The superstructure determination of displacive distortions via symmetry-mode analysis

By

Sean C. Kerman

Physics Department

Brigham Young University

August 2011

ABSTRACT

For any crystal structure that can be viewed as a low-symmetry distortion of some higher-symmetry parent structure, one can represent the details of the distorted structure in terms of symmetry-adapted distortion modes of the parent structure rather than the traditional list of atomic *xyz* coordinates. Symmetry-mode analysis (SMA) often simplifies the refinement of a distorted structure because most symmetry modes tend to be inactive, while a relatively small number of mode amplitudes are dominant in producing the observed distortion. Here, we demonstrate that a symmetry-mode refinement against powder-diffraction data, when combined with a global-search algorithm, enables one to directly detect the space-group symmetry of the distorted phase, which includes both the space-group type and the locations of its symmetry elements relative to the parent lattice. This is an important capability when peak splittings are small, superlattice intensities are weak, or systematic absences fail to distinguish between candidate symmetries. Because the symmetry-detection process automatically reveals which of the modes belonging to the detected symmetry are active, the subsequent determination of the superstructure (i.e. the phasing of the superlattice peaks) is greatly simplified

TABLE OF CONTENTS

ABSTRACT ..... **ii**

LIST OF FIGURES ..... **v**

LIST OF TABLES ..... **vi**

BACKGROUND **1**

Crystallography 1

Symmetry 1

Diffraction 3

Powder diffraction 5

Neutron diffraction 5

Structure solution 6

Refinements 6

INTRODUCTION **7**

METHODS **13**

Rietveld refinement 13

Pawley refinement 14

ISODISTORT 14

TOPAS Academic/J-Edit 14

EXPERIMENTAL PROCEDURE 14

RESULTS AND DISCUSSION **17**

Symmetry-mode description of WO_3 17

WO_3 symmetry detection 19

WO_3 structure determination 26

$LaMnO_3$ 28

A procedure to eliminate phantom modes 31

Shortcuts 34

Signal-to-noise limitations 35

Randomization range 36

CONCLUSIONS 38

REFFERENCES 40

LIST OF FIGURES

Figure 1 1

Figure 2 3

Figure 3 4

Figure 4 20

Figure 5 21

Figure 6 23

Figure 7 24

Figure 8 25

Figure 9 28

Figure 10 30

Figure 11 30

Figure 12 33

Figure 13 36

Figure 14 37

LIST OF TABLES

Table 1 18

Table 2 29

BACKGROUND

Crystallography

Crystallography is a field of material science that deals with the study of the arrangement of atoms in a crystal (i.e. crystal structures). A crystal is a solid that can be defined relative to a periodic set of points called a lattice. Fig. 1 shows a 2-D representation of a crystal lattice. The repeating unit for the lattice is marked in the figure by a square. The square contains one atom (only count the part of the atom that is actually inside the box) and is called the *unit cell* of the crystal.

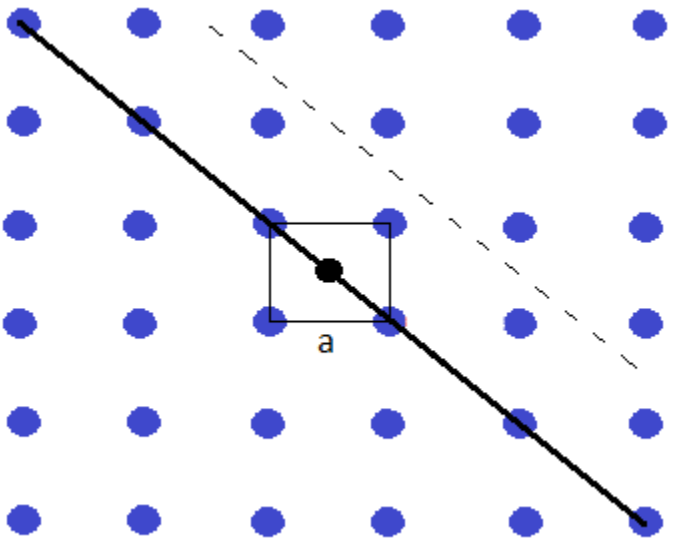

Figure 1: Example of a 2D cubic crystal lattice. The crystal has translational symmetry where the unit cell (the repeating unit) is shown by the black box. The crystal also has a 4 fold center of symmetry about the black dot in the middle of the unit cell; a line of reflection represented by the solid black line; and a glide plane represented by the dotted black line.

Symmetry

One way to describe a crystal is to choose one of the points in the crystal as the origin, and define all of the atoms in the crystal in relation to that point. This would

correctly depict the location of the atoms in the lattice, but would be very tedious. Instead, we can employ symmetry to simplify the description of the crystal. A symmetry operator is any systematic operation that can be performed on the crystal that leaves the crystal invariant. There are a number of different types of symmetry operators that can be possessed by crystals. Take, for example, Fig. 1. If each atom in Fig. 1 is translated to the right by a distance *a*, then each atom will map to an equivalent position in the lattice. This type of symmetry is called *translational symmetry*. A crystal, by definition, is an arrangement of atoms that has translational symmetry. In addition to translational symmetry, some crystals also have *point symmetry*. Point symmetry is broken up into two categories, rotational symmetry and reflective symmetry. Referring again to Fig. 1, the black dot in the center of the figure represents a 4-fold center of symmetry. If the crystal is rotated 90,180, 270 or 360 degrees about this point, all atoms will map to equivalent positions in the lattice. This is an example of *rotational symmetry*. The diagonal line drawn in the figure represents a plane of reflection. If all the atoms in the crystal are reflected across this line, the crystal, again, remains invariant. This is an example of *reflective symmetry*. Crystals can also have symmetry that combines translational symmetry and point symmetry, such as a translation followed by a rotation. Such an example is also seen in Fig. 1 and is represented by the diagonal dashed line. If the entire crystal is translated to the right by ½ *a* and then up by ½ *a*, and is then reflected across this line, all atoms will map to equivalent positions in the lattice. This type of symmetry is called a *hybrid symmetry*. The complete set of symmetry operators possessed by the crystal is called the *space group symmetry* of the crystal. Space groups are convenient for describing crystal structures because the space group, along with a choice of origin and

lattice basis, describes exactly how the atoms in a crystal are arranged. The space group can also give useful information about the material properties of the crystal, such as magnetism and electrical conductivity, because the material properties of the crystal are closely related to its symmetry (Stokes 2007).

Diffraction

In crystallography an experimental method called diffraction is used to discover the detailed atomic structure of a crystal. In diffraction experiments, subatomic probe particles (eg. electrons, x-ray photons, neutrons, etc.) that exhibit wave particle duality are fired at a small sample of the crystal (Fig. 2). As the particles scatter from the atoms in the crystal, their scattered waves interfere with each other. At most diffraction angles, the scattered waves interfere destructively with one another causing them to cancel. At certain specific angles, however, the waves interfere constructively with one another to

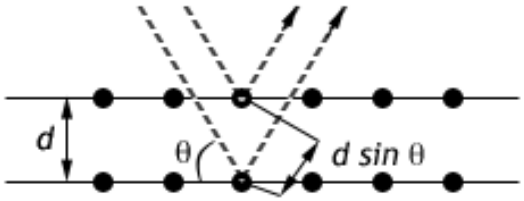

Figure 2: Diagram of diffraction . The black circles represent atoms and the dotted lines represent diffracted particles. The Bragg angle, and d spacing are also shown.

produce highly directed emissions. The angles at which particles scatter with constructive interference are called Bragg angles and the associated high intensity peaks in the diffraction pattern are called Bragg peaks. Bragg angles are related to the wavelength of

the scattered particles and the spacing between atoms in the crystal. This relationship is expressed in Bragg's law:

$$2d\sin\theta = \lambda$$

where d is the spacing between planes of atoms for fundamental reflections or some integral fraction of the spacing for higher order harmonics, λ is the wavelength of the scattered particles, and θ is the Bragg angle. The result of a diffraction experiment is a diffraction pattern like that shown in Fig. 3. Since Bragg's law only allows scattering at specific angles, the diffraction pattern for a signal crystal is discrete.

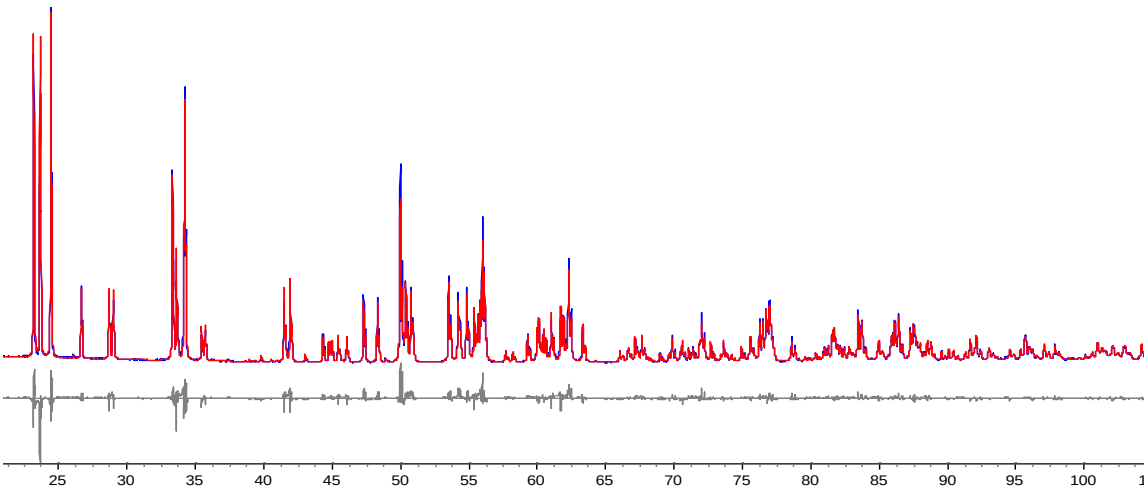

Figure 3: Example x-ray diffraction pattern. The calculated, observed, and difference patterns are shown.

The intensity, $I(\vec{h})$, of any Bragg peak in reciprocal space can be calculated as:

$$I(\vec{h}) = \left|F(\vec{h})\right|^2$$

$$F(\vec{h}) = \sum_m f_m e^{2\pi i \vec{h}\vec{x}_m}$$

where f_m and $\vec{x}_m$ are the scattering length and position of the m^{th} atom in the crystal respectively, and the structure factor, $F(\vec{h})$, is the Fourier transform of the scattering density

of the crystal. Note that $F(\vec{h})$ is a complex valued function with a magnitude and phase. Unfortunately, we can only measure $I(\vec{h})$ which gives us the magnitude of $F(\vec{h})$ but not its phase. If the phase information were known, one would merely need to perform the inverse Fourier transform to determine the scattering density of the crystal. This lack of phase information is one of the central problems of crystallography and is known as the *phase problem.*

Powder diffraction

Powder diffraction is a method of diffraction used when a single crystal is difficult or impossible to synthesize. A powder sample can be thought of as a group of small, randomly oriented single crystal samples. Because the crystals are randomly oriented, each Bragg angle produces a ring in the diffraction pattern instead of the discrete point. This causes the information contained in the spatial distribution of the pattern to be lost and often causes multiple Bragg peaks to map to the same diffraction angle in the powder diffraction pattern. This loss of information makes the crystal more difficult to solve. A powder pattern like the one shown in Fig. 3 is obtained from powder data by integrating around the rings to obtain the total intensity produced at the Bragg angle (Masa 2004).

