Gaussian and Augmented Plane Wave Method

In this tutorial we perform all-electron calculations with GAPW to compute the near-edge X-ray absorption spectra of ice-1h. When performing all-electron calculations GAPW accounts for electron density redistribution of the core electrons, and therefore can be used for the computation of X-ray absorption spectra resulting from the absorption of X-rays exciting the core-shell electrons.
For more information see 10.1007/s002140050523, 10.1039/B615522G
This tutorial is adapted from the cp2k online tutorial at https://www.cp2k.org/exercises:2017_uzh_cp2k-tutorial:gapw

Short Intro

In near edge X-ray absorption fine structure (NEXAFS), the X-rays absorbed by each individual atom excite the core electrons, which therefore leave a hole in the core shell (i.e. a core hole). In the decay process the photo-electron can scatter from electrons of neighboring atoms. This scattering process affects the absorption coefficient that is measured for different energies of the incoming synchrotron radiation. In the case of ice we are interested in the spectra with a range of transition energies roughly between 530 and 550 eV.
The absorption coefficient is proportional to the transition probability, which can be computed by applying Fermi's Golden rule, and where the initial unperturbed state is that of the core orbital and the final state is that of the excited photo-electron that is affected by the presence of neighboring atoms.
The transition probability within DFT is computed by modifying the Kohn-Sham equations to include the core-hole potential on the absorbing atom. In practice the transition probability can be computed under the electric dipole approximation and using the full-core hole or half-core hole scheme.

Task 1: compute the XAS spectrum of ice-1h

Go into your folder and copy the directory "`gapw_ice1h_xas`" there:

```
cp -r $WS/00_workshop_files/day02/gapw_ice1h_xas/ .
```

Run the XAS simulation using 4 cores.
After this calculation is finished you should see several files called `ice1h-xas_at*.spectrum`, which contain the bare XAS spectra. In the first column there is the index of the virtual KS state, the second column is the transition energy, the third, fourth and fifth column are the transition probabilities projected onto X, Y and Z, and finally the 6th column is the norm of the transition probability.

Now run the following command to convolute the spectrum so we can compare it to previously published calculations and NEXAFS experiments for ice 1h:

```
./get_average_spectrum.sh
```

Open `gnuplot` and plot, respectively, the bare and the convoluted spectrum by typing

```
plot "./spectrum.inp" using 2:6 with l,  "./spectrum.out" using 1:5 with l
```

Hartree-Fock exchange

The purpose of this section is to explain how to perform hybrid functional calculations (or Hartree-Fock exchange,HFX) with CP2K in condensed phase systems. Hartree-Fock exchange in CP2K is based on four center electron repulsion integrals, these are computed with an external library (libint).

This approach has a computational cost that depends strongly on the nature of the basis. Unless combined with ADMM (see Task 5 below), do not use MOLOPT basis sets with HFX. We use basis sets from HFX_BASIS, which are suitable. This tutorial is adapted from the online cp2k tutorial (https://www.cp2k.org/exercises:2017_uzh_cp2k-tutorial:hybrid)

Truncated Coulomb operator

To enable HFX in the condensed phase CP2K employs a truncated Coulomb operator for the exchange part. The physical picture is that we do not want to have 'self-exchange interactions' of an electron with its image in neighboring unit cells. As a rule of thumb, the maximum range (truncation radius) is L/2 where L is the smallest edge of the unit cell. The convergence of the exchange energy is exponential wrt. this radius. Typically, 5-6 A provides good results, but this depends on the nature of the system, i.e. the band gap or the range of the maximally localized Wannier orbitals.

Task 2 : GGA restart wfn

We will perform a single point PBE-D3 calculation to generate an initial wavefunction (wfn) restart. HFX calculations benefit from this. You will find an input file pbed3.inp for a single point (energy) calculation with CP2K under $WS/00_workshop_files/day02/hfx. The folder contains an xyz file with the structure of a box containing 64 water molecules. Copy the folder into your directory, where you will run the calculation.

