

Small-angle scattering tensor tomography algorithm for robust reconstruction of complex textures

Leonard C. Nielsen,^a Paul Erhart,^a Manuel Guizar-Sicairos^{b,c} and Marianne Liebi^{a,b,c,*}

^aDepartment of Physics, Chalmers University of Technology, Gothenburg, Sweden, ^bPaul Scherrer Institute (PSI), Villigen, Switzerland, and ^cÉcole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. *Correspondence e-mail: marianne.liebi@psi.ch

Received 16 June 2023

Accepted 1 October 2023

Edited by A. Altomare, Institute of Crystallography - CNR, Bari, Italy

Keywords: tensor tomography; SAXS; small-angle X-ray scattering; scattering; reciprocal-space maps; optimization.

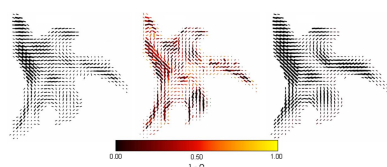
Supporting information: this article has supporting information at journals.iucr.org/a

The development of small-angle scattering tensor tomography has enabled the study of anisotropic nanostructures in a volume-resolved manner. It is of great value to have reconstruction methods that can handle many different nanostructural symmetries. For such a method to be employed by researchers from a wide range of backgrounds, it is crucial that its reliance on prior knowledge about the system is minimized, and that it is robust under various conditions. Here, a method is presented that employs band-limited spherical functions to enable the reconstruction of reciprocal-space maps of a wide variety of nanostructures. This method has been thoroughly tested and compared with existing methods in its ability to retrieve known reciprocal-space maps, as well as its robustness to changes in initial conditions, using both simulations and experimental data. It has also been evaluated for its computational performance. The anchoring of this method in a framework of integral geometry and linear algebra highlights its possibilities and limitations.

1. Introduction

Small-angle X-ray scattering (SAXS) probes the nanometre-scale variations in the electron density of materials averaged over areas typically in the range of $1 \times 1 \mu\text{m}$ to $1000 \times 1000 \mu\text{m}$, depending on the size of the X-ray beam. The data from a SAXS experiment carry information about the nanostructure of a sample, including characteristic length scales and orientation, and have been used to study numerous materials (Fratzl *et al.*, 1996; Georgiadis *et al.*, 2016; Lichtenegger *et al.*, 1999). Scanning SAXS can be performed across a sample to yield a two-dimensional map, with each scanned pixel associated with a corresponding two-dimensional cut through the reciprocal-space map (Fratzl *et al.*, 1997; Pabisch *et al.*, 2013; Paris, 2008). Since a SAXS measurement with an area detector gives two-dimensional data, measurements must be performed at several angles to obtain three-dimensional data from a sample (Liu *et al.*, 2010; Seidel *et al.*, 2012; Georgiadis *et al.*, 2016). By rotating a three-dimensional sample around a single axis, standard tomographic reconstruction can be used in cases where the sample scattering is isotropic or when the scattering is symmetric about the axis of rotation (Stribeck *et al.*, 2006, 2008; Feldkamp *et al.*, 2009; Schroer *et al.*, 2006; Álvarez-Murga *et al.*, 2012; Jensen *et al.*, 2011).

With the use of a tilt angle in addition to rotation, recent works by Schaff *et al.* (2015), Liebi *et al.* (2015, 2018) and Gao *et al.* (2019) present tomographic methods for the reconstruction of the three-dimensional reciprocal-space map using



OPEN ACCESS

Published under a CC BY 4.0 licence

scanning SAXS projections, or small-angle X-ray scattering tensor tomography (SAXSTT).

