JOURNAL OF APPLIED CRYSTALLOGRAPHY
ISSN 1600-5767

Theoretical and computational improvements to the algebraic method for discovering cooperative rigid-unit modes

Branton J. Campbell,* Harold T. Stokes, Tyler B. Averett, Shae Machlus and Christopher J. Yost

Physics and Astronomy, Brigham Young University, Provo, Utah 84602, USA. *Correspondence e-mail: branton_campbell@byu.edu

Received 20 May 2021
Accepted 7 September 2021

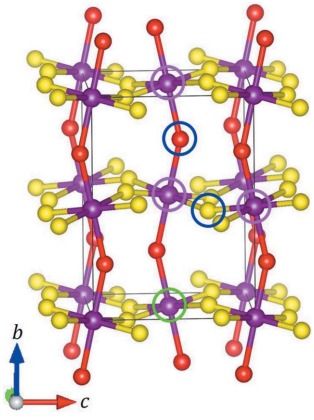
Edited by A. H. Liu, HPSTAR and Harbin Institute of Technology, People's Republic of China

Keywords: rigid-unit modes; symmetry modes; group representation theory; singular value decomposition; Gaussian elimination.

A linear-algebraic algorithm for identifying rigid-unit modes in networks of interconnected rigid units has recently been demonstrated. This article presents a series of enhancements to the original algorithm, which greatly improve its conceptual simplicity, numerical robustness, computational efficiency and interpretability. The improvements include the efficient isolation of constraints, the observation of variable-block separability, the use of singular value decomposition and a quantitative measure of solution inexactness.

1. Introduction

In crystal structures containing an interconnected network of rigid units (*e.g.* polyhedra, molecules, clusters *etc.*), cooperative patterns of internal distortion that leave the rigid unit undistorted are often referred to as rigid-unit modes, or RUMs (Giddy *et al.*, 1993; Dove *et al.*, 1993, 2000; Saint-Gregoire & Smirnov, 2021). A linear-algebraic algorithm for identifying small-amplitude cooperative RUMs has recently been demonstrated (Campbell *et al.*, 2018; Phillips, 2018). The algorithm allows interconnectedness along one, two or all three crystal dimensions, accommodates multiple connection points (*e.g.* edge or face sharing rather than vertex sharing only) between rigid units, and allows a single atom to be shared amongst more than two rigid units. When applying the algorithm to a complicated crystal structure, some of the difficulties encountered include (1) the creation and storage of very large matrices, (2) the slow speed of computer operations on very large matrices, (3) the accumulation of round-off errors during the row reduction of large matrices, and (4) the treatment of numerical tolerances during row reduction, especially for inexact/approximate solutions associated with quasi-RUMs (*i.e.* RUMs that slightly distort some of the affected rigid bodies). In the present work, we introduce several theoretical simplifications to the original algorithm that address these issues.



2. Notation and terminology

Here, we assume a familiarity with the notation and terminology of Campbell *et al.* (2018), where $N_{\rm pivot}$ is the number of symmetry-unique pivot atoms (SUPAs) in the child structure, $N_{\rm shared}$ is the number of symmetry-unique shared atoms (SUSAs) in the child structure, $N_{\rm free}$ is the number of free and independent rotational and displacive parameters belonging to rigid-unit pivot atoms in the child structure, n_α is the number of directly connected pivot atoms (DCPAs) of the αth

SUSA, $N_{\rm pairs} = \sum_\alpha n_\alpha$ is the total number of SUSA–DCPA pairings, and $N_{\rm eqs} = 3N_{\rm pairs}$ is the number of equations that define the cooperative RUM-search problem. In the common case that every SUPA has exactly two DCPAs, we have $N_{\rm pairs} = 2N_{\rm shared}$, though the network connectivity can be more complicated than this in general. The relationships between SUPAs, SUSAs and DCPAs is illustrated in Fig. 1.

With three vector components per shared atom, the total number of shared-atom displacement parameters (SADPs) is $3N_{\rm shared}$. We then define the $N_{\rm eqs} \times 3N_{\rm shared}$ matrix $\mathbf{T}_U$ to have three rows (for the SUSA-displacement vector) per SUSA–DCPA pair and one column per SADP. Each row of $\mathbf{T}_U$ contains only one nonzero element (value equal to 1), located in the column of the corresponding SUSA-displacement component. Note that if each shared atom were connected to only one pivot (not really 'shared' in that case), $\mathbf{T}_U$ would simply be a square identity matrix. But when each SUSA is connected to multiple DCPAs, each SUSA-displacement component appears in multiple equations, each unique row of $\mathbf{T}_U$ is duplicated at least once and each column of $\mathbf{T}_U$ has multiple nonzero elements. $\mathbf{T}_S$ is an $N_{\rm eqs} \times N_{\rm free}$ numerical matrix, which captures the effect of each free and independent pivot-atom parameter (rotational or displacive) of the child structure on the displacement of each SUSA via each of its DCPAs.

$\mathbf{T}_A$ is an $N_{\rm free} \times N_{\rm free}$ numerical matrix derived from group representation theory, wherein each column indicates the pattern of free and independent pivot-atom rotations and displacements induced by a given symmetry mode. In this context, symmetry modes are patterns of atomistic rotations and displacements that prove to be basis functions of the irreducible representations of the symmetry group of the undistorted parent structure. These symmetry modes, which are derived using group representation theory, have elegant orthogonality and completeness properties (Bradley & Cracknell, 1972, Section 1.2; Campbell *et al.*, 2006) that are essential for the algebraic RUM-search method (Campbell *et al.*, 2018) and to the present work. The number of symmetry modes must always be equal to $N_{\rm free}$. Here, for convenience, we further define the $(N_{\rm eqs} \times N_{\rm free})$-dimensional matrix $\mathbf{T}_{SA} \equiv -\mathbf{T}_S\mathbf{T}_A$. Note that we abbreviate the term 'irreducible representation' as 'irrep' for convenience and refer to the abstract order-parameter vector of a given irrep as its order-parameter direction (OPD).

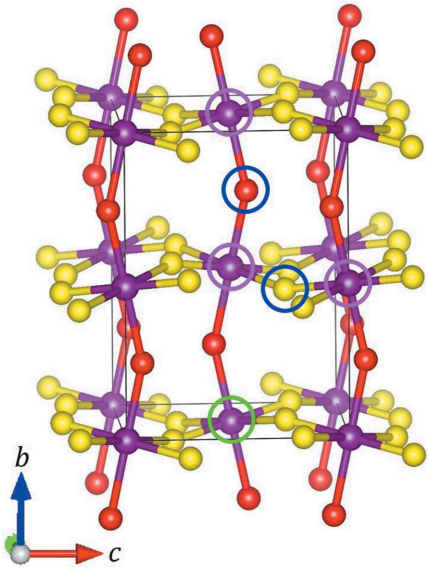


Figure 1
In this illustration of a distorted *Pnma*-symmetry ABO_3 perovskite ($2^{1/2} \times 2 \times 2^{1/2}$ supercell), the set of symmetry-equivalent *B*-site cations is purple, the two sets of symmetry-equivalent oxygen atoms are red and yellow, respectively, and the *A*-site cations are omitted for visual clarity. We can arbitrarily select any *B*-site atom as the symmetry-unique pivot atom (circled in green) and two inequivalent oxygen atoms as the symmetry-unique shared atoms (circled in blue), both of which have two directly connected pivot atoms (circled in purple).

The $N_{\rm eqs} \times (3N_{\rm shared} + N_{\rm free})$-dimensional matrix $\mathbf{M} = \{\mathbf{T}_U \,|\, \mathbf{T}_{SA}\}$ is constructed so that the columns of $\mathbf{T}_U$ are positioned before the columns of $\mathbf{T}_{SA}$. When identical rows of $\mathbf{T}_U$ have non-identical counterparts in $\mathbf{T}_{SA}$, a SUSA is simultaneously displaced in different directions by its DCPAs, which introduces a constraint on the pattern of pivot-atom rotations and displacements.

If $\mathbf{v} = \{\mathbf{v}_U \,|\, \mathbf{v}_A\}$ is a $(3N_{\rm shared} + N_{\rm free})$-dimensional vector of SUSA-displacement parameters and symmetry-mode amplitudes, then each non-trivial solution to the equation

$$\mathbf{M}\mathbf{v} = 0 \qquad (1)$$

is a small-angle RUM. In fact, the null space of $\mathbf{M}$ is the vector space of all possible small-angle RUMs. The last $N_{\rm free}$ components (*i.e.* the $\mathbf{v}_A$ part) of each null-space basis vector provide the coefficients of the linear combination of pivot-atom symmetry modes (rotational/displacive) that define a basis RUM.

Often, even a highly complicated network of interconnected rigid units will have only a small number of RUMs. In other words, for an adequately interconnected network, even a very large $\mathbf{M}$ matrix tends to have a low-dimensional null space.

3. Examples

Consider the tungsten bronze structures (M_xWO_3), which are composed of 3D networks of corner-sharing WO_6 octahedra and charge-balancing metal (M) atoms. The aristotype tetragonal tungsten bronze (TTB) structure has space group *P*4/*mbm* (No. 127) and ten tungsten atoms per unit cell (Wachsmann & Jacobs, 1995; Magnéli, 1949) and exhibits a variety of lower-symmetry octahedral WO_6 tilting patterns (Takusagawa & Jacobson, 1976; Goreaud *et al.*, 1980; Triantafyllou *et al.*, 1997; Haydon & Jefferson, 2002; Smirnov & Saint-Grégoire, 2014; Whittle *et al.*, 2015; Campbell *et al.*, 2018). In order to include contributions from all special $\mathbf{k}$ vectors, Campbell *et al.* (2018) considered a *P*1-symmetry supercell of TTB with relative basis {(2, 0, 0), (0, 2, 0), (0, 0, 2)}, which is eight times larger than the parent, providing $N_{\rm pivot} = 80$ tungsten SUPAs and $N_{\rm shared} = 240$ oxygen SUSAs with two DCPAs each, so that $N_{\rm free} = 3N_{\rm pivot} = 240$ and $N_{\rm eqs} = 3 \times 2 \times 240 = 1440$. As a result, the matrices $\mathbf{T}_U$, $\mathbf{T}_{SA}$ and $\mathbf{M}$ have respective dimensions of 1440×720, 1440×240 and 1440×960.

