

ShirleyXAS & MBXASPY

A Density-Functional Theory + Determinant Approach
for
Simulating Near-Edge X-Ray Spectra

Yufeng Liang and David Prendergast

A Software Project at
the Molecular Foundry,
Lawrence Berkeley National Laboratory

Abstract

ShirleyXAS is a first-principles software package that performs the one-body core-hole calculation for x-ray absorption spectra. It is modified based on Quantum Espresso 4.3 thus it can perform the typical self-consistent-field (SCF) and band-structure calculation. Additionally, it also performs a fast interpolation of band structure and wavefunctions using the Shirley optimal basis set. The package is written in Fortran together with some automatic scripts written in Bash that facilitates the entire workflow.

MBXASPY is a post-processing simulation package written in Python that calculates core-level spectra using the determinant approach. It takes the eigen-energies and eigen-wavefunctions from a typical DFT calculation, or simply a tight-binding calculation. MBXASPY is interfaced seamless with ShirleyXAS and therefore using ShirleyXAS + MBXASPY can handle the determinant calculation for realistic materials from first-principles. There is another set of Bash scripts that carry ShirleyXAS + MBXASPY (DFT + determinant) calculation so that the users only need to know a bit about first-principles calculations and just get their structures ready !

Chapter 1

Introduction

1.1 ShirleyXAS

The original ShirleyXAS before the development of MBXASPY is designed to efficiently computing the x-ray absorption spectra (XAS) using the one-body core-hole approach [1, 2]. “One body” means that we are not using the many-body wavefunctions in Fermi’s Golden rule. Instead, single-body orbitals obtained from Density-functional Theory (DFT) calculation will be directly used to calculate the transition matrix elements. The x-ray core-hole effects are captured by an impurity atom that carries an amount of positive charge. Then a delta-SCF (Δ SCF) procedure is used to simulate the charge density relaxation after the core hole is generated. This part is accomplished by the precursor of ShirleyXAS, Quantum Espresso 4.3, that takes care of all the DFT calculations.

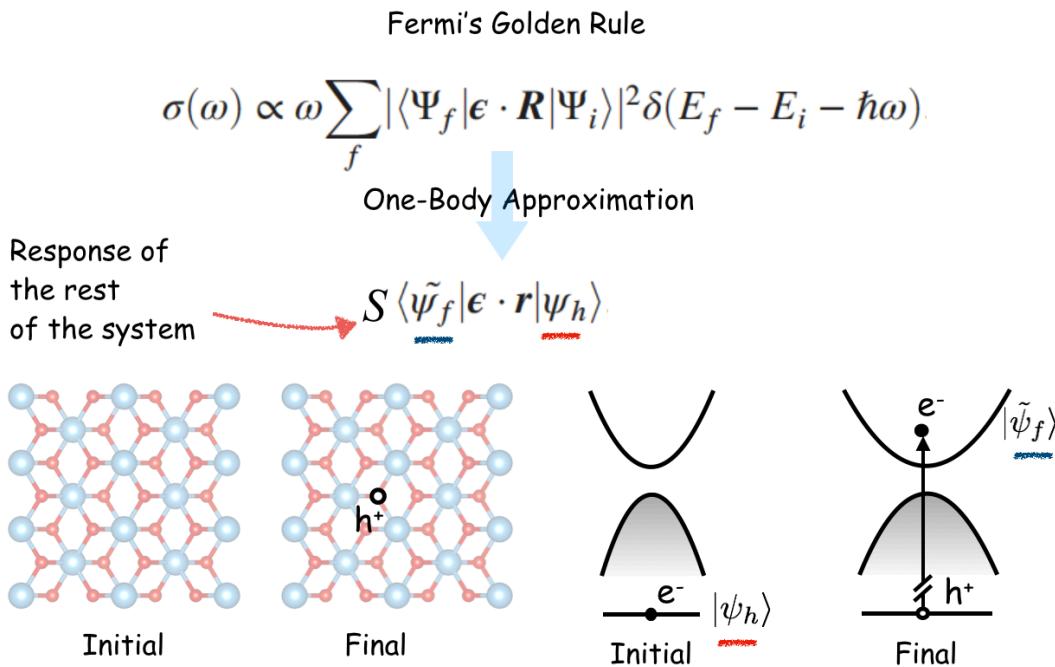


Fig. 1.1 Basic concepts about one-body core-hole approach based on Δ SCF.

Fig. 1.1 illustrates some basic concepts about Δ SCF core-hole approach.

The *initial state* is simply the pristine system before x-ray excitations. But because the x-ray-excited core hole is treated as an impurity, the initial state is typically a supercell of the pristine system.

The *final state* has exactly the same size as the initial state, but the atom whose core level is excited will be replaced by core-excited atom. This is usually realized by replacing the pseudopotential of the ground-state atom with the excited one. For instance, if we are looking at O K edge, the excited-O pseudopotential will be generated with *only one* electron on the 1s level (using the configuration $1s^1$ instead of $1s^2$).

It is usually a formidable task to find the exact many-body wavefunctions for a realistic system with many electrons. One way to get around this is to use one-body orbitals instead of many-body wavefunctions when obtaining the matrix element in Fermi's Golden rule, leading to the *one-body approximation* [1, 2]

$$\langle \Psi_f | \epsilon \cdot R | \Psi_i \rangle \rightarrow S \langle \tilde{\psi}_f | \epsilon \cdot r | \psi_h \rangle$$

where $|\psi_h\rangle$ is the core-level orbital in the *initial-state* system, which is simply the core-level orbital of the ground-state atom (e.g., 1s orbital of the ground-state O), whereas $|\tilde{\psi}_f\rangle$ is one empty orbital in the *final-state* system (final-state orbitals are annotated with a hat tilde). Then the matrix element $\langle \tilde{\psi}_f | \epsilon \cdot r | \psi_h \rangle$ represents the transition amplitude from the initial-state core level to the final-state empty orbital, for the photo-electron. Lastly, the prefactor S represents the response of rest of the many-electron systems and is typically taken as a constant.

Decomposing the matrix element into $\langle \tilde{\psi}_f | \epsilon \cdot r | \psi_h \rangle$ and S implies that the transition of the photon-electron and the response (charge-relaxation) of the other electrons are uncorrelated and disentangled. Alternatively, it can be understood as the two processes occur at a particular moment which does not take any time. While this is true for photo-electron transition (an electron requires no time to hop after absorbing a photon), the response / relaxation of the other electrons do occur *over a finite period of time*, simply because

each electron can hop at different times in response to the core-hole potential.

The functionalities of the *original ShirleyXAS* include:

1. Generate the pseudopotential that simulates the x-ray core-hole effects, i.e., core-hole pseudopotential
2. Provide a set of Bash scripts that automatizes the XAS calculation
 - a. Parse the user input using a Bash script file, `Input_Block.in`
 - b. Carry out the DFT calculation for the final-state system
 - c. Produce the eigenenergy and eigenwavefunction at the Γ -point, for a good number of orbitals to cover the energy range of interest
 - d. Extrapolate the band structure using the information at the Γ -point using the Shirley optimal basis sets
 - e. Calculate the one-body matrix element (oscillator strength) for each empty orbital
 - f. Produce XAS using the oscillator strengths of the empty orbitals
3. Obtain relative spectral alignment using DFT formation energy calculation.

1.2 MBXASPY

MBXASPY is a python package that is designed to perform the determinant calculation [3, 4], either using the input from ShirleyXAS or a user-defined Hamiltonian. MBXASPY is short for Many-Body X-ray Absorption Spectra with Python. Because we are not (yet) doing configuration interaction for the many-body wavefunctions used in XAS, the method would still be better referred to as the “determinant method” instead of a “many-body” method. The determinant method is essentially equivalent to the so-called MND theory [3, 4, 5] for x-ray excitations, except that we can now use DFT orbitals as input.

What MBXASPY seeks to calculate a determinant form of transition amplitude [3, 4]. Within the independent-electron approximation, the many-body wavefunction of the initial/final state is just one Slater determinant of the relevant orbitals. It turns out the transition amplitude is no longer a prefactor S multiplying by one-body matrix element, but, the summation of a series of prefactors A multiplying by one-body matrix element, as in Fig. 1.2. It is the prefactor A that takes the form of a $(N + 1) \times (N + 1)$ determinant, where N is number of electrons in the systems (core-electron excluded!). This implies that the transition of the photo-electron is now entangled with the response of the rest of the system, in contrast with the disentangled form in the one-body approximation [3, 4].

$$\langle \Psi_f | \boldsymbol{\epsilon} \cdot \mathbf{R} | \Psi_i \rangle = \sum_c A_c^{f*} \langle \psi_c | \boldsymbol{\epsilon} \cdot \mathbf{r} | \psi_h \rangle$$

Single Slater determinant for $|\Psi_i\rangle$ and $|\Psi_f\rangle$

Transition Amplitude:

$$A_c^f = \det \begin{bmatrix} \xi_{f_1,1} & \xi_{f_1,2} & \dots & \xi_{f_1,N} & \xi_{f_1,c} \\ \xi_{f_2,1} & \xi_{f_2,2} & \dots & \xi_{f_2,N} & \xi_{f_2,c} \\ \vdots & & & \vdots & \vdots \\ \xi_{f_{N+1},1} & \xi_{f_{N+1},2} & \dots & \xi_{f_{N+1},N} & \xi_{f_{N+1},c} \end{bmatrix}$$

Fig. 1.2 Basic formula for the determinant approach.