Whereas Schaff *et al.* (2015) demonstrated a two-step procedure of first reconstructing the reciprocal-space map without any data-reducing assumptions and then analyzing this reconstruction to find orientations, the approaches of Liebi *et al.* (2015, 2018) (referred to as the zonal harmonics or ZH method) and Gao *et al.* (2019) (referred to as the iterative reconstruction or IR method) use reduced models of the reciprocal-space maps. The ZH method of Liebi *et al.* (2015, 2018) models the reciprocal-space map of each voxel using squared band-limited spherical functions expressed in spherical harmonics. The demonstrations of the method by Liebi *et al.* (2015, 2018) used a reduced model, employing only zonal harmonics (spherical harmonics symmetric about an axis of rotation) and two Euler angles, with the angles parameterizing the main nanostructural orientation. The IR method of Gao *et al.* (2019) uses the symmetric rank-2 tensor as the basis for its model. The work of Gao *et al.* (2019) primarily describes the reconstruction of a proposed orientation distribution function derived heuristically for fiber scattering. The proposed orientation distribution function approach is not compatible with reciprocal space map based models of SAXSTT, because it does not yield the necessary invariance with respect to rotation of the sample (*e.g.* De Falco *et al.*, 2021). However, as noted in a footnote by Gao *et al.* (2019), it is also possible to configure the IR algorithm to reconstruct the reciprocal-space map directly, which is the way IR was employed in this work.

The use of a complete band-limited basis of even-ordered spherical harmonics described by Liebi *et al.* (2018) has to the best of our knowledge not been implemented or tested for SAXSTT as of the writing of this work. Moreover, the approach of Liebi *et al.* (2015, 2018) requires fitting the measured reciprocal-space map to sums of squared polynomials, which is a difficult class of optimization problems to solve (Ahmadi *et al.*, 2017), as it results in a non-linear system of equations. Here, we present spherical integral geometric tensor tomography (SIGTT) as an improvement on the complete basis approach proposed by Liebi *et al.* (2015, 2018). SIGTT eliminates the squaring of the polynomial which was used in the ZH approach, resulting in a linear system, and its implementation does not rely on Euler angles.

2. Results

2.1. Theory

The equation for the measured reciprocal-space map (RSM) of SAXS may be written as

$$\text{RSM}(\mathbf{q}) = \int \int \int dV [\tilde{\rho}(\mathbf{r}) \exp(-i\mathbf{q} \cdot \mathbf{r})], \quad (1)$$

where $\tilde{\rho}(\mathbf{r})$ is the auto-correlation function of the electron density of the probed volume, \mathbf{r} is the position and \mathbf{q} is the reciprocal-space vector. Consider this integral over a region of space (a voxel) and at a fixed $\|\mathbf{q}\|$. Then, $\text{RSM}(\mathbf{q})$ reduces to $\text{RSM}(\theta, \phi)$, a function on the unit sphere, which may by theorem be represented by an infinite series of spherical

harmonics (Kosmann-Schwarzbach & Singer, 2010). As discussed in previous work by Liebi *et al.* (2018), the summation is reduced to the even orders as a result of Friedel symmetry. The measured reciprocal-space map at a single \mathbf{q} and at a particular position \mathbf{r} in space may then be written $\text{RSM}(\mathbf{r}, \theta, \phi)$ and expanded in spherical harmonics as

$$\text{RSM}(\mathbf{r}, \theta, \phi) = \sum_{\ell=0,2,\dots}^{\infty} \sum_{m=-\ell}^{\ell} a_m^{\ell}(\mathbf{r}) \hat{Y}_m^{\ell}(\theta, \phi), \quad (2)$$

where θ and ϕ are, respectively, the polar and azimuthal angles of the reciprocal-space map, $\hat{Y}_m^{\ell}(\theta, \phi)$ is the real-valued spherical harmonic basis function of order ℓ and degree m , and $a_m^{\ell}(\mathbf{r})$ is the coefficient of that basis function at position \mathbf{r} . Note that the summation over ℓ in equation (2) only includes even terms, as indicated by the subscript of the summation sign.