Neutron powder diffraction

Neutron diffraction is governed by Bragg's law just like x-ray diffraction. Neutrons are heavy and charge neutral causing them to ignore electrons and to be scattered primarily by the atomic nuclei. This difference in scattering causes neutron

powder diffraction patterns to give additional information that is not present in x-ray diffraction patterns (Masa 2004).

Structure solution

Structure solution is the process of determining the crystal structure from the intensities of the diffraction pattern. There are a number of clever methods that have been devised to do this. One is called the *Patterson method*, where a map is created using the squared intensities of the powder pattern. Vectors from the origin of the map to maxima in the map correspond to interatomic vectors in the crystal. For sufficiently simple problems, these vectors can be used to determine the crystal structure (Masa 2004). Another approach is called *direct methods,* which method uses sophisticated statistical techniques to determine a self-consistent set of phases for the crystal. While the details of the method are too complicated to describe in this paper, the main principle is that the phases of three strong reflections sum to 180 degrees. Another, recently developed, approach is called *charge flipping*, which uses iterative transforms between direct space and reciprocal space to determine the structure of the crystal *(Oszlanyi and Suto 2004)*. From the structure solution process a model of the crystal in which most of the atoms are close to their true atomic positions is obtained.

Refinements

After structure solution has been performed, crystallographic refinements are done to determine the exact locations of the atoms in the crystal. A refinement is a nonlinear least-squares optimization in which a mathematical model is fit against the

observed diffraction pattern. The refined variables include model parameters such as atomic positions, lattice parameters, thermal parameter ellipsoids, atomic occupancies, and peak-shape profile until the pattern calculated from the model almost perfectly matches the observed pattern (Masa 2004). When the refinement is finished, the final values for the refined parameters constitute the crystal structure.

INTRODUCTION

The characterization of structural distortions from powder diffraction data is a distinct subclass of the broader field known as "structure determination from powder data" (SDPD). A distorted structure, by definition, can be parameterized in terms of its deviations from a known "parent" structure, and has a space group symmetry that is a subgroup of the symmetry of the parent. As a rule, distortions lower symmetry and increase structural complexity. It is common to define a structural distortion relative to the experimentally observed parent structure from which it arises in nature. We note, however, that it is often convenient to define a distortion relative to a more distant parent separated by several phase transitions, or even a hypothetical parent structure. We use the term "distortion" quite generally here to indicate the presence of any type of physical order parameter such as atomic displacements, magnetic moments, compositional ordering, lattice strain, etc., which distinguish the parent and child structures. Distortions arising from second order (i.e. continuous) phase transitions tend to be of special interest, though arbitrary discontinuous transformations involving one or more coupled order parameters also fall within the scope of this work.

After removing the parent symmetries that are broken by the distortion, those symmetry operations that remain comprise the "distortion symmetry", which is simply the space-group symmetry of the distorted structure. Here, it is important to distinguish a space group from its "type". The 230 crystallographic space-group types are tabulated in the International Tables Vol. A (Hahn 2005), whereas a complete space group describes both the symmetry operators and their actual locations within the crystal. Thus, there can be multiple ways to remove a portion of the parent symmetry, each of which yields the same space-group type but different distortion symmetries by virtue of differences in the locations of the remaining operators. $Pm\bar{3}m$, for example, has an (a' = 3a, b' = 3b, c' = 3c) maximal subgroup of the same $Pm\bar{3}m$ type, but which clearly has a much lower overall symmetry. A given distortion symmetry can always be uniquely identified by its combination of space-group type, lattice basis (i.e. the size/shape of the supercell) and supercell origin relative to the parent cell, provided that standard origin-choices from ITA are employed (Campbell, Stokes et al. 2006). In general, the distortions allowed by different distortion symmetries are fundamentally distinct.

When the powder diffraction pattern contains sufficient information, one can treat an unknown distortion like any other unknown structure: extracting the integrated intensities and attempting a direct solution of the phase problem via direct methods, charge-flipping, real-space global search or some hybrid algorithm (David, Shackland et al. 2002). In many cases, however, the combination of small peak splittings, weak superlattice reflections and structural complexity make it difficult or impossible to extract reliable intensities, so that direct-space models must then be tested directly against the entire powder pattern. This problem reflects the fact that the number of strong and

clearly-resolvable Bragg reflections often fails to keep up with the increased structural complexity of a subtle distortion.

We assume that the cell parameters of the distorted structure can be refined and used as a starting point for further analysis. Ideally, the space-group type would also be identifiable from systematic absences or other considerations, so that the number of structural variables in the distorted phase can be restricted by symmetry to a manageable number with the help of reasonable constraints and/or restraints. For highly-complex structural distortions, the likelihood of success is much improved by the use of global optimization algorithms which explore a wide parameter space and are more robust against getting caught in local minima (Černý and Favre-Nicolin 2007). Monte Carlo (Tremayne, Kariuki et al. 1997; Evans, Howard et al. 2003), simulated annealing (Deem and Newsam 1989; Campbell, Bellussi et al. 1998; Coelho 2000), parallel tempering (Falcioni and Deem 1999; Favre-Nicolin and Černý 2002), genetic algorithms (Kariuki, H. et al. 1997; Shankland, David et al. 1997) and various hybrid approaches have been employed with good results.

The use of multiple sources of structural information has also proven helpful for highly complicated structures. The $M_2P_2O_7$ family of octahedral/tetrahedral framework compounds, for example, experience distortions that create large supercells with over 100 unique atoms and over 400 displacive degrees of freedom. For these compounds, electron diffraction patterns and solid-state NMR spectra have been essential for determining the distortion symmetry; and joint x-ray/neutron datasets have been needed for successful structural refinements (Gover, Withers et al. 2002; Fayon, King et al. 2003; Stinton, Hampsons et al. 2006).

Unfortunately, the same problems that inhibit the extraction of useful intensities (too few resolvable strong reflections) can also make it very difficult to conclusively detect systematic absences, placing the determined crystallographer in a tight spot. Lacking the space group symmetry of the distorted structure, and deprived of an easy solution to the phase problem, he/she reluctantly turns to the design and testing of candidate distortion models by trial and error, which requires a great deal of time, patience and intuition. The challenge is compounded by the need to customize constraints and restraints for each distortion model that are both physically reasonable and appropriate to the symmetry.

When the distortion symmetry is not known beforehand, the trial and error approach to model building can still be guided by symmetry principles rather than by brute force. A relatively simple first step is to try each of the maximal subgroups of the parent space group (tabulated in the International Tables of Crystallography Vol. A, Hahn 2005) to see if the resulting cell parameters and crystal family are consistent with the size and shape of the experimentally-identified supercell. If not, some successive chain of maximal subgroups is guaranteed to produce the correct result, provided that your structure really is a distortion of the chosen parent. But it may be necessary to explore a large number of such chains before uncovering the right one.

The development of sophisticated computational tools for mapping out symmetry-descent chains has helped to make the symmetry-descent process simpler and more robust (Howard and Stokes 1998; Evans, Howard et al. 2003; Howard and Zhang 2003; Zhao, Ross et al. 2009; Carpenter, McKnight et al. 2010). For example, Bi_2Sn2O_7 exhibits a high-temperature pyrochlore phase (γ) with space-group $Fd\bar{3}m$ as well as two

non-centrosymmetric distorted pyrochlore phases: an intermediate-temperature β phase with a 2×2×2 supercell, and a room-temperature α phase with a √2×√2×2 supercell (Shannon, Bierlein et al. 1980; Jones and Knight 1997). When the two maximal subgroup of $Fd\bar{3}m$ consistent with the α-phase metric failed to produce a good fit (Kennedy, Ismunandar et al. 1998), Evans *et al*. (2003) systematically generated a symmetry-descent tree containing all of the non-centrosymmetric subgroups of the parent γ-phase symmetry that are consistent with the observed α-phase metric, and which also descend from an intermediate subgroup consistent with the β-phase metric. They then created models for each candidate subgroup and performed combined simulated-annealing-mode refinements against synchrotron x-ray and neutron powder datasets in order to discern the correct monoclinic structure.

Group theoretical symmetry mode analysis (SMA) tools now exist which generate all of the candidate distortion symmetries of a parent structure that are consistent with an experimentally-observed supercell and point-group or space-group type (Campbell, Stokes et al. 2006). They also generate structural models for each candidate distortion, thus dramatically reducing effort required for trial and error exploration. Most importantly, the resulting distortion models can either be expressed in terms of traditional atomic parameters (e.g. atomic coordinates, moments, occupancies, etc.) or in terms of symmetry-mode amplitudes. The traditional and symmetry-mode parameter sets are related by a simple linear transformation, so that one atom can be influenced by many modes, and one mode can influence many atoms. Both parameter sets always provide the same number of degrees of freedom. In practice, the active mode amplitudes can either be calculated from the atomic positions after the refinement (Orobengoa, Capillas et al.

2009), or can be refined directly themselves and used to calculate the atomic positions at each step (Campbell, Evans et al. 2007). Muller et al. (2010) further demonstrated a parametric refinement in which the temperature dependencies of the key symmetry-modes were modeled via the direct refinement of their critical exponents. The importance of the symmetry-mode description of a distorted structure clearly lies in the fact that most of the modes available to a distortion tend to be inactive, so that the essential structural features can be expressed in terms of a relatively small number of non-zero mode amplitudes (Campbell, Stokes et al. 2006; Campbell, Evans et al. 2007; Perez-Mato, Orobengoa et al. 2010).

When the symmetry of the distortion is unknown, one might naively consider a global optimization without symmetry (i.e. in space group *P*1) that simply restrains each atom to stay close to its position in the high-symmetry parent structure. However, if a scarcity of strong and well-resolved superlattice reflections is the underlying difficulty, then a *P*1 model is very likely to be grossly underdetermined anyway. In the present work, we demonstrate one can readily "bootstrap" a *P*1-symmetry model that has been parameterized in terms of symmetry-mode amplitudes. Subsequently turning off all of the inactive symmetry modes is equivalent to constructing a network of structural constraints, and greatly simplifies the structure solution by allowing only the essential degrees of freedom to be explored. The fundamental issue that we address here is how to discover which symmetry modes to turn off without prior knowledge of the distorted structure.

METHODS

Rietveld refinement

A Rietveld refinement is a refinement method useful for refining powder diffraction data. It fits the full profile of the observed diffraction pattern to the profile of a calculated diffraction pattern using a structural model. Typically the model contains parameters such as lattice parameters, background parameters, thermal parameters, and atomic coordinates; however, in this research, atomic coordinates have been replaced by atomic distortion modes. These parameters are allowed to vary in an iterative least squares process to minimize the function:

$$M = \sum_i W_i \{y_{i(obs)} - y_{i(calc)}\}^2$$

where $y_{i(obs)}$ is the *i*th step in the observed pattern, $y_{i(calc)}$ is the *i*th step in the calculated pattern, and w_i is the weight at each point.