The matrix elements $\xi_{\mu\nu}$ used in the obtaining the determinantal A come from the transformation coefficients from the initial-state orbitals to the final-state ones (with “~”). These orbitals are taken from the supercell as shown in Fig. 1. 1.

Definition of Coefficient

$$\xi_{\mu\nu} = \langle \psi_\nu | \tilde{\psi}_\mu \rangle$$

$$|\tilde{\psi}_\mu\rangle = \sum_\nu \xi_{\mu\nu} |\psi_\nu\rangle$$

The second index of $\xi_{\mu\nu}$ in A runs over a (N+1)-tuple $(1, 2, \dots, N, c)$, where $1, 2, \dots, N$ are the indices of the lowest N occupied levels and c indicates an empty level the core electron would be promoted to by the x-ray photon.

The first index of $\xi_{\mu\nu}$ in A runs over a (N+1)-tuple $(f_1, f_2, \dots, f_{N+1})$, which defines (N+1) **final-state** energy levels (orbitals) the (N+1) electrons will occupy (with that core electron included !).

You may imagine there are exponentially a lot ways of putting these (N+1) electrons onto all the M orbitals (of course, M has to be larger than N+1). We call one way of occupying these orbitals with these electrons a “configuration”. Fortunately, not all the configurations will contribute to the spectrum equally. Here, we borrow the convention from quantum chemistry to define the order of configuration [3, 4].

Electronic Configurations of the Final States

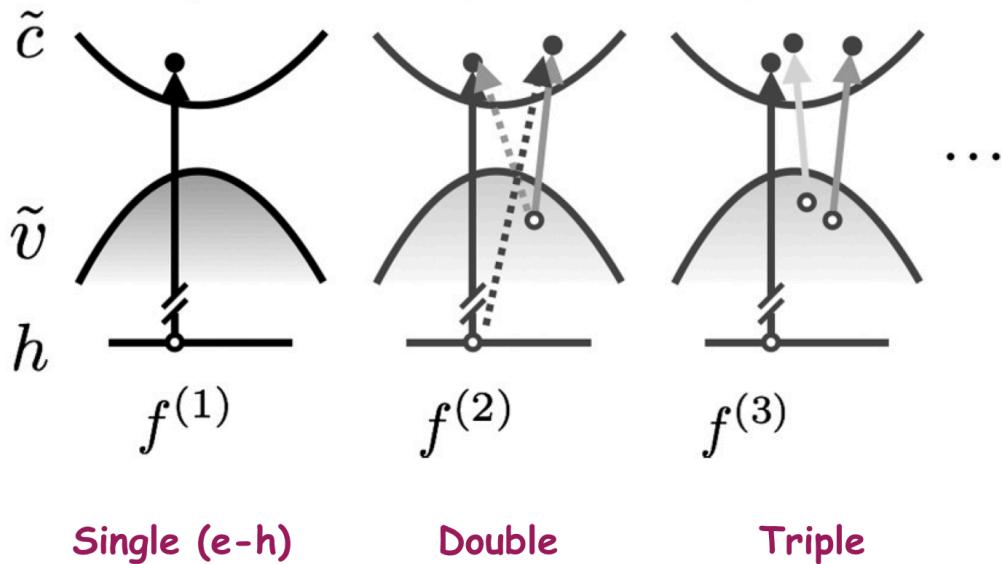


Fig. 1.3 Definition of final-state electronic configurations.

* Definitions:

A single or $f^{(1)}$ term has only one e-h pair (electron-core-hole pair);

A double or $f^{(2)}$ term has only two e-h pairs (including the electron-core-hole pair)

...

Here, n in the bracket of $f^{(n)}$ is referred to as the (excitation) order of the configuration.

As one can imagine, the higher the excitation order n , the less likely that configuration will contribute to the spectrum, because its overlap with the initial-state (ground-state) tends to be smaller. In fact, the XAS/XPS will quickly converge at $n = 2$ in many practical situations.

In order to interface MBXASPY with ShirleyXAS, the original ShirleyXAS is modified to output eigenvalue, eigenwavefunction, and projector (within the projector-augmented-wave, i.e., PAW method) information for both the initial-state and the final-state systems.

1.3 Where to download the codes?

1.3.1 ShirleyXAS

Check out the latest commit on the overlap branch if you have access to
<http://trac-foundry2.lbl.gov/git/Shirley>

```
$ git checkout overlap
$ git pull
$ make shirley
```

Or you may do module-load if you have access to local clusters of High-Performance Computing at Berkeley Lab:

```
$ module load shirley_overlap/QE4.3
```

1.3.2 Pseudopotentials for ShirleyXAS

Since ShirleyXAS is modified based on Quantum Espresso 4.3, the pseudopotentials will be in the Unified Pseudopotential Format (UPF). An automatic XAS calculation will rely on a library of pseudos. There are currently three sources of the pseudo library:

- a. XAS Webtool available locally on NERSC
- b. LBNL local cluster (subject to change)

`/global/home/groups-sw/nano/software/sl-
7.x86_64/MBXAS/XCH_pseudos`

- c. Yufeng's Github (subject to updates):
https://github.com/yufengliang/XCH_pseudos

1.3.3 MBXASPY

Checkout the code on github (<https://github.com/yufengliang/mbxaspy>):

```
git clone https://github.com/yufengliang/mbxaspy ~/mbxaspy
```

From now on you can follow my updates by using “git pull”. You will see many

python source codes with postfix *.py are cloned to your specified directory.

Python 2.7+ is required. Anaconda is recommended.

*** If you don't want to install or read the codes, or simply want to press and go with your structures, you may just use the installed module for the entire ShirleyXAS + MBXASPY calculation:

```
module load MBXAS
```

Please see the next chapter for some examples.

1.4 Contributors and References

Contributors:

David Prendergast (ShirleyXAS)

Eric Shirley (ShirleyXAS)

Yufeng Liang (ShirleyXAS + MBXASPY)

Han Wang (HSE, TDDFT)

References:

[1] Mathieu Taillefumier, Delphine Cabaret, Anne-Marie Flank, and Francesco Mauri, Phys. Rev. B **66**, 195107 (2002), X-ray absorption near-edge structure calculations with the pseudopotentials: Application to the K edge in diamond and α -quartz.

[2] David Prendergast and Giulia Galli, Phys. Rev. Lett. **96**, 215502 (2006), X-Ray Absorption Spectra of Water from First Principles Calculations.

[3] Yufeng Liang, John Vinson, Sri Pemmaraju, Walter S. Drisdell, Eric L. Shirley, and David Prendergast, Phys. Rev. Lett. **118**, 096402 (2017), Accurate X-Ray Spectral Predictions: An Advanced Self-Consistent-Field Approach Inspired by Many-Body Perturbation Theory.

[4] Yufeng Liang and David Prendergast, Phys. Rev. B **97**, 205127 (2018), Quantum many-body effects in x-ray spectra efficiently computed using a basic graph algorithm

[5] Chapter 9 Optical Properties of Solids, Gerald D. Mahan, Many-Particle Physics (2000)

Chapter 2

Quick Start

If you would like to skip the installation and just see how the codes work, you may quickly kick off by loading the MBXAS module on LBNL local cluster (<http://scs.lbl.gov>):

```
module load MBXAS
```

What you will be seeing is a brief introduction of how to run the codes for the entire ShirleyXAS + MBXASPY calculation:

```
Congratulations! You have loaded the MBXAS module
What should you do next?
If you have a set of coordinates (XYZ or POSCAR) and you are ready to
compute MBXAS in the current directory, then...
$ poscar2xyz.sh n m l POSCAR-file
This will generate an XYZ file with the desired supercell definition (n
x m x l)
and will create input in Cell_Block.in for use in the Input_Block.in
(other programs will read this)
OR
$ xyzmol2cell.sh XYZFILE
This will make a supercell for your molecule (with free space around
it)
Otherwise, if you have an XYZ file for a condensed phase system
already, then
you will need to define its unit cell and excited atoms in
Cell_Block.in,
for example:
IBRAV=14
A=10.5
B=12.3
C=8.9
XYZUNIT='angstrom'
XASELEMENTS='Mg6'
Check the template Input_Block.tpl to see what variables might be
overwritten,
or make your own MBXAS_ROOT and place a template there in
${MBXAS_ROOT}/xas_scripts/Input_Block.tpl
```

Now, set up your calculation:

```
$ setup_mbxas_lrc.sh name machine machine time procs
```

```
For example - setup_mbxas_lrc.sh mysystem etna etna 1:00:00 24  
(you can learn a little about your proposed workload by typing  
$ setup_arvid.sh  
note that this is called in most of the scripts that follow anyway).
```

Now run your calculations:

```
$ sbatch arvid.sh
```

Wait for this to finish

```
$ sbatch mbxas.sh
```

Then combine multiple/specific spectra if you want (this is called in
mbxas.sh)

```
$ combine_spectra.sh
```

You are done

When you have prepared your structures in the format of **VASP POSCAR**,
you may follow the above instruction (focus on the lines starting with “\$”) to
run the code. Here are some important sources of the structures:

Crystal Structures from Experiments

<http://www.crystallography.net/cod/search.html>

Crystal Structures from First-Principles Predictions

<https://materialsproject.org>

We will provide a few samples in this chapter to help get you started as
quick and easy as possible.