The following projection model is based on the discrete model of Liebi *et al.* (2018), cast in terms of line integrals within the sample coordinate system, which map it onto a projection space. The projection space is spanned by four coordinates (j, k, α, β) , which can be mapped to experimental parameters. The linear coordinates, (j, k) , map to the vertical and horizontal positioning of the sample during scanning SAXS. The angular coordinates (α, β) map to rotations of the sample about laboratory axes orthogonal to the beam direction during the measurement. Rotations about axes which are not orthogonal to the beam direction can be handled by decomposition into beam-orthogonal rotations that map to (α, β) , and beam-parallel rotations that map to transforms of (j, k) . The projection of a scalar function $f(\mathbf{r})$ from three to two dimensions at an arbitrary angle for a narrow beam is described by the John transform, also known as the X-ray transform, which may be expressed

$$\mathbb{P}[f](j, k, \alpha, \beta) = \int_{-\infty}^{\infty} f[\mathbf{v}(j, k, \alpha, \beta) + s\mathbf{u}(\alpha, \beta)] ds,$$

where α and β are the azimuthal and polar angles, respectively, with respect to a fixed plane in the sample coordinate system, for a line of integration which intersects with the system's origin. Then, j and k give the line's offset from the origin in the plane of projection, which is orthogonal to the line's direction. In this representation, parameterized by s , \mathbf{v} gives the position of the beam in the plane of projection, whereas \mathbf{u} gives its direction. See Note 1 in the supporting information for more details on how to map the sample coordinate system to an experimental coordinate system.

To keep our equations compact, we will use the shorthand notation

$$\bar{\nu} \equiv (\nu_1, \nu_2, \nu_3, \nu_4) \equiv (j, k, \alpha, \beta), \quad (3)$$

such that $\bar{\nu}$ represents simultaneously a line of integration in three-dimensional space and a measured point in projection space. By inserting equation (2) into the John transform $\mathbb{P}[f]$, we obtain an expression for the projection of the spherical harmonics from three to two dimensions using two projection angles:

$$\mathbb{P}[\text{RSM}](\theta, \phi, \bar{\nu}) = \sum_{\ell=0,2,\dots}^{\infty} \sum_{m=-\ell}^{\ell} \hat{Y}_m^{\ell}(\theta, \phi) A_m^{\ell}(\bar{\nu}) \quad (4)$$

with

$$A_m^{\ell}(\bar{\nu}) = \int_{-\infty}^{\infty} a_m^{\ell}[\mathbf{v}(\bar{\nu}) + s\mathbf{u}(v_3, v_4)] ds.$$

From equation (4) it is apparent that both the reciprocal-space map at a particular point in space and the projection of it may be represented using an even-ordered spherical harmonic expansion with a one-to-one correspondence between the representations. In other words, each order and degree in the projected harmonic can be regarded as the projection of a single order and degree in the harmonics distributed in space. However, small-angle scattering does not permit us to probe the entirety of the reciprocal-space map at a single projection angle; instead, it probes a set of points which lie approximately on a great circle spanned by the set of unit vectors orthogonal to the direction of the beam. We describe this using a parametric curve $C(\varphi, \alpha, \beta)$, where for fixed α and β , $C(\varphi)$ is a great circle orthogonal to the direction of \mathbf{u} .

We can then write the model (which we will denote \mathcal{I}) for a single measured reciprocal-space map as

$$\mathcal{I}(\varphi, \bar{\nu}) = \sum_{\ell=0,2}^{\infty} \sum_{m=-\ell}^{\ell} \hat{Y}_m^{\ell}[C(\varphi, v_3, v_4)] A_m^{\ell}(\bar{\nu}), \quad (5)$$

which completes the forward model of SIGTT. Anisotropic scattering signals have previously been modeled using line integrals of spherical polynomials by Wieczorek *et al.* (2016) for dark-field tomography. The key difference of SIGTT from the spherical harmonic model described by Liebi *et al.* (2018) is the linearity of the system, which is achieved by not squaring the spherical harmonics. Unlike the implementation demonstrated by Liebi *et al.* (2015, 2018), SIGTT does not employ a local coordinate system for the reciprocal-space map in each voxel.

