.

5.0.1 scfac

The SCaling FACtor is a real number that modifies the calculated Slater-Condon parameters for the atomic case. For 2nd row elements this should be 1.d0 while for transition metals a factor of 0.83 is more appropriate. See de Groot, some others probably [3].

5.0.2 fill

The '`*.fill`' file determines the numerical parameters used for the OPF construction. Including the energy range to construct waves over, the cut-off radius to use, and the momentum grid for evaluating the Fourier transform of the projectors. In the table below on the left is a sample '`*.fill`', and on the right are the corresponding variable names for each item.

2	pow
0.30 2.00 0.0001	Epad, Emax, prec
3.5	cutoff radius
0.05 20	q step, q max

- pow: The matrix elements between the core state χ_α and local orbital ϕ_i are calculated up to $\langle \chi_\alpha | r^{pow} | \phi_i \rangle$.
- Epad and Emax: The OPF reconstruction starts at energy Epad below the lowest valence state for a given angular momentum channel and attempts to span the space to Emax (both in Ha.).
- Cutoff radius: The OPF cutoff radius is in Bohr and sets the radius of the sphere in which the local basis is calculated. Convergence of this must be checked.
- prec: This should be sufficient for a range of systems and not changed.
- q step and q max: These control the Fourier transform of the projectors from real-space. The defaults should be fine.

5.0.3 opts

The '`*.opts`' file contains information about the pseudopotential such as z, core-valence partitioning, and reference configuration. This information should match that used to construct the pseudopotential. In the case of pseudopotentials whose origin is unknown a reasonable guess must be made. Below (on the left) is a sample '`*.opts`' file for oxygen, while descriptors for each element are on the right.

008	Z
1 0 0 0	core states; s p d f
scalar rel	relativity
lda	functional
2.0 3.5 0.0 0.0	valence occupation; s p d f
2.0 3.5 0.0 0.0	repeat of above

- Z: The atomic number
- core states: The number of core levels for each angular momentum channel (s,p,d,f) that are filled. In this case only the $1s$ level is treated as core so only the s channel is non-zero. For sulphur this would likely read "2 1 0 0" relecting that both the $2s$ and $2p$ would be treated as core.
- relativity: Specifying how to treat the relativistic spin-orbit effects. The options are scalar or dirac and rel[ativistic] or nonrel[ativistic].
- functional: Specifies the exchange-correlation functional that is used for the atomic calculations that give the local orbitals and projectors. Currently only LDA (local density approximation) and HF (Hartree-Fock) are available.
- valence occupation: Unlike the core states, here we are interested in the number of electrons per angular momentum channel that are in the reference configuration for the pseudopotential. Many pseudopotentials are slightly positively ionic and so we choose a $2s^2$,$2p^{3.5}$ occupation.

For semi-core pseudopotentials, such as the $3d$ transition metals, some care should be taken. For potassium through copper better results may be obtained by including the $3s$ and $3p$ states as valence. This means that the pseudopotential will often be [Ne]$3s^2 3p^6 3d^n 4s^0 4p^0$. Where n will depend on which metal. The opts file may then read something like

022	Z
2 1 0 0	core states; s p d f
scalar rel	relativity
lda	functional
2.0 6.0 2.0 0.0	valence occupation; s p d f
2.0 6.0 2.0 0.0	repeat of above

where the occupations of the s and p channels reflects the $n = 3$ principle quantum level.

5.1 Outputs from the OPF and atomic program

Currently many of the outputs are placed in the '`zpawinfo`' directory.

5.1.1 OPFs

The all-electron and pesudo OPFs are stored by angular momenta and z in '`ae#z###`' and '`ps#z###`', respectively.

5.1.2 Fourier Coefficients

The '`psft#`' files contain the Fourier transform of the projectors. They are kept in the '`zpawinfo`' directory.

5.1.3 Core-level orbitals

The all-electron eigenfunctions from the atomic calculation are stored in the '`coreorbz###n##l##`' by z, n, and l on a radial grid.

5.1.4 Info files

The '`prjfilez###`' file, where # is the three digit Z value, contains info about the Fourier grid and angular momentum range of the projector augmented waves. The first line specifies the minimum angular momentum, maximum angular momentum, number of fourier grid points, and the spacing of the fourier grid. Each following line specifies the number of OPFs for each angular momentum channel.

the '`radfilez###`' files, where # is the three digit Z value, contains info about the real-space grid. The first line is cut-off radius, rounded from the input cut-off to the nearest grid point. The second is the number of real-space grid points, and the last is the index of the cut-off.

5.1.5 Matrix Elements

The '`melfilez###n##l##`' file contains the matrix elements between the specified core-hole and the different OPF functions. The first line specifies the maximum x for the operator r^x, always starting at 0. Starting with the fast index the file is then the value of the matrix elements looped over projectors, powers of x, and finally angular momentum.