2.1 Quartz SiO₂

Focus

- Making a supercell for your crystal structure
- Setup input files and scripts
- Choosing the excited atom
- Choosing k-grid density for the supercell SCF calculation
- Changing absolute spectral energy alignment and broadening
- Job submission
- Understanding the results

2.1.1. Supercell

Let's say you have your quartz SiO₂ structure prepared in `sio2.poscar.vasp` in a folder called `$SCRATCH/sio2` as:

```
O2 Si
1.0
 4.9114999771    0.0000000000    0.0000000000
 -2.4557499886   4.2534837509    0.0000000000
  0.0000000000    0.0000000000    5.4038000107
Si  O
 3  6
Direct
 0.469700009   0.0000000000    0.166666672
 0.0000000000   0.469700009     0.833333313
 0.530300021   0.530300021     0.5000000000
 0.413300008   0.267199993     0.285499990
 0.732800007   0.146100014     0.952166677
 0.853899956   0.586699963     0.618833303
 0.267199993   0.413300008     0.714500010
 0.146100014   0.732800007     0.047833342
 0.586699963   0.853899956     0.381166667
```

First, let's make a supercell for it. *A common criterion is to make each dimension $\sim 1 \text{ nm (10 Angstroms)}$.* The above cell is roughly $5\text{\AA} \times 5\text{\AA} \times 5\text{\AA}$ so we can make a $2 \times 2 \times 2$ supercell by

```
$ poscar2xyz.sh 2 2 2 sio2.poscar.vasp > sio2.xyz
```

Then you'll see the supercell is directed into the xyz file **sio2.xyz** (Note that coordinates are in the unit of **angstroms**) and a **Cell_Block.in** for the crystal parameters will show up.

2.1.2 Setup files

Next you may setup the input and job-scheduling files using a loaded script file called **setup_mbxas_lrc.sh**. Its usage is:

```
$ setup_mbxas_lrc.sh xyz_name account partition_name
wall_time nproc
```

By carrying out:

```
$ setup_mbxas_lrc.sh sio2 etna etna 06:00:00 32
```

You have setup:

- two job-scheduling files: **arvid.sh**, **mbxas.sh**
- one input file: **Input_Block.in**

“sio2” is simply the job name (you name it).

“etna” is the both the partition and queue name.

“06:00:00” sets the wall time to be 6 hours.

“32” means running the job with 32 processors.

Vaguely speaking, **arvid.sh** *is for DFT calculations and mbxas.sh is for the determinant calculation and relative spectral alignment.* A really cool thing about the two scripts is that they will restart from the last checkpoint. If your job halts because of exceeding the wall time, the scripts will detect at which step the calculation is, and then carry on the calculation.

* you may read this paragraph later

This detection is made possible by finding and scanning through the output files *.out (*.scf.out, *.nscf.out, *.ham.out, *.xas.out, *.mbxas.out, ...)

).Therefore, if you did something wrong in the previously halted calculation, please remember go into the “./XAS/xyz_name/GS” and “XAS/xyz_name/XATOM” folder and remove the problematic *.out files.

The input file `Input_Block.in` defines all the variables used in your core-level spectroscopy calculation. Please treat `Input_Block.in` as a bash script and specify variables in a bash way. You may check the first few lines to see where the ShirleyXAS (`$SHIRLEY_ROOT`), MBXASPY (`$MBXASPY_ROOT`), and pseudopotential library (`$PSEUDO_DIR`) come from, in case you want to make a change.

Below we focus on a few most-frequently edited variables in `Input_Block.in`.

2.1.3 Excited atom

Let's say if you'd like to calculate O *K* edge for quartz SiO₂. Search for the variable XASELEMENTS (it is in the comment lines and undefined when `Input_Block.in` is initialized). Since all the oxygen atoms in SiO₂ are equivalent, you only need to choose one of them as the core-excited atom. To do so, add this line below the “XASELEMENTS” comment block:

XASELEMENTS='O25'

By doing this, the ground-state oxygen pseudopotential will be replaced by the excited-atom pseudopotential in the final-state DFT calculation.

* Read this later

When the calculation is done, you can take a peek at the DFT calculation. You can tell from the “./XAS/sio2/O25/sio2.O25-XCH.scf.in” that the ground-state O pseudo “O.pbe-van_bm.UPF” is replaced by “O.pbe-van-dgp-1s1.UPF” !

2.1.4 K-points

You've already had a supercell, use your experience of DFT calculation to figure out what would appropriate for the k-grid in the SCF calculation. Here for SiO₂, we may use

K_POINTS="K_POINTS automatic

2 2 2 0 0 0"

This is exactly how k-points are defined in QE. This k-grid for the supercell corresponds to a $4 \times 4 \times 4$ k-grid for the unit cell.

2.1.5 Absolute Energy Alignment and Broadening

ESHIFT=3.5

SIGMA=0.6

ESHIFT controls the where the entire spectrum on the energy axis and SIGMA is the broadening. Both are in the unit of eV and fit from experiment.

2.1.6 Job Submission

With the above variable settings, you can just press and go ! Here are the commands you will need to execute:

```
$ sbatch arvid.sh
```

After this DFT step is done, you may do

```
$ sbatch mbxas.sh
```

2.1.7 Results

Should everything go on smoothly, you will have three files in the root working directory:

1. `spec0_i_all.dat`, the initial-state spectrum, or the PDOS
2. `spec0_f_all.dat`, the one-body final-state spectrum
3. `spec_xas_all.dat`, the determinant spectrum

So, just plot them out !

```
$ gnuplot
gnuplot> plot [0:25] "spec0_i_all.dat" u 1:2 w 1,
"spec0_f_all.dat" u 1:2 w 1, "spec_xas_all.dat" u 1:2 w 1
```

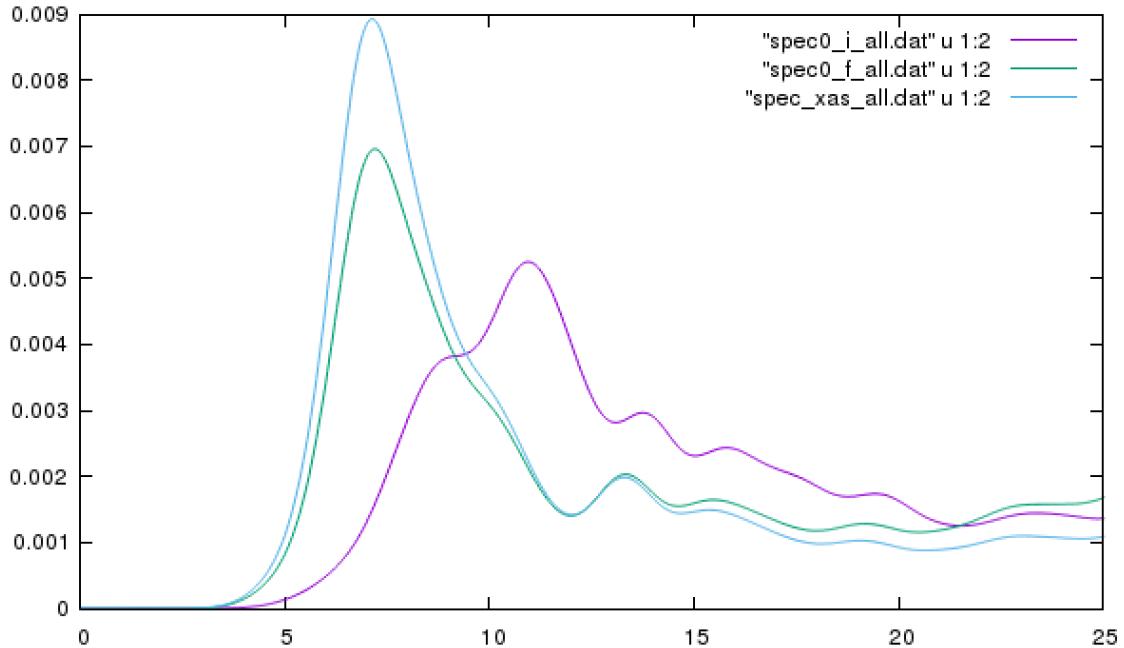


Fig. 2.1 The computed SiO_2 O K edge. The x-axis is the energy axis (in eV).

