Linear-Response Time-Dependent Density Functional Theory for X-Ray Absorption Spectroscopy in CP2K

(LR-TDDFT for XAS in CP2K)

Augustin Bussy

University of Zurich

CONEXS Summer School 2019

Presentation Outline

Theory

- Basics of LR-TDDFT

- Restriction to XAS

- Making it affordable

Using XAS_TDP in CP2K

- Setting up an appropriate ground state calculation

- Selecting the donor core orbitals

- Defining the Kernel

- Other keywords

- Typical workflow

- Tuning the precision

- Performance tricks

Basics of LR-TDDFT

Standard DFT:

$$H_{KS}^\sigma[n_\alpha, n_\beta](\mathbf{r}) \; \phi_{j\sigma}(\mathbf{r}) = \varepsilon_{j\sigma}\phi_{j\sigma}(\mathbf{r})$$

$$H_{KS}^\sigma[n_\alpha, n_\beta](\mathbf{r}) = -\frac{1}{2}\boldsymbol{\nabla}^2 + v_{eff}^\sigma[n_\alpha, n_\beta](\mathbf{r})$$

$$v_{eff}^\sigma[n_\alpha, n_\beta](\mathbf{r}) = v_{ext}(\mathbf{r}) + v_H[n](\mathbf{r}) + v_{xc}^\sigma[n_\alpha, n_\beta](\mathbf{r})$$
$$= v_{ext}(\mathbf{r}) + \int \mathrm{d}\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{xc}[n_\alpha, n_\beta]}{\delta n_\sigma(\mathbf{r})}$$

$$n_\sigma(\mathbf{r}) = \sum_{j=1}^{N_\sigma} \phi_{j\sigma}^*(\mathbf{r})\phi_{j\sigma}(\mathbf{r})$$

Basics of LR-TDDFT

Time-dependent DFT:

$$i\frac{\partial}{\partial t}\phi_{j\sigma}(t,\mathbf{r}) = H^{\sigma}[n_{\alpha}, n_{\beta}](t,\mathbf{r})\phi_{j\sigma}(t,\mathbf{r})$$

$$H^{\sigma}[n_{\alpha}, n_{\beta}](t,\mathbf{r}) = -\frac{1}{2}\boldsymbol{\nabla}^2 + v_{eff}^{\sigma}[n_{\alpha}, n_{\beta}](t,\mathbf{r})$$
$$= -\frac{1}{2}\boldsymbol{\nabla}^2 + v_{ext}(t,\mathbf{r}) + v_H[n](t,\mathbf{r}) + v_{xc}^{\sigma}[n_{\alpha}, n_{\beta}](t,\mathbf{r})$$

$$n_{\sigma}(t,\mathbf{r}) = \sum_{j=1}^{N_{\sigma}} \phi_{j\sigma}^*(t,\mathbf{r})\phi_{j\sigma}(t,\mathbf{r})$$

Basics of LR-TDDFT

Assume incoming photon of frequency $\omega \Rightarrow$ harmonic perturbation

$$v_{ext}(t, \mathbf{r}) = v_{ext}^0(\mathbf{r}) + v^-(\mathbf{r})e^{-i\omega t} + v^+(\mathbf{r})e^{i\omega t}$$

This leads to a harmonic perturbation in the density:

$$n_\sigma(t, \mathbf{r}) = n_\sigma^0(\mathbf{r}) + n_\sigma^-(\mathbf{r})e^{-i\omega t} + n_\sigma^+(\mathbf{r})e^{i\omega t}$$

And to the Kohn-Sham orbitals:

$$\phi_{j\sigma}(t, \mathbf{r}) = \phi_{j\sigma}^0(t, \mathbf{r}) + \phi_{j\sigma}^+(\mathbf{r})e^{i\omega t} + \phi_{j\sigma}^-(\mathbf{r})e^{-i\omega t}$$

Basics of LR-TDDFT

Basis set expansion:

$$\left|\phi_{j\sigma}^0\right\rangle = \sum_{q=1}^M c_{qj\sigma}^0 \left|\varphi_q\right\rangle \qquad \left|\phi_{j\sigma}^{\pm}\right\rangle = \sum_{q=1}^M c_{qj\sigma}^{\pm} \left|\varphi_q\right\rangle \qquad S_{pq} = \langle\varphi_p|\varphi_q\rangle$$