5.1.6 Slater Integrals

The various f and g integrals are also tabulated for all of the atomic, all-electron eigenstates (core levels) in the '`fk*`' and '`gk*`' files, respectively. The first three digits in these file names identify the core-level angular momentum, the OPF angular momentum, and the power of k in the expansion of the Coulomb interaction in spherical harmonics.

Chapter 6

Screening Calculation

The screening of the core-hole interaction is done in real space using the random phase approximation up to some radius around the core, 2-5 Bohrs, and then using an analytic function of the local density, distance, and static dielectric constant outside of this radius [9]. Convergence with respect to this radius should be checked and will change based on both the core and system of interest. The screening is only exact for an infinite number of bands, but should smoothly converge w.r.t. number of bands included. For the average system a 2^3 k-point grid is sufficient for convergence, while a very small system might require a slightly higher sampling or a large unit cell might be ok with just the gamma point.

Converging the screening calculation requires only checking four parameters; SCREEN.NKPT, SCREEN.NBANDS, SCREEN.SHELLS and DIEMAC. To check the convergence of the screening calculation you should graph the file '`ropt`' using columns 1 and 4. This gives you calculated valence response to the core-hole potential. The file '`rpot`' can also be graphed using column 2 versus column 1, and it shows the (negative) total screened core-hole potential. Both files give the radius in Bohr and potential in Hartree as well as assuming a spherical potential. The screening info for each atomic site will be stored in a directory labeled by the atomic symbol, the index of that atom, and the core level n and l.

The macroscopic dielectric constant, ϵ_∞, is normally taken from experiment, and so it is not really a free parameter. The number of k-points needed for the screening is small and most calculations should be fine with the default grid of $2 \times 2 \times 2$, possible exceptions are very tiny unit cells which may need more points and very large cells or supercells which may only require the gamma point.

Several different values of the screening cut-off radius can be chosen by listing them in the input SCREEN.SHELLS. For a calculation several different radii should be tried, for example, 3.5 and 4.0 Bohr. The calculation for each radius is stored in a subdirectory labeled with the radius, '`zRXT#.##/`'. The directories labeled '`zRXF#.##/`' should be ignored at present.

Lastly the number of bands included in the truncated sum for calculating the RPA dielectric function also requires checking. The code automatically computes two screening functions per screening calculation. One contains all of the bands in SCREEN.NBAND, while the second reduces the number of unoccupied bands included by about 80%. This allows a quick check of the convergence with bands and the results of the smaller calculation are stored in the subdirectory ‘`zRXT#.##_small/`’.

Chapter 7

Core-Level Shifts and Absolute Energies

Calculating the correct onset of absorption is important for both UV/vis and x-ray spectra. In the valence case we need to accurately calculate the relative band positions of the valence and conduction states, e.g., a correct band gap energy. While the alignment of x-ray edges is more complicated in practice, the necessary ingredients are the same as for low-energy spectra, and we will introduce the concepts and approach in this chapter. The calculation of core-level shifts in ammonium nitrate (NH_4NO_3) provides a good test and is featured in a tutorial (**??**).

7.1 Exciton energies in BSE

The Bethe-Salpeter equation (BSE) gives a diagrammatic expansion of the two-particle Green's function. Immediately we are making the approximation that multiple particle-hole excitations (bi-excitons, &c.) are not important for describing the electronic response of our systems, except in a mean-field, plasmon-like way. The next big approximation is that we truncate the BSE at the GW level, and therefore the interactions have two terms, the direct W and exchange V_x

$$H = H_0 + V_x - W \tag{7.1}$$

where H_0 are the two particles, electron and hole. This can also be written in the standard Dyson's equation way where the electron-hole Green's function is written in terms of the non-interacting Green's functions and the interaction vertex I

$$G^{(2)}(1,1';2,2') = G_0^{(2)}(1,1';2,2') + G_0^{(2)}(1,1';3,3')I(3,3';4,4')G^{(2)}(4,4';2,2') \tag{7.2}$$

$$I(1,1';2,2') = -\delta(3,4)\delta(3',4')W(3,3') + \delta(3,3')\delta(4,4')V(3,4) \tag{7.3}$$

where $1 = \mathbf{x}, \sigma, t$ and W and V are the screened and unscreened Coulomb operators, respectively (taken from Hanke and Sham [**?**]). The last big approximation we make is that the screening can be considered statically. Namely, that the exciton binding energies are small compared to the band gap.