This is a trivial case where the determinant spectrum (at the order of $f^{(1)}$) highly resembles the one-body final-state spectrum (except for a overall factor), which is discussed in Ref. [4] in Chapter 1. However, you may see a significant core-hole effects by comparing `spec0_i_all.dat` and `spec0_f_all.dat` (`spec_xas_all.dat`), which red-shifts the “main” peak from ~ 12 eV to ~ 7 eV!

For more discussion, please refer to Ref. [4].

2.2 Benzene

Focus

- Supercell for molecules
- Change the number of empty bands (orbitals) and energy range
- Polarization dependence

2.2.1 Supercell for molecules

The benzene molecular structure is prepared in an xyz file, `benzene.xyz`:

```
12
benzene example
C 0.00000 1.40272 0.00000
H 0.00000 2.49029 0.00000
C -1.21479 0.70136 0.00000
H -2.15666 1.24515 0.00000
C -1.21479 -0.70136 0.00000
H -2.15666 -1.24515 0.00000
C 0.00000 -1.40272 0.00000
H 0.00000 -2.49029 0.00000
C 1.21479 -0.70136 0.00000
H 2.15666 -1.24515 0.00000
C 1.21479 0.70136 0.00000
H 2.15666 1.24515 0.00000
```

Since this is a molecule defined in an xyz file using Angstrom as unit, and we are using plane-wave basis for the DFT calculation, we will need to define a bounding box in the XAS simulation so that spurious interactions between supercell images are small enough. This is easily made possible by the loaded command `xyzmol2cell.sh`:

```
$ xyzmol2cell.sh benzene.xyz
```

Then we can setup the calculation:

```
$ setup_mbxas_lrc.sh benzene etna etna 02:00:00 32
```

2.2.2 Number of bands

The variable **NBND_FAC** controls how many bands (orbitals) you want to include for calculating XAS. The number of bands is equivalent to the variable “nbnd” in Quantum Espresso. Let’s say if n bands ($\text{nbnd} = n$) are used in the SCF calculation (usually ShirleyXAS will determine n automatically), then the “nbnd” in the NSCF calculation will be set to

$$n \times \text{NBND_FAC}$$

NBND_FAC=2 or **3** is typically a good number of many solids, which will cover the near-edge structure up to several tens of eV. But one needs to increase this number to a few tens for molecules so as to cover orbitals up to the continuum, if necessary. Here, we set

NBND_FAC=30

In the SCF step, $\text{nbnd} = 19$ (see `./XAS/benzene/C1/benzene.C01-XCH.scf.in`). There $\text{nbnd} = 570$ in the NSCF step (see `./XAS/benzene/C1/benzene.C01-XCH.nscf.in`).

Other settings:

XASELEMENTS='C1' # all carbon atoms are equivalent

After the job finishes, you can see in `./XAS/benzene/C1/benzene.C01-XCH.nscf.out` that the maximal orbital energy is > 22.9 eV, whereas the HOMO is ~ -1.7 eV and LUMO is ~ -0.9 eV. This means you can expect the code to cover the energy range within $22.9 + 0.9 = 23.8$ eV above the absorption onset.

Note that **NBND_FAC** can be a fractional number like 1.5 and 1.23 ! Just so you can precisely control the band number.

2.2.3 Polarization

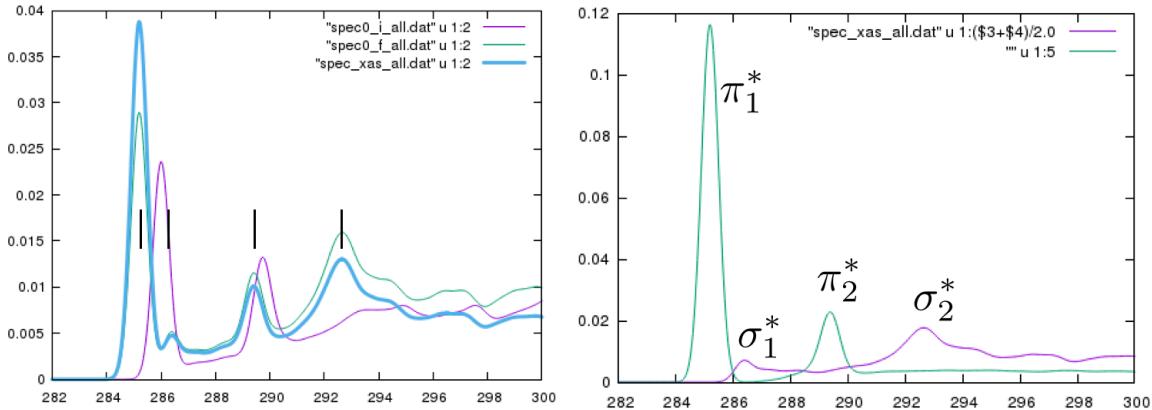


Fig. 2.2

Left panel: spherically (angularly) averaged of the XAS of a benzene molecule. This may be compared with the XAS of gas-phase benzene. The determinant spectra are at the $f^{(1)}$ order.

Right panel: XAS with polarization within the benzene plane (green) and perpendicular to the benzene plane (purple).

For spinless case, there are five (5) columns in all the spec*.dat files. *The meanings of each column are:*

- (1) energy (eV)
- (2) the spherically averaged spectrum
- (3)-(5) the spectrum with polarization along x, y, z direction respectively

From the benzene xyz file, we know that the benzene molecule is lying within the xy-plane. So plotting the 1st column against 3rd column + 4th column yields the XAS with in-plane polarization for benzene, reflecting the π^ orbitals, whereas plotting the 1st column against 5th column yields the XAS with polarization perpendicular to the benzene plane.*

2.3 TiO₂

Focus

- First nontrivial example using the determinant approach
- Full core-hole approach (FCH)

2.3.1 Parameter settings

Notable examples for the determinant approach to work are transition metal oxides (TMOs) in which the *d* levels split into *t*_{2g} and *e*_g levels in the octahedral crystal field. The rutile phase of TiO₂ could be the simplest example (see Ref. [3, 4] in Chapter 1 for more discussion)

Starting from tio2.poscar.vasp:

```
O2 Ti
1.0
  4.5887799263    0.0000000000    0.0000000000
  0.0000000000    4.5887799263    0.0000000000
  0.0000000000    0.0000000000    2.9575600624

Ti  O
  2  4
Direct
  0.0000000000    0.0000000000    0.0000000000
  0.5000000000    0.5000000000    0.5000000000
  0.302329987    0.302329987    0.0000000000
  0.697669983    0.697669983    0.0000000000
  0.197670013    0.802330017    0.5000000000
  0.802330017    0.197670013    0.5000000000
```

```
$ poscar2xyz.sh 2 2 3 tio2.poscar.vasp > tio2.xyz
$ setup_mbxas_lrc.sh tio2 etna etna 06:00:00 64
```

Parameters:

```
XASELEMENTS='O25'
ECUT_WFC='30'
ECUT_RHO='300'
K_POINTS="K_POINTS automatic"
```

```
2 2 2 0 0 0"
ESHIFT=535.7
SIGMA=0.6
CHAPPROX="FCH" # XCH → FCH
```

Here, we change the charge approximation (defined by **CHARPROX**) from the eXcited-electron and the Core-Hole approach “**XCH**” to the full core-hole approach “**FCH**”. For why we need to do this for TMOs please refer to Ref. [3, 4] in Chapter 1.

We also increase **ECUT_WFC** and **ECUT_RHO** here for TMOs.

