**SAPA Utilities**

The sapa\_utils.py contains some basic utility scripts for analysing the output of SAPA. To import, it must be located in the same directory as the SAPA output files you want to analyse.

* Import the script – e.g. import sapa\_utils as su
* Set up a class instance using an isodistort cif – e.g. a = su.sapa\_utils(“scf3\_iso.cif”) – this is done first so the irreps can be parsed from the cif file
* Import the SAPA output – a.import\_files(“sample”,temps) – sample is the sample designation given when writing the input file (i.e the string between irrep\_ and \_temp in the SAPA output files) and temps is a list of temperatures (or other variables) at which SAPA has been executed

Once the above has been done, the class instance stores the output files as pandas dataframes. To access them, use a.irrep[“temp”]. It is important to note here that due to restrictions in python, if the irrep label ends in a +/-, this gets changed to a p/m. e.g, to access the 350K X5+ output file, use a.X5p[“350”]. In addition, to access the temperatures, they must be strings. Another thing to be aware of is that the *import\_files* function discards the first refinement (it is often performed at the extremes of parameter ranges) and orders the refinements from lowest Rwp to highest.

In addition, the *import\_files* function does some basic processing of the data. For each refined parameter, it records the value of that parameter for each irrep at the best refinement for each temperature. This is stored for each irrep, so to view the values for e.g. the X5+ irrep, you would use a.X5p[“prmname”], where prmname is the name of the refined parameter. It also stores the best Rwp for each irrep at each temperature (a.irrep[“Rwp”]) and the difference between the best Rwp and the Rwp of a refinement with no modes active at that temperature (a.irrep[“delrwp”]). These can all be plotted using the *plot\_all()* function. It is recommended to look at some of these plots, in particular the scale plot, to ensure consistency between irreps.

The Boltzmann-Weighted Mode amplitudes can also be calculated. There are two functions that do this – one, *calc\_bwmas(sigma)* uses the global minimum Rwp to weight the mode amplitudes. This is ok if the Rwp vs temperature plot is fairly flat, however it can give misleading results if the Rwp changes with temperature. If this is the case, then use *calc\_dwmas(sigma)*, which uses a weighting based off the difference between the delrwp value and the best delrwp value for each temperature. The mode amplitudes are stored as a.irrep[“bwmma”] for the weighted mean mode amplitude, and a.irrep[“bwma”] for the weighted mean before dividing through by the sum of weights (for consistency with Senn *et al*, Phys. Rev. Lett., 2016). Once this function has been ran, the amplitudes can also be plotted with the plot\_all(bwmas=True) function, using the optional argument. NB: setting bwmas=True is necessary once calc\_bwmas or calc\_dwmas has been ran.

There are also some other utilities. These are *make\_histograms()*, which plots and saves histograms of the different rwps for each irrep and temperature (NB – this generates a lot of extra files) and *generate\_cif(irrep,temp),* which generates a cif for a given irrep and temp (NB – there is a function in ciftoiso, *findsym,* that can save a symmetrised version of the generated cif).