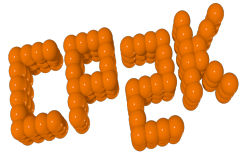
Running CP2K Calculations

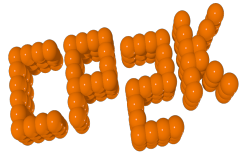
BASICS

- How to run CP2K
- CP2K Input File
  - Basic Structure
  - Method and System
  - Simulation Protocol
- Basis Sets and Pseudopotentials
- CP2K Output
  - Controlling the Output
  - Overview of Output Structure
- Restarting Calculations



How to run CP2K Calculations

- CP2K binaries
  - `sopt` – serial
  - `ssmp` – single process + multiprocessor (OpenMP)
  - `popt` – parallel (MPI)
  - `psmp` – parallel (MPI) + multiprocessor (OpenMP)
- Available from `http://www.cp2k.org/download`
  - Linux binaries (released versions)
  - Linux package managers, see also MAC and Windows versions, Docker images
  - Source code (all versions, current developer version) from GITHUB
  - materialscloud.org -> WORK -> Quantum Mobile
  - Preinstalled at your computer center



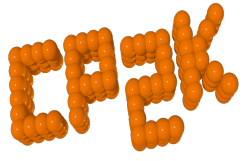
How to run CP2K Calculations

- Basic command line options:
  - `cp2k.sopt -i input_file -o output_file`

    By default, output goes to standard output

    Output to file appends (!!)

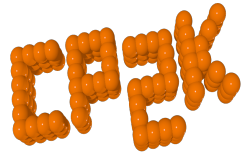
    Input file is the last argument if not otherwise specified
- Other useful options:
  - `cp2k.sopt -version`
  - `cp2k.sopt -check input_file`
  - `cp2k.sopt -html-manual`
  - `cp2k.sopt -help`



How to run CP2K Calculations

Typical files associated with a CP2K run:

- Input (required): e.g. `H2O-32.inp` (main input file, name and extension are arbitrary)
- Optional inputs: (default files are in `cp2k/data`
  - `POTENTIAL` (pseudopotential library)
  - `BASIS_SET` (basis set library)
  - Structure file (e.g. psf, xyz, crd ...)
- Outputs:
  - `PROJECT-1.restart` (input file to restart calculation)
  - `PROJECT-pos-1.xyz` (trajectory for MD or GEO_OPT)
  - `PROJECT-1.ener` (MD energies, temperature, cons. Q ...)
  - `PROJECT-1.cell` (cell parameters)
  - `PROJECT-RESTART.wfn` (orbitals for restart)



CP2K Input File: Basics

- Full documentation available online:

  `http://manual.cp2k.org`

  Or generate with `-html-manual`
- Sections - 13 (optional) top level sections

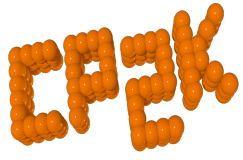
```
&BEGIN section_name [params]
...
&END [section_name]
```

- Keywords

```
KEYWORD value
KEYWORD [ON|OFF] [YES|NO] [TRUE|FALSE] ...
KEYWORD
```

- Nesting

  Sections may contain others sections and keywords

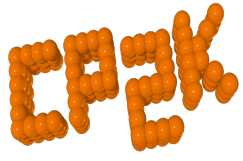


CP2K Input File: Basics

- Basic pre-processing syntax

`@INCLUDE 'filename'`	copy in text from file
`@SET VAR value`	define a variable
`@VAR`	replaced with variable value
`@IF / @ENDIF`	simple logic
or `#`	comments

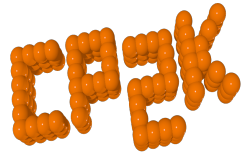
- Units
  - Numerical entries have a default unit (see manual)
  - Specify other units by hand e.g.
    `ABC [nm] 100 100 100` (or `bohr`, default is `angstrom`)
    `EMAX_SPLINE [eV] 50` (or `Ry`, default is `hartree`)
  - Also combinations e.g. `[hartree*bohr^2]`