2.3.2 Results

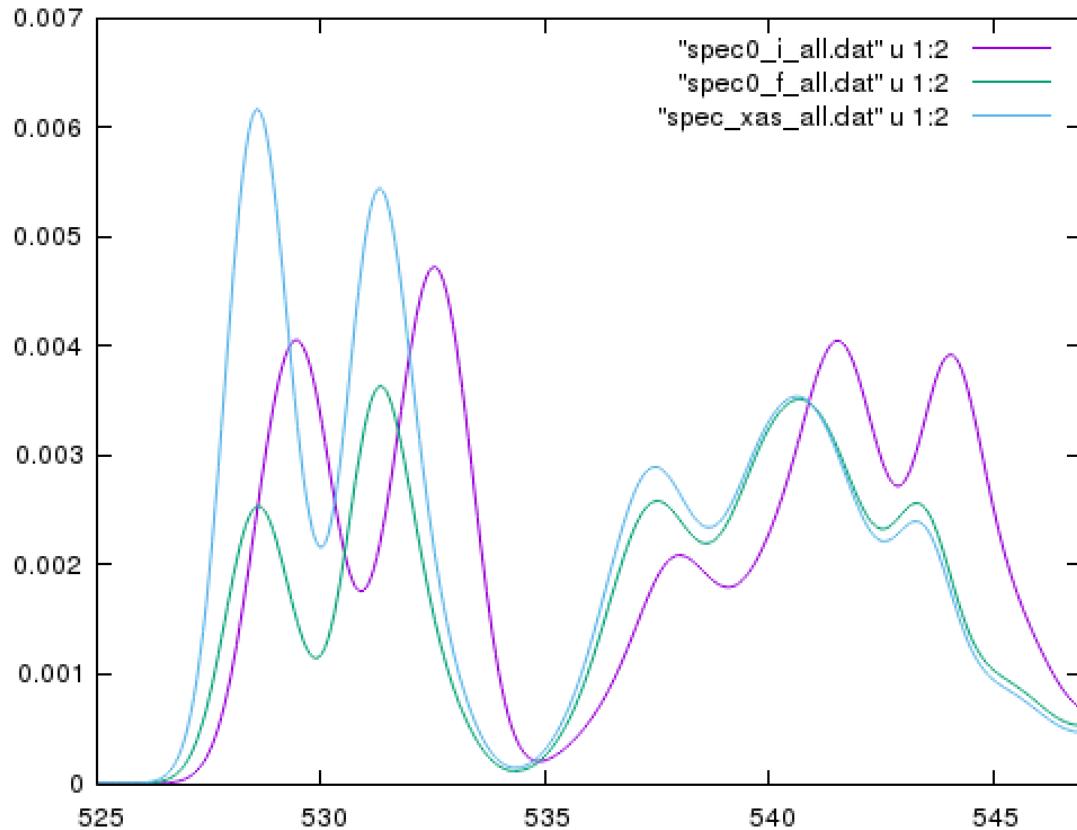


Fig. 2.3 Initial-state, one-body final-state, and the determinant XAS at the $f^{(1)}$ order for rutile TiO_2 .

The above figure is created by:

```
gnuplot> plot [525:546] "spec0_i_all.dat" u 1:2 w l,  
"spec0_f_all.dat" u 1:2 w l, "spec_xas_all.dat" u 1:2 w l
```

As one can see, the redshift from the initial-state spectrum to the final-state spectrum, the core-level excitonic effects are NOT negligible in TMOs like TiO₂, although the core-hole is created at the oxygen site but not the metal site. The core hole at O can significantly change the hybridization, leading to (1) a reverse of the peak intensities of the first two peak; (2) redshift of the oscillator strength at the near-edge region by ~ 1 eV, which is also large !

2.4 TiO₂ spin-polarized

Focus

- spin-polarized calculation
- high-order contribution from $f^{(n)}$ for $n = 2$
- XPS spectrum

2.4.1 Parameter settings

Take the Input_Block.in from Sec. 2.3 for the spinless calculation of TiO₂, and simply add the block (can be found in the comment)

```
SPIN=" nspin=2
      starting_magnetization(2)=0.01,
      starting_magnetization(3)=0.0,"
```

You may see this is just how spin-polarized DFT is initiated in QE. For detailed instruction, please go to:

https://www.quantum-espresso.org/Doc/INPUT_PW.html

The index i in starting_magnetization(i) refers to the atoms defined in the ATOMIC_SPECIES block. ShirleyXAS will automatically arrange all the atoms in *alphabetical order*, and add *the excited atom at the end* of the list.

For the case of TiO₂, the ATOMIC_SPECIES block looks like (as can be found in ./XAS/tio2/O25/tio2.O25-FCH.scf.in):

```
ATOMIC_SPECIES
O 15.999 O.pbe-van_bm.UPF
Ti 47.867 Ti.pbe-van-yufengl.UPF
OX 15.999 O.pbe-van-dgp-1s1.UPF
```

$i = 1, 2, 3$ refers to the atom O, Ti, and OX respectively. Here, starting_magnetization(2)=0.01 initialize tiny spin polarization on Ti, as required by QE, although the spin-up and spin-down manifold are equivalent for TiO₂. You may wonder why it is necessary to perform a spin-polarized calculation in this case. While it is true that the initial-state, or the one-body

final-state XAS does not dependent on spin for TiO_2 , you'll find in the following discussions that *the spinless and spin-polarized determinant XAS do differ* due to shakeup effects !

2.4.2 Spinless v.s. spin-polarized spectra

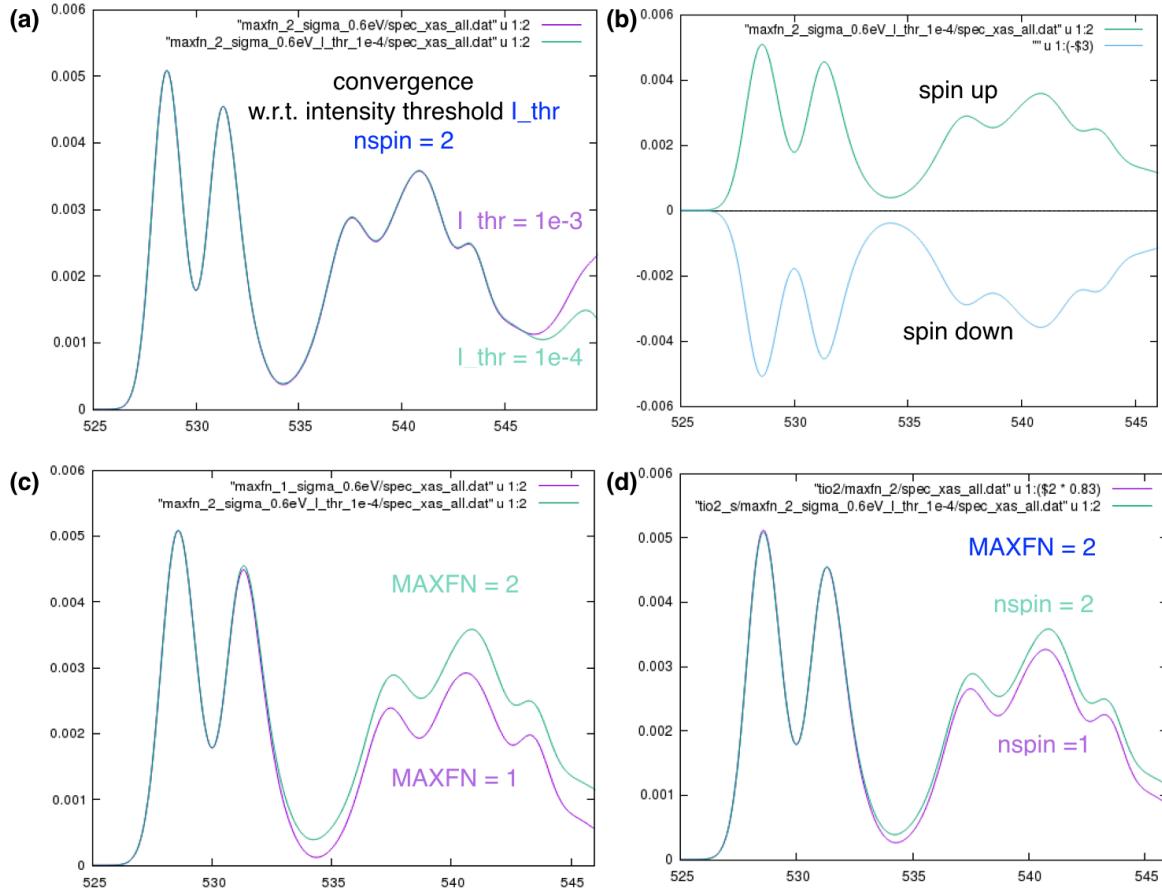


Fig. 2.4.1 (a) Convergence with respect to intensity threshold I_{THR} (I_{thr} in mbxaspy) for the determinant XAS up to $f^{(2)}$. (b) Spin-up and spin-down XAS up to $f^{(2)}$. (c) comparison of the XAS generated from MAXFN=1 (up to $f^{(1)}$) and MAXFN=2 (up to $f^{(2)}$). (d) spinless versus spin-polarized calculation up to $f^{(2)}$.