In summary:

1. $\chi(1,2) = G^{(2)}(1,1';2,2')$
2. $\Sigma = GW$
3. $W(\omega) \approx W(0)$

Having bravely decided not to consider the above sources of error, we have reduced the question of accurate excitation energies to adequately calculating our two interaction terms and the one-electron Green's functions that make up our two-electron, non-interacting Green's function. Undaunted we will also, for this chapter, ignore the calculation of the direct and exchange, except to say that one must, of course, converge these two terms.

7.2 One-electron Green's function

The non-interacting, two-particle Green's function above in Eq. 7.2 is made up of the two (interacting) one-electron Green's functions

$$G_0^{(2)}(1,1';2,2') = G(1,1') + G(2,2') \tag{7.4}$$

which can be written in their own Dyson's series

$$G(1,1') = G_0(1,1') + G_0(1,2)\Sigma(2,3)G(3,1') \tag{7.5}$$

where Σ is the proper self-energy. For an in-depth look at Σ we would suggest the reader look elsewhere **??**. Here we will only point out that, given a correct Σ, and hence a dressed one-electron Green's function G, the BSE will give exciton energies that will match experimental measurements, subject to the approximations at the end of section 7.1. In practice there are two widely used approximations for Σ, namely the GW method (really many different approximations with a range of costs and sophistication) and the DFT method where explicitly or implicitly it is assumed that solving the DFT Hamiltonian gives a good approximation to the one-electron orbitals and energies. It is not usually phrased this way, but the success and popularity of DFT is due to 1) its low computational cost; and 2) the extraordinary fact that the exchange-correlation potential is an adequate approximation to the true exchange and correlation of system.

For a valence calculation, UV/visible spectrum, the problem of aligning the absorption spectrum is then simply the problem of getting a good enough calculation of the self-energy. At present OCEAN has no automatic way to carry out a GW calculation, but many codes exist: ABINIT and BerkleyGW among others. The BSE section of OCEAN has several methods for incorporating these corrections ΔE^{GW} including full corrections $\Delta E^{GW}_{n,\mathbf{k}}$, band-by-band corrections ΔE^{GW}_{n}, energy-dependent corrections $\Delta E^{GW}(E)$, and a simple scissor operator.

Until such time as the developers write a better interface and document it, interested users should email and politely ask the devs for guidance.

In principle, this procedure of carrying out quasi-particle corrections could be done also for a core-level hole. However, the authors are not aware of a code capable of such calculations, leading us to recommend a different approach for alignment of core-level spectra.

7.3 Core-level shifts

The OCEAN code is built on top of pseudopotential DFT, making the calculation of the core-level energies tricky. Instead of attempting to be able to calculate the energy of each core level from first-principles we content ourselves with being able to calculate the energy shifts in the core levels between chemical environments. For instance, in the ammonium nitrate tutorial it is seen that the relative shift between the ammonium nitrogen site and the nitrate nitrogen site is around 4 eV. Our ultimate goal is to be able to align each element edge by comparing a single material to experiment, e.g., correcting all nitrogen edges by setting the N_2 XAS onset. The energy of the core-level hole is given by

$$\epsilon_h = V_{\mathrm{tot}} - {}^1\!/\!{}_2 W + X_{Znl} \tag{7.6}$$

where the offset X is fit to a specific element Z and edge nl. To arrive at the above expression we make two main approximations.

The first is that we assume that the shape of the orbital wavefunction is unchanged by changes in the chemical environment. In other words, the difference between the NH_4^+ and NO_3^- is captured by lowest order perturbation theory. The orbital remains fixed, and the change in energy is simply the change in the external potential exerted on the orbital in the ground state. Currently, this is done by taking the limit that the core orbital is a delta function, but in the future this will be done by integrating the potential over the core orbital density.

$$\begin{aligned} V_{\mathrm{tot}} &= \langle \phi_{Znl} | V_{\mathrm{tot}}(\mathbf{r}) | \phi_{Znl} \rangle = \int d^3r \rho_{Znl}(\mathbf{r}) V_{\mathrm{tot}}(\mathbf{r}) \\ &\approx V_{\mathrm{tot}}(\mathbf{r}) \delta(\mathbf{r} - \tau) \end{aligned} \tag{7.7}$$

where τ is the location of the excited atom, ϕ and ρ are the atomic orbital and its density, respectively, and V_{tot} is the total potential. The value of the total potential (evaluated at the core site) is determined from the self-consistent density, and is easily extracted from the DFT calculation, using a post-processing tool like PP.X or CUT3D.