The aristotype hexagonal tungsten bronze (HTB) structure has space group *P*6/*mmm* (No. 191) and nine tungsten atoms per unit cell (Magnéli, 1953), and also exhibits a variety of lower-symmetry octahedral WO_6 tilting patterns (Prinz *et al.*, 1992; Smirnov & Saint-Grégoire, 2014; Whittle *et al.*, 2015; Campbell *et al.*, 2018). A *P*1-symmetry supercell of HTB with relative basis $\{(4, 2, 0), (\bar{2}, 2, 0), (0, 0, 2)\}$ is 24 times larger than the parent, providing $N_{\rm pivot} = 72$ tungsten SUPAs and $N_{\rm shared} = 216$ oxygen SUSAs with two DCPAs each, so that $N_{\rm free} = 3N_{\rm pivot} = 216$ and $N_{\rm eqs} = 3N_{\rm shared} \times 2 = 1296$. As a result, the matrices $\mathbf{T}_U$, $\mathbf{T}_S\mathbf{T}_A$ and $\mathbf{M}$ have respective dimensions of 1296×648, 1296×216 and 1296×864.

Whereas each oxygen vertex is shared between exactly two WO_6 octahedra in TTB or HTB, the corner-sharing network of $(Al/Zn)O_4$ tetrahedra in $Ca_3Al_4ZnO_{10}$ (CAZO) offers more structural variety. Its orthorhombic structure contains 20 Al/Zn atoms and 40 O atoms per unit cell. In space group *Pbcm* (No. 57), four of the tetrahedral (T = Al/Zn) sites and seven of the O sites are symmetry unique. While most of the symmetry-unique oxygen atoms are shared between two tetrahedra, two are not: the multiplicity-four O1 site is connected only to a single T site, and the multiplicity-four O3 site is shared between three T sites (Kahlenberg, Hejny & Krüge, 2019). CAZO has been reported to exhibit tetrahedral tilting under high pressure (Kahlenberg, Hejny & Krüge, 2019), and the isostructural $Ca_3Al_4MnO_{10}$ compound might be expected to do the same (Kahlenberg, Albrecht *et al.*, 2019). In order to include contributions from all special **k** vectors, we consider a *P*1-symmetry supercell of the *Pbcm*-symmetry parent cell with relative basis $\{(2, 0, 0), (0, 2, 0), (0, 0, 2)\}$, which is eight times larger than the parent, providing $N_{\rm pivot} = 160$ T-site SUPAs and $N_{\rm shared} = 320$ oxygen SUSAs (including the O1 sites). As for the β–α phase transition in quartz, we find that the TO_4 tetrahedra of CAZO must be permitted to simultaneously rotate and translate in order to remain rigid (Campbell *et al.*, 2018), which doubles the number of pivot degrees of freedom, so that $N_{\rm free} = 2 \times 3N_{\rm pivot} = 2 \times 480 = 960$. Because every neglected O1-site equation is balanced by an extra O3-site equation, the number of equations will be as if the oxygen vertices were all two-connected, so that $N_{\rm eqs} = 3 \times 2 \times 320 = 1920$. As a result, the matrices $\mathbf{T}_U$, $\mathbf{T}_{SA}$ and $\mathbf{M}$ have respective dimensions of 1920×960, 1920×960 and 1920×1920.

4. Isolating the constraints on the RUM space

The matrices $\mathbf{M}$ in these examples are so large as to be somewhat unwieldy. And in general, the generation, manipulation, operations and storage of very large matrices require considerable computational power, time and storage space. Because we anticipate the need to apply the algebraic RUM-search algorithm to systems much larger than these in the future, improvements are needed. Fortunately, the $\mathbf{M}$ matrix has some special internal structure that can be exploited towards this end.

We can use elementary row operations to permute the rows of $\mathbf{M}$ into two row blocks, such that the upper block contains only a single instance of each unique row of $\mathbf{T}_U$ (*i.e.* the information about the relationship between the given shared atom and the first pivot atom to which it is connected). The second row block of $\mathbf{M}$ then contains any duplicate copies of the unique rows of $\mathbf{T}_U$. The result is a matrix $\mathbf{M}$ of the form

$$\mathbf{M} = \begin{Bmatrix} \mathbf{T}_{U1} & \mathbf{T}_{SA1} \\ \mathbf{T}_{U2} & \mathbf{T}_{SA2} \end{Bmatrix}, \tag{2a}$$

where $\mathbf{T}_{U1}$ is the $3N_{\rm shared} \times 3N_{\rm shared}$ identity matrix, $\mathbf{T}_{U2}$ is a $3(N_{\rm pairs} - N_{\rm shared}) \times 3N_{\rm shared}$ matrix of zeros and ones, the $3N_{\rm shared} \times N_{\rm free}$ matrix $\mathbf{T}_{SA1}$ contains those rows of $\mathbf{T}_{SA}$ associated with the first DCPA of each SUSA, and the $3(N_{\rm pairs} - N_{\rm shared}) \times N_{\rm free}$ matrix $\mathbf{T}_{SA2}$ contains those rows of $\mathbf{T}_{SA}$ associated with all other DCPAs of each SUSA. Then, we apply additional elementary row operations to zero the $\mathbf{T}_{U2}$ block without altering the $\mathbf{T}_{U1}$ or $\mathbf{T}_{SA1}$ blocks in order to achieve the form

$$\mathbf{M}' = \begin{Bmatrix} \mathbf{1} & \mathbf{T}_{SA1} \\ \mathbf{0} & \mathbf{T}_0 \end{Bmatrix}, \tag{2b}$$

where the matrix $\mathbf{T}_0$ has the same dimensions as $\mathbf{T}_{SA2}$.

In the common case where each SUSA has exactly two DCPAs, $\mathbf{T}_{U2}$ and $\mathbf{T}_{U1}$ are identical (both identity matrices), $\mathbf{T}_{SA2}$ has the same dimensions as $\mathbf{T}_{SA1}$, and $\mathbf{T}_0 = \mathbf{T}_{SA2} - \mathbf{T}_{SA1}$. This is the case with the tungsten bronze examples, where the dimensions of $\mathbf{T}_0$ are $3N_{\rm shared} \times N_{\rm free}$. But in general, the structure of $\mathbf{T}_0$ can be more complicated and will depend on the number of DCPAs connected to each individual SUSA. In any case, because the locations of the nonzero values in $\mathbf{T}_U$ are prescribed in advance, it is easy to generate $\mathbf{T}_0$ directly without the need to perform any row operations.

Recall the $N_{\rm free}$ columns' of $\mathbf{T}_{SA1}$ and $\mathbf{T}_0$ correspondence to the $N_{\rm free}$ rotational and displacive symmetry modes of the pivot atoms. The identity matrix in the upper left-hand corner of $\mathbf{M}'$ makes it clear that every SADP is either dependent on one or more symmetry modes or fixed at zero; none of the SADPs are independent parameters. This special structure of $\mathbf{M}'$ ensures that $\mathbf{T}_0$ alone shapes the interdependencies amongst symmetry modes that define the vector space of independent RUMs, the determination of which is the essence of the algebraic RUM-search algorithm. In other words, there is a one-to-one correspondence between the solutions of equation (1) and the solutions of

$$\mathbf{T}_0\mathbf{v}_A = 0. \tag{3}$$

Thus, only the null space of $\mathbf{T}_0$ (rather than all of $\mathbf{M}$) needs to be determined. The upper portion of $\mathbf{M}'$ places no additional restrictions on the RUMs of the system, but merely reveals the pattern of SADPs associated with each basis RUM of the null space of $\mathbf{T}_0$.

The computational complexity of a typical algorithm for determining the null space of a matrix with $n_{\rm r}$ rows and $n_{\rm c}$ columns is of order $O(n_{\rm r}n_{\rm c}^2)$. Because $\mathbf{T}_0$ is generally much smaller than $\mathbf{M}$, an algorithm that determines only the null space of $\mathbf{T}_0$ rather than of the full matrix $\mathbf{M}$ is computationally more efficient. The size of $\mathbf{T}_0$ is 648×216 for HTB, 720×240 for TTB and 960×960 for CAZO. Owing to the reduced row and column counts relative to the full matrix $\mathbf{M}$, the expected

complexity reduction factors offered by using $\mathbf{T}_0$ are 32 for HTB and TTB, and eight for CAZO.

5. Singular value decomposition

An arbitrary $n_r \times n_c$ numerical matrix $\mathbf{A}$ always has a unique singular value decomposition (SVD) (Strang, 2016, ch. 7), which can be expressed as a product $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^t$, where $\mathbf{U}$ is an $n_r \times n_r$ unitary matrix, $\mathbf{V}$ is an $n_c \times n_c$ unitary matrix, and $\mathbf{\Sigma}$ is an $n_r \times n_c$ matrix such that elements on the main diagonal values are all non-negative and appear in order of decreasing magnitude and such that all elements off the main diagonal are zero. In our application of SVD, the matrix $\mathbf{A}$ is real valued, so that $\mathbf{U}$ and $\mathbf{V}$ are also orthogonal (a special case of unitary). The diagonal values of $\mathbf{\Sigma}$ are called the singular values of $\mathbf{A}$. The orthonormal columns of $\mathbf{V}$ (or rows of $\mathbf{V}^t$) are called the right-singular vectors (RSVs) of $\mathbf{A}$, and the columns of $\mathbf{U}$ are called the left-singular vectors (LSVs) of $\mathbf{A}$.

Each row of $\mathbf{\Sigma}\mathbf{V}^t$ is an RSV $\mathbf{v}_i$ of $\mathbf{A}$ multiplied by the corresponding singular value σ_i, so that $\mathbf{\Sigma}\mathbf{V}^t\mathbf{v}_i = \sigma_i\mathbf{e}_i$, where $\mathbf{e}_i$ is the n_r-dimensional unit vector with a one in the ith component and zeros everywhere else. It is then obvious that the basis vectors of the null space of $\mathbf{A}$ are those rows of $\mathbf{V}^t$ that correspond to zero singular values in $\mathbf{\Sigma}$. Approximate null-space vectors (*i.e.* those of approximate RUMs) can also be identified as rows of $\mathbf{V}^t$ corresponding to acceptably small but nonzero singular values. The matrix $\mathbf{U}$ has no impact on the null space of $\mathbf{A}$ and so can be neglected in the present context.