Starting from 533 eV, the contribution from the $f^{(2)}$ terms becomes noticeable. Here, we can perform the determinant calculation up to $f^{(2)}$, i.e., $f^{(1)} + f^{(2)}$ by setting:

MAXFN='2' # string type

in `Input_Block.in`.

If you have successfully run `MAXFN='1'` before, please rename/relocate/remove the file

```
./XAS/tio2/O25/tio2.O25-FCH.mbxas.out
```

because the code will detect this file to judge whether `mbxas.sh` has been finished or not (by finding the keyword “`mbxaspy done`” at the end of the file). If the keyword is found, then the code will simply skip `mbxas.sh` for that excited atom (`O25` in this case).

In general, you may also relocate

```
./spec*_all.dat
./XAS/xyz_name/excited_atom/*.*
```

into a separate folder to allow the next `mbxas.sh` run and for future record.

If performing higher order calculation, then the convergence with respect to the intensity threshold, i.e., `I_THR` in `Input_Block.in` needs to be tested. Assume the intensity (oscillator strength) of the strongest transition I_{\max} (typically within the $f^{(1)}$ terms). Then the transitions (configurations) with an intensity less than:

$$I_{\max} \times I_THR$$

will be neglected so as to speed up the calculation. For a detailed discussion of `I_THR`, please refer to Sec. II. E of Ref. [4] in Chapter 1. `I_THR` is equivalent to R_{th} in the article.

A good number for TiO2 up to 546 eV is (also the default value):

$$I_THR='1e-3'$$

As shown in Fig. 2.4.1 (a), the XAS of TiO2 are identical for `I_THR=1e-3` and `I_THR=1e-4` up to 546 eV, which means the XAS has already converged at `I_THR=1e-3`. For the spectrum at higher-energy, one may need

to use lower `I_THR` like 5e-4 or even 1e-4, although rarely this part of XAS will be of interest.

To further speed up the determinant calculation (and to even avoid out-of-memory !), one may also use fewer orbitals than what have been generated in the NSCF step. If `NBND_FAC=3` in this case, then Kohn-Sham orbitals will be generated up to 58 eV, whereas the Fermi energy is ~ 10 eV. But truly we may just be interested in the first 20 eV above the absorption onset. Here we can also play with the energy range in the `mbxas.sh` step:

`ELOW=-5`
`EHIGH=-26`

which means the determinant calculation will cover from -5 eV to $58 - 26 = 22$ eV with respect to the *lowest-energy transition*. You may seek more information in `mbxas.out`. Reducing the energy window to the range we are interested in can substantially prune down the search space because it grows fast with respect to the number of orbitals.

*

For the collinear spin-polarized case, there are nine (9) columns in all the `spec*.dat` files. *The meanings of each column are:*

- (1) energy (eV)
- (2) the spherically averaged spectrum, spin up
- (3) the spherically averaged spectrum, spin down
- (4)-(9) the spectrum with polarization along x up, x down, y up y down, z up, and z down direction respectively.

As shown in Fig. 2.4.1 (b), the spin-up (the 2nd column) and spin-down (the 3rd column) XAS of rutile TiO2 are *identical*.

*

As shown in Fig. 2.4.1 (c), the contribution from the $f^{(2)}$ terms becomes discernable above 533 eV, which increases the intensity of the absorption hump between 535 to 545 eV to $\sim 25\%$.

*

As shown in Fig. 2.4.1 (d), the determinant XAS up to $f^{(2)}$ for spinless and spin-polarized calculation are slightly different at the higher-energy hump between 535 to 545 eV. This is because in the spin = 2 case, there are more shakeup channels than in the spin = 1 case. If a core hole is created, shakeup effects may occur in either the spin up or spin down channel for spin=2, but there is only one channel for spin = 1. Stronger shakeup effects will transfer more spectral weight to higher energy as in spin = 2. Including the shakeup effects make the intensity ratio between the pre-edge (527~535 eV) and the main edge (535~545 eV) even closer to experiment !

2.4.3 Shakeup effects: contribution from $f^{(2)}$

The XPS is output to spec_xps_all.dat at the root working directory. This is the final XPS spectrum with two spin-wise XPS convoluted so there are only two columns in this file: the energy axis and the intensity.

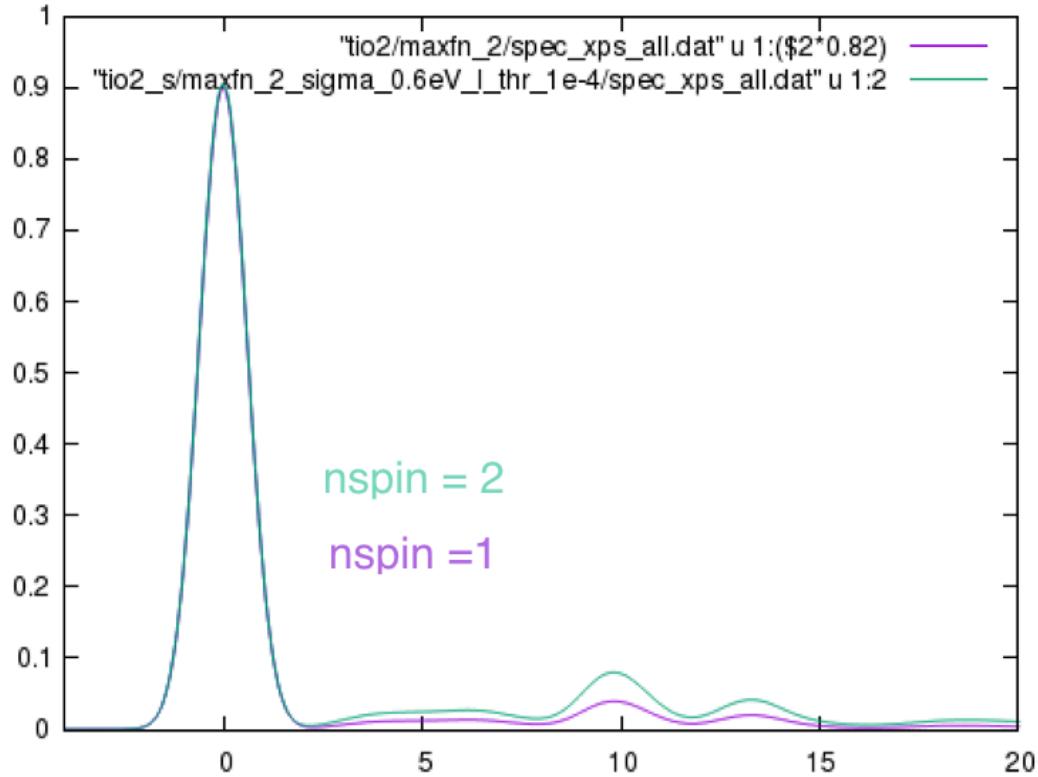


Fig. 2.4.2 XPS for spin = 1 and 2. Note that the first peak intensity is normalized to the same intensity by multiplying the XPS for spin=1 by 0.82.

If **MAXFN='1'** then no shakeup effects will be included in the determinant calculation, and there is only one peak (defined as $f^{(0)}$ of XPS) at $E = 0$,

which reflects the overlap of the two ground many-body states of the pristine and core-excited system.

If **MAXFN='2'** then shakeup effects from $f^{(1)}$ (of XPS) begin to kick in at $E > 0$. The $f^{(1)}$ contribution has two discernable peaks at ~ 10 eV and 14 eV, with some minor features between $0 \sim 10$ eV and near 20 eV. The shakeup effects at 10 eV represent the electron-hole (valence, non-core-level) excitations at ~ 10 eV induced by the core-hole generation. When coupled with the near-edge core-level excitation, the $f^{(1)}$ of XPS contribute to the shakeup features between 533 to 546 eV of $f^{(2)}$ of XAS. Note that the most significant $f^{(2)}$ contribution of XAS is also located around 10 eV above the onset !

2.5 CrO₂

Focus

- spin-polarized DFT + U calculation
- higher-order contribution from $f^{(n)}$ for $n = 2$
- parallelization for MBXASPY

2.5.1 DFT + U

Rutile phase of CrO₂ stored in cro2.poscar.vasp:

```

Cr O2
1.0
 4.4210000038    0.0000000000    0.0000000000
 0.0000000000    4.4210000038    0.0000000000
 0.0000000000    0.0000000000    2.9170000553

Cr  O
 2  4
Direct
 0.0000000000    0.0000000000    0.0000000000
 0.5000000000    0.5000000000    0.5000000000
 0.3009999999    0.3009999999    0.0000000000
 0.6990000001    0.6990000001    0.0000000000
 0.1990000001    0.8009999999    0.5000000000
 0.8009999999    0.1990000001    0.5000000000

```

```

$ poscar2xyz.sh 2 2 3 cro2.poscar.vasp > cro2.xyz
$ setup_mbxas_lrc.sh cro2 etna etna 06:00:00 64

```

Setting the following lines in Input_Block.in will initiate the spin-polarized DFT + U calculation.

```

SPIN=" nspin=2
      starting_magnetization(1)=1.0,
      starting_magnetization(2)=0.0,"
LDAU="lda_plus_u=.true., Hubbard_U(1)=3.5"

```

Here, index 1 = Cr, 2 = O, 3 = OX. Therefore, `Hubbard_U(1)=3.5` means using a U value of 3.5 eV for Cr. For more details related to indexing and ordering of the element, please refer to Sec. 2.4 for spin-polarized TiO₂ and also the manual for pw.x.

