Scripts for XR postprocessing with USPEX

Michele Galasso

August 7, 2019

1 split_CIFs.py

The script takes the following arguments from command line:

1. the output file `Individuals` from USPEX
2. the output file `gatheredPOSCARS` from USPEX
3. the value of the external pressure used for the USPEX run
4. a boolean parameter `remove_hydrogens=YES/NO`

The script performs the following operations:

1. for each structure, it reads the parameters *enthalpy* and *fitness* from `Individuals` and the geometry from `gatheredPOSCARS`
2. if `remove_hydrogens=YES`, it deletes all hydrogen atoms
3. it selects, for each reduced formula, the 5 structures with lowest enthalpy
4. it outputs the selected structures as CIF files in a new folder *results*

The name of each CIF file has the format `i_fitness_enthalpy_reducedformula_pressure_symmetry.cif`, where:

i is a natural number which orders the output with increasing *fitness*

fitness is the *fitness* of the structure

enthalpy is the *enthalpy* of the structure

reducedformula is the *reduced formula* of the structure, with hydrogens always shown

pressure is the pressure used for the USPEX run

symmetry is the space group number, determined with a tolerance of 0.2 and without hydrogens if `remove_hydrogens=YES`

Example: `python3 split_CIFs.py Individuals gatheredPOSCARS 50GPa remove_hydrogens=NO`

2 exclusion.py

The script takes the following arguments from command line:

1. the wavelength of the incident radiation in Å
2. the cut-off in % of the maximum intensity
3. a number of exclusion regions in degrees, expressed as two angles separated by a hyphen (-)

The script works in a folder with many CIF files, and performs the following:

1. it opens, one by one, all CIF files and it predicts the XRD pattern of the structure according to the given wavelength
2. if the predicted pattern contains any peak in the exclusion regions that is bigger than the given cut-off, it deletes the CIF file

Example: `python3 exclusion.py 0.6199 25 25-28 31-32`

3 fix_comp.py

The script performs postprocessing for USPEX calculations with *fixed composition*. It takes the following arguments from command line:

1. the output file `Individuals` from USPEX
2. the output file `gatheredPOSCARS` from USPEX
3. the value of the external pressure used for the USPEX run

The script performs the following operations:

1. for each structure, it reads the parameter *enthalpy* from `Individuals` and the geometry from `gatheredPOSCARS`
2. it computes `real_fitness` = `enthalpy` / `total_number_of_atoms`
3. it outputs the structures as CIF files in a new folder *results*

The name of each CIF file has the format `i_realfitness_iupacformula_symmetry_ID_pressure.cif`, where:

i is a natural number which orders the output with increasing `real_fitness`

realfitness is the `real_fitness` of the structure

iupacformula is the *IUPAC formula* of the structure

symmetry is the space group number, determined with a tolerance of 0.2

ID is the *structure ID* read from the input files

pressure is the pressure used for the USPEX run

Example: `python3 fix_comp.py Individuals gatheredPOSCARS 50GPa`