

WPPF Fitting of Multiple Data Sets: Sequential and Parametric Refinement

Prof. John S.O. Evans, Department of Chemistry, Durham University, Lower Mount Joy, South Rd., Durham, DH1 3LE, UK.

1. Introduction/overview

Previous lectures have shown how huge numbers of powder patterns can be collected on samples as a function of time, temperature, pressure or other variables (*in situ* studies), or on a system such as a battery or catalytic bed under working conditions (*operando* studies). Sources such as XFELs can provide thousands of patterns per second and, using pump-probe methods, at time resolution down to a few 100 femtoseconds. Later lectures cover topics such as XRD-CT which again produce enormous bodies of powder data. These notes contain some tips on automating the independent analysis of multiple patterns, which I'll call *sequential refinement*. They also outline the method of *parametric refinement*, where a large body of data is simultaneously fitted using a single evolving structural model. This technique can allow extraction of more reliable structural information, or the direct extraction of non-crystallographic information that one wouldn't expect to derive from a crystallographic analysis.

2. Sequential Rietveld refinement

Sequential Rietveld refinement, or the serial independent analysis of multiple patterns, just applies the fitting methods outlined in the chapter on Whole Powder Pattern Fitting multiple times. Many Rietveld packages allow the user to automate this by setting up an initial Rietveld refinement, then using the same model to analyse a series of similar patterns. Some packages do this via the gui; others work via a set of software-specific instructions in an input file; others can be controlled via a series of commands in *.bat* or *.sh* files of the computer operating system, or through Python scripts.

Some tips and tricks to control sequential refinements are given below. They are written in terms of a variable temperature (T) experiment, but apply to most other situations:

1. Refinements can be performed by using the results of one refinement (e.g. at T_1) as the starting model for refinement T_2 , or by always starting from the same model. The former approach means the starting model updates through the analysis, which can help convergence. The latter approach can be helpful if parameters diverge at any point.
2. The refinement should ideally be configured such that the same starting model will converge for every pattern recorded. If this can be achieved, either approach in (1) should converge to the same answer.
3. Individual phases can appear or disappear during an experiment. Some software lets you flag the data sets where you want specific phases included in the model.
4. Refinements are likely to diverge when a phase is included in the model, but not observed in the data. For example, peak shape parameters describing a missing phase may refine to give "infinitely broad" peaks which merely fit the background. This slows the refinement, makes quantitative analysis of all phases incorrect, and can prevent that phase from contributing to the fit when its peaks appear. Consider tricks such as: limiting parameters to sensible values; restricting the parameters refined; equating parameters between phases; using restraints on derived quantities such as overall peak width. Consider parametric methods.
5. Consider setting minimum and maximum limits on parameters such that convergence always occurs. These limits can be made temperature-dependent.
6. Consider performing a few cycles of simulated annealing on each pattern analysed in which parameters are reset to values believed to be close-to-sensible after each convergence.
7. When cell parameters change substantially (e.g. at a discontinuous phase transition) it can be helpful to start refinements with a broad peak shape. This allows some overlap

between observed and calculated profiles and helps find the global minimum. Consider restricting the q -range in initial refinement cycles to get an initial estimate of cell parameters from data regions where peak overlap is lowest.