These changes result in the preservation of the orthogonality of each component of the per-voxel model, and simplify gradient calculations. See Note 2 in the supporting information for how the SIGTT representation maps to Cartesian tensors, as in the model used by IR. The inverse problem of obtaining $a_m^{\ell}(\mathbf{r})$ is solved by using a regularized least-squares approach (Section 4).

2.2. Simulations

In order to compare and evaluate the different methods, a simulation framework was developed (Section 4). In total, three simulations were created, labeled ‘M’, ‘T’ and ‘mammoth’; see Fig. 1 for an overview. Sample ‘M’ is intended to provide a simulated reciprocal-space map with an intensity distribution that possesses the zonal symmetry assumed by ZH, meaning that the distribution has an axis of rotational symmetry. This is done by constraining its simulated reciprocal-space map to be approximately described by zonal harmonics (which are defined by an expansion of spherical harmonics with $m = 0$ and an axis of symmetry) up to $\ell_{\max} = 12$. In the simulation of ‘T’, the reciprocal-space map is described entirely by symmetric rank-2 tensors, which is the model employed by IR. Finally, the simulation of ‘mammoth’ employed spherical harmonics up to $\ell_{\max} = 8$, with a weak $\ell = 2$ component, to model reciprocal-space maps with more complicated symmetries. The reconstructions of each method are compared with the simulated samples by calculating the squared Pearson correlation coefficient R^2 [equation (15)] of the simulated and reconstructed reciprocal-space maps on a voxel-by-voxel basis.

In Fig. 2, the results from reconstructing simulated data for sample ‘M’, which possesses an intensity distribution with zonal symmetry, are shown. See Section 4 for details on the box plots. SIGTT performs the best in the comparison, with a peak correlation centered around $R^2 = 0.8$ at the highest signal-to-noise ratio (SNR), decaying down to about 0.75 at

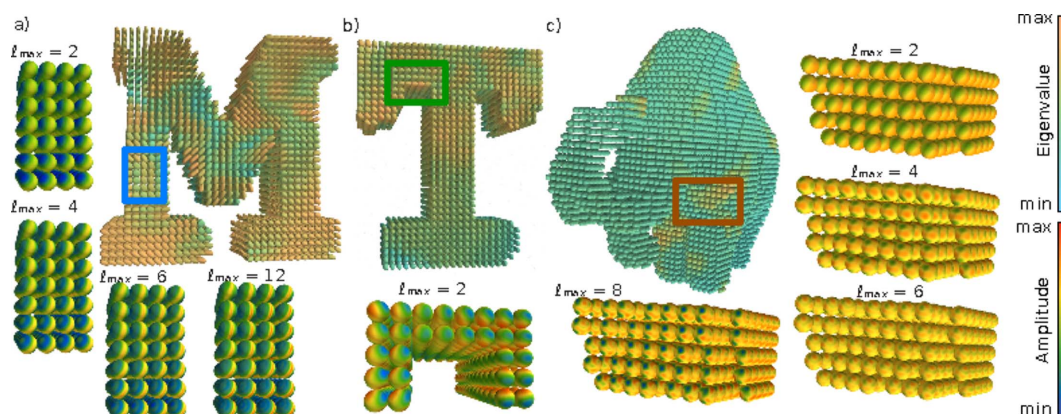


Figure 1

Superquadric glyph render of simulations. (a) Sample ‘M’ with its zonally symmetric reciprocal-space maps from one region (blue square), truncated at $\ell = 2, 4, 6, 12$. (b) Sample ‘T’ with its rank-2 tensor reciprocal-space maps from one region (green square). (c) Sample ‘mammoth’ with its unrestricted $\ell_{\max} = 8$ reciprocal-space maps from one region (orange square), truncated at $\ell = 2, 4, 6, 8$. The colors of the superquadric glyph renders show the largest eigenvalue of the rank-2 tensor component of each simulated sample, and the upper and lower bounds of the color mapping are individually set to reveal the texture of each sample. The colors of the reciprocal-space map renders show their amplitudes and are scaled to the maximum and minimum amplitude in the selected region of each sample.