Other parameter settings:

```
XASELEMENTS='O25'
ESHIFT=535.7
NBND_FAC=3      # energies up to 67.1 eV,  $E_{\text{fermi}} \sim 10.6$  eV
MAXFN='2'        # Do MAXFN='1' first
I_THR='1e-4'     # 1e-3 is not low enough
CHAPPROX='FCH'   # Important because of doing spin-polarized DFT+
U
SIGMA=0.5
EHIGH=-35        # up to 67 - 35 = 32 eV, ~ 22 eV above onset
ECUT_WFC='30'
ECUT_RHO='300'
```

Before running the mbxas.sh calculation for `MAXFN='2'`, remember to relocate `./XAS/cro2/O25/cro2.O25-FCH.mbxas.out` for `MAXFN='1'`.

2.5.2 $f^{(2)}$

For CrO₂, it is essential to perform the determinant calculation up to the $f^{(2)}$ order.

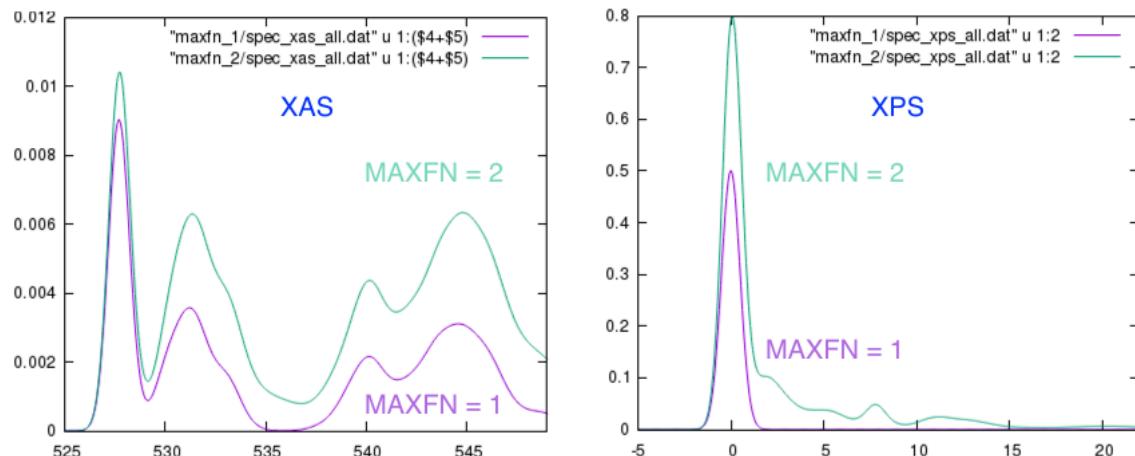


Fig. 2.5.2 XAS and XPS for MAXFN=1 and 2 for CrO₂. For XAS, x up + x down is plot (the 4th column = x up, the 5th column = x down).

The XPS of CrO₂ has some significant continuous features near E=0, which extend all the way up to 15 eV. This suggests shakeup effects (valence e-h excitations induced by the core hole) are important in x-ray excitations of a metallic system like CrO₂.

2.5.3 Parallelization

Although we have been using supercell for generating x-ray spectra, a *k*-grid is also necessary for sampling the Brillouin Zone of the supercell. This *k*-grid density is defined by a sole parameter

[XAS_ARG=5](#)

in the Input_Block.in, which means a $5 \times 5 \times 5$ *k*-grid. Separate determinant calculations will be carried out for each *k*-grid and eventually the spectra will be combined. If you would like to even out the calculation, it is recommended to choose the number of processors (for each excited atom) to be the multiple of 125. Normally 125 and 250 will serve the purpose.

2.6 Acetic acid

Focus

- Multiple excited atoms

2.6.1 the acetic acid molecule

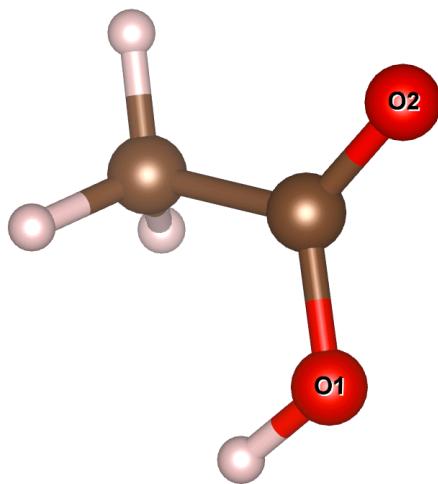


Fig. 2.6.1 An acetic acid molecule has two nonequivalent O atoms. O1 is bonded to a carbon atom and a hydrogen whereas O2 is simply bonded to one carbon.

acetic.xyz

8

acetic acid

O	0.3034	-1.2890	-0.0002
O	0.9801	0.8876	-0.0002
C	-1.3743	0.3517	-0.0002
C	0.0907	0.0496	0.0005
H	-1.8394	-0.0566	-0.9009
H	-1.8413	-0.0611	0.8976
H	-1.5186	1.4361	0.0026
H	-0.5077	-1.8414	-0.0002

In this section, we will perform a XAS calculation that combines the XAS from the two nonequivalent oxygen atoms. This is made possible by DFT formation energy calculation, i.e., *how difficult is it for the lowest electron-core-hole pair (the 1st core-excited state) to form at each site*. The relative formation energy difference between the excited O1 and O2 will be used to align the spectra.

2.6.2 Parameter settings

```
$ xyzmol2cell.sh acetic.xyz
$ setup_mbexas_lrc.sh acetic etna etna 01:00:00 32

XASELEMENTS='O1 O2'
NBND_FAC=30
ESHIFT=526.5
```