Run the input and rename the generated wfn file (`WATER-RESTART.wfn`) to `WATER-RESTART-GGA.wfn`. Also make a note of the `HOMO - LUMO gap [eV]` which is reported (the experimental water band-gap is of about 6.9 eV).

Task 3: PBE0-D3 water

To do a hybrid calculation, we just change the &XC section. Various examples can be found in the regtests folder distributed with cp2k, here we employ settings equivalent to PBE0-D3. One can also find more functionals like meta-GGA etc using the library of exchange-correlation functionals (using input section &LIBXC), and mix with different ratio of HF contribution to formulate various hybrid functionals.

Change the pbed3.inp input as follows, and save as pbe0d3.inp:

- `SCF_GUESS RESTART` in the `&SCF` section
- `WFN_RESTART_FILE_NAME WATER-RESTART-GGA.wfn` in the `&DFT` section
- Replace the `&XC…` section with the following one:

```
    # specify the exchange and correlation treatment
    &XC
      # use a PBE0 functional
      &XC_FUNCTIONAL
       &PBE
         # 75% GGA exchange
         SCALE_X 0.75
         # 100% GGA correlation
         SCALE_C 1.0
```

```
             &END PBE
          &END XC_FUNCTIONAL
          &HF
            # 25 % HFX exchange
            FRACTION 0.25
            # Important to improve scaling from O(N^4) to O(N)
            &SCREENING
              # important parameter to get stable HFX calcs (contributions to hfx
smaller than EPS_SCHWARZ are not considered)
              EPS_SCHWARZ 1.0E-6
              # needs a good (GGA) initial guess
              # screening on the product between maximum of density matrix elements and
ERI
              SCREEN_ON_INITIAL_P TRUE
            &END
            &INTERACTION_POTENTIAL
              # for condensed phase systems
              POTENTIAL_TYPE TRUNCATED
              # should be less than half the cell
              CUTOFF_RADIUS  6.0
              # data file needed with the truncated operator
              T_C_G_DATA ${LIB}/t_c_g.dat
            &END
            &MEMORY
              # In MB per MPI rank.. use as much as need to get in-core operation
              MAX_MEMORY 4000
              EPS_STORAGE_SCALING 0.1
            &END
          &END
          # adding Grimme's D3 correction (by default without C9 terms)
          &VDW_POTENTIAL
            POTENTIAL_TYPE PAIR_POTENTIAL
            &PAIR_POTENTIAL
              PARAMETER_FILE_NAME ${LIB}/dftd3.dat
              TYPE DFTD3
              REFERENCE_FUNCTIONAL PBE0
              R_CUTOFF [angstrom] 16
            &END
          &END VDW_POTENTIAL
       &END XC
```

- EPS_PGF_ORB, EPS_SCHWARZ, SCREEN_ON_INITIAL_P and EPS_FILTER_MATRIX are parameters to guarantee stable SCF. You can consult the online cp2k input manual to remember the meaning of these parameters.

Have a look at the output, in the section where CP2K is performing the SCF loop using the OT method. The first iteration should look something like this:

```
-------------------------------- OT --------------------------------------

 Step     Update method      Time    Convergence          Total energy     Change
 --------------------------------------------------------------------------

 Trace(PS):                                512.0000000000
 Electronic density on regular grids:     -511.9999881229        0.0000118771
 Core density on regular grids:            511.9999857029       -0.0000142971
 Total charge density on r-space grids:     -0.0000024200
 Total charge density g-space grids:        -0.0000024200


 HFX_MEM_INFO| Est. max. program size before HFX [MiB]:                   388

 *** WARNING in hfx_energy_potential.F:600 :: The Kohn Sham matrix is not  ***
 *** 100% occupied. This may result in incorrect Hartree-Fock results. Try ***
 *** to decrease EPS_PGF_ORB and EPS_FILTER_MATRIX in the QS section. For   ***
 *** more information see FAQ: https://www.cp2k.org/faq:hfx_eps_warning     ***

 HFX_MEM_INFO| Number of cart. primitive ERI's calculated:      11551170770
 HFX_MEM_INFO| Number of sph. ERI's calculated:                  2681710894
 HFX_MEM_INFO| Number of sph. ERI's stored in-core:              1115419808
 HFX_MEM_INFO| Number of sph. ERI's stored on disk:                       0
 HFX_MEM_INFO| Number of sph. ERI's calculated on the fly:                0
 HFX_MEM_INFO| Total memory consumption ERI's RAM [MiB]:              1075
 HFX_MEM_INFO| Whereof max-vals [MiB]:                                  44
 HFX_MEM_INFO| Total compression factor ERI's RAM:                    7.92
 HFX_MEM_INFO| Total memory consumption ERI's disk [MiB]:                0
 HFX_MEM_INFO| Total compression factor ERI's disk:                  0.00
 HFX_MEM_INFO| Size of density/Fock matrix [MiB]:                      10
 HFX_MEM_INFO| Size of buffers [MiB]:                                   2
 HFX_MEM_INFO| Number of periodic image cells considered:              27
 HFX_MEM_INFO| Est. max. program size after HFX  [MiB]:               664

    1 OT DIIS     0.80E-01  121.7     0.00096329     -1102.1377186463 -1.10E+03
```