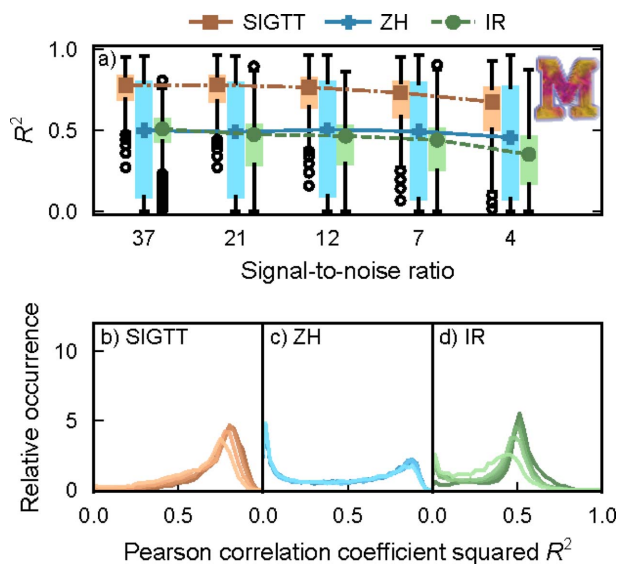


Figure 2
Correlations for sample 'M'. (a) Box plots of R^2 as defined in equation (15), with lines and symbols indicating the respective median of each box plot. Outlier dots each represent 100 voxels. (b)–(d) Correlation coefficient distribution for each of the three methods. The signal-to-noise ratio goes from 37 (darkest lines) down to 4 (lightest lines). The image inset shows a volume render of the simulated sample.

the lowest SNR, with a greater interquartile range. The best possible performance of SIGTT in this comparison is constrained by the fact that the chosen discretization of the reciprocal-space map permits the fitting of harmonics only up to $\ell = 6$, whereas the sample contains harmonics up to $\ell = 12$. The reconstruction is also affected by the missing wedge problem, a common limitation in tomography when only a subset of projection space is sampled. The performance of ZH is almost unaffected by the SNR, with its correlation centered around $R^2 = 0.5$ with a large interquartile range. In the lower middle plot, a peak can be seen both around 0.8 and around 0, suggesting a mixture of good and poor performance across the sample volume. The trend in the performance of IR is more similar to that of SIGTT in that its correlation decays and its interquartile range increases with the decrease in SNR. However, its maximum possible correlation to most of the sample is bounded at around 0.5 by the fact that it can only correlate to the $\ell = 2$ component of the model, due to being restricted to the symmetric rank-2 tensor.

Volume renders of the errors of each method for 'M' can be seen in Fig. 3. It is clear from the figure that all methods have larger errors in the middle regions of the model, but that these are larger for ZH and IR. Moreover, outside these regions, the error for SIGTT is closer to 0.

In Fig. 4 are virtual slices with glyphs showing the orientation of each reciprocal-space map for (a) the simulation 'M', (b) SIGTT, (c) ZH and (d) IR. See Note 3 (supporting information) for details on the orientation analysis. The comparison shows that the orientation of the reciprocal-space maps in the simulation is overall reasonably well followed by all reconstructions. However, from (c) it is clear that the ZH reconstruction contains many deviations from the simulated

orientation. In terms of the relative anisotropy [equation (14)], indicated by the color of each glyph, it is generally well followed by SIGTT, as seen in (b), but poorly followed by ZH, seen in (c). It can be inferred from this that the numerous deviations in the ZH reconstruction are the cause of the large dispersion in R^2 seen in Fig. 2(a). IR, in (d), gives a reconstruction that follows the orientations and relative anisotropy nearly as well as (b), with the exception of certain regions. This is to be expected, since while the anisotropy in fiber-like symmetry is generally dominated by the rank-2 tensor component, this will not be the case everywhere.