The quality of the Reitveld refinement is measured using one among several possible *R* factors. The one used in the analysis presented here is called R_{wp}, or "*R*-weighted pattern" given by:

$$R_{wp} = \frac{\sum_i w_i \{y_{i(obs)} - y_{i(calc)}\}^2}{\sum_i w_i y_{i(obs)}{}^2}.$$

R_{wp} measures the weighted squared distances between the measured and observed diffraction patterns (Masa 2004).

Pawley refinement

A Pawley refinement fits a calculated diffraction pattern against an observed diffraction pattern without the use of a structural model (Stokes 2007). Instead a model composed only of cell parameters, peak shape parameters, sample height, and zero error is used. An iterative least squares process is performed in which the parameters are varied until the calculated diffraction pattern fits very well with the observed diffraction pattern. The resulting values for lattice parameters, peak shape parameters, sample height, and zero error can then be used as a starting point for further structural analysis.

ISODISTORT

ISODISTORT is an online software package for exploring distortions in crystal structures (Campbell, Stokes et al. 2006). The software was used to generate distortion mode decomposed structural models in *P*1 symmetry. Models were generated by uploading a high symmetry model of the material and using method 3 to decompose the high symmetry model into a low symmetry model with the appropriate supercell basis. Models were then saved as TOPAS format .str files for use in refinement software.

TOPAS Academic/ J-Edit

Topas Academic is a structure refinement software capable of performing both Pawley refinements and distortion-mode Rietveld refinements (Coelho 2007). The Topas input (.inp) files were generated using John Evans' Topas Academic macros. Topas was also used for generating simulated x-ray diffraction data for $LaMnO_3$.

EXPERIMENTAL PROCEDURE

Rietveld refinements were performed on room-temperature (RT) powder-diffraction data from $LaMnO_3$ and WO_3 using the Topas Academic software package (Coelho 2007). The refinements reported here employed a global optimization strategy best described as "repeated local minimization from random starting values" (RLM), which is similar to approach described by Shankland *et al*. (2010). This was accomplished with the *continue_after_convergence* keyword of TOPAS, and custom macros for variable randomization after each convergence cycle. Displacive symmetry-mode amplitudes, rather than traditional atomic *xyz* coordinates, were refined directly (Campbell, Evans et al. 2007). The ISODISTORT software package was used to compute the atomic coordinates as linear functions of the symmetry-mode amplitudes and to export these symbolic relationships in TOPAS .str format.

A WO_3 sample was prepared according to the method described by Woodward *et al* (1997). Approximately 12g of reagent grade WO_3 (Alpha Aesar, 99.8% pure) was ground with a pestle and mortar and placed in a platinum crucible. This was heated in air to 1273K and allowed to cool to 298K at a cooling rate of 0.25K/min. The sample was a green crystalline powder with no visible yellow impurities. We found no impurities belonging to the lower-temperature $P\bar{1}$ -symmetry phase.

Laboratory x-ray powder data from WO_3 were collected over the range 10-150º 2θ on a Bruker D8 Advance diffractometer in Cu-Kα_1 mode with soller slits and a Lynx-Eye detector. The data were subsequently corrected to account for the variable divergence slit, which maintained a 6 mm beam footprint on the sample. The samples were finely ground and passed through an 80-mesh sieve onto a zero-background silicon

disk coated in petroleum jelly to ensure good orientation sampling. Room temperature (RT) x-ray powder patterns were taken from variable-temperature sequences of approximately one-minute scans over the range from 1285k to 290k, which were collected using an Anton-Parr HTK1200 furnace attachment. These x-ray powder diffraction patterns did not have exceptional counting statistics because each single-temperature pattern was collected quickly. The decision to use modest-quality data for the current study was strategic because it better demonstrates the limits of the novel analysis method that we present here. Neutron powder diffraction data on a second sample from the same synthetic batch of WO_3 was collected at the HRPD beamline at ISIS over a time-of flight range of 40-114 ms using its back-scattering detector bank. The RT neutron powder pattern was taken from a variable-temperature sequence of approximately one-minute scans (300 K to 85 K) collected in a CCR cryostat.

The $LaMnO_3$ x-ray diffraction patterns were simulated from a published model (Rodriguez-Carvajal, Hennion et al. 1997) and included various levels of artificial noise. This simulated patterns had zero background, a simple peak shape and zero thermal parameters. All $LaMnO_3$ refinements discussed here employed the simulated pattern with artificial noise added, whereas all WO_3 refinements discussed here employed only the experimental diffraction pattern. The WO_3 refinements were performed separately against x-ray and neutron data, and also jointly against both x-ray and neutron data.

The analysis of each powder pattern began with a Pawley fit in order to obtain reasonable estimates for the cell parameters, the peak profile, and the sample height and zero-angle corrections. In some cases, we further improved the peak shape and background via Rietveld refinement of a structural model. The high-temperature high-

symmetry phase above the transition is known and will usually provide a very similar peak shape. All non-structural parameters were then fixed for symmetry-detection work. Because no significant stoichiometric deviations were expected, all occupancies were fixed at 1. For WO_3, one isotropic thermal parameter was permitted for W and another for O. These thermal parameters were estimated from the best fit amongst several simulated-annealing convergences in *P*1 symmetry and then fixed. Only the scale factor and any displacive parameters were allowed to refine when generating mode-amplitude histograms for symmetry detection. All displacive-amplitude parameters were randomized relative to zero after each convergence using a uniform distribution of preset width in order to ensure good statistical sampling. By leaving TOPAS running overnight or over the weekend in RLM mode, we were reliably able to obtain over 1000 convergences for each powder pattern studied, even when refining a large number of displacive degrees of freedom.

RESULTS AND DISCUSSION

Symmetry-mode description of WO_3

The temperature phase diagram of WO_3 is quite complicated (Diehl, Brandt et al. 1978; Woodward, Sleight et al. 1995; Vogt, Woodward et al. 1999). At room temperature, it is reported to have a monoclinic $P2_1/n$ supercell with $\mathbf{a}_o = 2\mathbf{a}_c$, $\mathbf{b}_o = 2\mathbf{b}_c$ and $\mathbf{c}_o = 2\mathbf{c}_c$ (Tanasaki 1960; Howard, Luca et al. 2002). The supercell origin is located at (1/2,0,1/2) within the parent cell. Being 8 times larger than the cubic parent cell, this supercell contains a total of 32 atoms and possesses either 24 free displacive variables in $P2_1/n$ symmetry or 96 free displacive variables in *P*1 symmetry. Of the 24 displacive

symmetry modes available to RT-WO_3 in space group $P2_1/n$, Campbell *et al*. (2007) demonstrated that only five of these modes have large amplitudes. They further showed that a refinement including only these 5 modes produced nearly as good a fit as a refinement that included all 24 modes.

Atom	Mode # ($P1$)	Mode name ($P1$)	Amplitude (Å)($P1$)	Mode name ($P2_1/n$)
W	14	$X_5^-(a,b,c,d,e,f)[\mathrm{W}]T_{1u}(e)$	–0.47168	$X_5^-(0,0,a,a,b,-b)[\mathrm{W}]T_{1u}(b)$
W	15	$X_5^-(a,b,c,d,e,f)[\mathrm{W}]T_{1u}(f)$	0.47168	
W	16	$M_3^-(a,b,c,d,e,f)[\mathrm{W}]T_{1u}(a)$	0.73279	$M_3^-(a,b,0)[\mathrm{W}]T_{1u}(a)$
O	37	$R_4^+(a,b,c)[\mathrm{O}]E_u(a)$	–1.10338	$R_4^+(a,b,0)[\mathrm{O}]E_u(a)$
O	38	$R_4^+(a,b,c)[\mathrm{O}]E_u(b)$	0.83820	$R_4^+(a,b,0)[\mathrm{O}]E_u(b)$
O	78	$M_3^+(a,b,c)[\mathrm{O}]E_u(c)$	–1.08814	$M_3^+(0,0,a)[\mathrm{O}]E_u(a)$

Table 1: Important distortion modes in room temperature WO_3 (space group symmetry $P2_1/n$)

If reducing 24 structural variables down to 5 important active-mode amplitudes provides an advantage in $P2_1/n$, the advantage of using symmetry modes will be far greater in $P1$ symmetry where there are 96 displacive degrees of freedom. Due to low-symmetry mode splitting, however, there will be 6 rather than 5 $P1$-symmetry counterparts to the 5 important $P2_1/n$-symmetry modes (listed in Table 1). To illustrate mode splitting, consider that X_5^- is a six-dimensional irrep at the X=[0,1/2,0] point in k-space, and therefore has a general order-parameter direction (OPD) with 6 degrees of freedom, which is indicated as (a,b,c,d,e,f). Order parameter directions are described in more detail elsewhere.(Stokes and Hatch 1987; Campbell, Stokes et al. 2006) A distortion along this OPD would result in a rather complicated superstructure having $P1$ symmetry. The RT distortion with $P2_1/n$ symmetry, on the other hand, is achieved via

the action of a special six-dimensional OPD containing only two free parameters: (0,0,*a*,*a*,*b*,–*b*). This OPD provides the parent W atom with two displacive modes (one for *a* and one for *b*), only one of which (the *b* branch) is shown to be important (i.e. has a substantially non-zero amplitude and a significant effect on the diffraction pattern) in RT WO_3. In *P*1 symmetry, this $X_5^-(0,0,a,a,b,-b)[\mathrm{W}]T_{1u}(b)$ mode splits into two modes, $X_5^-(a,b,c,d,e,f)[\mathrm{W}]T_{1u}(e)$ and $X_5^-(a,b,c,d,e,f)[\mathrm{W}]T_{1u}(f)$, which must maintain equal and opposite amplitudes in order to describe a structure that actually has $P2_1/n$ symmetry.

WO_3 symmetry detection

When the RT structure of WO_3 is represented with 96 displacive modes in *P*1 symmetry, the ability to experimentally identify the important modes (those in Table 1) would enable us to establish the $P2_1/n$ symmetry of the structure without making any *a priori* assumptions. The identification of these modes is the principle objective of the present work. Rather than attempting to achieve a "best fit" of a *P*1 model to experimental diffraction data, which would not likely succeed because of the large number of parameters, we instead statistically sample a large number of convergence cycles from randomized starting points, and accumulate the individual parameter values from each convergence. From these results, we are able to compile a separate histogram for each of the 96 free displacive mode amplitudes. While a mode may converge to an incorrect value on occasion, it statistically tends toward its true value, which is manifested as a peak in its histogram. A peak whose position clearly deviates from zero is evidence that the mode is active.

When detecting active modes in *P*1 symmetry, the global structural origin needs to either be constrained or restrained so as not to wander randomly. We constrained the origin by fixing the three ferroelectric tungsten Γ_4^- modes at zero amplitude. Note that in the symmetry-mode description, only ferroelectric modes can translate the origin. A more sophisticated approach would be to restrain the center of mass to stay close to zero.