8. Consider using non-standard restraints. For example, if a sample undergoes a 1st order transition from orthorhombic to monoclinic with $\beta \approx 91^\circ$, automated fitting of a two phase model close to the transition can be hard as the lower-symmetry monoclinic phase can distort to fit the peaks of the orthorhombic phase. A soft restraint on the monoclinic angle of the form $(91 - \beta)^2$ can prevent the monoclinic angle reaching 90° and stabilise the refinement.
9. For data collected sequentially, try fitting the patterns from the start-to-end of the experiment and also end-to-start. If any aspect of the models differ, it suggests refinements aren't correctly converged.
10. Plot all refined quantities to look for unexpected changes in parameters or agreement factors through the experiment. Do this in an automated way (e.g. using Python scripts) so there's no activation barrier to reanalysing everything with a different model. A good plotting package will let you automatically colour-code data points to highlight relationships with other parameters (e.g. parameter value colour-coded by R_{wp}) or shade the background to highlight different experimental conditions. Packages like `seaborn` in python can produce remarkably informative plots with just a couple of commands (see the images at <https://seaborn.pydata.org/tutorial/introduction.html>); see also Figure 1.
11. Ideally feed metadata from the experiment through the Rietveld pipeline. This will help with automatic plotting, and lets you control the refinement according to the experimental conditions. For example, cell parameters in each fit could be started at temperature-dependent values, or temperature-dependent limits could be placed on parameters.
12. Consider using internal standards during the experiment. The Rietveld analysis can then use peak intensities of the standard to help calibrate scale factor and weight % changes, or use the peak positions to calibrate sample temperature.
13. In quantitative work don't just plot the Rietveld-derived weight %s of each phase as these will be meaningless if a portion of the sample goes amorphous. Output scale factors or $scale \times ZMV$ values. The sum of $scale \times ZMV$ s across all phases should be constant throughout the experiment if everything remains crystalline, or just show a smooth evolution if scale factors correlate with, for example, atomic displacement parameters.
14. Quantities such as R_{wp} or χ^2 should have similar values for each pattern and vary smoothly if all data are recorded under comparable conditions. Any major fluctuations suggest the model is not appropriate for all patterns.
15. Check predictable parameters show the temperature dependence expected. For example, if atomic displacement parameters don't increase with temperature in a variable temperature experiment then either the model is incorrect or there's interesting physics or chemistry at play. If a cell parameter decreases rather than increases with temperature, then you've rediscovered negative thermal expansion. Unexpected "humps" in cell parameters or site occupancies can reveal interesting and important things about kinetically-controlled processes in your sample.¹⁻⁴

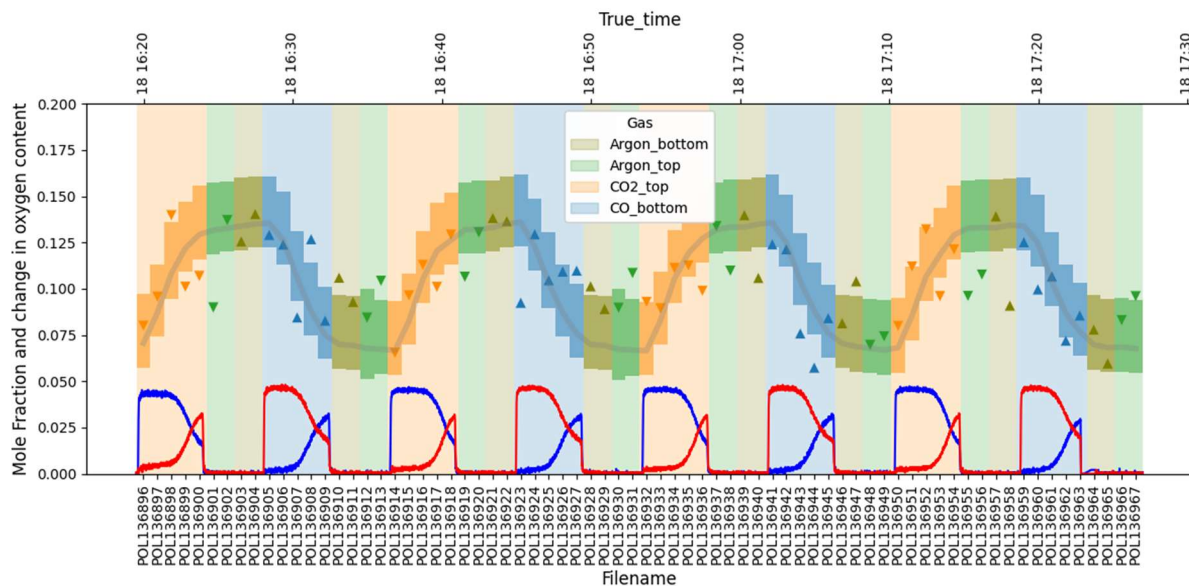


Figure 1: A useful information-rich plot from sequential analysis (unpublished). A few lines in a python script produce the plot above which was part of a study on a working catalyst bed under different conditions. Data points show oxygen site occupancies in the catalyst bed extracted by Rietveld refinement. The points are colour-coded according to the gas flowing over the bed. Arrows pointing up and down reflect the direction of gas flow through the bed. The grey line shows the unit-cell parameter changes that occur, and the shaded bars on that line shows the standard uncertainties of oxygen occupancy. The red and blue lines show the gases flowing off the reactor as analysed by mass spectrometry. The background shading highlights different gases used. During the operando neutron experiment thousands of diffraction patterns were analysed in real time and reaction progress followed using this type of plot.