The SVD of the symmetric matrix $\mathbf{A}^t\mathbf{A} = (\mathbf{U}\mathbf{\Sigma}\mathbf{V}^t)^t\mathbf{U}\mathbf{\Sigma}\mathbf{V}^t = \mathbf{V}\mathbf{\Sigma}^2\mathbf{V}^t$ shows that its LSVs are identical to its RSVs, and also identical to the RSVs of $\mathbf{A}$. In fact, since $\mathbf{\Sigma}^2$ is a square diagonal matrix, the SVD of $\mathbf{A}^t\mathbf{A}$ is also its eigenvalue decomposition, so that the RSVs and squared singular values of $\mathbf{A}$ are also the eigenvectors and eigenvalues of $\mathbf{A}^t\mathbf{A}$.

The algebraic RUM-search algorithm of Campbell *et al.* (2018) determined the null space of the matrix $\mathbf{M}$ by first bringing it to reduced row-echelon form through the process of Gaussian elimination using the `RowReduce` function of the *Mathematica 11.1* (Wolfram, 2017) programming language with a user-specified tolerance factor. A Gaussian-elimination algorithm must decide, on the basis of a binary comparison against a user-specified tolerance factor, whether to rescale a small matrix pivot up to 1 by multiplying the entire pivot row by a large number or to round it down to 0 and leave the rest of the row unchanged. Rescaling to 1 declares the mode either to be dependent on some other mode or to be impossible (not a RUM), whereas rounding to 0 identifies the corresponding mode as a RUM. The decision to scale a value that should be zero up to one, or to round a truly nonzero value down to zero, has a nonlinear effect on subsequent pivot decisions and on each null-space vector.

As a result, Gaussian elimination tends to be unstable when small but nonzero 'matrix pivots'[1] are encountered (Golub & Van Loan, 1996, Sections 3.3 and 5.5), though partial pivoting is known to help (Higham & Higham, 1989). For exact null-space vectors (exact RUMs), small but nonzero matrix pivots often arise because of uncertainties in the atomic positions and lattice parameters of the parent structure, and the subsequent accumulation of round-off errors during row operations, though both `RowReduce` and an optimized in-house code for Gaussian elimination with partial pivoting tend to handle this scenario well. For approximate null-space vectors (quasi-RUMs), however, Gaussian elimination proved to be a significant hindrance; when the tolerance factor is raised high enough to accept one or more quasi-RUMs, sudden changes in the other null-space vectors are often quite noticeable.

For determining the exact or approximate null space of $\mathbf{T}_0$, SVD is superior to Gaussian elimination (Trefethen & Bau, 1997, pp. 28, 84 and 143), though the two approaches have comparable computational complexity. The process of computing the SVD is strictly linear, even when dealing with small but nonzero singular values. The only nonlinear decision to be made comes after the decomposition, when the singular values are compared against the tolerance factor to determine which RSVs are accepted as RUMs. Because the singular values and vectors have already been determined by that point, the decision to accept or reject one RSV has no impact on the accuracy of the corresponding RSV and no impact at all on the acceptance or accuracy of any other RSV, hence the stability of the SVD approach.

6. Block separability

The determination of the null space of $\mathbf{T}_0$ would be greatly simplified if it could be broken up into smaller pieces that could be treated separately without loss of generality. Owing to the special properties of symmetry modes, the parameter set of choice in the algebraic RUM-search algorithm, this is possible. In Appendix *A*, we prove that the pattern of multi-pivot shared-atom displacements induced by a symmetry mode of pivot-atom rotations is a basis function of the same irrep to which the original symmetry mode belongs, which means that the shared-atom displacement patterns also have elegant orthogonality and completeness properties. Note that the term 'displacement pattern' here refers to split-atom displacements wherein the non-cooperative displacements of the same SUSA by distinct DCPAs are distinct components of the overall multi-atom 'displacement' vector.

Imagine starting with a special matrix whose columns are these multi-pivot shared-atom-displacement basis functions. Column vectors belonging to different irreps, or even to different branches of the same irrep, are strictly orthogonal and therefore independent. The matrix $\mathbf{T}_{SA}$ can be constructed from this special matrix via row operations. Additional row operations further separate out the matrix $\mathbf{T}_0$ at the bottom, which inherits and fully encodes any column dependencies from the original matrix of irrep basis functions. This implies that columns of $\mathbf{T}_0$ associated with distinct irreps or distinct OPD branches of the same irrep are mutually independent, though not necessarily orthogonal. Thus, if we partition the columns (also known as modes) of $\mathbf{T}_0$ into mutually exclusive

[1] The word 'pivot' is used here in the context of Gaussian elimination rather than object rotation.

column blocks by irrep and OPD branch, there can be no inter-dependencies between the column spaces of different blocks. Because each column of $\mathbf{T}_0$ belongs to the symmetry mode of a specific branch of a specific order parameter of a specific irrep, we sometimes use the terms 'column' and 'symmetry mode' interchangeably.

The key result here can be restated in slightly different terms. Symmetry modes from different irreps or irrep branches must be independent owing to the properties of irrep basis functions. Thus, only symmetry modes from different order parameters, but belonging to the same irrep and branch, can be interdependent. Potentially interdependent symmetry modes get grouped within a block. The null spaces of distinct symmetry-mode blocks can be determined separately and one at a time without loss of generality. We refer to this property as 'block separability'. The block separability property provides a tremendous simplification of the RUM-search algorithm.

The columns within a block of $\mathbf{T}_0$ correspond in a one-to-one fashion to the distinct order parameters contributed to the child structure by that irrep, and may have interdependencies, so that some columns of a block with a non-trivial null space can be independent while others depend on the independent columns of that block. If a block with n columns has one independent RUM (*i.e.* one zero or near-zero singular value), we can view one participating column or symmetry mode (it does not matter which one) as the independent variable and view the other $n-1$ columns as depending on it in fixed proportions. However, it is more helpful to view the RUM as a linear combination of symmetry modes whose coefficients are the components of the associated normalized RSV multiplied by an overall variable amplitude. In general, a symmetry-mode block can have multiple zero or near-zero singular values, and therefore multiple RUMs that span a multi-dimensional null space. RSVs with finite singular values are non-cooperative or non-RUM like. The larger the singular value, the less cooperative the RSV.

Table 1 shows how the 240 rotational symmetry modes of the TTB example are distributed over six wavevector stars and 36 irreps of space group No. 127. For a given irrep, the second number in parentheses is the number of independent OPD branches, which is also the number of symmetry-mode blocks belonging to the irrep. The first number in parentheses is the number of order parameters contributed by the irrep, which is also the number of symmetry modes per block for that irrep. Irrep R_1, for example, provides four blocks with ten symmetry modes each, whereas irrep R_2 has four blocks of five symmetry modes each. The $R = [0, \frac{1}{2}, \frac{1}{2}]$ star then has eight blocks and 60 symmetry modes altogether. Summed over each of its **k** stars, the TTB example has 64 symmetry-mode blocks. However, only the eight blocks corresponding to Z_5^+, A_5^- and R_1 prove to have non-trivial null spaces. These are the eight basis RUMs described by Campbell *et al.* (2018). Table 1 also shows how the 216 rotational symmetry modes of the HTB example are distributed over eight wavevector stars and 42 irreps of space group No. 191. Summed over each of its **k** stars, the HTB example has 132 symmetry-mode blocks, though only the 11 blocks corresponding to A_3^+, A_6^+ and L_2^- prove to have non-trivial null spaces. These are the 11 basis RUMs described for HTB by Campbell *et al.* (2018).

Table 1
Irreps capable of contributing rotational order parameters to the child structures defined above for HTB, TTB and CAZO, grouped according to **k** vector.

In parentheses next to each irrep, we include the block size s_Γ and the block repetition d_Γ. In parentheses next to each **k** vector, we give the total number of symmetry modes contributed by irreps at that **k** vector ($\sum_\Gamma d_\Gamma s_\Gamma$) and the total number of independent OPD branches available at that **k** vector ($\sum_\Gamma d_\Gamma$). Irreps that produce RUMs are indicated with an asterisk (*). The data for TTB and HTB are modified with permission from Tables 3 and 5 of Campbell *et al.* (2018).

TTB (s_Γ, d_Γ).

Γ(30, 12)	**M**(30, 12)	**Z**(30, 12)	**A**(30, 12)
$\Gamma_1^+(1,1)$	$M_1^+M_4^+(2,2)$	$Z_1^+(1,1)$	$A_1^+A_4^+(2,2)$
$\Gamma_2^+(2,1)$	$M_2^+M_3^+(2,2)$	$Z_2^+(2,1)$	$A_2^+A_3^+(2,2)$
$\Gamma_3^+(2,1)$	$M_5^+(2,2)$	$Z_3^+(2,1)$	$A_5^+(2,2)$
$\Gamma_4^+(1,1)$	$M_1^-M_4^-(2,2)$	$Z_4^+(1,1)$	$A_1^-A_4^-(2,2)$
$\Gamma_5^+(6,2)$	$M_2^-M_3^-(1,2)$	$Z_5^+(6,2)$*	$A_2^-A_3^-(1,2)$
$\Gamma_1^-(2,1)$	$M_5^-(6,2)$	$Z_1^-(2,1)$	$A_5^-(6,2)$*
$\Gamma_2^-(2,1)$		$Z_2^-(2,1)$	
$\Gamma_3^-(2,1)$	**X**(60, 8)	$Z_3^-(2,1)$	**R**(60, 8)
$\Gamma_4^-(2,1)$	$X_1(10,4)$	$Z_4^-(2,1)$	$R_1(10,4)$*
$\Gamma_5^-(2,2)$	$X_2(5,4)$	$Z_5^-(2,2)$	$R_2(5,4)$

HTB (s_Γ, d_Γ).