CP2K Input File: Basics

The GLOBAL section (required)

```
&GLOBAL
    PROJECT H2O-32
    RUN_TYPE MD
    PRINT_LEVEL HIGH
    &TIMINGS
       THRESHOLD 0.000001
    &END
    WALLTIME 3600
&END GLOBAL
```

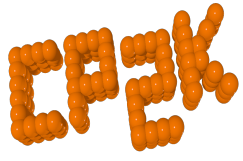


CP2K Input File: Method and System

The `FORCE_EVAL` section (required)

```
&FORCE_EVAL
      METHOD QS (or FIST, QMMM ...)
      &DFT
            ...
      &END DFT
      &SUBSYS
            ...
      &END SUBSYS
&END FORCE_EVAL
```

Definition of Simulation Method and System
(atomic coordinates and cell size)

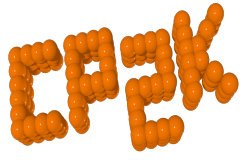


CP2K Input File: Simulation Protocol

The `MOTION` section

```
&MOTION
   &MD
      ENSEMBLE NVE
      STEPS 10
      TIMESTEP 0.5
      TEMPERATURE 300.0
   &END MD
&END MOTION
```

Used to control MD, Geometry Optimisation, NEB, Monte Carlo



Basis Sets and Pseudopotential Libraries

- CP2K uses seperable dual-space pseudopotentials

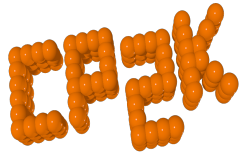
  Several sets of PPs and corresponding optimised basis sets are available, See `cp2k/data` or online at GITHUB
- `POTENTIAL, GTH_POTENTIALS`

  Wide range of PPs for many elements
  Optimised with different XC functional: LDA (PADE), PBE, BLYP ...
- `BASIS_SET, GTH_BASIS_SET, BASIS_MOLOPT`

  Contracted Gaussian basis sets; Various qualities / size of basis
  Make sure Basis and PP match (number of electrons and functional)
  Some documentation and references at head of each file

Controlling the Output

- The `PRINT_LEVEL` keyword in `&GLOBAL`
  - `SILENT, LOW, MEDIUM` (default), `HIGH, DEBUG`
  - `HIGH` can give more information if you are interested, also gives some per-process logging in parallel jobs
  - For long MD runs (e.g. classical), recommend using `LOW`
- Fine grained control is available via print-keys
  - Most input sections contain a `&PRINT` sub-section
  - Each `&PRINT` sub-section has further subsections for each quantity that may be printed



Controlling the Output

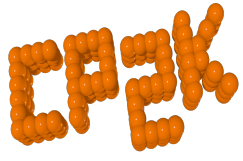
- For example, the `&PRINT` section in `&MOTION` contains

```
&CELL
&FORCES
&TRAJECTORY
&VELOCITIES
...
```

- Each section has parameters (and defaults) for which print level it is output

`&TRAJECTORY` defaults to `LOW`

`&VELOCITIES` defaults to `HIGH`

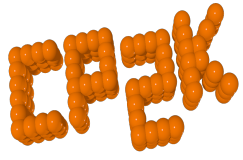


Controlling the Output

- Can also specify frequency of printing via `&EACH` sub-section, e.g.

```
&PRINT
  &CELL
    &EACH
      MD 100
    &END EACH
  &END CELL
&END PRINT
```

- Control over filenames, file formats etc. in each `&PRINT` section



Restarting a Calculation

- If you need to restart your job
  - Hardware failure
  - Batch system time limit
  - Failed to converge
  - Need more MD sampling
  - ...
- CP2K dumps a restart input file which can be directly re-run
  - `cp2k.sopt -i PROJECT-1.restart`
  - Continuous numbering of MD steps
  - Stores all state variables (incl. extended system)
  - Use `SCF_GUESS RESTART`