3. Parametric refinement

Sequential or independent Rietveld refinement is normally the best method for analysing multiple powder patterns. However, there are some cases where analysing a collection of patterns with a single constrained but evolving model is the best or only way to obtain an unambiguous description of a changing system. This is called *parametric refinement*.⁵ More details are given elsewhere⁶ and there are practical examples on how to do this online (https://topas.webspace.durham.ac.uk/tutorial_surface_new).

The basic idea behind parametric refinement is that you know that some quantities influencing a powder pattern must vary in a smooth or predictable way through a series of data collections.^{5,7} For example, in a laboratory variable-temperature Bragg-Brentano experiment, the diffractometer zero-shift is unlikely to change during the experiment, whereas a quantity like the sample height will vary in a smooth way as the furnace heats and cools. If these quantities correlate with parameters of interest (such as unit-cell parameters) then it might be best to refine their value (a single parameter for the zero-shift) or a simple function describing their temperature dependence (for the height) from all the data.

The parametric approach is particularly useful for analysing complex data sets where phases appear and disappear without having to impose *a priori* assumptions on parameter values. This is exemplified in the WO_3 quantitative analysis example in Section 3.1. Finally, it is also possible to set up refinement models in which certain parameters are described by specific well-known physical models. This can lead to quantities such as sample temperature, kinetic rate constants or activation energies being directly refinable from powder diffraction experiments. Examples of quantities that might be constrained, and the type of functions that might be used are given in Table 1; these can be applied to a range of different in situ and operando experiments. Other applications of parametric fitting include the simultaneous fitting of patterns from multiple detectors as part of the data reduction on synchrotron beam lines.⁸

Table 1: Examples of functions used in parametric refinements.

Quantity	Expression	Comment
Cell parameters or adps	$a(T) = a_0 + \frac{c_1 \theta_1}{e^{(\theta_1/T)} - 1}$	Simple Einstein-like model with a_0 , c_1 and θ_1 as refinable parameters. Ensures a physically-sensible zero gradient at $T = 0$ K.
Fractional coordinates	$x(T) = x_0(1 + c_1 T + c_2 T^2 + c_3 T^3)$	Simple polynomial form.
Critical behaviour	$occ(T) = c_1 \left(1 - \frac{T}{T_c}\right)^\beta$	Site occupancy, magnetic moment or other quantity related to an order parameter approaching a phase transition.
Kinetic parameters	$\begin{aligned} frac(t) &= c_1(1 - e^{-k_{frac}t}) + c_2 \\ cell(t) &= c_1(1 - e^{-k_{cell}t}) + c_2 \end{aligned}$	Simple rate expression. $k(t - t_0)$ or $k(t - t_0)^n$ describe more complex evolution.
Zero shift	$zero(t) = const$	Zero shift correction unchanging over all data sets.
Sample height	$height(T) = h_0(1 + c_1 T + c_2 T^2)$	Sample height as a smooth function of temperature.
Temperature error	$\Delta T = c_0(1 + c_1 T_{set} + c_2 T_{set}^2)$	Offset between furnace set point and sample temperature as a smooth function.
Equation of state	$a(P) = a_0[(1 + K'_0 P/K_0)]^{-1/(3K'_0)}$	Pressure dependence of cell parameters e.g. Murnaghan EoS for a cubic material.
Energy dispersive profile	$I(E) = \frac{a}{E} \exp\left(-\frac{1}{2} \left\{ \frac{\ln[(E - b)/c]}{d} \right\}^2\right)$	Lognormal function to describe energy profile for energy-dispersive diffraction.