Γ(9, 7)	**A**(9, 7)	**H**(18, 14)	**K**(18, 14)
$\Gamma_2^+(1,1)$	$A_2^+(1,1)$	$H_2(1,2)$	$K_2(1,2)$
$\Gamma_3^+(1,1)$	$A_3^+(1,1)$*	$H_3(1,2)$	$K_3(1,2)$
$\Gamma_4^+(1,1)$	$A_4^+(1,1)$	$H_4(1,2)$	$K_4(1,2)$
$\Gamma_5^+(1,2)$	$A_5^+(1,2)$	$H_5(1,4)$	$K_5(1,4)$
$\Gamma_6^+(2,2)$	$A_6^+(2,2)$*	$H_6(2,4)$	$K_6(2,4)$
L(27, 21)	**M**(27, 21)	**Λ**(54, 24)	**Q**(54, 24)
$L_2^+(1,3)$	$M_2^+(1,3)$	$\Lambda_1(1,6)$	$Q_1(1,6)$
$L_3^+(1,3)$	$M_3^+(1,3)$	$\Lambda_2(3,6)$	$Q_2(3,6)$
$L_4^+(1,3)$	$M_4^+(1,3)$	$\Lambda_3(3,6)$	$Q_3(3,6)$
$L_1^-(2,3)$	$M_1^-(2,3)$	$\Lambda_4(2,6)$	$Q_4(2,6)$
$L_2^-(2,3)$*	$M_2^-(2,3)$		
$L_3^-(1,3)$	$M_3^-(1,3)$		
$L_4^-(1,3)$	$M_4^-(1,3)$		

CAZO (s_Γ, d_Γ).

Γ(120, 8)	**X**(120, 8)	**S**(120, 4)	**Y**(120, 4)
$\Gamma_1^+(14,1)$	$X_1^+(14,1)$	$S_1(30,2)$*	$Y_1(30,2)$*
$\Gamma_2^+(16,1)$	$X_2^+(16,1)$*	$S_2(30,2)$	$Y_2(30,2)$
$\Gamma_3^+(14,1)$	$X_3^+(14,1)$*		
$\Gamma_4^+(16,1)$	$X_4^+(16,1)$*	**U**(120, 4)	**Z**(120, 4)
$\Gamma_1^-(14,1)$*	$X_1^-(14,1)$	$U_1(30,2)$*	$Z_1(30,2)$
$\Gamma_2^-(16,1)$*	$X_2^-(16,1)$	$U_2(30,2)$	$Z_2(30,2)$
$\Gamma_3^-(14,1)$*	$X_3^-(14,1)$		
$\Gamma_4^-(16,1)$*	$X_4^-(16,1)$	**R**(120, 4)	**T**(120, 4)
		$R_1R_2(30,4)$	$T_1T_2(30,4)$

Recall that the computational complexity of the determination of the null space of an $n_r \times n_c$ matrix is of order $O(n_r n_c^2)$. We partition the columns of $\mathbf{T}_0$ into blocks, where s_Γ is the block size, which is the number of columns in any block associated with irrep Γ (*i.e.* the number of instances of the irrep in the overall pattern), and d_Γ is the block repetition, which is the number of blocks contributed by the irrep Γ (*i.e.* the number of independent variables in the OPD, which is $\leq$ the irrep dimension), so that $n_c = \sum_\Gamma d_\Gamma s_\Gamma$. The reduced complexity is of order $O(n_r n_{rms}^2)$, and the complexity-

Table 2
Singular values from selected symmetry-mode blocks of matrix $\mathbf{T}_0$ for the HTB, TTB and CAZO examples.

All blocks with exact RUMs or suspected quasi-RUMs (*i.e.* having zero or near-zero singular values as indicated in bold) are presented, along with those of a few other blocks for contrast. Blocks from branches of the same irrep having identical singular values share a single row for compactness. The number s_Γ of symmetry modes per block is shown in the second column. The block repetition d_Γ is evident from the number of free variables in the OPD of each irrep. Up to six singular values are listed for each block in descending order. When $s_\Gamma > 6$, only the lowest five and the maximum singular values are given; those not listed are indicated by '...'.

Irrep/OPD (HTB)	s_Γ	$\sigma_{\max}$	$\sigma_{s\Gamma-4}$	$\sigma_{s\Gamma-3}$	$\sigma_{s\Gamma-2}$	$\sigma_{s\Gamma-1}$	$\sigma_{s\Gamma}$
$A_3^+(a)$	1	0.00000	...	...	...	...	**0.00000**
$L_2^-(a,b,c)$	2	0.34213	...	...	...	0.34213	**0.00000**
$A_6^+(a,b)$	2	0.48333	...	...	...	0.48333	**0.00000**
$A_5^+(a,0)$	1	0.09720	...	...	...	...	0.09720
$A_5^+(0,b)$	1	0.15625	...	...	...	...	0.15625
$\Gamma_5^+(a,0)$	1	0.09720	...	...	...	...	0.09720
$\Gamma_5^+(0,b)$	1	0.15625	...	...	...	...	0.15625
$\Gamma_4^+(a)$	1	0.51664	...	...	...	...	0.51664

Irrep/OPD (TTB)	s_Γ	$\sigma_{\max}$	$\sigma_{s\Gamma-4}$	$\sigma_{s\Gamma-3}$	$\sigma_{s\Gamma-2}$	$\sigma_{s\Gamma-1}$	$\sigma_{s\Gamma}$
$A_5^-(a,b)$	6	0.47680	0.42946	0.38615	0.25554	0.15626	**0.00000**
$Z_5^+(a,b)$	6	0.47663	0.42930	0.38613	0.25552	0.15695	**0.00000**
$R_1(a,b,c,d)$	10	0.46201	0.28393	0.27598	0.14625	0.13564	**0.00292**
$\Gamma_4^+(a)$	1	0.08603	...	...	...	...	0.08603
$Z_4^+(a)$	1	0.08603	...	...	...	...	0.08603
$X_2(a,b,c,d)$	5	0.20551	0.20551	0.17541	0.15330	0.12664	0.09626
$R_2(a,b,c,d)$	5	0.20551	0.20551	0.17541	0.15330	0.12664	0.09626
$M_2^+M_3^+(a,b)$	2	0.43852	...	...	...	...	0.37926

Irrep/OPD (CAZO)	s_Γ	$\sigma_{\max}$	$\sigma_{s\Gamma-4}$	$\sigma_{s\Gamma-3}$	$\sigma_{s\Gamma-2}$	$\sigma_{s\Gamma-1}$	$\sigma_{s\Gamma}$
$S_1(a,b)$	30	0.32480	...	0.02016	0.01110	0.00635	**0.00001**
$Y_1(a,b)$	30	0.32732	...	0.01907	0.01179	0.00731	**0.00001**
$\Gamma_4^-(a)$	16	0.28317	...	0.02638	0.01059	**0.00001**	**0.00000**
$X_2^+(a)$	16	0.28444	...	0.04256	0.01785	**0.00002**	**0.00001**
$\Gamma_3^-(a)$	14	0.27858	...	0.03111	0.02525	0.01321	**0.00000**
$X_3^+(a)$	14	0.31946	...	0.04550	0.02936	0.01413	**0.00001**
$\Gamma_1^-(a)$	14	0.32220	...	0.04890	0.02381	0.01935	**0.00000**
$X_4^+(a)$	16	0.31957	...	0.04264	**0.00001**	**0.00001**	**0.00000**
$\Gamma_2^-(a)$	14	0.32228	...	0.04821	**0.00002**	**0.00001**	**0.00000**
$U_1(a,b)$	30	0.32247	...	0.02297	0.02104	0.01193	**0.00060**
$Z_1(a,b)$	30	0.32497	...	0.02181	0.01272	0.01197	0.00189
$Z_2(a,b)$	30	0.32507	...	0.02206	0.01707	0.00737	0.00193
$Y_2(a,b)$	30	0.29390	...	0.01895	0.01263	0.00569	0.00481
$\Gamma_3^+(a)$	30	0.32697	...	0.07634	0.04432	0.03149	0.02145

reduction factor is $(n_c/n_{\rm rms})^2$, where $n_{\rm rms} = (\sum_\Gamma d_\Gamma s_\Gamma^2)^{1/2}$. The values of (s_Γ, d_Γ) are given in parentheses in Table 1 for each irrep contributing to the HTB, TTB and CAZO examples. According to these data, the complexity-reduction factor provided by block separability should be roughly $(216/20.78)^2 = 108$ for HTB, $(240/37.74)^2 \simeq 40$ for TTB and $(960/158.80)^2 \simeq 37$ for CAZO. HTB notably benefits from a large number of relatively small blocks.

An analysis of computational complexity alone fails to capture the greatest advantage of block separability, which is improved robustness against false-positive detection. When symmetry-mode blocks are not separated, the numerous row operations in one column propagate errors and uncertainties into all subsequent columns, whether in the same block or not, an effect that gets compounded as the pivot column advances. For large systems with many symmetry modes and symmetry-mode blocks, we observe that mutually independent blocks can still erroneously interfere with one another. This problem is entirely eliminated when block separability is exploited.

For TTB and HTB, each symmetry-mode block of each RUM-capable irrep contributes a single RUM. But in general, a block with multiple columns can give rise to multiple independent RUMs, as seen for the CAZO example in Table 2.

7. Singular values and tolerance factors

By treating each symmetry-mode block separately, we are no longer required to employ a single tolerance to simultaneously differentiate all quasi-RUMs and false positives according to their singular values, which completely resolves a key difficulty reported by Campbell *et al.* (2018). The tolerance can be chosen separately for each block if desired, which can be helpful. Recall that singular values are non-negative and appear in $\mathbf{\Sigma}$ in descending order. For a given block, each RSV with a singular value equal to zero or differing from zero only because of propagated uncertainties can be viewed as an independent exact RUM (*i.e.* it distorts no rigid units), and each RSV with a nonzero singular value below the block-specific tolerance can be viewed as an independent quasi-RUM (*i.e.* it introduces only minor distortions of the rigid units).

It is common for distinct branches of the OPD to produce symmetry-equivalent patterns of rigid-unit rotations and displacements within the child structure (*i.e.* different domains of the single-branch order parameter that are related by child symmetry operations). Naturally, the singular values of symmetry-mode blocks arising from such branches of the same irrep will be identical. When distinct branches of the OPD do not give rise to symmetry-equivalent patterns of rigid-unit rotations and displacements, their singular values will normally (but not necessarily) be different.

