2FAST2CAR - User Manual

Oscar Gayraud, Hugo Geindre, Frédéric Guégan

December 12, 2024

1 Foreword and new stuff

2FAST2CAR is the second version of the program FASTCAR (see *Phys. Chem. Chem. Phys.* **2024**, *26*, 25780-25787). In this, we added some functionalities:

- running by default in iterative mode: one a first run is complete (CREST search, geometries pruning, DFT calculations and second stage pruning), a new search is initiated starting from the lowest energy conformer at the DFT level;
- post-SCF analyses: user may require the calculation of conceptual DFT indices (finite difference approximation), as well as the production of wfn/wfx files for further jobs (QTAIM, ELF...);

and some further things are underway (including and interface with ORCA 6).

2 Dependencies

FASTCAR is currently designed to operate under Python 3 (tested against Python 3.8.18). It relies on the use of the following dependencies:

- CREST;
- SpyRMSD;
- python libraries: `sys`, `re`, `os`, `shutil`, `math`, `numpy`, `unicodedata`

The present implementation is designed to manage Gaussian 16 calculations under Slurm (tested against Slurm 23.11.4).

3 Typical usage

3.1 Before the first run.

Users need to execute the `configure.py` script to update the path to the fastcar directory in all files. They also need to update the file `config.txt` with the pathes where the quantum chemistry softwares that will be used are located on their cluster. Users can also add the program folder to their path to be able to easily call the starting script (`sparkplug.py`).

3.2 Running a fastcar calculation.

In the desired folder, containing the output of a Gaussian calculation (at least frequency), the sequence is initiated by calling `sparkplug.py`. This script requires as argument the name of the output file, and uses a second file called "parameters.txt" containing all the run parameters. Its content is divided into 4 sections, discussed hereafter.

3.2.1 FASTCAR parameters.

In this section, the user should indicate the maximum number of iterations of FASTCAR.

3.2.2 CREST parameters.

In this section 7 elements of information are required:

- `CREST version`: this keyword indicates the crest version to be used. User may either provide a specific version (*e.g.* 2.12) or `default` (default version in the python environment of the user).
- `CREST solvent`: indicates whether a solvent model (using the analytical linearized Poisson-Boltzmann model) should be used during the CREST search.[1] The user should here provide the appropriate xtb keyword, for instance

  ```
  CREST solvent --alpb THF
  ```

  or indicate `none` if no solvent is to be used.
- `NCI`: indicates whether the non-covalent biased search should be used (argument: `selected` or `selected with scale factor`) or not (argument: `none`).
- `EWIN`: energy window for the conformers selection, in kcal/mol. By default this value is 6 kcal/mol.
- constraints: constraints on bonds, angles and dihedrals can be enforced. Each constraint should be written on a separate line starting by the appropriate label ("Bond constrained", "Angle constrained", "Dihedral angle constrained"[2]) followed by the number of atoms involved in the constraint.[3]
- `Force constant`: provides the value of the force constant used in the constraint;
- `Active atoms`: In case of a transition state, specifies which atoms are moving the most in the vibration mode associated with the imaginary frequency. FASTCAR will check if these atoms are among the most active in the transition states optimised during the run to decide if they describe the right transition state

3.2.3 RMSD parameters.

In this section 1 element is required:

- `RMSD threshold`: the RMSD threshold for the rejection of duplicate geometries, provided after the keyword.

[1] https://xtb-docs.readthedocs.io/en/latest/gbsa.html
[2] https://gist.github.com/hypnopump/30d6bfdb8358d3a57d010c9a501fda56
[3] https://crest-lab.github.io/crest-docs/page/examples/example_3.html#{}sampling-of-non-covalent-complexes-and-aggregates-

3.2.4 DFT parameters.

In this section 6 elements are required:

- `functional`: this keyword is used to provide the functional for the computation.
- basis set: the basis set to use (single keyword only). For single basis set calculations, the information is passed through the `Unique base` keyword. For 2-stage calculations, the first basis set is given by the `First base` keyword, and the second by the `Unique base`.
- dispersion model: the dispersion model to be used is provided through the `Dispersion` keyword. User should provide the desired keyword in G16/orca syntax. If no dispersion model is required, the user should indicate "None" (case independent).
- solvent: the solvent model to use it provided through the `DFT solvent` keyword. Similarly, user should provide the full keyword (*e.g.* `solvent scrf=(solvent=methanol)`).
- `additional calculations`: The following additional calculations can be carried out on the lowest energy structure: irc (IRC forward and reverse calculation, only for TS), cdft (compute several conceptual dft descriptors), nbo (Natural Bond Orbital analysis with NBO7)
- exclusion of nodes: if some computing nodes should be excluded for the computation (to avoid overfilling), it should be indicated in the `Node(s) excluded` keyword. The argument of this keyword can be "None" (if no exclusion) or any valid argument for the slurm sbatch option `exclude`: a single node identifier, a list or a range.

Hereafter is reproduced an example of a FASTCAR parameters file.

```
------------------------------------------------
----------------FASTCAR parameters--------------
------------------------------------------------
Check 5

------------------------------------------------
-----------------CREST parameters---------------
------------------------------------------------
CREST version default
CREST solvent --alpb THF
EWIN 4
NCI none
Bond constrained 1 10
Bond constrained 1 38
Force constant 1
Active atoms 1

------------------------------------------------
-----------------RMSD parameters----------------
------------------------------------------------
RMSD threshold 0.3

------------------------------------------------
-----------------DFT parameters-----------------
------------------------------------------------
Functional  b3lyp
Unique base  6-311++G(d,p)
Dispersion  empiricaldispersion=gd3
DFT solvent  SCRF=(Solvent=THF)
Additional calculation none
Excluding nodes 01-08
```