One criticism of parametric refinement is that imposing a constrained model across multiple diffraction patterns may be a good idea with an appropriate function, but may be a disaster with an inappropriate one. In fact, using an incorrect model can be an advantage of the method. If the R_{wp} of an individual pattern (or series of patterns) in a parametric fit is significantly worse than others (or worse than in a sequential refinement), it indicates that the model is missing an essential descriptor of the chemistry, physics or crystallography. This can easily be missed in sequential work where aspects of an unconstrained model can “distort” to “mop up” unfitted features in the data. We’ll see examples of this below.

3.1 WO_3 phase transitions – quantitative parametric refinement through phase transitions

WO_3 is a compound that undergoes a series of phase transitions on cooling that involve small distortions of the corner-sharing $WO_{6/2}$ octahedra, and their tilting relative to the cubic structural arisotype. A figure showing all the structural changes is given in the notes on symmetry refinement. Figure 2 shows a series of 100 diffraction patterns recorded on cooling WO_3 from 300 to 90 K. Over this temperature range the sample changes first from a monoclinic $P2_1/n$ structure to triclinic $P\bar{1}$, then to monoclinic Pc at the lowest temperature. As the approximate structures of all three phases are known, determining the amount of each present as a function of temperature should be straightforward using a three-phase Rietveld analysis. Unfortunately, the similarity of the three structures makes this difficult. For example, if we look at the zoomed 2θ range of 22–25 degrees in Figure 2c we see that all three structures predict hkl reflections of similar intensity in similar places. If we were to try and fit the room temperature data (where the sample is predominantly $P2_1/n$), it turns out that the lower symmetry $P\bar{1}$ structure can always give an equivalent or better fit (it has all the structural degrees of freedom of $P2_1/n$ plus some extra ones). Figure 3a shows the apparent weight % extracted from a series of sequential independent refinements in which 56 parameters were refined against each pattern. Each refinement was deliberately started from the correct (determined

parametrically) minimum for that temperature, and several cycles of simulated annealing were performed to ensure the best fit. The extracted weight %'s show a temperature evolution that doesn't make physical sense: weight %s increase and decrease abruptly with temperature. In addition, the unit-cell parameters of each individual phase (Figure 3b) depart significantly from the smooth evolution with temperature expected.