You can see a `WARNING` that is related to the tags `EPS_PGF_ORB` and `EPS_FILTER_MATRIX`. This particular warning is related to the stability of the SCF cycle. For more information see https://www.cp2k.org/faq:hfx_eps_warning.

- Look at the output where the HOMO-LUMO gap has been printed out. How does this compare to the result?
- Adjust the fraction of exchange (modify the input in two places!) to 20% and/or 30%, how does this influence the gap?
- Optional You can check if the cycle is stable by decreasing the values of `EPS_PGF_ORB`, `EPS_FILTER_MATRIX` as well as `EPS_SCHWARZ`. Re-run the calculation and see how this affects the `ENERGY`.
- Optional CP2K tries to store the ERI in-core and avoid to calculate them at each SCF step. Especially for large systems that can be run on large HCP machines it is important to run in-core operation and fit the calculations of the ERI into memory. To see the effect of not having enough memory on the time for the SCF cycle, modify the tag of `MAX_MEMORY` to 40 and rerun the calculation. How do the timings compare with those where `MAX_MEMORY` was

larger? **It is recommended to run calculations in-core, either by setting larger `MAX_MEMORY` or using more processes.**

Truncated Coulomb operator with long range correction

In screened hybrid functionals, the 1/r exchange interaction potential can be separated into a long range and a short range part, where only the short range part is computed using exact exchange and the long-range part is computed using GGA exchange. This is what is done also when using a truncated Coulomb operator. This can be used to speedup the calculation, while retaining the benefits of HFX. The functional employed in this way smoothly goes from PBE to PBE0 as the `CUTOFF_RADIUS` goes from 0 to Infinity.

Task 4:

Add to the &XC_FUNCTIONAL section (i.e. in addition to &PBE) the following section:

```
    &PBE_HOLE_T_C_LR
      CUTOFF_RADIUS 2.5
      SCALE_X 0.25
    &END
```

and employ the same `CUTOFF_RADIUS` for the `INTERACTION_POTENTIAL`. (i.e. you need to modify this value also). Rerun the single point energy calculation and note the band gap.

- Is such a short range sufficient to have a sizable effect on the band gap?
- Is `HFX_MEM_INFO| Number of cart. primitive ERI's calculated` very different for calculations with 2.5 and 6.0 A truncation radius? Remember to use the same cutoff radius under the `&_HOLE_T_C_LR` and `&INTERACTION_POTENTIAL` sections.

Auxiliary Density Matrix Methods (ADMM)

ADMM is an approach to mitigate the cost of HFX for large basis sets. In particular, if MOLOPT basis sets are used, standard HFX becomes too expensive (CP2K can not deal efficiently with highly contracted AOs). In ADMM, an `AUX_FIT_BASIS_SET` is introduced, which is used to create an auxiliary density matrix (ADM) by projection. HFX is evaluated for this ADM, while the error introduced by using an ADM is corrected for with a GGA exchange functional.