This approach was first tested against the room-temperature laboratory x-ray power diffraction data shown in Fig. 4. All non-displacive parameters except the scale factor were fixed, as well as the three ferroelectric tungsten mode amplitudes, which left

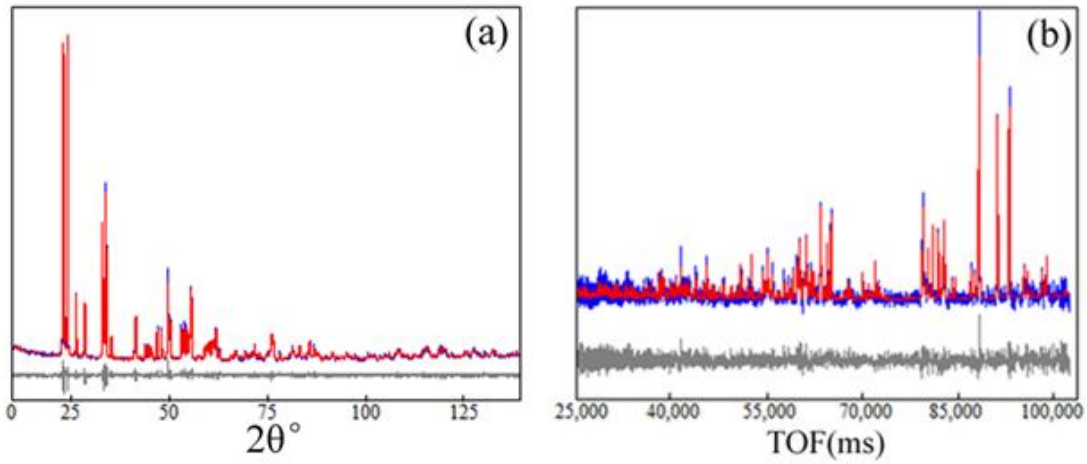

Figure 4: $P2_1/n$-symmetry fits to room temperature (a) lab x-ray data and (b) time-of-flight neutron data from WO_3. Calculated, observed and difference patterns are shown.

93 refinable atomic-displacement mode amplitudes. More than 1000 convergence cycles were then run in Topas in RLM mode, and histograms were generated for each mode. We found it helpful to keep the number of convergence cycles greater than ten times the number of free parameters. As can be seen from Fig. 5, the histograms of the three important tungsten modes belonging to the X_5^- irrep (#14 and #15) and the M_3^- irrep (#16) have very well defined peaks at nonzero values, clearly showing that they are

active, while all other tungsten modes have distinct peaks at zero. Because the combination of these X_5^- and M_3^- tungsten modes are insufficient for detecting the actual $P2_1/n$ symmetry, oxygen mode information is essential.

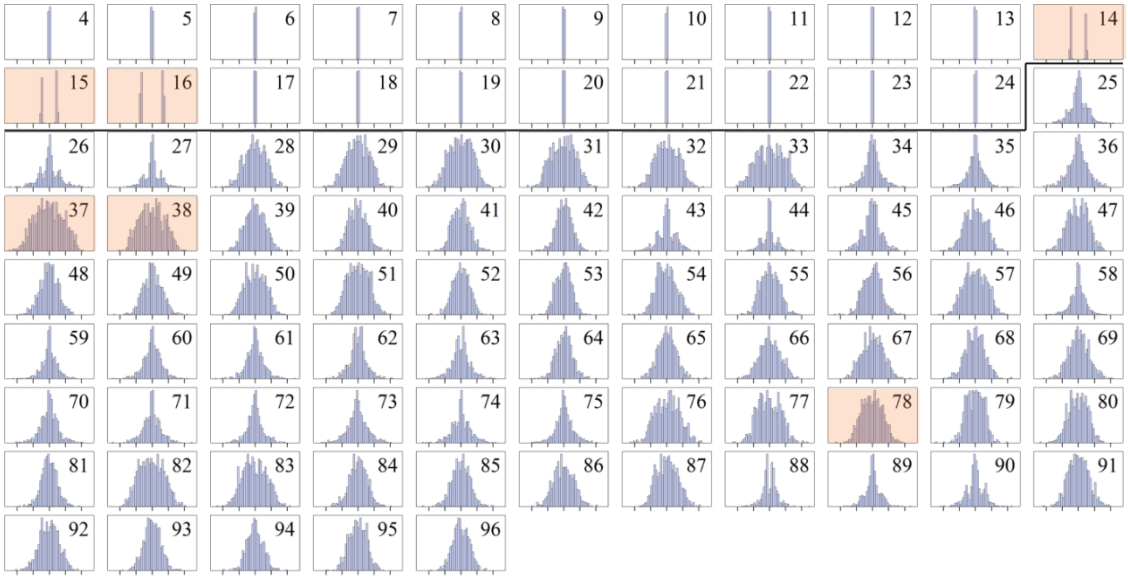

Figure 5: Multi-convergence mode-amplitude histograms for each of the displacive WO_3 model parameters in *P*1 symmetry based on a refinement against lab x-ray data. Histograms above the horizontal black line correspond to tungsten modes, while those below the black line correspond to oxygen modes. The horizontal axis of each plot runs from -2.5 Å to 2.5 Å, with tick-marks placed every 1.0 Å (mode amplitude is defined as root-summed-squared displacement of all affected atoms). A histogram with a clearly split peak is evidence that the mode is active. The histograms of several important active modes have been highlighted.

The oxygen modes, both active and inactive, proved to be less sensitive to x-rays, and did not yield well-defined peaks, tending instead to display a relatively uniform distribution over a broad range of amplitudes. The oxygen-mode histograms belonging to the R_4^+ irrep (#37 and #38) are precisely what we need to complete the symmetry, but yield no information. Even the large-amplitude oxygen M_3^+ mode (#78), which is not helpful for identifying the actual symmetry, is not well-defined. One oxygen X_5^- mode

(#88) is clearly split; and two oxygen M_3^- modes (#62 and #63) are slightly split; but these modes don't add any new symmetry information to the tungsten X_5^- and M_3^- modes discussed above, as they belong to corresponding branches of same irreps. Because of the relative insignificance of the large-amplitude R_4^+ oxygen modes in these modest-quality x-ray powder patterns, we were unable to clearly detect the $P2_1/n$ symmetry of RT WO_3, which underscores the importance of obtaining high-quality data that is sensitive to all of the important order parameters. We did, however, subsequently find that the important R_4^+ oxygen modes could be readily detected from significantly higher-quality x-ray data. It is, of course, possible to introduce additional chemical information to help identify the oxygen modes. For example, in this case we can use restraints to keep the bond valence sums around 6.0 for W sites and 2.0 for O sites and include a simple "anti-bump" penalty function to prevent O–O distances shorter than ~2.4 Å. Both of these are relatively unbiased pieces of information designed to restrict the RLM process to finding chemically plausible structures. Introduction of these restraints leads to a significant narrowing of the majority of the histograms of Fig. 5 and gives clear indication of the importance of R_4^+ and M_3^+ modes.

Similar refinements were also performed against modest-quality room-temperature time-of-flight neutron powder diffraction data (Fig. 4) in order to achieve greater sensitivity to oxygen relative to tungsten. In the neutron-based histograms of Fig. 6, the oxygen R_4^+ (#37 and #38) and M_3^+ (#78) modes are clearly active, while two of the oxygen modes belonging to the X_5^- irrep (#46 and #47) also display very slight peak splittings, which turn out to be to be real. Unfortunately, these modes alone are once

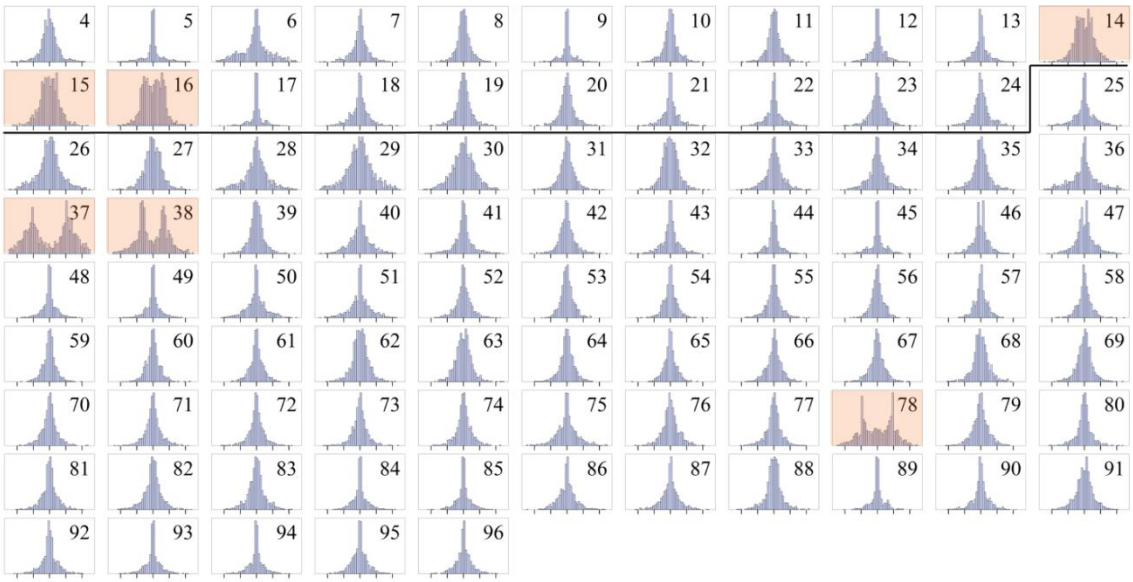

Figure 6: Multi-convergence mode-amplitude histograms for each of the displacive WO3 model parameters in *P*1 symmetry based on a refinement against neutron time-of-flight data. Histograms above the horizontal black line correspond to tungsten modes, while those below the black line correspond to oxygen modes. The horizontal axis of each plot runs from -2.5 Å to 2.5 Å, with tick-marks placed every 1.0 Å. The histograms of several important active modes have been highlighted.

again insufficient to detect the $P2_1/n$ symmetry. The active W modes could resolve the matter but are not conclusively split. And the oxygen M_3^- mode (#88) that showed up clearly in its x-ray histogram is not at all observable in the neutron histogram. As a result, one should judge its appearance in the x-ray histogram as unreliable (possibly a phantom mode).

Some of the mode histograms in Figs. 5 and 6 that we have not yet discussed are not nicely shaped as single peaks. Some modes, for example, have small satellite peaks, indicating that they tend towards an alternate amplitude in a significant fraction of the convergence cycles. Due to the nature of multivariable optimization problems, this is to be occasionally expected, even for modes that actually have zero amplitude. But by collecting a sufficiently large statistical sample, we are usually able to distinguish between amplitudes associated with global vs local minima. While this outcome cannot

always be guaranteed, it is necessary for the success of the method. The probability of detecting active vs phantom mode amplitudes increases with the quality of the data and the simplicity of the distorted structure. We will discuss the observation and elimination of phantom modes in more detail below.