Matrix formulation:

$$\omega \begin{pmatrix} -\mathbf{G} & 0 \\ 0 & \mathbf{G} \end{pmatrix} \begin{pmatrix} \mathbf{C}^+ \\ \mathbf{C}^- \end{pmatrix} = \begin{pmatrix} \mathbf{A} + \mathbf{B} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} + \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{C}^+ \\ \mathbf{C}^- \end{pmatrix}$$

$$A_{pi\sigma,qj\tau} = \left(F_{pq}^\tau - \varepsilon_{j\tau} S_{pq}\right)\delta_{ij}\delta_{\sigma\tau} \qquad G_{pi\sigma,qj\tau} = S_{pq}\delta_{ij}\delta_{\sigma\tau}$$
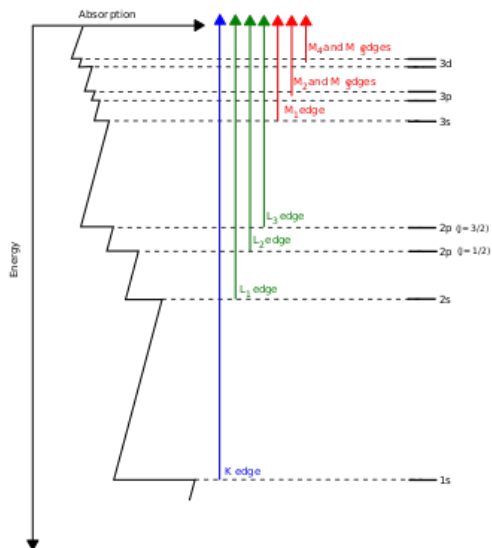
$$B_{pi\sigma,qj\tau} = \sum_{stuv} Q_{ps}^\sigma \, c_{ti\sigma}^0 \, K_{st\sigma,uv\tau}^{Hxc} \, c_{uj\tau}^0 \, Q_{vq}^\tau$$

Basics of LR-TDDFT

- Huge matrices. Size: $2 \times N_{MO} \times N_{sgf} \times N_{spin}$
  $\rightarrow$ Iterative solvers

- Many integrals to compute, $\mathcal{O}(N_{sgf}^4)$
  $\rightarrow$ Exploit symmetry and locality

- The kernel is frequency-dependent
  $\rightarrow$ Adiabatic approximation: $f^{Hxc} = f_{\sigma,\tau}^{Hxc}(\mathbf{r}, \mathbf{r'})$

- Non-hermitian eigenvalue problem
  $\rightarrow$ Tamm-Dancoff approximation, $c_{qj\tau}^+ = 0$:

$$\omega \ \mathbf{G} \ \mathbf{C}^- = (\mathbf{A} + 2\mathbf{B}) \, \mathbf{C}^-$$

Restriction to XAS

Restriction to XAS

Weak coupling between core and valence excitations

$\rightarrow$ Only consider MOs below a given energy[1]

$\rightarrow$ Drastically reduces the size of the matrix to diagonalize: indices i, j of $A_{pi\sigma,qj\tau}$ only span a small subset of core MOs

Very localized core orbitals

$\rightarrow$ Sudden approximation[2]: total absorption is the sum of single electron excitations contributions

$\rightarrow$ In TDDFT[3], reduction to a single donor state. Reduces the matrix size to that of the KS one

---

[1] M. Stener, G. Fronzoni & M. de Simone, Chem. Phys. Let. **373**, 115-123 (2003)

[2] J. Rehr, E. Stern, R. Martin & E. Davidson, Phys. Rev. B **17**, 560 (1978)

[3] S. DeBeer George, T. Petrenko & F. Neese, Inorg. Chim. Acta **361**, 965-972 (2008)

