CASTEP convergence automation tool 1.0 User Guide

Simone Sturniolo

October 30, 2013

1 Introduction

The CASTEP convergence automation tool (CASTEPconv) is a Python script designed to automate the process of calculating the convergence of energy and forces in DFT calculations with CASTEP, as a function of the cutoff energy and the number of kpoints used. It works with Python 2.6 or higher. Given a .cell file, the script is able to create a set of folders containing input files for the required simulations, run the jobs, and process the output to produce tabulated ASCII data files and graphs. A .param file can be introduced as well to define any additional options for the DFT calculation, while a .conv file can be used to provide further options for the automated convergence calculation itself.

2 Installation

To install the script on Linux or CygWin just open a terminal, navigate to the folder containing the files, and type:

```
user@machine:~$ sudo python setup.py install
```

Enter your password as prompted and the script will be installed and available as a command.

3 Usage

CASTEPconv requires only a .cell file to be present in the folder when it's used. Considering a file named "<*seedname*>.cell", where <*seedname*>represents the name of the job, the syntax to run the convergence job is simply

```
user@machine:~$ castepconv.py <seedname>
```

while in the folder where the file is kept. If one wants the convergence DFT simulations to have additional parameters (e.g. redefine the convergence criteria for SCF iterations, add dispersion correction etc.), a <*seedname*>.param file can be added in the same folder with these options. To control the convergence job itself, instead, a new file, <*seedname*>.conv, is needed.

4 Convergence parameters

The syntax of the .conv file is similar to the one of the CASTEP .param file:

```
<parameter name 1>:    <parameter value 1>
<parameter name 2>:    <parameter value 2>
...
```

The accepted parameter names are listed in the subsections below.

4.1 String parameters

convergence_task: describes the task that is required from the convergence script. This can be INPUT (creation of input files and folders), INPUTRUN (same as INPUT, plus actually runs the jobs), OUTPUT (processes some already finished calculations and creates output files) and ALL (does all of the previous things, waiting for the jobs to finish before processing the output). Default is INPUT.
Legal values: INPUT, INPUTRUN, OUTPUT, ALL

running_mode: describes the mode in which the calculations should be ran. If PARALLEL, all calculations will be launched at the same time (ideal for job submission on a cluster). If SERIAL, the program will wait for one calculation to finish before the next one begins (and reuse the .check file from the previous calculation as a starting point). In the latter case, all files will be created in a single folder. Default is PARALLEL.
Legal values: PARALLEL, SERIAL

output_type: plotting program for which an output script should be created. For now only GNUPLOT is supported. Default is GNUPLOT.
Legal values: GNUPLOT

running_command: command line that should be used to run jobs. This should be expressed by replacing the name of the job with the generic token *<seedname>*. Default is *castep <seedname>-dryrun*.
Legal values: any string containing the token *<seedname>*

4.2 Float parameters

cutoff_min: Minimum value for the cutoff range explored in eV. Default is 400.0 eV.
Legal values: Any positive float.

cutoff_max: Maximum value for the cutoff range explored in eV. Default is 800.0 eV.

Legal values: Any positive float greater than cutoff_min.

cutoff_step: Step between the values of the cutoff range explored in eV. Default is 100.0 eV.
Legal values: Any positive float.

displace_atoms: Displacement in Angstroms to introduce in atom positions - necessary when the cell is equilibrated and it is not possible to converge forces because they are zero. Default is 0.0 Ang.
Legal values: Any float

final_energy_delta: Tolerance on final energy for the estimate of convergence. The system will be deemed converged in cutoff/kpoints when increasing the respective value won't cause a change in final energy greater than this value (normalized per atom). Default is 0.01 eV/atom.
Legal values: Any positive float

forces_delta: Tolerance on forces for the estimate of convergence. See above. Default is 0.001 eV/Ang.
Legal values: Any positive float

stresses_delta: Tolerance on stresses for the estimate of convergence. See above. Default is 0.01 GPa.
Legal values: Any positive float

4.3 Integer parameters

kpoint_n_min: Minimum value for the k-point range explored. This applies to the shortest side of the kpoint_mp_grid: depending on the size of the other cell parameters, there might be proportionally more k-points along other sides. Default is 1.
Legal values: Any positive integer.

kpoint_n_max: Maximum value for the k-point range explored. Default is 4.
Legal values: Any positive integer greater than kpoint_n_min.

kpoint_n_step: Step between the values of the k-point range explored. Default is 1.
Legal values: Any positive integer.

4.4 Boolean parameters

converge_stress: Apply calculation of stresses to the simulations and then estimate convergence on stresses as well as energy and forces.
Legal values: Anything. The word "true", regardless of the case, means the stresses are calculated. Anything else will be interpreted as False.

5 Output

When the calculations are over, CASTEPconv will have produced the following files:

<*seedname*>.conv_tab: this file is created during the INPUT phase of the run (when folders and input files for the calculations are created) and sums up the values used for cutoff and the kpoint grids employed in the various files. This is just useful as a memo, and mostly used to resume analysis of previously ran calculations (by using the OUTPUT convergence_task) - if it's not present, the ranges will be recalculated from the .conv file.

<*seedname*>_cut_conv.dat: generated in the OUTPUT phase of the run, will contain a tabulated ASCII of cutoff values (in eV), final energies (in eV) and total forces (in eV/Ang) for the various calculations ran. The total forces will be calculated as:

$$F = \sum_{i}^{atoms} |\mathbf{f}_i| \tag{1}$$

where $\mathbf{f}_i$ is the force applied on the atom i.

<*seedname*>_kpn_conv.dat: generated in the OUTPUT phase of the run, it is similar to the above, except that it expresses its quantities as a function of the total number of kpoints in the grid.

<*seedname*>_conv.<variable extension>: this file is the script meant for generation of graphic output. By default it will be a Gnuplot script (extension .gp). When other forms of output will be supported, choosing the appropriate value for the output_type option will replace it with a different format.