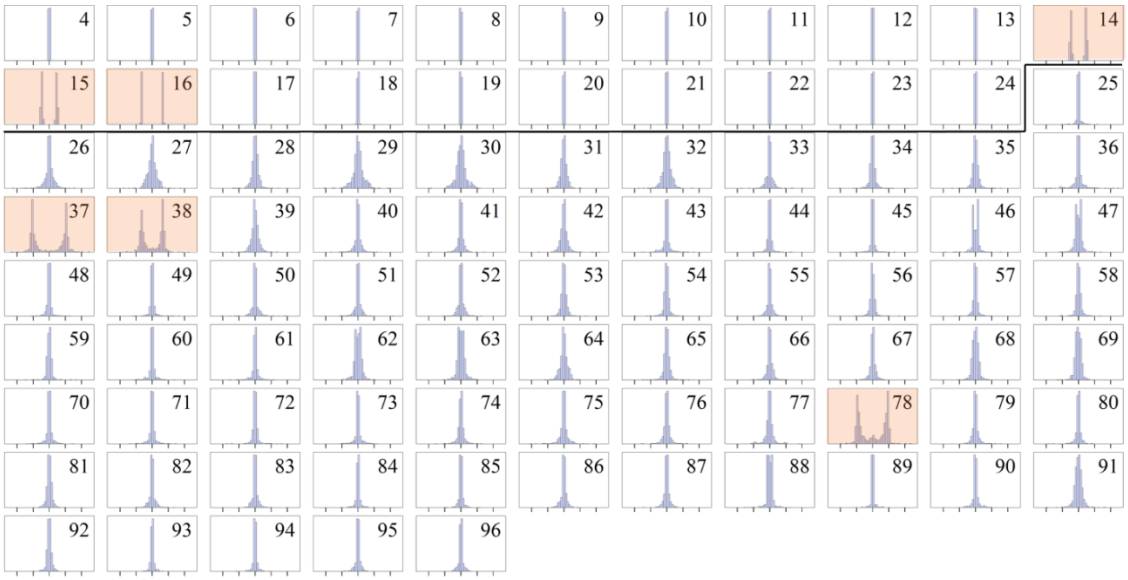

Figure 7: Multi-convergence mode-amplitude histograms for each of the displacive WO_3 model parameters in *P*1 symmetry based on a combined refinement against both lab x-ray data and neutron time-of-flight data. Histograms above the horizontal black line correspond to tungsten modes, while those below the black line correspond to oxygen modes. The horizontal axis of each plot runs from -2.5 Å to 2.5 Å, with tick-marks placed every 1.0 Å. The histograms of several important active modes have been highlighted.

The separate x-ray and neutron refinements were followed by a joint refinement against both datasets, which combined good sensitivity to both oxygen and tungsten modes. The refined parameters were as described above. All of the important active oxygen and tungsten modes are readily detected in this case, along with a number of less important modes that yielded slightly split histograms (Fig. 7). No phantom modes were observed.

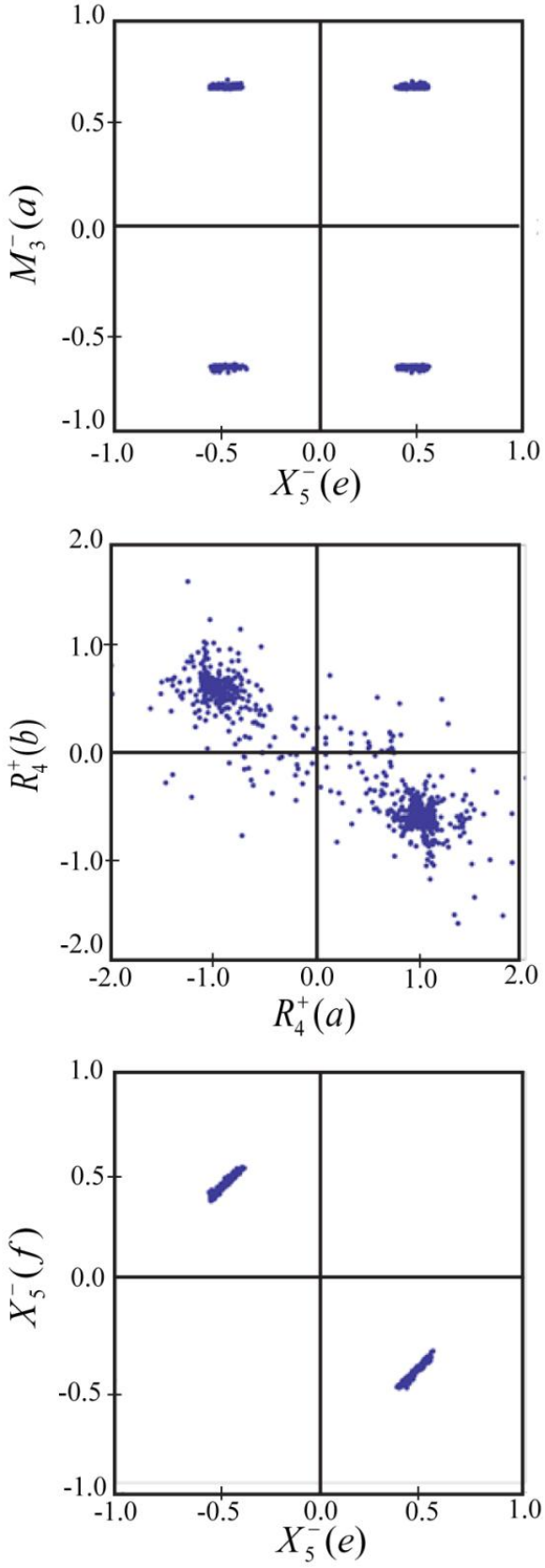

Figure 8: Multi-convergence mode-correlation plots from the combined refinement of a *P*1-symmetry WO_3 model against both lab x-ray data and neutron time-of-flight data. In the top panel, the tungsten M3-(a,0,0) and X5-(0,0,0,0,*e*,0) modes (#15 and #16) are both sharply split, but not correlated to one another because they belong to separate irreps. In the middle panel, the oxygen R4+(*a*,0,0) and R4+(*b*,0,0) modes (#37 and #38), on the other hand, are strongly anti-correlated, but still independent. In the bottom panel, the tungsten X5-(0,0,0,0,*e*,0) and X5-(0,0,0,0,0,*f*) modes (#14 and #15) are perfectly anti-correlated so as to represent only one independent parameter within the true symmetry.

As discussed above, the 6 important modes identified from the joint x-ray/neutron histograms in Fig. 7 must be merged into a smaller set of 5 parameters in order to correspond to the correct $P2_1/n$ symmetry. In other words, one of these modes must be related to one of the others by symmetry. In Fig. 7, we can see at a glance that $X_5^-(a,b,c,d,e,f)[\mathrm{W}]T_{1u}(e)$ and $X_5^-(a,b,c,d,e,f)[\mathrm{W}]T_{1u}(f)$ have equal amplitudes. The two dimensional correlation plot in Fig. 8 further reveals that the two mode amplitudes are perfectly anticorrelated. These modes prove to be the *b* and *–b* components of the $X_5^-(0,0,a,a,b,-b)[\mathrm{W}]T_{1u}(b)$ mode of the $P2_1/n$ distortion. In contrast, the top panel shows that modes belonging to different irreps (i.e. M_3^- and X_5^-) are uncorrelated. The two oxygen R_4^+ modes shown in the middle panel, $R_4^+(a,b,c)[\mathrm{O}]E_u(a)$ and $R_4^+(a,b,c)[\mathrm{O}]E_u(b)$, are clearly correlated, yet still have different amplitudes, which tells us that the *a* and *b* branches of this order parameter are not related by symmetry. Because these observations are all consistent with the details of Table 1, the $P2_1/n$ distortion symmetry can be clearly and unambiguously determined.

WO_3 structure determination

Once the important active modes have been identified, the structure has already been solved by virtue of the fact that we have refined a model containing only these modes. It is a simple matter to then detect the symmetry of the resulting structure and free up all of the less important modes that are still permitted by the correct symmetry. These small-amplitude contributions polish the final product provided that the

information content of the powder pattern is sufficient to support them. Identifying the active modes is the hard part.

Multiple modes that are simultaneously active in a distortion are said to be *coupled*. One can think about mode coupling as a sequential process of lowering the symmetry of a structure by degrees. Upon invoking the first mode (i.e. order parameter component) that appears to be important (i.e. primary), the resulting distortion symmetry will include this mode as a symmetry-allowed degree of freedom. Additionally, there may be several secondary modes that also belong to that distortion symmetry that come along for the ride. If the associated secondary modes already include all of the important modes that are observed, we're done. If not, add another important mode that has not yet been included, which will further lower the symmetry and gather in additional secondary modes. Again, check the list of modes permitted by the new distortion symmetry to see if any of the observed important modes are not yet included in the list. If so, continue coupling in this way until all of the important modes have been included. The resulting distortion symmetry will be the correct one.

In practice, there is a much easier way to determine the distortion symmetry after identifying the active modes. After refining the model that includes only the apparently active modes, simply employ a symmetry detection tool like MISSYM or ADDSYM (Le Page 1987) or FINDSYM (Rodriguez-Carvajal, Hennion et al. 1997). This approach readily identified the correct $P2_1/n$ symmetry of RT-WO_3.

$LaMnO_3$

In order to further explore the limitations of our approach, we next attempted to detect the symmetry of RT $LaMnO_3$ while controlling statistical variables such as data signal-to-noise level and mode-randomization amplitude. RT $LaMnO_3$ has an orthorhombic supercell with space-group symmetry *Pnma* that is related to its cubic perovskite parent lattice by $\mathbf{a}_o = \mathbf{a}_c+\mathbf{b}_c$, $\mathbf{b}_o = -\mathbf{a}_c+\mathbf{b}_c$ and $\mathbf{c}_o = \mathbf{c}_c$. Its supercell has the same origin as the parent structure. Being 4 times larger than the parent cell, the RT-$LaMnO_3$ supercell contains a total of 20 atoms and possesses 7 free displacive variables in *Pnma* symmetry or 60 free displacive variables in *P1* symmetry.

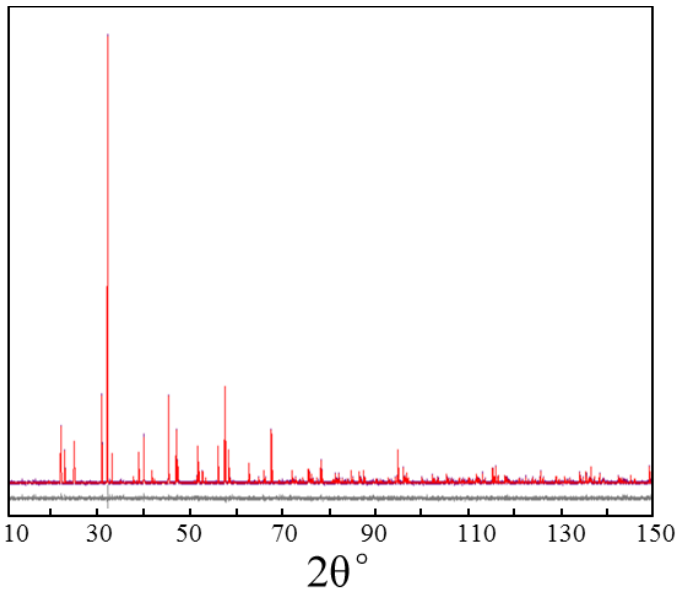

Figure 9: *Pnma*-symmetry fit to low-noise simulated x-ray diffraction data from $LaMnO_3$. Calculated, observed and difference patterns are shown.