Making it affordable

The kernel matrix involve 4-center 2-electrons integrals ($\mathcal{O}(N^4)$):

$$K_{pI\sigma,qJ\tau}^{Hxc} = \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r'} \varphi_p(\mathbf{r})\phi_{I\sigma}(\mathbf{r}) f_{\sigma,\tau}^{Hxc}(\mathbf{r}, \mathbf{r'}, \pm\omega)\varphi_q(\mathbf{r'})\phi_{J\tau}(\mathbf{r'})$$
$$= (pI_\sigma | f_{\sigma,\tau}^{Hxc} | qJ_\tau)$$

where (adiabatic approximation)

$$f_{\sigma,\tau}^{Hxc}(\mathbf{r}, \mathbf{r'}) = \frac{1}{\mathbf{r} - \mathbf{r'}} + \frac{\delta^2 E_{xc}[n_\alpha, n_\beta]}{\delta n_\sigma(\mathbf{r})\delta n_\tau(\mathbf{r'})}\bigg|_{n^0(\mathbf{r})}$$

Making it affordable: Coulomb kernel

Make use of a resolution of the identity (RI) in the V-approximation[4]:

$$(qI|pJ) \approx \sum_{\mu\nu} (pI|\mu) \, (\mu|\nu)^{-1}(\nu|qJ) = \sum_{\mu\nu} (qI|\mu) \, V_{\mu\nu}^{-1} \, (\nu|qJ)$$

$|\mu)$ form a RI basis that should span the $|pI)$ product.

$\rightarrow |pI)$ is local, can take a small RI basis centered on excited atom.

$\rightarrow$ Drastically reduce the number of 3-center integrals to compute !

$\rightarrow V_{pq}^{-1}$ is precomputed for each excited kind.

---

[4]Vahtras *et. al.*, Chem. Phys. Lett. **213**, 5-6 (1993)

Making it affordable: exchange-correlation kernel

Use the V-approximation twice:

$$(pI|f^{xc}|qJ) \approx \sum_{\mu\nu\xi\chi} (pI|\mu) \, V_{\mu\nu}^{-1} \, (\nu|f^{xc}|\xi) \, V_{\xi\chi}^{-1} \, (\chi|qJ)$$

$\rightarrow$ 3-centers integrals $(pI|\mu)$ and $V_{\mu\nu}^{-1}$ already computed

$\rightarrow$ $(\nu|f^{xc}|\xi)$ is a simple 2-centers integral. Only need $f^{xc}[n]$ in the vicinity of the excited atom. Numerical integration on small radial grid.

Making it affordable: exchange-correlation kernel

Need to evaluate the density at each atomic grid point.

Projection of $n(\mathbf{r})$ on the local RI basis ($\{\chi_\mu\}$):

$$n(\mathbf{r}) = \sum_{pq} P_{pq} \; \varphi_p(\mathbf{r})\varphi_q(\mathbf{r})$$
$$\approx \sum_{pq} \sum_{\mu\nu} P_{pq} \; (\varphi_p\varphi_q\chi_\mu) \; S_{\mu\nu}^{-1} \; \chi_\nu(\mathbf{r})$$
$$= \sum_\nu d_\nu \; \chi_\nu(\mathbf{r})$$

$(\varphi_p\varphi_q\chi_\mu)$ is a 3-center overlap between the orbital and RI basis functions and $S_{\mu\nu}^{-1} = (\chi_\mu\chi_\nu)^{-1}$ is the inverse overlap between two RI basis functions.

Making it affordable: exact exchange kernel

Necessary for high quality hybrid functionals. Use the same RI as for Coulomb:

$$(pq|IJ) \approx \sum_{\mu\nu} (pq|\mu) \; V_{\mu\nu}^{-1} \; (\nu|IJ)$$

$\rightarrow$ Despite locality of μ, ν, I, J all overlapping p, q of the system must be considered

$\rightarrow$ This is impossible in periodic boundary conditions. One must resort to short range potentials such as truncated Coulomb or $\mathrm{erfc}(\omega r)/r$

$\rightarrow$ Usually the bottleneck in large-scale simulations

Using XAS_TDP in CP2K

Using XAS_TDP in CP2K

All examples will refer to Trimethylaluminium, aka $Al_2(CH_3)_6$:
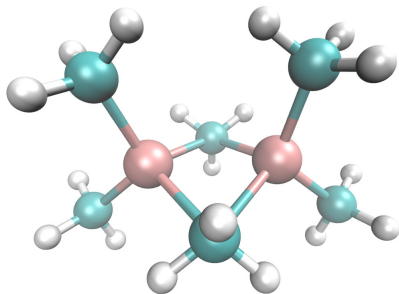
Setting up the ground state calculation