Singular values from selected symmetry-mode blocks of matrix $\mathbf{T}_0$ are given in Table 2 for each of the HTB, TTB and CAZO examples. For HTB, we detect six exact RUMs belonging to $A_3^+(a)$, $A_6^+(a,b)$ and $L_2^-(a,b,c)$, as expected. The A_3^+ irrep has only one block and one symmetry mode, which is a RUM, as identified by a singular value of zero. The three symmetry-mode blocks of irrep L_2^- share the same row of the table because they are equivalent by symmetry and have identical singular values; the two blocks of A_6^+ similarly share a row. Each block of L_2^- and A_6^+ has two singular values, only one of which is zero, and therefore has only one independent RUM. The blocks of A_5^+, Γ_4^+ and Γ_5^+ do not have RUMs and are provided merely for contrast. The blocks of irreps A_5^+ and Γ_5^+ have the smallest last singular value (0.09720) that does not qualify for RUM status. The block of irrep Γ_4^+ has the largest last singular value (0.51664) seen for HTB. Note that the *a* and *b* branches of $A_5^+(a,b)$, which are not symmetry equivalent, yield different singular values, as is the case for $\Gamma_5^+(a,b)$; and yet, the *a* branches (or the *b* branches) of A_5^+ and Γ_5^+, which

cannot be related by symmetry, have blocks with the same singular values, indicating that their non-cooperative RSVs split the shared atoms in very similar ways. In summary, a tolerance of 0.00001 is adequate for differentiating the singular values of RUMs from those of non-cooperative RSVs of HTB.

For TTB, we detect four exact RUMs (as seen in Table 2) belonging to irreps $Z_5^+(a,b)$ and $A_5^-(a,b)$, and four quasi-RUMs belonging to irrep $R_1(a,b,c,d)$, as expected. The four exact RUMs belonging to $A_5^-(a,b)$ and $Z_5^+(a,b)$ are clearly evident as symmetry-mode blocks containing a single zero singular value. We further identify the four blocks of $R_1(a,b,c,d)$ as quasi-RUMs owing to the smallness of their last singular value (0.00292). The blocks of $\Gamma_4^+(a)$ and $Z_4^+(a)$ have the next-smallest last (and only) singular value (0.08603), which is roughly 30 times larger than that of R_1, making it clear that the R_1 RSVs are far more cooperative or RUM like. The other TTB symmetry-mode blocks in the table have no RUMs and are presented merely for contrast. Irreps X_2 and R_2 were selected because they have curiously identical singular values, indicating that their blocks have some interesting structural similarities – only different in their action along the direction perpendicular to the TTB layers. There were also several other such pairs of irreps. The blocks of $M_2^+M_3^+$ have the largest last singular value (0.37926) seen for TTB. In summary, a tolerance of 0.003 is adequate for differentiating the singular values of RUMs from those of non-cooperative RSVs of TTB.

Owing to the presence of some non-shared polyhedral vertices in its framework, the CAZO example has a fairly large number of RUMs (as seen in Table 2). The symmetry-mode blocks belonging to irreps $\Gamma_1^-(a)$, $\Gamma_2^-(a)$, $\Gamma_3^-(a)$, $\Gamma_4^-(a)$, $X_2^+(a)$, $X_3^+(a)$, $X_4^+(a)$, $S_1(a,b)$ and $Y_1(a,b)$ have exact RUMs Some of these exact RUMs correspond to very small nonzero singular values of order 0.00001, which we judge to be the result of accumulated error propagation. Some of these blocks have multiple independent RUMs: two RUMs for each block of $\Gamma_4^-(a)$ and $X_2^+(a)$ and three RUMs for each block of $\Gamma_2^-(a)$ and $X_4^+(a)$. Each instance of a zero singular value within a block indicates a distinct exact RUM.

We now call attention to an important caveat. Because the RUMs of CAZO simultaneously require rigid-unit rotations and translations, it was necessary to include pivot-atom-displacement symmetry modes in the analysis. This allowed each of the ferroelectric irreps (Γ_2^-, Γ_3^- and Γ_4^-) to produce one purely displacive symmetry mode wherein the entire crystal shifts along a given crystal axis. Naturally, because such ferroelectric displacements do not distort any of the rigid bodies, they will have zero singular values in the present analysis. But we cannot refer to them as RUMs. After eliminating the ferroelectric displacement modes, the total number of exact RUMs from the CAZO example is 12 rather than 15.

The CAZO blocks belonging to irrep U_1 have a slightly larger last singular value of 0.00060, which still suggests a quasi-RUM. The next-smallest last singular values of 0.00189 and 0.00193 belong to the symmetry-mode blocks of irreps Z_1 and Z_2, respectively, followed by several irreps (including Y_2) with smallest last singular values of around 0.005. Although these last singular values of Z_1 and Z_2 are slightly smaller than that of the quasi-RUMs of TTB, we have other means of judging their RSVs to be inadequately RUM like, as we will explain in the next section. The symmetry-mode block of Γ_3^+ has the second-largest last singular value (0.32697) seen for CAZO.

8. Relative r.m.s. deviation

In addition to judging the smallness of a singular value, one should also visually validate the pattern of rigid-unit rotations and displacements arising from the corresponding RSV before declaring it to be a quasi-RUM. A quantitative measure of how far a given pattern deviates from exactness [*i.e.* the magnitude of $\mathbf{Mv}$ in equation (1)], or rather to what extent the rigid bodies become distorted, would also be helpful. We now introduce such a measure.

For an exact RUM, each of the DCPAs of a given SUSA induce precisely the same total displacement of the SUSA (via DCPA rotation and translation). For a quasi-RUM, however, the various DCPA-induced displacements of a given SUSA are generally different, so that one can define a distribution of two or more SUSA displacements, from which a mean displacement and a standard deviation from the mean can be computed. Then, considering all of the SUSAs of the child cell, we can define the Wyckoff-multiplicity-weighted r.m.s. average of the mean-displacement magnitudes and standard deviations.

Let $\mathbf{v}_A$ be an RSV of $\mathbf{T}_0$ (having $n_{\rm free}$ components), consider the vector $\mathbf{T}_{SA}\mathbf{v}_A$ (having $n_{\rm eqs}$ components), and let the three-dimensional vector $\mathbf{u}_{\beta_\alpha}$ be formed by taking the three elements of $\mathbf{T}_{SA}\mathbf{v}_A$ corresponding to the vector displacement of the αth SUSA induced by its βth DCPA (among n_α DCPAs for that SUSA) and transforming them from lattice coordinates into Cartesian coordinates using the child-cell parameters. We define the average and difference vectors

$$\mathbf{a}_\alpha = \frac{1}{n_\alpha}\sum_{\beta_\alpha}\mathbf{u}_{\beta_\alpha} \quad \text{and} \quad \mathbf{d}_{\beta_\alpha} = \mathbf{u}_{\beta_\alpha} - \mathbf{a}_\alpha, \tag{4}$$

and use them to construct the r.m.s. split-atom displacement and the standard deviation of the distribution of difference vectors for a given SUSA:

$$u_\alpha = \left(\frac{1}{n_\alpha}\sum_{\beta_\alpha}|\mathbf{u}_{\beta_\alpha}|^2\right)^{1/2} \quad \text{and} \quad d_\alpha = \left(\frac{1}{n_\alpha}\sum_{\beta_\alpha}|\mathbf{d}_{\beta_\alpha}|^2\right)^{1/2}. \tag{5}$$

Next, define h_α to be the Wyckoff multiplicity of the αth SUSA, define $h = \sum_\alpha h_\alpha$ to be the total number of SUSAs in the child unit cell, define

$$u = \left(\frac{1}{h}\sum_\alpha h_\alpha u_\alpha^2\right)^{1/2} \quad \text{and} \quad d = \left(\frac{1}{h}\sum_\alpha h_\alpha d_\alpha^2\right)^{1/2} \tag{6}$$

to be the respective multiplicity-weighted r.m.s. averages of the mean displacements and deviations from all of the SUSAs, and define the 'relative r.m.s. deviation' (RrmsD) to be

Table 3
The intermediate quantities used to calculate the relative r.m.s. deviations of the quasi-RUMs of the pre-distorted perovskite example from Campbell *et al.* (2018).

Mode	α	β	$\mathbf{u}_{\beta_\alpha}$	u_α (Å)	d_α (Å)	u (Å)	d (Å)	r
R_4^+	1	1	(0.00000, −0.20955, −0.97700)	0.99922	0.20955	0.82498	0.18108	0.21949
		2	(0.00000, +0.20955, −0.97700)					
	2	1	(0.00000, −0.63451, +0.15004)	0.72227	0.16501			
		2	(0.00000, −0.77183, −0.15004)					
M_3^+	1	1	(+0.1983, 0.0000, +0.0348)	0.20137	0.00000	0.80379	0.13827	0.17202
		2	(+0.1983, 0.0000, +0.0348)					
	2	1	(−0.6005, 0.0000, +0.8472)	0.97408	0.16934			
		2	(−0.7305, 0.0000, +0.5345)					

Table 4
Relative r.m.s. deviations (r) of the RSVs corresponding to selected singular values (σ) from Table 2 for the HTB, TTB and CAZO examples.

Values small enough to clearly indicate a RUM or quasi-RUM are indicated in bold.