For this work, we employed a simulated x-ray diffraction pattern which was free of systematic experimental errors (Fig. 9). As with the WO_3 data, the oxygen modes are more difficult to detect than the heavy-atom modes. In this case, however, the modes of

the heavy lanthanum atom break enough symmetry by themselves to generate the observed distortion symmetry. Luckily, the new oxygen degrees of freedom don't lower the symmetry any further. In the discussion that follows, we will abbreviate the names of the active modes as shown in Table 2.

Atom	Mode # ($P1$)	Mode name ($P1$)	Amplitude (Å)($P1$)	Mode name (*Pnma*)
La	4	$R_5^+(a,b,c)[\mathrm{La}]T_{1u}(a)$	0.06196	$R_5^+(a,a,0)[\mathrm{La}]T_{1u}(a)$
La	5	$R_5^+(a,b,c)[\mathrm{La}]T_{1u}(b)$	0.06196	
La	9	$X_5^+(a,b,0,0,0,0)[\mathrm{La}]T_{1u}(b)$	0.54731	$X_5^+(a,0,0,0,0,0)[\mathrm{La}]T_{1u}(a)$
O	37	$R_4^+(a,b,c)[\mathrm{O}]E_u(a)$	–0.83865	$R_4^+(a,-a,0)[\mathrm{O}]E_u(a)$
O	38	$R_4^+(a,b,c)[\mathrm{O}]E_u(b)$	0.83865	
O	40	$R_5^+(a,b,c)[\mathrm{O}]E_u(a)$	–0.01886	$R_5^+(a,a,0)[\mathrm{O}]E_u(a)$
O	41	$R_5^+(a,b,c)[\mathrm{O}]E_u(b)$	–0.01886	
O	45	$X_5^+(a,b,0,0,0,0)[\mathrm{O}]E_u(b)$	0.14768	$X_5^+(a,0,0,0,0,0)[\mathrm{O}]E_u(a)$
O	53	$M_2^+(0,0,a)[\mathrm{O}]A_{2u}(a)$	0.36055	$M_2^+(0,0,a)[\mathrm{O}]A_{2u}(a)$
O	54	$M_3^+(0,0,a)[\mathrm{O}]E_u(a)$	-0.89971	$M_3^+(0,0,a)[\mathrm{O}]E_u(a)$

Table 2: Important distortion modes in room temperature $LaMnO_3$ (space group symmetry *Pnma*).

Starting with the simulated $LaMnO_3$ data described above, we fixed the lattice parameters, background shape, peak shape and all other non-displacive parameters except the scale factor at the appropriate values. The three ferroelectric tungsten mode amplitudes that set the origin were also fixed, which left 57 refinable atomic-displacement mode amplitudes. More than 1000 convergence cycles were then run in Topas in simulated-annealing mode, and histograms were generated for each of these modes. The La-mode histograms in Fig. 10 have sufficiently well-defined peaks to determine which ones are active and which are not, whereas only two oxygen-mode

histograms (#45 and #53) have well-defined peaks. One inactive lanthanum R_5^+ mode (#6) and one inactive manganese M_5^- mode (#23) also appeared to split (Fig. 11 provides a closer view of mode #6). Such phantom modes point us to the wrong space group symmetry and must be eliminated if we are to detect the true space group.

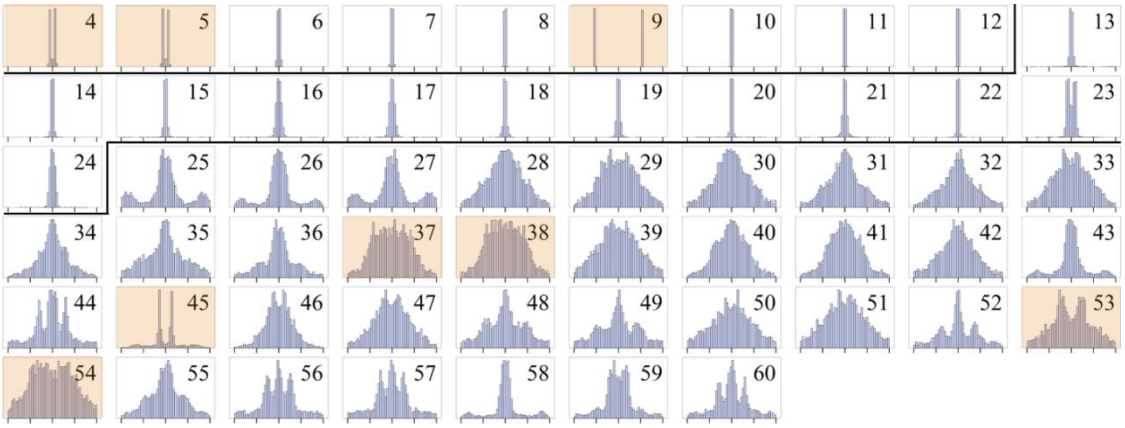

Figure 10: Multi-convergence mode-amplitude histograms for each of the displacive $LaMnO_3$ model parameters in *P*1 symmetry based on a refinement against simulated low-noise x-ray data. The horizontal axis of each plot runs from 0.5 Å to 0.5 Å, with tick-marks placed every 0.25 Å. The two horizontal black lines separate the well-defined La modes, the Mn modes and the O modes. The histograms of several important active modes have been highlighted.

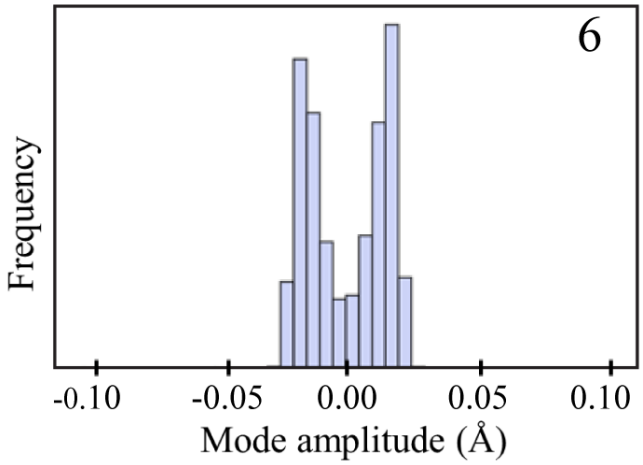

Figure 11: Zoomed-in view of the multi-convergence mode-amplitude histogram for phantom mode #6, for which the splitting wasn't apparent in the previous figure.

A procedure to eliminate phantom modes

A *P*1-symmetry refinement of a higher-symmetry structure is inherently underdetermined. When one atom is incorrectly placed, the fit may be improved if a correlated atom is likewise incorrectly placed. By creating an overall tendency towards misplacement, this correlated behavior can give rise to phantom modes, which are inactive in reality, but still give rise to split mode-amplitude histograms. Fortunately, phantom modes have certain characteristics that allow us to distinguish them from truly active modes. Most importantly, fixing the amplitude of an active mode to zero should noticeably increase the *R* value of the fit, whereas fixing the amplitude of a phantom mode at zero would not be expected to increase *R* under normal circumstances. Because phantom-modes are inherently correlated to one another, however, simple *R*-value tests don't always make their identification easy. But with some care, phantoms can be identified and eliminated by iterative testing. The procedure that we employed requires that we define some terms.

(1) Initial mode set: An initial collection of modes whose mode-amplitude histograms are sufficiently well split to be interesting candidates.
(2) Current mode set: The set of candidate modes that we are considering at any given step of the procedure.
(3) "Measure *R*": Deactivate all modes (i.e. fix the amplitudes to zero) that are not in the mode set, and determine the minimum number (*N*) of least-squares cycles needed to achieve a representative (but underdetermined) best fit involving these modes. *N* should be large enough to be effective, but no larger than necessary, as this process will be repeated many times. For this example, we obtained a reasonably good fit in under one minute on a desktop PC with $N = 100$ least-squares cycles, which yielded approximately 10 RLM convergences. Record the minimum R_{wp} value achieved during the course of these *N* cycles. This is what it means to "measure *R*" for a given mode set.

(4) Mode inclusion run: Perform an inclusion test on each mode that is not part of the current mode set one at time in a convenient order (e.g. lexicographical order by name). This means that for each mode, we activate it, measure *R* for the new mode set (adjust *N* as needed), and deactivate it again. Make a bar chart of measured *R* values from these tests, sorted according to ascending *R* value, and use it to determine which modes, if any, should be included in the current mode set. The best candidates will be those that lower *R* the most relative to the current mode set (i.e. the modes on the left-hand side of the graph). Then activate the included modes.

(5) Mode exclusion run: Perform an exclusion test on each mode in the current mode set one at time in a convenient order. This means that for each mode, we note the original amplitude, deactivate it, measure *R* for the new mode set (adjust *N* as needed), activate the mode again (i.e. allow it to be refined), and restore its original amplitude. Make a bar chart of the measured *R*-values from these tests, sorted according to descending *R* value, and use it to determine which, if any, modes should be excluded from the current mode set. The best candidates will be those that raise *R* the least relative to the current mode set (i.e. the modes on the right-hand side of the graph). Then deactivate the excluded modes.

Using Python scripts to run TOPAS from the command line, we applied the above method to our simulated $LaMnO_3$ data, beginning with 1000 RLM convergence cycles, from which we created amplitude histograms. Based on the histograms, we selected modes #[4, 5, 9, 23, 27, 33, 45, 46, 47, 53, 54, 59] as our initial mode set, though we could just as easily selected a few more or a few less. We treated this step as an effective mode inclusion run, and followed it with a mode exclusion run, which led to the elimination of modes #[23, 33, 46]. Next, we performed a mode inclusion run (Fig. 12a) in which we added all modes that had an impact on the *R* value greater or equal to that of mode #60 (i.e. those that brought R_{wp} below 9.6%), namely modes #[37, 38, 40, 41, 30, 31, 28, 32, 44, 60]. Another mode exclusion run (Fig. 12b) led us to eliminate all modes with an impact less than or equal to that of mode #40 (i.e. those that took R_{wp} above

8.8%), namely modes #[40, 30, 31, 60, 47, 44, 27, 41, 28, 32, 59]. Additional inclusion-exclusion iterations required no changes to the mode set, which gave us considerable confidence in the result. The final mode set, #[4, 5, 9, 37, 38, 45, 53, 54], consisted only of active modes and was sufficient to identify the correct *Pnma* symmetry. One can, in principle, repeat the inclusion and exclusion steps as many times as desired.