The following keywords should be set outside of the &XAS_TDP section:

- In &GLOBAL, set RUN_TYPE to ENERGY

- In &FORCE_EVAL, &DFT:

    - Set AUTO_BASIS RI_XAS to SMALL, MEDIUM or LARGE
      To automatically generate a RI basis set based on the orbital one

    - In &QS, set METHOD to GAPW
      This allows for all-electron treatment of excited atoms

    - In &SCF, set SCF_GUESS to RESTART
      In case the XAS_TDP calculation is done separately, the ground state
      is readily available

- In &FORCE_EVAL, &SUBSYS:
  Make sure to use all-electron basis sets and to set POTENTIAL to
  ALL when defining an atomic &KIND

Setting up the ground state calculation

Input snippets for XAS_TDP compatible ground state calculation:

```
&GLOBAL
  PROJECT trimethylAl
  RUN_TYPE ENERGY
&END GLOBAL

&FORCE_EVAL
  &DFT                              &SUBSYS
    AUTO_BASIS RI_XAS MEDIUM          &KIND Al
                                        BASIS_SET 6-31G*
    &QS                                 POTENTIAL ALL
      METHOD GAPW                     &END KIND
    &END QS
                                      &KIND C
    &SCF                                BASIS_SET 6-31G*
      SCF_GUESS RESTART                 POTENTIAL ALL
    &END SCF                          &END KIND

    !+ functional, cutoff, poisson, etc.    !+ hydrogen, cell, coordinates, etc.
                                      &END SUBSYS
  &END DFT                        &END FORCE_EVAL
```

Selecting the donor core orbitals

XAS calculations are triggered by the &XAS_TDP subsection of &DFT

The first step is to define which are the core donor MOs.

```
&XAS_TDP
  &DONOR_STATES
    DEFINE_EXCITED BY_KIND
    KIND_LIST Al
    STATE_TYPES 1s
    N_SEARCH 2
    LOCALIZE
  &END DONOR_STATES

  !+ kernel, grid size, etc.
&END XAS_TDP
```

The above snippet will look for Aluminium 1s orbitals among the 2 lowest eigenvalue MOs. The donor orbitals will be localized, which is a necessary step if different excited atoms are equivalent under symmetry.

```
&XAS_TDP
  &DONOR_STATES
    DEFINE_EXCITED BY_INDEX
    ATOM_LIST 1 2
    STATE_TYPES 1s 1s
    N_SEARCH 2
    LOCALIZE
  &END DONOR_STATES

  !+ kernel, grid size, etc.
&END XAS_TDP
```

Here the code looks for the 1s orbital of atoms 1 and 2. This works assuming the Aluminiums are listed before other atoms in the coordinates definition.

Defining the kernel

The &KERNEL subsection defines which functionals are used. It entirely relies on the LIBXC library[5] for the xc part:

```
&XAS_TDP
  &KERNEL
    &XC_FUNCTIONAL
      &LIBXC
        FUNCTIONAL GGA_C_PBE
      &END LIBXC
      &LIBXC
        FUNCTIONAL GGA_X_PBE
      &END LIBXC
    &END XC_FUNCTIONAL
  &END KERNEL

  !+ donor_states, grid size, etc.
&END XAS_TDP
```

The very popular GGA functional PBE for the kernel. It is explicitly made of a correlation and an exchange part

```
&XAS_TDP
  &KERNEL
    &XC_FUNCTIONAL
      &LIBXC
        FUNCTIONAL HYB_GGA_XC_B3LYP
      &END LIBXC
    &END XC_FUNCTIONAL
    &EXACT_EXCHANGE
      OPERATOR COULOMB
      SCALE 0.2
    &END EXACT_EXCHANGE
  &END KERNEL

  !+ donor_states, grid size, etc.
&END XAS_TDP
```

The B3LYP hybrid functional with 20% of exact exchange. Note: never use the Coulomb operator in periodic boundary conditions, but TRUNCATED or SHORTRANGE instead

---

[5]M. Marques, M. Oliveira & T. Burnus, Comp. Phys. Comm. **183**, 2272-2281 (2012)

Other keywords

Some other stand alone keywords define XAS_TDP:

- CHECK_ONLY   this keyword triggers a trial run to check the quality of the donor core MOs. None of the heavy computation is done.