Task 5: ADMM

Make the following changes to the input:
- insert an additional line `BASIS_SET_FILE_NAME ${LIB}/BASIS_ADMM`
- insert for each `&KIND` a line `AUX_FIT_BASIS_SET cFIT3`
- insert a secion `&AUXILIARY_DENSITY_MATRIX_METHOD`

```
    # use ADMM
    &AUXILIARY_DENSITY_MATRIX_METHOD
      # recommended, i.e. use a smaller basis for HFX
      # each kind will need an AUX_FIT_BASIS_SET.
      METHOD BASIS_PROJECTION
      # recommended, this method is stable and allows for MD.
      # can be expensive for large systems
      ADMM_PURIFICATION_METHOD MO_DIAG
    &END
```

N.B. In this tutorial we combine using a very small basis set (cFIT3) with a small primary basis (DZVP-), so gains are small at best (and the results not very accurate). ADMM is most useful with good quality primary basis sets, such as e.g. MOLOPTs

Run the input, what's the `HOMO - LUMO` gap?

^{1}H Nuclear Magnetic Shieldings

GAPW all-electron calculations can be used to calculate absolute nuclear magnetic shieldings

Task 6: Calculating 1H nuclear magnetic shieldings

● Here, we calculate the magnetic shielding of the three groups of protons in ethanol.

● One of the earliest detections of the chemical shift, Stanford 1951: (http://demonstrations.wolfram.com/NuclearMagneticResonanceSpectrumOfEthanol)
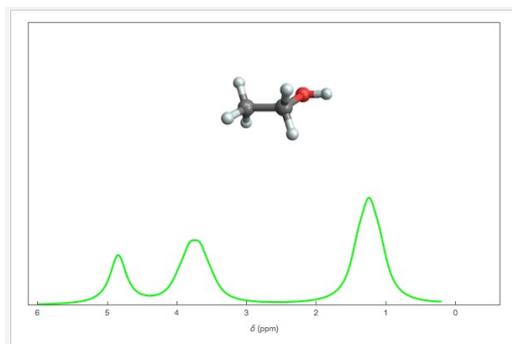
● Copy the $WS/day02/nmr folder into your working folder and run the calculation.

Note the following:

● We are using the all-electron aug-pcSseg-2 basis sets from Jensen. (pcSseg-2 for the carbons)

● The calculation type under the &GLOBAL section is "LR" for "Linear Response"

● The section &PROPERTIES&LINRES... gives the input parameters for the response calculation. In CP2K the magnetic shielding is calculated from the induced current density, hence the &CURRENT section.

Results:

● The shieldings are output in "EtOH_nmr-SHIFT-1.data". For each atom the total shielding tensor is printed under "SIGMA TOTAL", the rank-0 isotropic average (which is the value observed in high resolution NMR in liquids) is then printed as "ISOTROPY= ".

● These values are absolute shieldings, i.e. they are not referenced to some standard, like is typically done in experimental NMR (^{1}H shielding is usually referenced to TMS). Easiest way to compare to experiment is by comparing the differences in the shieldings.

● Except at very low temperatures, the chemically identical protons in ethanol give one resonance frequency, thus one should average over the shieldings of the two CH_2 protons (atoms 4 and 5), and over the three CH_3 protons (atoms 6, 7, 8).

● Comparing experimental and theoretical NMR data is particularly tricky. Magnetic resonance is highly sensitive to the slightest variations in the local environment around the spin, including solvation effects and temperature fluctuations (this is one main reason why magnetic resonance is a very useful tool!). NMR is also a slow spectroscopy (resonance in the MHz frequency range), and vibrational averaging is of paramount importance to obtain a quantitative agreement with experiments.

● Two experimental gas phase measurements of ethanol report a CH3-CH2 magnetic shielding difference of 2.48 and 2.75 ppm, and for the CH3-OH difference the values -0.65 and 0.73 ppm are reported. Note that the exchangeable protons (the OH proton) are tricky and they could give resonances at widely varying frequencies depending on the conditions. How do these experimental values compare to your calculated values?