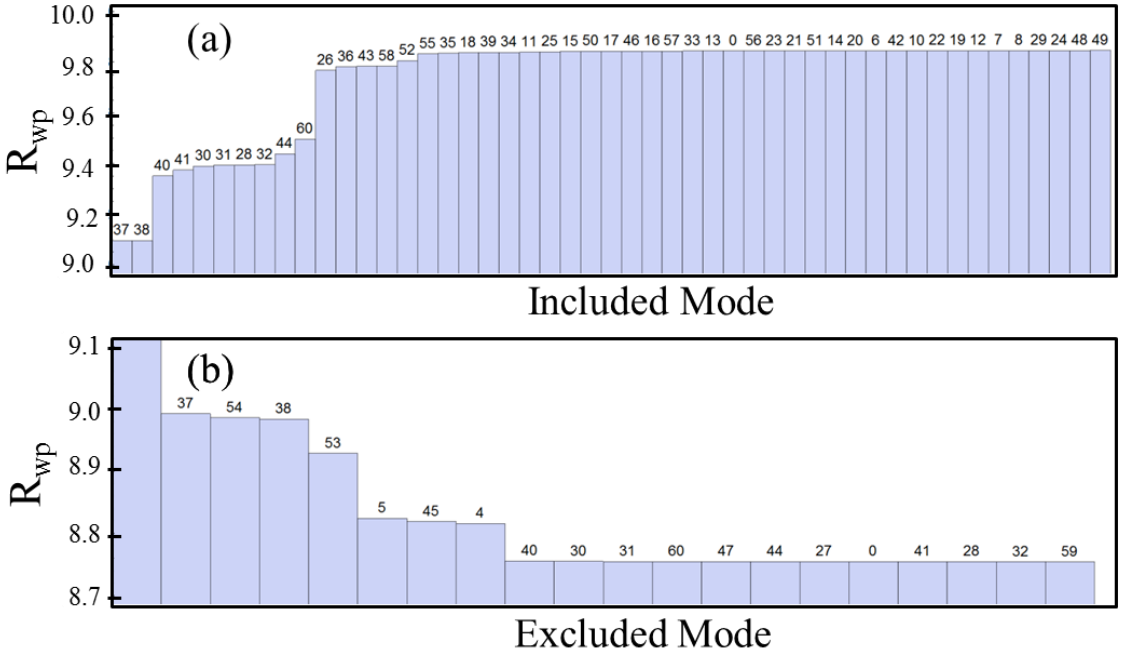

Figure 12: Mode inclusion and exclusion charts for the 57-parameter $LaMnO_3$ refinement against simulated x-ray data. (a) The inclusion run began with an initial set consisting of 9 modes (see text for details). *R* was measured and plotted for the other 48 modes alongside with the initial run with no modes included (labeled 0). This motivated the inclusion of 10 more modes(up through #60 in the graph). (b) The exclusion run started with these 19 modes and led to the exclusion of all but the first 8 (up through #4 in the graph). The first mode in the exclusion-run graph is mode #9, which has its top cropped off because it had a much larger impact than any of the others. It has amplitude 27.22.

The same approach was also applied separately to the x-ray and neutron data from WO_3 described above. For the x-ray data, the final mode set consisted of three WO_3 modes: #[14, 15, 16], while for the neutron data, it consisted of #[14, 15, 16, 37, 38, 46, 47, 78]. All of these modes are real – any phantom modes were eliminated. As per the

previous discussion of WO_3 histogram analysis, the modes identified from the x-ray data were insufficient to fully detect the correct $P2_1/n$ symmetry. In the case of the neutron data, however, the iterative mode inclusion and exclusion runs enhanced the sensitivity of the analysis so that the correct $P2_1/n$ symmetry was conclusively established without using the complementary x-ray information.

Shortcuts

If we were to initially include all modes, a single mode exclusion run might be sufficient to correctly identify the most important active modes without picking up any phantom modes. With such a large number of refined amplitudes, however, each least squares cycle can be very slow, and the number of cycles required to achieve a single convergence can become very large. We tried this approach on the x-ray data from WO_3 and found it to be computationally prohibitive. Using a small initial mode set based on mode histograms proved far more efficient. We also tried starting with an empty mode set which worked quite well for our WO_3 and $LaMnO_3$ examples, ultimately yielding the same final mode sets with much lower overall computational expense. This is, perhaps, the best option of all, though a greater number of inclusion/exclusion iterations are needed due to the lack of initial mode information.

The generation and use of mode histograms has both pros and cons. On one hand, histograms deliver a superior initial mode set that already includes most of the important modes that can be detected, so that further analysis is relatively easy. On the other hand, the histograms are computationally expensive to generate for large diffraction datasets and require user interaction. Using a simplistic RLM approach, without any effort to

improve speed, it took about 72 hours to achieve 10×96 convergences in the WO_3 example above. We thus recommend trying an empty initial mode set first. If unsuccessful, one can then generate and visualize mode-amplitude histograms to obtain a higher-quality initial mode set.

Signal-to-noise limitations

We also tested the limit to which active-mode detection is sensitive to noisy diffraction data by refining the $LaMnO_3$ model against simulated x-ray data to which artificial Gaussian noise was added. For each noise-level studied, over a range between zero and the approximate average Bragg-peak height, histogram sets comparable to those of Fig. 10 were generated. Because La is a strong scatterer, the active lanthanum modes proved to be robust against modest noise levels, but eventually disappeared (i.e. lost their distinctive histogram splittings) at higher noise levels. The large-amplitude $R_5^+(a,b,c)[\mathrm{La}]T_{1u}(b)$ (#5) mode persisted until the noise level exceeded the heights of all but the strongest Bragg peak. The weaker oxygen modes were difficult to detect, even with zero noise, and washed out completely at intermediate noise levels. It seems intuitive that when the magnitude of the noise becomes comparable to a mode's strongest contribution to the diffraction pattern, the mode's histogram splitting will disappear.

Phantom modes tend to become more prevalent at higher noise levels. This happens because inactive modes can take on larger amplitudes at higher noise levels without significantly impacting the *R* value of the fit. They are then more likely to be pulled to consistently non-zero amplitudes via correlations with other modes. After completing a fit that included phantom mode number #6 of La, we manually varied its

amplitude while holding all others constant and observed that the minimum *R* value occurred at a distinctly non-zero value (see Fig. 13) for this particular refinement model. The lowering of *R* was, however, very small and this effect was correlated with other trivial aspects of the fitting model such as peak shape or background description.

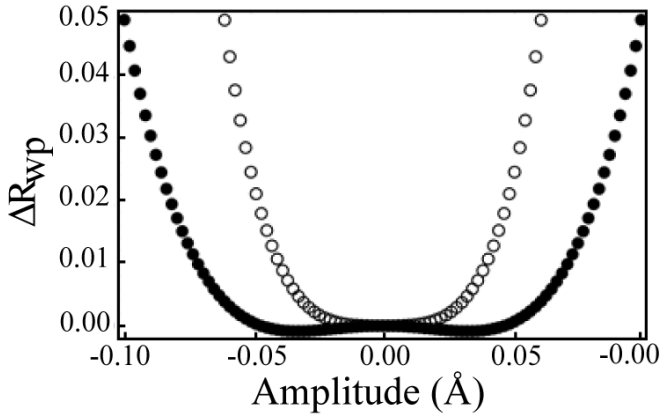

Figure 13: Plot of ΔR_{wp} vs mode amplitude for phantom mode #6 of the $LaMnO_3$ distortion. After allowing the active mode amplitudes to refine to their optimal values, they were all fixed, at which point, only mode #6 was varied over a narrow range around zero in steps of 0.002 Å. This analysis was performed at two different simulated noise levels: 50 (open circles) and 150 (filled squares). Deviations from the expected parabolic behavior are more apparent at higher noise levels.

Randomization Range

We next explored the optimal width for the uniform distribution of initial mode-amplitude values. We tested a variety of widths between 0.01 Å and 2.0 Å. For each candidate width, we ran a series of 1000 or more RLM-mode convergences against a low noise $LaMnO_3$ dataset and generated a histogram of converged R_{wp} values, as shown in Fig. 14. The R_{wp} axis on these histograms has been shifted so that the lowest R_{wp} value recorded in the refinement is defined to be zero. The histograms tend to be bimodal, possessing a low-*R* peak of relatively good fits that are likely to be close to the correct structure, and a broader high-*R* peak of relatively poor fits. At 0.75 Å, the size of the low-*R* peak indicates that about 1/5 of the convergence cycles end with a good fit. A

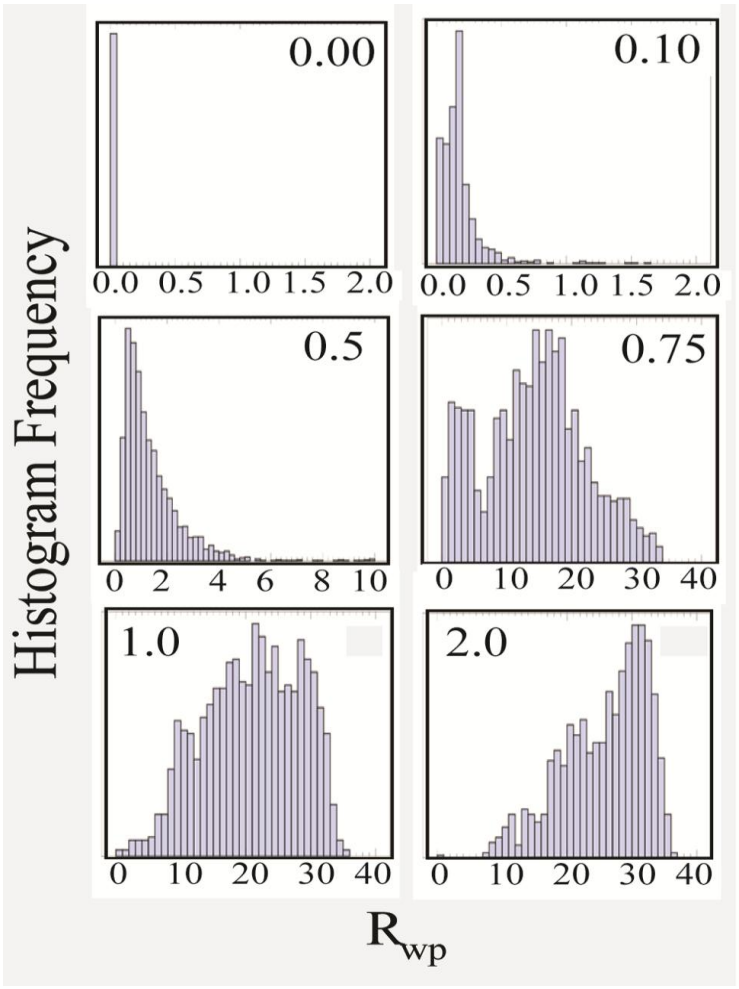

Figure 14: Histograms of *R* values for approximately 1000 convergence cycles of simulated annealing against simulate $LaMnO_3$ x-ray data at room temperature. Randomization amplitudes of 0, 0.1, 0.5, 0.75, 1.0, and 2.0 are shown.

width of 0.75 Å is probably the highest tolerable value in this example – reasonable structures are rare beyond that point. As the distribution width decreases towards zero, the frequency of a good fit increases. If the width of the distribution is too small, however, it may be impossible for the fit to escape a common local minimum. For this example, we determined that a randomization range of 0.1 Å provided sufficient randomness for effective parameter-space exploration, but not so much that the fits can't find their way home.