- TAMM_DANCOFF   enables the Tamm-Dancoff approximation (cheaper, still high quality). It is on by default.

- GRID   defines the density of the integration grid of a given kind for the XC kernel. For our molecule: **GRID Al 100 150** $\Rightarrow$ for kind Al, the angular grid has a 100 points and the radial one has 150

There are other keywords, but the default are in principle good enough

Typical workflow

1. Get the structure(s): from geometry optimization, MD trajectory snapshots, experimental data, etc.

2. Make a compatible ground state calculation (GAPW, all electron potential and basis set, etc.)

3. Add the &XAS_TDP section in the input with the CHECK_ONLY keyword. Make sure you set SCF_GUESS RESTART to avoid recomputing the ground state. In the ouput, check that the donor MOs have a well defined type (1s, 2s, 2p) and are local in space

4. Remove the CHECK_ONLY keyword and rerun. The spectral data (energies and dipole oscilaltor strengths) is in the *.spectrum file

Note: steps 2 and 3 can be combined

Typical workflow

Checking the donor MOs quality in the *.out file:

Almost perfect overlap of the MO
with a pure Aluminium 1s
=> we are indeed doing K-edge XAS

```
# Check the donor states for their quality. They need to have a well defined type
  (1s, 2s, etc) which is indicated by the overlap. They also need to be localized,
  for which the Mulliken population analysis is one indicator (must be close to 1.0)

 -Donor state of type 1s for atom   1 of kind Al:
  The following canonical MO(s) have been associated with the donor state(s)
  based on the overlap with the components of a minimal STO basis:
                                    Spin   MO index    overlap(sum)
                                     1        1          0.99939

  The next best overlap for spin 1 is 0.00000 for MO with index   1

  Mulliken population analysis retricted to the associated MO(s) yields:
                                    Spin  MO index     charge
                                     1       1         1.000
```

100% of the charge carried by
this MO is on the excited atom
=> it is well localized

Typical workflow

Getting the spectrum (*.spectrum file):

```
=============================================================
XAS TDP singlet excitations (no SOC) for DONOR STATE: 1s,
from EXCITED ATOM:    1, of KIND (index/symbol):    1/Al
=============================================================
    Index      Excitation energy (eV)    fosc dipole (a.u.)
      1              1526.770467             0.004640
      2              1526.986501             0.002163
      3              1527.262972             0.002232
      4              1527.611562             0.000805
      5              1527.691835             0.000313
      6              1528.124279             0.001084
      7              1528.992356             0.000094
      8              1529.254595             0.000167
      9              1529.698447             0.000027
     10              1529.783112             0.000401
     11              1530.265039             0.000392
     12              1530.457520             0.000671
     13              1530.806057             0.000147
     14              1530.895715             0.001086
     15              1531.388507             0.000510
     16              1531.672528             0.000102
     17              1531.739339             0.000321
     18              1531.863481             0.000367
     19              1532.154992             0.000824
     20              1532.418933             0.000126
```

Tuning the precision

- Changing the **GRID** density, *e.g.* GRID AI 300 300 for a very fine integration

- Using **higher quality RI basis**, setting AUTO_BASIS RI_XAS to LARGE or HUGE

- Using the **RI_RADIUS** keyword of the &KERNEL section. RI basis elements of all atoms within the sphere of radius RI_RADIUS centered on the excited atom contribute to the density description.

  Very useful for **hybrid functionals**, which are very sensitive to the quality of the XC kernel. Including nearest neighbors is usually enough.

Performance tricks: exploiting symmetry

In many molecules and solids, some atoms are **equivalent under symmetry**, *i.e.* they see and feel the same environment. In this case, there is **no need to compute excitations from all of them** as they should all be the same.