In Fig. 5 the performance between the three models is compared in reconstructing sample 'T', which consists of reciprocal-space maps with $\ell_{\max} = 2$. SIGTT and IR perform fairly similarly, with the performance of IR consistently being somewhat worse. Since these models are very similar for the special case of $\ell_{\max} = 2$, this difference in performance is likely related to the fact that the IR implementation uses a less precise projection algorithm, as well as its use of fixed step size steepest descent gradient optimization. It is likely the case that SIGTT performs better due to modeling the measured reciprocal-space map as a line integral on the reciprocal-space sphere, rather than as a point, as well as using continuity-enforcing regularization. ZH performs very poorly regardless of noise level in the case of 'T'. This is likely both because it cannot directly represent coefficients of the $\ell = 2$ harmonics due to its use of squared polynomials, and because 'T' does not follow zonal symmetry, meaning it does not have an axis of rotational symmetry, and cannot be represented by a rotated spherical harmonic expansion with only $m = 0$ components.

The results from the reconstruction of the sample 'mammoth' are seen in Fig. 6. SIGTT has similar performance to the case of sample 'M', with its correlation starting around 0.8 and decaying to about 0.65 as the SNR decreases, with an increase in the interquartile range. ZH performs very poorly, which is to be expected, as the 'mammoth' sample does not follow zonal symmetry at all. IR performs better than ZH, but is bounded by the fact that the reciprocal-space maps in this simulation only have a weak rank-2 tensor component. Thus, it does not reach a median R^2 above 0.3. See Note 4 and Figs. S1 and S2 (in the supporting information) for volume renders of the errors of 'T' and 'mammoth', as well as Fig. S3 and

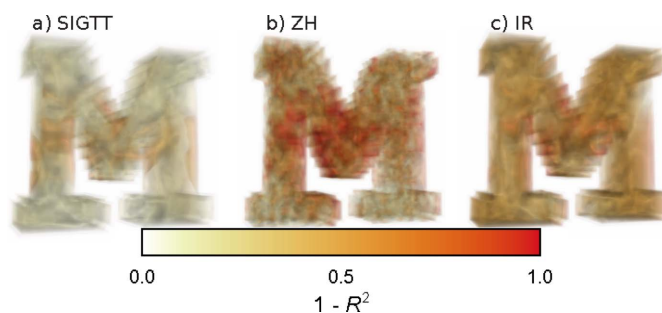


Figure 3
Volume renders of errors for reconstructions of 'M'. (a) Errors for SIGTT. (b) Errors for ZH. (c) Errors for IR. The error is defined as $1 - R^2$, where R^2 is given by equation (15). Larger errors are rendered with greater opacity and are thus visible even if they are in the interior.

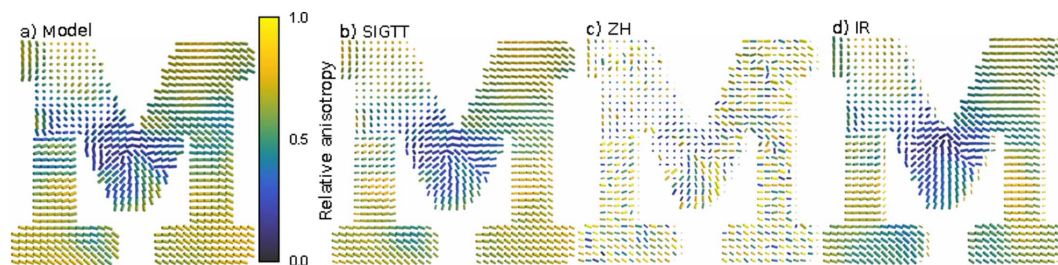


Figure 4

Comparison of virtual slices of 'M'. (a) Virtual slice of simulated sample 'M'. (b) SIGTT reconstruction. (c) ZH reconstruction. (d) IR reconstruction. The glyphs are colored according to the relative anisotropy [equation (14)] and scaled according to the isotropic component [equation (13)] of each reciprocal-space map. All three reconstructions follow the orientations of the model reasonably well on average. The SIGTT reconstruction follows both the orientations and relative anisotropy of the model closely. The ZH reconstruction has a lot of variation in the relative anisotropy, as well as many orientations deviating from the overall tendency to follow the model. The IR reconstruction follows the model almost as well as the SIGTT reconstruction.