CONCLUSIONS

In contrast to general structure-determination work, the problem of determining a distorted structure lends itself to a symmetry-mode parameterization, where the key lies in determining which modes are active in the distortion. Assuming that the number of active modes is relatively small (which it usually is), the subsequent steps of detecting the distortion symmetry and refining the active mode amplitudes is simple and straightforward. The symmetry-mode description is ideal here because the interactions responsible for a distortion tend to activate a relatively small number of modes in order to break the parent symmetry in a specific way. It is also helpful that the mode amplitudes themselves tend to be small (deviations from zero).

We used modest-quality experimental laboratory x-ray and time-of-flight neutron powder data from RT WO_3, and also noisy simulated x-ray powder data from RT $LaMnO_3$, to test several different methods of identifying active modes and distinguishing them from phantom modes. In each case, the symmetry was lowered to *P*1 within the experimentally observed supercell and the structural degrees of freedom were parameterized in terms of symmetry-modes. A "repeated local minimization" (RLM) global optimization strategy was used to achieve a number of convergences equal to roughly ten times the number of free parameters, and separate histograms were generated for each symmetry-mode amplitude. Active modes were then identified by histograms that peaked at distinctly non-zero values. Even without generating histograms, we found that turning one mode on at a time and examining the effect on *R* allowed us to select a reasonably good initial mode set, from which the iterative use of inclusion and exclusion cycles eventually produced a high-quality final mode set.

Data noise was seen to increase the occurrence of phantom modes (inactive modes that tend toward non-zero amplitudes), which arise due to inter-mode correlations. Phantom modes are a problem because their inclusion will often result in a distortion symmetry that is too low. While it goes without question that good signal to noise levels are important, we did find that phantom modes can be identified and eliminated by iteratively including and excluding modes while monitoring the impact on the *R* value of the fit.

The displacive symmetry-mode analysis that we demonstrate here is highly analogous to the magnetic symmetry-mode approach of Wills (2000; Oszlanyi and Suto 2004; Stokes 2007). But there are also differences. The symmetry-mode tools employed here are equally applicable to the determination of other types of distortions, including atomic displacements, compositional order-disorder, and macroscopic strains. Rather than coupling every mode belonging to every irrep defined at a particular *k*-point in the Brillioun zone, we couple all of the modes capable of contributing to a *P*1-symmetry distortion within the observed supercell. And after identifying the modes that are active, we endeavor to identify the resulting distortion symmetry and to employ it in the final refinement.

The symmetry-mode approach to solving distorted structures has general utility, though its impact is greatest for challenging problems where other methods fail, particularly for powder diffraction data from subtle distortions that yield small peak splittings and weak superlattice reflections, and especially when the symmetry of the distortion is not known. When combined with a good global-optimization algorithm, we

believe that the symmetry-mode approach to whole profile fitting is the best way to squeeze all possible information out of a powder diffraction dataset.

REFERENCES

Campbell, B. J., G. Bellussi, L. Carluccio, G. Perego, A. K. Cheetham, D. E. Cox and R. Millini (1998). "The synthesis of the new zeolite, ERS-7, and the determination of its structure by simulated annealing and synchrotron X-ray powder diffraction." Chemical Communications(16): 1725-1726.

Campbell, B. J., J. S. O. Evans, F. Perselli and H. T. Stokes (2007). Rietveld refinement of structural distortion-mode amplitudes. IUCr Computing Commission Newsletter, International Union of Crystallography**:** 81-95.

Campbell, B. J., H. T. Stokes, D. E. Tanner and D. M. Hatch (2006). "ISODISPLACE: a web-based tool for exploring structural distortions." Journal of Applied Crystallography **39**: 607-614.

Carpenter, M. A., R. E. A. McKnight, C. J. Howard and K. S. Knight (2010). "Symmetry and strain analysis of structural phase transitions in $Pr_{0.48}Ca_{0.52}MnO_3$." Physical Review B **82**(9).

Černý, R. and V. Favre-Nicolin (2007). "Direct space methods of structure determination from powder diffraction: principles, guidelines and perspectives." Zeitschrift Fur Kristallographie **222**: 105-113.

Coelho, A. A. (2000). "Whole-profile structure solution from powder diffraction data using simulate annealing." Journal of Applied Crystallography **33**: 899-908.

Coelho, A. A. (2007). TOPAS Academic: General Profile and Structure Analysis Software for Powder Diffraction Data.

David, W. I. F., K. Shackland, L. B. McCusker and C. Baerlocher (2002). Structure Determination from Powder Diffraction Data, Oxford University Press.

Deem, M. W. and J. M. Newsam (1989). "Determination of 4-Connected Framework Crystal-Structures by Simulated Annealing." Nature **342**(6247): 260-262.

Diehl, R., G. Brandt and E. Salje (1978). "Crystal-Structure of Triclinic WO_3." Acta Crystallographica Section B **34**(Apr): 1105-1111.

Evans, I. R., J. A. K. Howard and J. S. O. Evans (2003). "α-$Bi_2Sn_2O_7$ – a 176 atom crystal structure from powder diffractiondata." Journal of Materials Chemistry **13**: 2098-2103.

Falcioni, M. and M. W. Deem (1999). "A biased Monte Carlo scheme for zeolite structure solution." Journal of Chemical Physics **110**(3): 1754-1766.

Favre-Nicolin, V. and R. Černý (2002). "FOX, 'free objects for crystallography': a modular approach to ab initio structure determination from powder diffraction." Journal of Applied Crystallography **35**: 734-743.

Fayon, F., I. J. King, R. K. Harris, R. K. B. Gover, J. S. O. Evans and D. Massiot (2003). "Characterization of the Room-Temperature Structure of SnP_2O_7 by ^{31}P Through-Space and Through-Bond NMR Correlation Spectroscopy." Chemistry of Materials **15**: 2234-2239.

Gover, R. K. B., N. D. Withers, S. Allen, R. L. Withers and J. S. O. Evans (2002). "Structure and Phase Transitions of SnP_2O_7." Journal of Solid State Chemistry **166**: 42-48.

Hahn, T. (2005). International Tables for Crystallography Vol. A, Kluwer Academic.

Howard, C. J., V. Luca and K. S. Knight (2002). "High-temperature phase transitions in tungsten trioxide - the last word?" Journal of Physics: Condensed Matter **14**(3): 377-387.

Howard, C. J. and H. T. Stokes (1998). "Group-theoretical analysis of octahedral tilting in perovskites." Acta Crystallographica Section B **54**: 782-789.

Howard, C. J. and Z. M. Zhang (2003). "Structures and phase transition in the layered perovskite $La_{0.6}Sr_{0.1}TiO_3$: a new orthorhombic structure solved from high-resolution diffraction in combination with group theoretical analysis." Journal of Physics: Condensed Matter **15**(26): 4543-4553.

Jones, R. H. and K. S. Knight (1997). "The structure of gamma-$Bi_2Sn_2O_7$ at 725 degrees C by high-resolution neutron diffraction: implications for bismuth(III)-containing pyrochlores." Journal of the Chemical Society-Dalton Transactions(15): 2551-2555.

Kariuki, B. M., S.-G. H., R. L. Johnston and K. D. M. Harris (1997). "The application of a genetic algorithm for solving crystal structures from powder diffraction data." Chemical Physics Letters **280**(3-4): 189-195.

Kennedy, B. J., Ismunandar and M. M. Elcombe (1998). "Structure and bonding in $Bi_2Sn_2O_7$." Materials Science Forum **278-2**: 762-767.

Le Page, Y. (1987). "Computer Derivation of the symmetry elements implied in a structure description." Journal of Applied Crystallography **20**: 254-269.

Masa, W. (2004). Crystal Structure Determination. New York, Springer.

Müller, M., R. E. Dinnebier, N. Z. Ali, B. J. Campbell and M. Jansen (2010). "Direct access to the order parameter: parameterized symmetry modes and rigid body movements as a function of temperature." Materials Science Forum **651**: 79-95.

Orobengoa, D., C. Capillas, M. I. Aroyo and J. M. Perez-Mato (2009). "AMPLIMODES: symmetry-mode analysis on the Bilbao Crystallographic Server." Journal of Applied Crystallography **42**(820-833).

Oszlanyi, G. and A. Suto (2004). "Ab initio structure solution by charge flipping." Acta Crystallographica Section A **60**(2): 134-141.

Perez-Mato, J. M., D. Orobengoa and M. I. Aroyo (2010). "Mode crystallography of distorted structures." Acta Crystallographica Section A **66**: 558-590.

Rodriguez-Carvajal, J., M. Hennion, F. Moussa and A. H. Moudden (1997). "Neutron-diffraction study of the Jahn-Teller transition in stoichiometric $LaMnO_3$." Physical Review B **57**(6): 3189-3192.

Shankland, K., W. I. F. David and T. Csoka (1997). "Crystal structure determination from powder diffraction data by the applictaion of a genetic algorithm." Zeitschrift Fur Kristallographie **210**(8): 550-552.

Shankland, K., A. J. Markvardsen, C. Rowlatt, N. Shankland and W. I. F. David (2010). "A benchmark method for global optimization problems in structure determination from powder diffraction data." Journal of Applied Crystallography **43**: 401-406.

Shannon, R. D., J. D. Bierlein, J. L. Gillson, G. A. Jones and A. W. Sleight (1980). "Polymorphism in $Bi_2Sn_2O_7$." Journal of Physics and Chemistry of Solids **41**(2): 117-122.

Stinton, G. W., M. R. Hampsons and J. S. O. Evans (2006). "The 136-Atom Structure of ZrP_2O_7 and HfP_2O_7 from Powder Diffraction Data." Inorganic Chemistry **45**: 4352-4358.

Stokes, H. T. (2007). Solid State Physics for Advanced Undergraduate Students. Provo, BYU Academic Publishing.

Stokes, H. T. and D. M. Hatch (1987). Isotropy subgroups of the 230 crystallographic space groups. Singapore, World Scientific.

Tanasaki, S. (1960). "Crystal structure of monoclinic tungsten trioxide at room temperature " Journal of the Physical Society of Japan **15**(4): 573-581.

Tremayne, M., B. M. Kariuki, K. D. M. Harris, K. Shankland and K. S. Knight (1997). "Crystal Structure Solution from Neutron Powder Diffraction Data by a New Monte Carlo Approach Incorporating Restrained Relaxation of the Molecular Geometry." Journal of Applied Crystallography **30**: 968-974.

Vogt, T., P. M. Woodward and B. A. Hunter (1999). "The high-temperature phases of WO_3." Journal of Solid State Chemistry **144**(1): 209-215.

Wills, A. S. (2000). "A new protocol for the determination of magnetic structures using simulated annealing and representational analysis (SARAh)." Physica B **276**: 680-681.

Woodward, P. M., A. W. Sleight and T. Vogt (1995). "Structure refinement of triclinic tungsten trioxide." Journal of Physics and Chemistry of Solids **56**(10): 1305-1315.

Woodward, P. M., A. W. Sleight and T. Vogt (1997). "Ferroelectric tungsten trioxide." Journal of Solid State Chemistry **131**(1): 9-17.

Zhao, J., N. L. Ross, R. J. Angel, M. A. Carpenter, C. J. Howard, D. A. Pawlaka and T. Lukasiewicz (2009). "High-pressure crystallography of rhombohedral $PrAlO_3$ perovskite." Journal of Physics-Condensed Matter **21**(23).