	σ	r		σ	r
HTB			CAZO		
$A_3^+(a)$	**0.00000**	**0.00000**	$S_1(a,b)$	**0.00001**	**0.00004**
$L_2^-(a,b,c)$	**0.00000**	**0.00000**	$Y_1(a,b)$	**0.00001**	**0.00005**
$A_6^+(a,b)$	**0.00000**	**0.00000**	$\Gamma_4^-(a)$	**0.00000**	**0.00000**
$A_5^+(a,0)$	0.09720	0.53294		**0.00001**	**0.00003**
$A_5^+(0,b)$	0.15625	0.53294	$X_2^+(a)$	**0.00001**	**0.00004**
$\Gamma_5^+(a,0)$	0.09720	0.53294		**0.00002**	**0.00005**
$\Gamma_5^+(0,b)$	0.15625	0.53294	$\Gamma_3^-(a)$	**0.00000**	**0.00000**
$\Gamma_4^+(a)$	0.51664	1.00000	$X_3^+(a)$	**0.00001**	**0.00007**
			$\Gamma_1^-(a)$	**0.00001**	**0.00006**
TTB			$X_4^+(a)$	**0.00000**	**0.00003**
$A_5^-(a,b)$	**0.00000**	**0.00000**		**0.00001**	**0.00003**
$Z_5^+(a,b)$	**0.00000**	**0.00000**		**0.00001**	**0.00003**
$R_1(a,b,c,d)$	**0.00292**	**0.00423**	$\Gamma_2^-(a)$	**0.00000**	**0.00000**
$\Gamma_4^+(a)$	0.08603	0.38641		**0.00001**	**0.00003**
$Z_4^+(a)$	0.08603	0.38641		**0.00002**	**0.00006**
$X_2(a,b,c,d)$	0.09626	0.43504	$U_1(a,b)$	**0.00060**	**0.00443**
$R_2(a,b,c,d)$	0.09626	0.43504	$Z_1(a,b)$	0.00189	0.01326
$M_2^+M_3^+(a,b)$	0.37926	0.87147	$Z_2(a,b)$	0.00193	0.01644
	0.43852	0.94427	$Y_2(a,b)$	0.00481	0.04062
			$\Gamma_3^+(a)$	0.02145	0.15099
				0.32697	0.92987

$$r = d/u. \tag{7}$$

The RrmsD provides a quantitative and easily interpretable measure of the inexactness of the RSV (or any other linear combination of symmetry modes) as a solution of equation (3). It has a maximum value of 1, which applies to the case of a maximally non-cooperative mode pattern (*i.e.* the DCPAs of every SUSA act in directions that cancel to zero so that $\mathbf{d}_{\beta_\alpha} = \mathbf{u}_{\beta_\alpha}$), and a minimum value of 0 for an exact or perfectly cooperative solution (*i.e.* $\mathbf{u}_{\beta_\alpha} = \mathbf{a}_\alpha$ so that $\mathbf{d}_{\beta_\alpha} = 0$). The denominator u of the RrmsD can only be zero for a mode that displaces no shared atoms, which cannot happen since we have only considered modes that actually contribute to a distortion. For a slightly inexact solution (a quasi-RUM), r will have a small but positive value ($0 < r \ll 1$).

Consider the hypothetical pre-distorted perovskite example from Campbell *et al.* (2018), which was based on the octahedral rotations present in a *Pnma*-symmetry $2^{1/2} \times 2 \times 2^{1/2}$ supercell (a = 5.74731, b = 7.69287, c = 5.53667) of the cubic ($Pm\bar{3}m$) parent, and three rotational symmetry modes $R_4^+(a,-a,0)$, $X_5^+(a,a,0,0,0,0)$ and $M_3^+(0,a,0)$, of which both R_4^+ and M_3^+ prove to be quasi-RUMs (Rodríguez-Carvajal *et al.*, 1998; Howard & Stokes, 1998, 2005). Campbell *et al.* (2018) present the relevant matrix **M** towards the end of Section 2.7. The 12 numbers in each mode column can be partitioned into the three-dimensional vectors $\mathbf{u}_{\beta_\alpha}$ (α runs over two SUSAs of multiplicity $h_1 = 4$ and $h_2 = 8$, and β_α runs over $n_\alpha = 2$ DCPAs for each SUSA). The RrmsDs for this distorted perovskite example, and the intermediate quantities used to obtain them, are given in Table 3.

For each irrep listed in Table 2, the corresponding entry in Table 4 presents the RrmsDs of the RSVs corresponding to selected singular values. Each near-zero singular value is presented, along with the last singular values from symmetry-mode blocks of irreps that do not support cooperative RUMs, and even one very large singular value from each example. We observe that the RrmsD scales loosely with, but is not proportional to, the singular value. The last irrep presented for each example has a very large RrmsD and illustrates an extreme example of non-cooperative character.

For the HTB and TTB examples, the distinctions in Table 4 between the RrmsDs of RUMs and non-cooperative RSVs are striking. For the CAZO example, the distinction is less clear, though the RrmsDs still inform the decision on where to draw the line between RUMs and non-cooperative RSVs. Depending on the problem at hand, one could be either strict or lenient in deciding what level of rigid-unit distortion to tolerate. The RSVs from the last singular values of irreps Z_1 and Z_2 have RrmsDs below 2% but are still three to four times less cooperative than that of irrep U_1.

9. A symmetric matrix formulation

In discussing the substantially inexact quasi-RUMs of the R_1 irrep of TTB, Campbell *et al.* (2018) pointed out that, when the tolerance factor is raised sufficiently high to detect them, a number of false positives are detected for other irreps. They reported finding a way to mitigate this difficulty, but did not explain it in detail. The approach employed was to row reduce the symmetric matrix $\mathbf{M}^t\mathbf{M}$ rather than **M**. The use of $\mathbf{M}^t\mathbf{M}$ arises naturally in the context of minimizing the magnitude of a nonzero deviation **Mv**.

Consider the square matrix

$$\mathbf{Y} = \mathbf{M}'^t\mathbf{M}' = \begin{Bmatrix} \mathbf{1} & \mathbf{T}_{SA1} \\ \mathbf{T}_{SA1}^t & \mathbf{T}_{SA1}^t\mathbf{T}_{SA1} + \mathbf{T}_0^t\mathbf{T}_0 \end{Bmatrix}. \tag{8}$$

If we use elementary row operations to eliminate the lower-left block of **Y**, we obtain

$$\mathbf{Y}' = \begin{Bmatrix} \mathbf{1} & \mathbf{T}_{SA1} \\ \mathbf{0} & -\mathbf{T}_{SA1}^t\mathbf{T}_{SA1} + \mathbf{T}_{SA1}^t\mathbf{T}_{SA1} + \mathbf{T}_0^t\mathbf{T}_0 \end{Bmatrix} = \begin{Bmatrix} \mathbf{1} & \mathbf{T}_{SA1} \\ \mathbf{0} & \mathbf{T}_0^t\mathbf{T}_0 \end{Bmatrix}. \tag{9}$$

Because only row operations were involved in generating $\mathbf{M}$ from $\mathbf{M}'$, they have the same null space. For the same reason, $\mathbf{Y}$ and $\mathbf{Y}'$ also have the same null space. If $\mathbf{M}'\mathbf{v} = 0$ for some vector $\mathbf{v}$, then $\mathbf{Y}\mathbf{v} = \mathbf{M}'^{\mathrm{t}}(\mathbf{M}'\mathbf{v}) = 0$. If $\mathbf{M}'\mathbf{v} \neq 0$ for some vector $\mathbf{v}$, then $\mathbf{M}'\mathbf{v}$ is a nonzero vector in the column space of $\mathbf{M}'$, which also implies that $\mathbf{Y}\mathbf{v} = \mathbf{M}'^{\mathrm{t}}(\mathbf{M}'\mathbf{v}) \neq 0$. Thus, $\mathbf{M}'\mathbf{v} = 0$ if and only if $\mathbf{Y}\mathbf{v} = 0$, so that $\mathbf{M}'$ and $\mathbf{Y}$ have the same null space. Clearly, $\mathbf{M}$, $\mathbf{M}^{\mathrm{t}}\mathbf{M}$, $\mathbf{M}'$, $\mathbf{Y} = \mathbf{M}'^{\mathrm{t}}\mathbf{M}'$ and $\mathbf{Y}'$ all have the same null space.

Just as the constraints on the symmetry modes imposed by matrix $\mathbf{M}$ or $\mathbf{M}'$ are determined entirely by the null space of matrix $\mathbf{T}_0$, the constraints on the symmetry modes imposed by matrix $\mathbf{Y}$ or $\mathbf{Y}'$ are determined entirely by the null space of the $N_{\mathrm{free}} \times N_{\mathrm{free}}$ matrix $\mathbf{T}_0^{\mathrm{t}}\mathbf{T}_0$. By the same arguments made above, $\mathbf{T}_0$ and $\mathbf{T}_0^{\mathrm{t}}\mathbf{T}_0$ have the same null space and therefore seem to provide interchangeable routes to the desired constraints. In general, $\mathbf{T}_0^{\mathrm{t}}\mathbf{T}_0$ tends to be a much smaller matrix than $\mathbf{T}_0$. For the example structures presented above, the dimensions of $\mathbf{T}_0^{\mathrm{t}}\mathbf{T}_0$ are 216×216 for HTB, 240×240 for TTB and 960×960 for CAZO. However, because the computational expense of computing the product $\mathbf{T}_0^{\mathrm{t}}\mathbf{T}_0$ is comparable to the expense saved by analyzing the smaller of the two matrices, $\mathbf{T}_0^{\mathrm{t}}\mathbf{T}_0$ proves to be more useful as a conceptual tool than as a computational tool.

The conceptual utility of the symmetric matrix approach is best understood in the context of singular values. If the ith singular value σ_i of $\mathbf{M}$ is small relative to the largest singular value σ_1 of $\mathbf{M}$, then the corresponding relative singular value σ_i^2 of $\mathbf{M}^{\mathrm{t}}\mathbf{M}$ will be even smaller relative to σ_1^2. In other words, $\sigma_i^2/\sigma_1^2 < \sigma_i/\sigma_1$, so that the singular value of a quasi-RUM is more likely to be differentiated as relatively small when using the symmetric $\mathbf{M}^{\mathrm{t}}\mathbf{M}$ matrix.

Because the present work exploits block separability and singular value decomposition, the difficulty initially encountered by Campbell *et al.* (2018) is no longer an issue. However, the differentiation of quasi-RUMs from false positives can be improved when needed by simply squaring the singular values of matrix $\mathbf{T}_0$. In the CAZO example, the smallest relative singular values of irreps U_1 and Z_1 are 1.86×10^{-3} and 5.86×10^{-3}, respectively, which differ by a multiplicative factor of 3.15; after squaring, they differ by an order of magnitude.

10. Conclusions

The identification of possible cooperative rigid-unit modes in a crystalline material is important for understanding and controlling its phase transitions, phonon dynamics and other physical properties. Campbell *et al.* (2018) developed a general algebraic approach to discovering RUMs that employs rotational symmetry-mode patterns in the small-angle approximation. Here, we present several critical improvements to the original implementation.