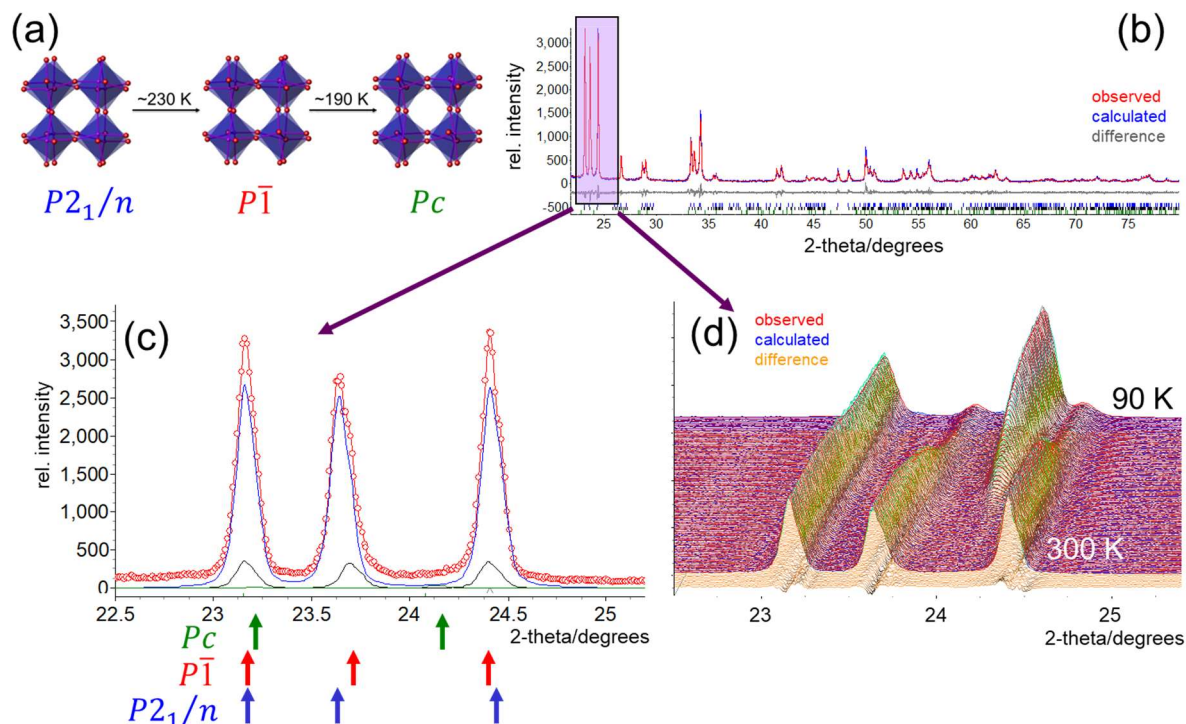


Figure 2: (a) Phase transitions of WO_3 . (b) room temperature X-ray Rietveld fit. (c) Zoom of fit; blue solid line is $P2_1/n$ and gray $P\bar{1}$. (d) Parametric fit of the same region across all 100 powder patterns.

Figure 3c and 3d demonstrate how these problems can be avoided in a parametric refinement. In this fit we make three straightforward assumptions: (1) the cell parameters of each phase vary smoothly with temperature (though we don't impose values); (2) fractional atomic coordinates of each phase vary smoothly with temperature; and (3) each phase has a temperature-independent peak shape. All 100 patterns can then be fitted in a single Rietveld refinement with 1167 parameters (far fewer than the 5600 parameters refined sequentially). The extracted weight %'s in Figure 3c then show a sensible temperature dependence. In essence what's happening is that the portions of the data where a phase is present, and its behaviour is therefore well-defined by the data, are controlling the overall model in regions where that phase is not present. For example, if the $P\bar{1}$ structure were to distort to artificially fit patterns close to room temperature, it would no longer be able to fit the ~ 200 K patterns where it is present. Similarly, we avoid problems such as peak shapes becoming infinitely broad and just fitting the background as they are controlled by regions of the data where the phase is really present. Note that we are imposing a simple physical behaviour on parameters of minor interest (the unit-cell parameters) to get better information on the parameter of interest (weight %). This is done without any assumption about the actual values of parameters, just that they follow a sensible temperature-dependence.

