

## Discovering features in Sr14Cu24O41

Neutron single crystal diffraction is a powerful way to examine the atomic structure of technologically interesting materials. The particular material represented in this data is  $\text{Sr}_{14}\text{Cu}_{24}\text{O}_{41}$ <sup>i,ii</sup>, a semiconducting solid with industrial application potential for optoelectronic and thermoelectric devices.  $\text{Sr}_{14}\text{Cu}_{24}\text{O}_{41}$  has two intertwined structures. One is a  $\text{Cu}_2\text{O}_3$  two-leg ladders, the other is  $\text{CuO}_2$  chains. The ladder and chain structures are incommensurate. That means there is not a simple ratio of the number of chain links to the number of ladder rungs.

Neutron diffraction is highly sensitive to the internal structure or atomic arrangement of the material to be studied, which is built from a periodic arrangement of closely connected atoms as structural units, repeating on a 3D translational grid in  $(h, k, l)$  units.

This data set has 4 dimensions: Intensity ( $I$ ), and three spatial dimensions  $(h, k, l)$  which describe three directions within the material. If there is a large, sharp intensity,  $I$ , at a given  $(h, k, l)$  coordinate, that means there are many pairs of atoms spaced at that characteristic length scale. If there is a broad distribution of atom spacing broader features associated with disorder in the system appear.

The data itself has already been corrected for experimental effects, normalized, and put onto a regular three-dimensional grid in  $(h, k, l)$ . However, not all of the regular grid has been measured. At these points the intensity ( $I$ ) has been set to not-a-number (NaN). The data is in an HDF5 file.

H5py is a useful set of libraries for accessing HDF5 files from python.

If you are not familiar with hdf5 and want to browse the data structure want to browse the data structure we recommend [hdfview](#) or [nexpy](#). However, the developed algorithms should run independently of these codes.

Do these:

1. Segment the data – There are features that vary in sharpness and intensity as well as background. Separate these into various categories
2. There are multiple sets of sharp features (known as Bragg peaks) that are 3D arrays. Pick out these arrays and give the minimum distance between peaks for each array in each direction.

Do at least two of the following:

3. Characterize the sharp features and relationships between them
  - a. Identify regular intensity variations, directionality, widths etc.
  - b. Identify how many collections of these features are in the data
4. Characterize the broad features and relationships between them
  - a. Identify regular intensity variations, directionality, etc.

- b. Identify how many collections of these features are in the data
5. Are there relationships between the sharp and broad features
6. Visualize the above results

Our preference is for the algorithms to be implemented in Python with use of the numpy library, other necessary libraries are welcome. If the participant prefers a compiled code, C/C++, is our preferred solution with the output being in a python or numpy data structure.

This data set is moderately sized. Larger data sets are possible so a design with an eye toward scalability is desired.

---

<sup>i</sup> J. Etrillarda, M. Braden, A. Gukasov, U. Ammerahl, A. Revcolevschi; *Physica C* 403 (2004) 290–296; DOI: 10.1016/j.physc.2004.01.003

<sup>ii</sup> X. Chen, D. Bansal, S. Sullivan, D.L. Abernathy, A.A. Aczel, J. Zhou, O. Delaire, L. Shi; *Phys. Rev. B* 94, 134309 (2016); DOI:<https://doi.org/10.1103/PhysRevB.94.134309>