Considerable computational expense can be saved by determining the null space of the matrix $\mathbf{T}_0$ rather than of the full constraint matrix $\mathbf{M}$. All constraints on the space of allowed symmetry-mode combinations are contained in the smaller $\mathbf{T}_0$ matrix, which has a column for each rotational/displacive symmetry mode but no columns for shared-atom displacement parameters; it also has $3N_{\mathrm{shared}}$ fewer rows than $\mathbf{M}$. We further prove that the multi-pivot shared-atom displacement pattern induced by a rotational/displacive pivot-atom symmetry mode of a given irrep is also a basis function of the same irrep, and that this implies that the columns of $\mathbf{T}_0$ are separable into strictly independent blocks by irrep and OPD branch, so that the null space of each column block can be determined separately without loss of generality. This leads to another very significant reduction in computational complexity. Relative to the original algorithm based on matrix $\mathbf{M}$, the complexity reduction expected from the use of a block-separated $\mathbf{T}_0$ matrix is over three orders of magnitude for HTB and TTB, and over two orders of magnitude for CAZO. Block separability saves computation time, greatly reduces the accumulation of propagated uncertainties and round-off errors, and prevents numerical interference between independent variables.

In cases with approximate RUMs, the null-space determination by Gaussian elimination tends to be unstable owing to nonlinear treatment of small matrix pivots, whereas an SVD is both stable and robust. Each RSV corresponding to zero singular values is a RUM, while each RSV corresponding to near-zero singular values is a quasi-RUM. Squaring the singular values of $\mathbf{T}_0$ and presenting them relative to the largest singular value tends to better differentiate the relatively small singular values of quasi-RUMs from those of false positives.

We define relative r.m.s. deviation as a quantitative measure of the extent to which a pattern of rotations and displacements distorts the rigid units of the interconnected network and demonstrate its utility in differentiating RUMs from non-RUM-like patterns.

These theoretical and computational improvements to the algebraic RUM-search algorithm have now been employed in a new web-based software tool for RUM identification called *ISOTILT*, which is a new component of the *ISOTROPY Software Suite* (https://iso.byu.edu). A companion article on the *ISOTILT* software (Campbell *et al.*, 2021) will contain more detail for each of the HTB, TTB and CAZO examples, and will show the reader how to reproduce the data in Tables 1–4 and to visualize the resulting RUMs.

APPENDIX A
Proof of block separability

In the discussion that follows, all rotation and displacement vectors are expressed in Cartesian coordinates.

The i component of the vector rotation of pivot atom β in normalized rotational symmetry mode m is called $r_{(\beta i)m}$. In joint row index (βi), β runs over all pivot atoms in the child cell and i runs over each of the three vector components of atom β. The column index m runs over each of the rotational symmetry modes allowed by the symmetry of the child structure. More accurately, $r_{(\beta i)m}$ should be viewed as a numerical matrix of coefficients which, when acting on the

vector of variable symmetry-mode amplitudes, yields the actual pivot-atom rotations.

The vector displacement of shared atom α due to pivot atom β in rotational symmetry mode m is called $u_{(\beta_\alpha k)m}$, which lives in a space where each shared atom has as many displacement vectors as it has connected pivots. In the joint row index $(\beta_\alpha k)$, α runs over all shared atoms in the child cell, β_α runs over all DCPAs, *i.e.* pivot atoms that are directly connected to shared atom α, and k runs over the vector components of the corresponding shared-atom displacement. The subscript in β_α is employed because the β index is defined separately for each shared atom α. Here, $u_{(\beta_\alpha k)m}$ should also be viewed as a matrix of numerical coefficients which acts on the vector of symmetry-mode amplitudes to yield the actual shared-atom displacements.

Because the displacements of symmetry-related shared atoms are redundant, it is often convenient to consider only the displacements of the SUSAs in the child cell. The quantity $\overline{u}_{(\beta_{\overline{\alpha}} k)m}$ is a truncation of $u_{(\beta_\alpha k)m}$ which includes only the rows corresponding to SUSA $\overline{\alpha}$. In a similar fashion, the quantity $\overline{r}_{(\overline{\beta}\ell)m}$ includes only the independent variable components of r from the SUPAs in the child cell. In the joint index $(\overline{\beta}\ell)$, $\overline{\beta}$ runs only over SUPAs and ℓ runs over the nonzero independent vector components of atom $\overline{\beta}$.

The j component of the bond vector reaching to shared atom α from pivot atom β is $b_{(\beta_\alpha j)}$. We also define the Levi–Cevita tensor ϵ_{ijk} in the usual way: +1 when ijk is an even permutation of 123, −1 for an odd permutation and 0 otherwise.

Without its index, the symbol a represents a vector containing the a_m amplitudes of all rotational symmetry modes. Similarly the matrices r, b, u, $\overline{r}$ and $\overline{u}$ represent all of the respective components of $r_{(\beta i)m}$, $b_{(\beta_\alpha k)m}$, $u_{(\beta_\alpha k)m}$, $\overline{r}_{(\overline{\beta}\ell)m}$ and $\overline{u}_{(\beta_{\overline{\alpha}} k)m}$, where in each case the joint index in parentheses runs over matrix rows while the last index runs over matrix columns. The matrices $\overline{u}$ and $\overline{r}$ were, respectively, referred to as $\mathbf{T}_S$ and $\mathbf{T}_A$ in the main body of this paper.

The multi-pivot shared-atom displacement pattern associated with a rotational symmetry mode is then computed as a vector product of rotation vectors and bond vectors:

$$u_{(\beta_\alpha k)m} = \sum_{ij\beta'} \delta_{\beta_\alpha \beta'} r_{(\beta' i)m} b_{(\beta_\alpha j)} \epsilon_{ijk}, \tag{10}$$

where β' runs over each of the pivot atoms in the child cell and where the Kronecker δ is 1 if β_α and β' indicate the same pivot atom (modulo a lattice vector) or 0 otherwise.

The index m is a compound index that specifies an irrep (Γ^m), a branch (B^m) of the order-parameter direction of the irrep and an instance (I^m) of the irrep. We define a square matrix $\Gamma(g)_{mn}$ such that the row = m and column = n element of the matrix is equal to the row = B^m and column = B^n element of the irrep matrix of parent space group element g for irrep Γ^m when $\Gamma^m = \Gamma^n$, or else zero if $\Gamma^m \neq \Gamma^n$. By definition, the rotational symmetry modes must transform under the influence of space-group element g according to the coefficients of the contributing irreps as

$$(g \cdot r)_{(\beta i)m} = \sum_n r_{(\beta i)n} \Gamma(g)_{nm}, \tag{11}$$

where we have used the ' $\cdot$ ' symbol to indicate the action of the space-group symmetry element.

In the three-dimensional space of the crystal, the action of the symmetry operation is directly expressed as

$$(g \cdot r)_{(\beta i)m} = \sum_{\beta' i'} |P^g_{\beta\beta'}| P^g_{(\beta i)(\beta' i')} r_{(\beta' i')m}, \tag{12}$$

where indices i and i' run from 1 to 3, $P^g_{\beta\beta'}$ is the point portion of the three-dimensional vector transformation enacted by g when that action takes atom β' to atom β (modulo a lattice vector) or zero otherwise, and the determinant factor accounts for the axial vector nature of rotations. Putting equations (11) and (12) together, we see that

$$\sum_{\beta' i'} |P^g_{\beta\beta'}| P^g_{(\beta i)(\beta' i')} r_{(\beta' i')m} = \sum_n r_{(\beta i)n} \Gamma(g)_{nm}. \tag{13}$$

Before the rotational order parameter is introduced, the pattern of bond vectors already respects the parent symmetry group. So when we reach back to the SUSA–DCPA pair β'_α that will transform under g into atom pair β_α, the corresponding bond vector $b_{(\beta'_\alpha j')}$ transforms into $b_{(\beta_\alpha j)}$. Mathematically, we express this invariance as follows:

$$(g \cdot b)_{(\beta_\alpha j)} = \sum_{\beta'_\alpha j'} P^g_{(\beta_\alpha j)(\beta'_\alpha j')} b_{(\beta'_\alpha j')} = b_{(\beta_\alpha j)}. \tag{14}$$

We are now ready to consider the action of g on the full multi-pivot shared-atom displacement pattern induced by a rotational symmetry mode. By combining equations (10), (11) and (14), we obtain

$$\begin{aligned}
(g \cdot u)_{(\beta_\alpha k)m} &= \sum_{ij\beta'} \delta_{\beta_\alpha \beta'} (g \cdot r)_{(\beta' i)m} (g \cdot b)_{(\beta_\alpha j)} \epsilon_{ijk} \\
&= \sum_{ij\beta'} \delta_{\beta_\alpha \beta'} \left(\sum_n r_{(\beta' i)n} \Gamma(g)_{nm} \right) b_{(\beta_\alpha j)} \epsilon_{ijk} \\
&= \sum_n \left(\sum_{ij\beta'} \delta_{\beta_\alpha \beta'} r_{(\beta' i)n} b_{(\beta_\alpha j)} \epsilon_{ijk} \right) \Gamma(g)_{nm} \\
&= \sum_n u_{(\beta_\alpha k)n} \Gamma(g)_{nm}.
\end{aligned} \tag{15}$$

Thus we conclude that the multi-pivot shared-atom displacement pattern induced by a rotational symmetry mode of a given irrep is also a basis function of the same irrep.

The two distinct rotational symmetry modes $r_{(\beta i)m}$ and $r_{(\beta i)n}$ are irrep basis functions, and are therefore orthogonal (as βi runs over all Cartesian components of all pivot atoms in the child cell), because that is how we defined them. Thus, the columns of r are pairwise orthogonal. Now, because the rotation-induced shared-atom displacive modes are also irrep basis functions, we know that two distinct rotation-induced displacive modes $u_{(\beta_\alpha k)m}$ and $u_{(\beta_\alpha k)n}$ must be orthogonal when they correspond to different irreps or irrep branches, *i.e.* when $\Gamma^R(m) \neq \Gamma^R(n)$ or when both $\Gamma^R(m) = \Gamma^R(n)$ and $B^R(m) \neq B^R(n)$. Note, however, that there is no guarantee that distinct modes of the same irrep and branch will be orthogonal or even independent.