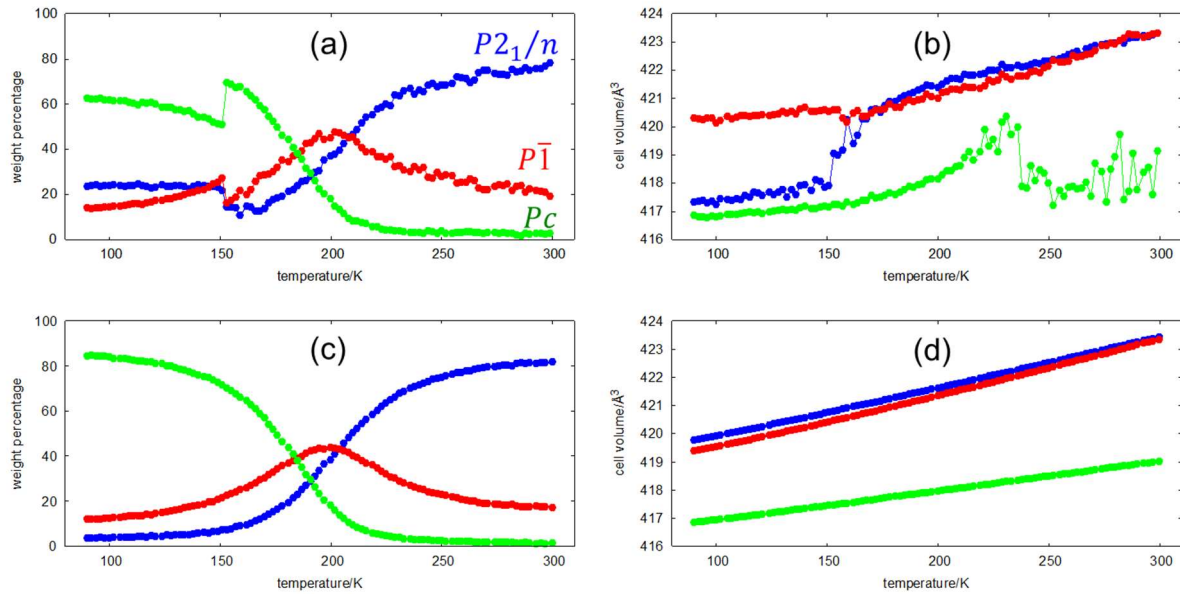


Figure 3: Evolution of weight % and unit cell volumes from (a), (b) sequential fitting and (c), (d) parametric fitting.

3.2 Non-crystallographic information from parametric refinement

The parametric approach can also be used to gain information on non-crystallographic parameters. One example is in determining the true temperature of a sample during an in-situ powder experiment where it might differ from the furnace set point (T_{set}) due to its position relative to the thermocouple, due to the presence of flowing gases, or due to heat evolved during a chemical process (Figure 4a). One trick in these cases is to use an internal standard whose thermal expansion has been measured by diffraction or dilatometric techniques (which can be very high precision), and to use its cell parameter to infer the true temperature. In a lab experiment things are still not straightforward as peak positions of both the standard and sample are influenced by temperature offsets from the furnace set point $\Delta T(T_{set})$, zero-shift errors and sample height errors. In a parametric refinement the cell parameter(s) of the standard can be described using published thermal expansion data of the form $a_0(1 + b(T_{set} + \Delta T) + c(T_{set} + \Delta T)^2)$ where a_0 is known approximately, b and c known coefficients and ΔT is described by a low order polynomial as in Table 1. The parameters of the ΔT polynomial, sample height, zero-shift and a_0 can be derived from all the data simultaneously to obtain a smooth temperature calibration curve $\Delta T(T_{set})$. It's also possible to use two internal standards, one with low expansion and one with high expansion, to derive the temperature from the difference in expansion, such that absolute knowledge of the standard cell parameters isn't required.

For materials undergoing a structural change as a function of time, quantities such as cell parameters or site occupancies can be described with kinetic expressions such as $c_1(1 - e^{-kt})$ and rate constants k can then be refined directly from the data. In experiments where material evolution as a function of time and temperature has been measured, the temperature dependence of the rate constant can be expressed with an Arrhenius-like expression $k = Ae^{-E_A/k_B T}$ (k_B = Boltzmann constant) and activation energies E_A can be extracted directly from a parametric refinement.