For $Al_2(CH_3)_6$:
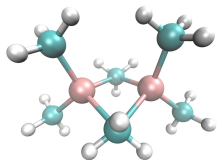
The Aluminiums are the same under symmetry. It is hence necessary to LOCALIZE two donor MOs, but only 1 atom can be treated under XAS_TDP

```
&XAS_TDP
  &DONOR_STATES
    DEFINE_EXCITED BY_INDEX
    ATOM_LIST 1
    STATE_TYPES 1s
    N_SEARCH 2
    LOCALIZE
  &END DONOR_STATES

  !+ kernel, grid size, etc.
&END XAS_TDP
```

Performance tricks: distributing the workload

By default, the integration grids for the XC kernel are distributed over the processors with **one grid per processor**. If there are more processor being used than excited atoms, some of them will be idling. The keyword NPROCS_GRID allows to distribute integration grids over batches of processors for **improved speed and load balance**.

```
&XAS_TDP
  NPROCS_GRID 2

  &XC_FUNCTIONAL
    &LIBXC
      FUNCTIONAL GGA_C_PBE
    &END LIBXC
    &LIBXC
      FUNCTIONAL GGA_X_PBE
    &END LIBXC
  &END XC_FUNCTIONAL

  !+ donor_states, grid size, etc.
&END XAS_TDP
```

Assuming we simulate both Aluminiums of $Al_2(CH_3)_6$ on **4** processors:

- With NPROCS_GRID 2, grids are split over 2 processors, integration is simultaneous for both atoms

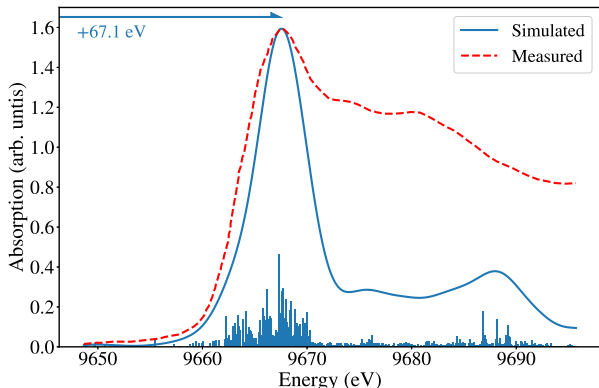- With NPROCS_GRID 4, grids are split over 4 processors and atoms are treated in a row

The denser the grid, the greater the benefits of increasing NPROCS_GRID

Showcase: Zn-MOF-5

- $(C_{48}H_{24}O_{26}Zn_8)_4$

- 424 atoms in the unit cell

- 25.8 Å$\times$ 25.8 Å$\times$ 25.8 Å

- Periodic boundary conditions
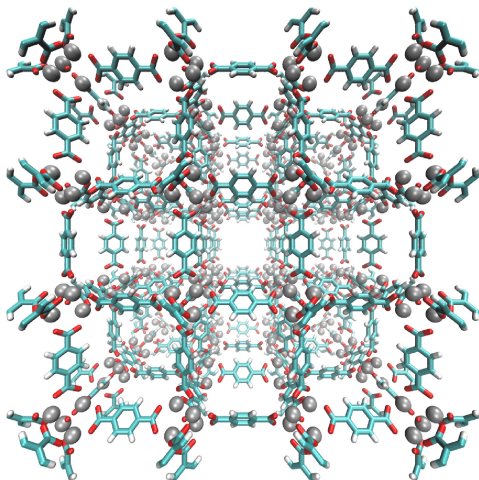
Showcase: Zn-MOF-5[6]

- BHandHLYP hybrid functional
- TZP-DKH basis set on 8 Zn from which excitations occur

- DZVP-MOLOPT-SR-GTH + pseudo potential
- DKH2 relativistic corrections

[6]J. Hafizovic *et. al.*, J. Am. Chem. Soc, **129**, 3612-3620 (2007)

Summary

Remember that:

- LR-TDDFT is a perturbative treatment of a ground state DFT calculation

- No good will come out of a XAS_TDP calculation if the ground state is not well converged

- The quality of the calculation can be controlled by using finer integration grids and/or better orbital and RI basis sets

- You can exploit symmetry and parallelism for faster calculations

- Hybrid functionals are usually the way to go in LR-TDDFT XAS simulations

- Absolute energies are usually not predictive (shift is needed), but spectral features are well reproduced

- Only the first few absorption peaks of a LR-TDDFT generated spectrum are physically meaningful

Questions