2.6.3 Results and discussions

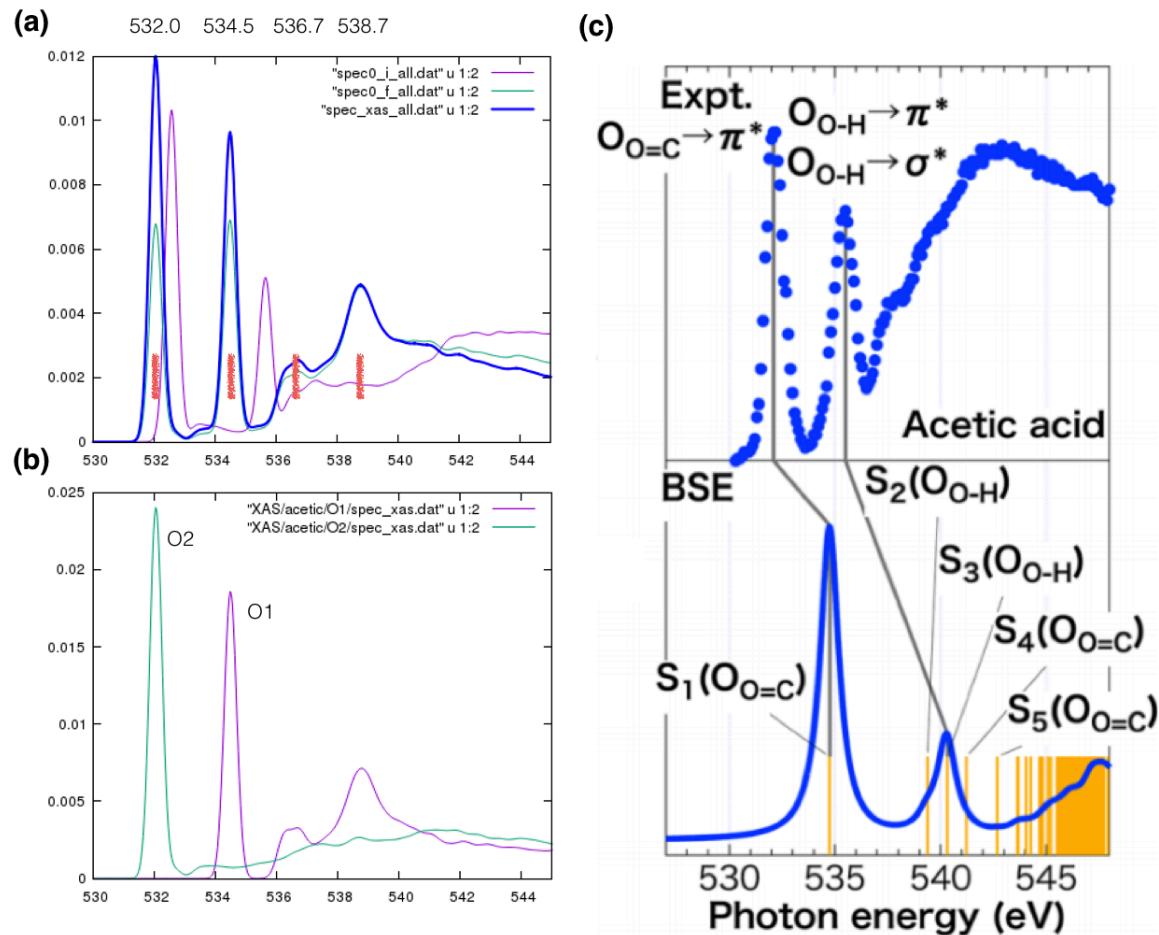


Fig. 2.6.2 (a) combined spectra (b) separate determinant spectra (c) Experimental and BSE spectra from *J. Chem. Theory Comput.*, **2015**, *11* (4), pp 1668–1673

The above determinant calculation (bolded blue curve in Fig. 2.6.2 (a)) produces the peak-intensity ratio better than the one-body final-state approach (green curve in Fig. 2.6.2 (b)) for the first two peaks, which originates from O2 and O1.

The high-energy part above 536.0 eV is less satisfactory and the spectral intensity is lower than in experiment. The extended region of XAS is defined by scattering states and we don't expect the near-edge theory to be able to efficiently compute spectrum at this part.

The orbital energies used in generating the spectra are extracted directly from DFT calculation with simple pbe functionals. We do not expect the peak positions to be precise. Dilation of the energy axis (due to many-body corrections) may also be applied to the simulated spectra in order to align each peak with the experimental spectrum.

2.7 V₂O₅

Focus

- Multiple-atom calculation in crystals
- Relative energy alignment

2.7.1 O species in layered V₂O₅

v2o5.poscar.vasp:

```
v2o5 my pwscf
1.0
11.503 0.000 0.000
0.000 3.557 0.000
0.000 0.000 4.369
V O
4 10
Direct
0.148326851 -0.000000000 0.892054942
0.351673108 0.500000000 0.107943786
0.648326952 0.500000000 0.107942280
0.851673089 0.000000000 0.892056112
0.181336098 0.500000000 0.999527467
0.318663861 0.000000000 0.000472618
0.681336171 0.000000000 0.000472625
0.818663918 0.500000000 0.999527561
0.855273582 0.000000000 0.527885576
0.644726381 0.500000000 0.472115749
0.355273601 0.500000000 0.472114575
0.500000024 0.500000000 0.997557899
-0.000000016 -0.000000000 0.002442214
0.144726378 0.000000000 0.527886512
```

The unit cell is shown as below

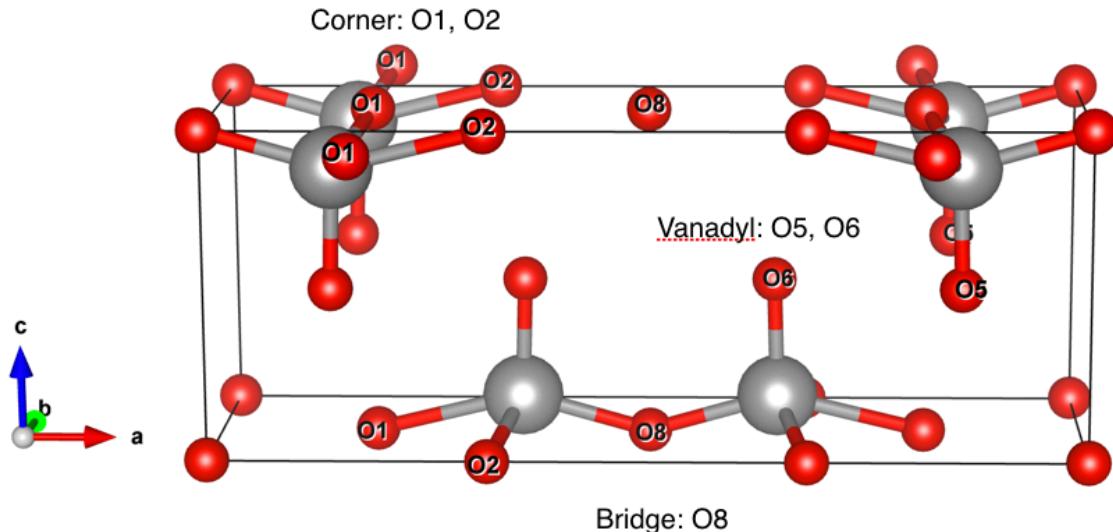


Fig. 2.7.1 Unit of layered V_2O_5 . There are 10 oxygen atoms in total, 4 of which belong to the “corner” oxygen, the other 4 belong to the dangling “vanadyl” oxygen, and the rest 2 belong to the “bridge” oxygen.

In order to keep this stoichiometry, we can take 2 “corner” oxygen ($\text{O}1, \text{O}2$), 2 “vanadyl” oxygen ($\text{O}5, \text{O}6$), and 1 “bridge” oxygen ($\text{O}8$) in the XAS calculation. So these are 5 oxygen atoms in total.

```
$ poscar2xyz.sh 1 3 2 v2o5.poscar.vasp > v2o5.xyz
$ setup_mbxas_lrc.sh v2o5 etna etna 24:00:00 64
```

Here, we have made a $1\times3\times2$ supercell. So which 5 atoms should we pick ? Since O coming after V in the POSCAR file above, then there are 24 V atoms coming in front of O. poscar2xyz will first visit each atom in the unit cell one by one and make 6 duplicates. So we can pick:

XASELEMENTS='O25 O31 O49 O55 O67'

which correspond to $\text{O}1, \text{O}2, \text{O}5, \text{O}6$, and $\text{O}8$ as in Fig. 2.7.1.

Other parameter settings:

```
ESSHIFT=535.7
NBND_FAC=2
I_THR='1e-4'
SIGMA=0.4
EHIGH=-8
```

```

CHAPPROX="FCH"
ECUT_WFC='30'
ECUT_RHO='300'
SPIN="" nspin=2
    starting_magnetization(2)=0.01,
    starting_magnetization(3)=0.0,"
LDAU="lda_plus_u=.true., Hubbard_U(2)=3.1" # 1=O, 2=V, 3=OX

```

2.7.2 Results

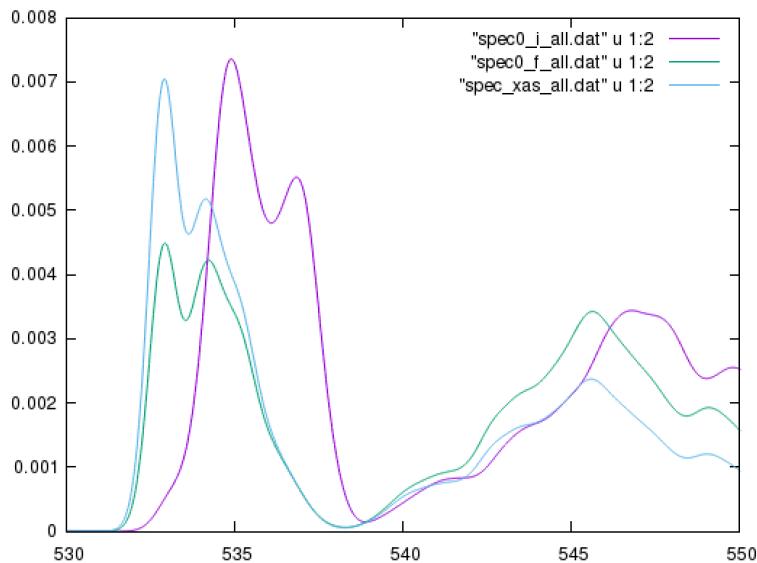


Fig. 2.7.2 the *combined* initial-state, one-body final-state, and the determinant spectrum of layered V_2O_5 .

The separate spectra (`spec0_i.dat`, `spec0_f.dat`, and `spec_xas.dat`) for individual oxygen are stored in `./XAS/v2o5/O*/`.

Note that `ESHIFT=535.7` is no longer a good alignment energy for the XAS in this case. For the experimental spectrum, one could refer to Nat. Comm. 7, 12022 (2016).

Chapter 3

Workflow