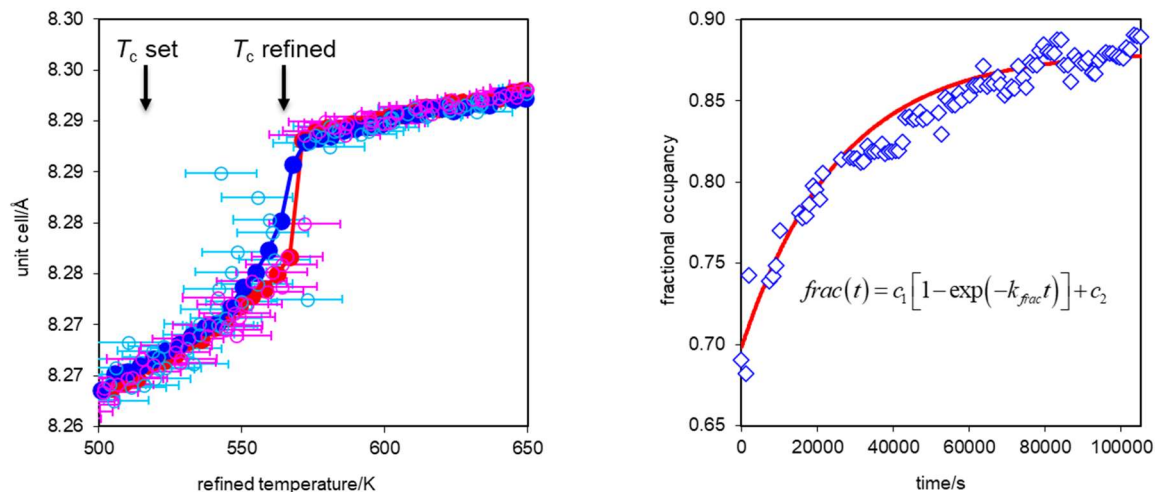


Figure 4: Left: unit cell parameters of ZrP_2O_7 close to a phase transition derived either sequentially (open points) or parametrically (closed points) on warming (red) and cooling (blue).⁹⁻¹⁰ A parametric temperature calibration gives significantly lower uncertainty in cell parameters. The uncorrected phase transition temperature, T_c , from the furnace set point was ~ 520 K whereas the parametrically refined T_c of 567–571 K agreed perfectly with DSC data. Right: a comparison between parametrically (red solid line) and sequentially (blue points) refined site occupancies for an order-disorder phase transition in $ZrWMoO_8$.² The parametric approach allows direct Rietveld refinement of the rate constant, k_{frac} .

4. Conclusions and what next

Information from multiple powder patterns can give enormous insights in many areas of materials science. In most cases information can be extracted using sequential Rietveld refinement, but there are cases where a parametric approach is essential. Software-specific methods can be explored in the workshops. Tutorials showing how this can be done in TOPAS are available online (https://topas.webspace.durham.ac.uk/topas_user_menu/).

5. Acknowledgements

Thanks to students and collaborators for the interesting materials challenges that led to the approaches described becoming core underpinning methods in our research. Thanks to Alan Coelho for the TOPAS developments that allowed them to be practically possible.

6. References

1. Allen, S.; Evans, J. S. O., Negative thermal expansion and oxygen disorder in cubic $ZrMo_2O_8$. *Physical Review B* **2003**, *68* (13), 134101 - 134103.
2. Allen, S.; Evans, J. S. O., The kinetics of low-temperature oxygen migration in $ZrWMoO_8$. *Journal of Materials Chemistry* **2004**, *14* (2), 151-156.
3. David, W. I. F.; Ibberson, R. M.; Matsuo, T., High-Resolution Neutron Powder Diffraction - a Case-Study of the Structure of C-60. *Proceedings of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences* **1993**, *442* (1914), 129-146.
4. Evans, J. S. O.; Evans, I. R., Beyond classical applications of powder diffraction. *Chem. Soc. Rev.* **2004**, *33* (8), 539-547.
5. Stinton, G. W.; Evans, J. S. O., Parametric Rietveld refinement. *Journal of Applied Crystallography* **2007**, *40*, 87-95.

6. Dinnebier, R. E.; Leineweber, A.; Evans, J. S. O., *Rietveld Refinement, Practical Powder Diffraction Pattern Analysis using TOPAS*. De Gruyter: 2018; p 331.
7. Berar, J.-F.; Baldinozzi, G., XND code: from X-ray laboratory data to incommensurately modulated phases. *IUCr-CPD Newsletter* **1998**, *20*, 3-5.
8. Fitch, A.; Dejoie, C.; Covacci, E.; Confalonieri, G.; Grendal, O.; Claustre, L.; Guillou, P.; Kieffer, J.; de Nolf, W.; Petitedemange, S.; Ruat, M.; Watier, Y., ID22 - the high-resolution powder-diffraction beamline at ESRF. *Journal of Synchrotron Radiation* **2023**, *30* (5), 1003-1012.
9. Birkedal, H.; Andersen, A. M. K.; Arakcheeva, A.; Chapuis, G.; Norby, P.; Pattison, P., The room-temperature superstructure of ZrP_2O_7 is orthorhombic: There are no unusual 180 degrees P-O-P bond angles. *Inorganic Chemistry* **2006**, *45* (11), 4346-4351.
10. Stinton, G. W.; Hampson, M. R.; Evans, J. S. O., The 136-atom structure of ZrP_2O_7 and HfP_2O_7 from powder diffraction data. *Inorganic Chemistry* **2006**, *45* (11), 4352-4358.