We have proven that the columns of u are irrep basis functions and must therefore be mutually orthogonal when corresponding to distinct irreps or irrep branches. However, the algebraic RUM-search algorithm did not employ the columns of u, but rather the columns of $\overline{u}$. Whereas a column of u has three rows for each of the shared atoms in the child cell, a column of $\overline{u}$ only has rows for the SUSAs. What can we then say about the orthogonality or independence of the columns in $\overline{u}$?

First, it is obvious that there exists a matrix $\mathcal{R}$ such that $r = \mathcal{R}\,\overline{r}$. The matrix element $\mathcal{R}_{(\beta i)(\overline{\beta}\ell)}$ is 1 when pivot atom β is related by symmetry to SUPA $\overline{\beta}$ and the ith vector component of the rotation vector of atom β corresponds to the lth free rotational parameter associated with atom $\overline{\beta}$, or zero otherwise. It essentially transforms any rotational pattern from its SUPA representation to its all-pivot-atom representation. Because no row of $\mathcal{R}$ can have more than one nonzero element, it is also obvious that the columns of $\mathcal{R}$ are pairwise orthogonal. Thus there exists a well defined left inverse $\mathcal{R}_{\rm left}^{-1} = (\mathcal{R}^{\rm t}\mathcal{R})^{-1}\mathcal{R}^{\rm t}$ such that $\overline{r} = \mathcal{R}_{\rm left}^{-1} r$.

Second, in equation (16), we can gather the coefficients of $r_{(\beta' i)m}$ in $u_{(\beta_\alpha k)m}$ into a matrix $Z_{(\beta_\alpha k)(\beta' i)}$ so as to obtain the matrix equation $u = Zr = Z\mathcal{R}\,\overline{r}$. We can also do something similar with $\overline{u}$ and $\overline{r}$ to obtain $\overline{u} = \overline{Z}\mathcal{R}_{\rm left}^{-1} r = \overline{Z} r$. In component form, these relations are expressed as in equations (16) and (17):

$$\begin{aligned} u_{(\beta_\alpha k)m} &= \sum_{ij\beta'} \delta_{\beta_\alpha \beta'} r_{(\beta' i)m} b_{(\beta_\alpha j)} \epsilon_{ijk} \\ &= \sum_{i\beta'} Z_{(\beta_\alpha k)(\beta' i)} r_{(\beta' i)m}, \end{aligned} \tag{16}$$

$$\begin{aligned} \overline{u}_{(\overline{\beta}_\alpha k)m} &= \sum_{ij\beta'} \delta_{\overline{\beta}_\alpha \beta'} r_{(\beta' i)m} b_{(\overline{\beta}_\alpha j)} \epsilon_{ijk} \\ &= \sum_{ij\beta'\overline{\beta}\ell} \delta_{\overline{\beta}_\alpha \beta'} \mathcal{R}_{(\beta' i)(\overline{\beta}\ell)} \overline{r}_{(\overline{\beta}\ell)m} b_{(\overline{\beta}_\alpha j)} \epsilon_{ijk} \\ &= \sum_{\overline{\beta}\ell} \overline{Z}_{(\overline{\beta}_\alpha k)(\overline{\beta}\ell)} \overline{r}_{(\overline{\beta}\ell)m}. \end{aligned} \tag{17}$$

Clearly, it does not matter whether we use r or $\overline{r}$; they are interconvertible.

Third, the conversion from u to $\overline{u}$ merely involves the removal of dependent rows, which does not change the row space of the matrix and cannot therefore affect the independence of its columns. Thus, columns or sets of columns that were mutually independent in u must still be mutually independent in $\overline{u}$, though columns that were mutually orthogonal in u might not remain orthogonal in $\overline{u}$. The columns associated with different irreps or irrep branches are orthogonal and hence independent in u, and must therefore also be independent in $\overline{u}$, but not necessarily orthogonal.

Fourth, the reduction of a matrix to reduced row-echelon form isolates the dependencies between its columns. Groups of columns that have no mutual dependencies (*i.e.* no linear combination of the columns in one group can equal any linear combinations of another group) can be reduced separately with no loss of generality. Blocks of columns in $\overline{u}$ associated with distinct irreps or irrep branches are mutually independent and can therefore be row reduced separately. We can say that such a row-reduction problem is 'block separable'.

Because the computational complexity of the row-reduction process scales as $n_{\rm r} n_{\rm c}^2$, where $n_{\rm r}$ is the number of rows and $n_{\rm c}$ is the number of columns, the exploitation of block separability can offer a very significant reduction of computational expense, particularly when the matrix has a large number of small blocks.

In summary, the multi-pivot shared-atom displacement pattern (with separate displacements of the shared atom for each connected pivot atom) induced by a rotational symmetry mode of a given irrep is also a basis function of the same irrep, so that such patterns enjoy the special orthogonality and completeness properties generally associated with irrep basis functions. The RUM-search algorithm of Campbell *et al.* (2018) is based on the row-reduction of matrix $\overline{u}$ as defined above, which is obtained from the matrix u of multi-pivot shared-atom displacements via the removal of all dependent rows, so that two sets of columns in $\overline{u}$ are mutually independent if and only if the corresponding set of columns in u are mutually independent. Thus, the RUM-search algorithm is separable into blocks of rotational symmetry modes, one for each distinct irrep and order-parameter branch, which can then be row reduced separately. Because the rotational parameter space available to the pivot atoms of a structure is often finely divided across multiple irreps and order-parameter branches, block separation will tend to be very advantageous.

Acknowledgements

We acknowledge John Evans (University of Durham, UK) for suggesting the idea of a tolerance map (implemented here as a singular value search), and acknowledge both Chris Howard (University of Newcastle, Australia) and Volker Kahlenberg (Universität Innsbruck, Austria) for suggesting and illuminating the structural examples presented, as well as Chris Howard for his careful reading of the manuscript.

Funding information

This material is based upon work supported by the National Science Foundation under grant No. PHY-1757998.

References

Bradley, C. J. & Cracknell, A. P. (1972). *The Mathematical Theory of Symmetry in Solids*. Oxford: Clarendon.
Campbell, B., Howard, C. J., Averett, T. B., Whittle, T. A., Schmid, S., Machlus, S., Yost, C. & Stokes, H. T. (2018). *Acta Cryst.* A**74**, 408–424.
Campbell, B. J., Stokes, H. T., Averett, T. B., Machlus, S. & Yost, C. J. (2021). *J. Appl. Cryst.* **54**, 1847–1856.
Campbell, B. J., Stokes, H. T., Tanner, D. E. & Hatch, D. M. (2006). *J. Appl. Cryst.* **39**, 607–614.
Dove, M. T., Giddy, A. P. & Heine, V. (1993). *Trans. Am. Crystallogr. Assoc.* **27**, 65–74.
Dove, M. T., Trachenko, K. O., Tucker, M. G. & Keen, D. A. (2000). *Rev. Mineral. Geochem.* **39**, 1–33.
Giddy, A. P., Dove, M. T., Pawley, G. S. & Heine, V. (1993). *Acta Cryst.* A**49**, 697–703.

Golub, G. H. & Van Loan, C. F. (1996). *Matrix Computation*, 3rd ed. Baltimore, London: Johns Hopkins University Press.

Goreaud, M., Labbé, Ph., Montfort, Y. & Raveau, B. (1980). *Rev. Chim. Miner.* **17**, 79–87.

Haydon, S. K. & Jefferson, D. A. (2002). *J. Solid State Chem.* **168**, 306–315.

Higham, N. J. & Higham, D. J. (1989). *SIAM J. Matrix Anal. Appl.* **10**, 155–164.

Howard, C. J. & Stokes, H. T. (1998). *Acta Cryst.* B**54**, 782–789.

Howard, C. J. & Stokes, H. T. (2005). *Acta Cryst.* A**61**, 93–111.

Kahlenberg, V., Albrecht, R., Schmidmair, D., Krüger, H., Krüger, B., Tribus, M. & Pauluhn, A. (2019). *J. Am. Ceram. Soc.* **102**, 2084–2093.

Kahlenberg, V., Hejny, C. & Krüger, H. (2019). *J. Solid State Chem.* **276**, 319–330.

Magnéli, A. (1949). *Ark. Kemi*, **1**, 213–221.

Magnéli, A. (1953). *Acta Chem. Scand.* **7**, 315–324.

Phillips, A. E. (2018). *Acta Cryst.* A**74**, 406–407.

Prinz, H., Müller, U. & Ha-Eierdanz, M.-L. (1992). *Z. Anorg. Allg. Chem.* **609**, 95–98.

Rodríguez-Carvajal, J., Hennion, M., Moussa, F., Moudden, A. H., Pinsard, L. & Revcolevschi, A. (1998). *Phys. Rev. B*, **57**, R3189–R3192.

Saint-Gregoire, P. & Smirnov, M. (2021). *Perovskites and Other Framework Structure Crystalline Materials: New Trends and Perspectives*, edited by P. Saint-Gregoire & M. Smirnov, ch. 1 and 11. Collaborating Academics.

Smirnov, M. & Saint-Grégoire, P. (2014). *Acta Cryst.* A**70**, 283–290.

Strang, G. (2016). *Introduction to Linear Algebra*, 5th ed. Wellesley: Cambridge Press/SIAM.

Takusagawa, F. & Jacobson, R. A. (1976). *J. Solid State Chem.* **18**, 163–174.

Trefethen, L. N. & Bau, D. III (1997). *Numerical Linear Algebra*. Philadelphia: Society for Industrial and Applied Mathematics.

Triantafyllou, S. T., Christidis, P. C. & Lioutas, Ch. B. (1997). *J. Solid State Chem.* **130**, 176–183.

Wachsmann, C. & Jacobs, H. (1995). *Eur. J. Solid State Inorg. Chem.* **32**, 1023–1025.

Whittle, T. A., Schmid, S. & Howard, C. J. (2015). *Acta Cryst.* B**71**, 342–348.

Wolfram (2017). *Mathematica 11.1*, https://www.wolfram.com/mathematica/.