There is a second element needed to calculate the shifts in quasi-particle energy of the core levels with chemical environment, and that is the response of the rest of electrons in the system to the excitation. This response is already calculated in the code as the screening of the core-hole potential. The assumption is made that this can be captured within linear response. For the methodology see the chapter 6 on screening. Using the same δ-function approximation as Eq. 7.7 the screening reduces the energy required to create a core hole by

$${}^1\!/\!{}_2 W = {}^1\!/\!{}_2 W(\mathbf{r}) \delta(\mathbf{r} - \tau) \tag{7.8}$$

though a more accurate integration over the core-hole density could also be implemented. In practice the code uses only the valence screening. The other core electrons will also screen the core hole, but, because we already have assumed that the core-level orbitals are the same irrespective of chemical environments, their screening will also be the same and can be incorporated into the shift X.

7.4 Errors in the CLS and comparisons with ΔSCF methods

The core-level shifts have been used in OCEAN calculations of both liquid water as well as vibrational disordered crystals of ammonium nitrate and ammonium chloride. However, for these systems the alignment has been done anew each time, and the various disordered snapshots have all had the same volume and the same number and type of atoms. A more complete investigation of alignment, say between the O K-edge absorption onset of different oxides $\mathrm{M}_x\mathrm{O}_y$, has not been carried out. As a guide, however, we would like to speculate on sources of error that may hamper our desire of a single X_{Znl} for each element and edge.

1. Inaccurate band gaps, or HOMO or LUMO energies
2. Breakdown of linear response
3. Effects of periodic boundary conditions?

For each of these short-comings, our investigations so far, comparing disordered, but otherwise similar cells, would not be sensitive.

Issue #1 may be especially significant. The errors in DFT-derived band gaps vary widely, ranging from 0 eV for metals to above 6 eV for wide gap materials like LiF. Making the crude estimate that the GW corrections are split between conduction and valence bands, the error in absorption onset could be several electron-Volts.

Chapter 8

CNBSE

The CNBSE section solves for the spectra. The important output file is absspct; column 1 is the energy, column 2 the spectra using a plasmon-pole broadening model and column 3 the plain spectra.

Appendix A

Installation

- The code has been tested with both gcc and intel fortran compilers. Would love to hear people's problems/successes using other vendors.
- The code requires perl5
- Either ABINIT or QUANTUMESPRESSO (QE) may be used as the DFT code
- Optionally FFTW3 may be used in place of the internal FFT routines

The bare bones installation instructions are as follows.

1. Make sure you have a working DFT code installed first and know where the binaries, object files and compiled modules for the code are located (see note on qexml below)
2. Copy '`Makefile.arch.example`' to '`Makefile.arch`' and edit as necessary. Basically comment/un-comment shell variables in '`Makefile.arch`' corresponding to features you wish to deactivate/activate. Helpful comments are included in '`Makefile.arch.example`'
3. Mind the INSTDIR variable which will set where the installation is placed.
4. Run 'make' and then 'make install'

Note on qexml: To work with QE, OCEAN requires the qexml library of QE to be compiled. This libarary is built by deafult in QE-5.1 and higher and is located in the Modules directory of the QE distribution. For earlier versions such as QE-4.3 and QE-5.0, the user needs to navigate into the PP (for QE-4.3) or PP/src (for QE-5.0) directroy of QE and invoke 'make qexml.o' to build the object file qexml.o . The location of qexml.o can then be indicated in the '`Makefile.arch`' of OCEAN.

Assuming success so far you can now go to the '`EXAMPLE/OBF`' directory. Here you will find what you need to do a preliminary calculation of the Ti $L_{2,3}$ edge in strontium titanate. To run OCEAN run the command *`/path/to/ocean/ocean.pl OBF.in`* and wait for the code to finish. The code will create several directories and the end results are in the file '`CNBSE/absspct`' which can be compared to the '`absspct`' file that is distributed in the OBF directory.

Bibliography

[1] The ABINIT code is a common project of the Université Catholique de Louvain, Corning Incorporated, and other contributors; `www.abinit.org`. 4, 4.1

[2] P. E. Blöchl. *Physical Review B*, 50:17953–17979, Dec 1994. 5

[3] F. de Groot and A. Kotani. *Core Level Spectroscopy of Solids (Advances in Condensed Matter Science)*. CRC Press, Boca Raton, FL, 2008. 5.0.1

[4] P. Giannozzi and *et al. Journal of Physics: Condensed Matter*, 21:395502, 2009. 4

[5] X. Gonze and *et al. Computational Materials Science*, 25:478, 2002. 4

[6] X. Gonze and *et al. Zeitschrift fur Kristallographie*, 220:558, 2005. 4

[7] X. Gonze and *et al. Computer Physics Communications*, 180:2582, 2009. 4

[8] C. Gougoussis, M. Calandra, A. P. Seitsonen, and F. Mauri. *Physical Review B*, 80:075102, 2009. 4

[9] E. L. Shirley. Local screening of a core hole: A real-space approach applied to hafnium oxide. *Ultramicroscopy*, 106(11-12):986–993, 2006. 6