equation (S1) for a comparison of the anisotropic power distribution of 'mammoth' and 'M'.

In Fig. 7 we show timing information for the reconstructions of (a) 'M', (b) 'T' and (c) 'mammoth'. Note the logarithmic y scale. In all cases, SIGTT is the fastest method, followed by IR, with ZH being more than an order of magnitude slower than either method. For example, one of the slowest cases for all methods is 'mammoth' with SNR 30 [panel (c)] – in this case, SIGTT requires approximately 7 min, IR requires approximately 25 min (3.5 times as long as SIGTT) and ZH requires approximately 10 h (85 times as long as SIGTT). In the case of 'M' [panel (a)], IR is nearly as fast as SIGTT, although it should be noted that in this case SIGTT is fitting 28 coefficients, whereas IR is fitting six coefficients. ZH is slower than both methods by approximately two orders of magnitude. While ZH also only uses six degrees of freedom in this case,

two of these degrees of freedom are polar and azimuthal angles, and it models the reciprocal-space map using the squared amplitude of spherical functions. Both of these features lead to an optimization problem which is more difficult to solve due to its non-linearity (*e.g.* Hochbaum, 2007; Ahmadi *et al.*, 2017). For 'T' [panel (b)], both SIGTT and IR are fitting six coefficients, as 'T' has only rank-2 tensor components, and in this case SIGTT is several times faster. The sample 'mammoth' [panel (c)] takes substantially longer to fit for all of the methods, as it has a greater volume ($60 \times 60 \times 80$ voxels, whereas 'M' and 'T' have $50 \times 50 \times 50$ voxels each), increasing the amount of time it takes to compute each projection. SIGTT is several times faster than IR here, and approximately two orders of magnitude faster than ZH. This comparison should only be understood as a broad guideline to the performance of the methods; each

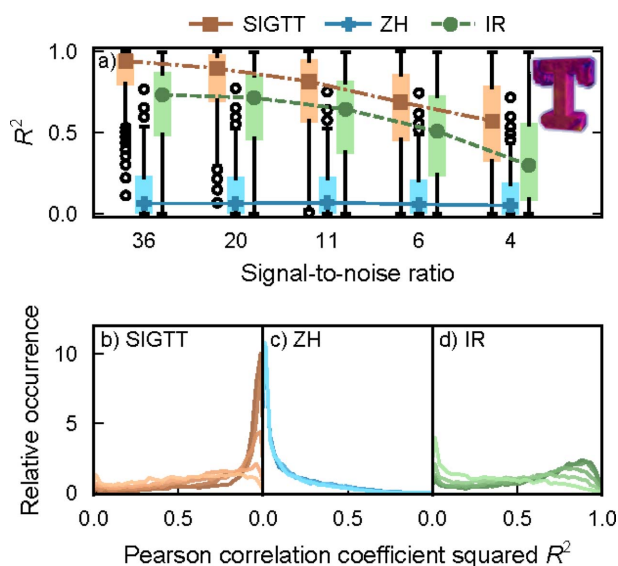


Figure 5

Correlations for sample 'T'. (a) Box plots of R^2 as defined in equation (15), with lines and symbols indicating the respective median of each box plot. Outlier dots each represent 100 voxels. (b)–(d) Correlation coefficient distribution for each of the three methods. The signal-to-noise ratio goes from 36 (darkest lines) down to 4 (lightest lines). The image inset shows a volume render of the simulated sample.

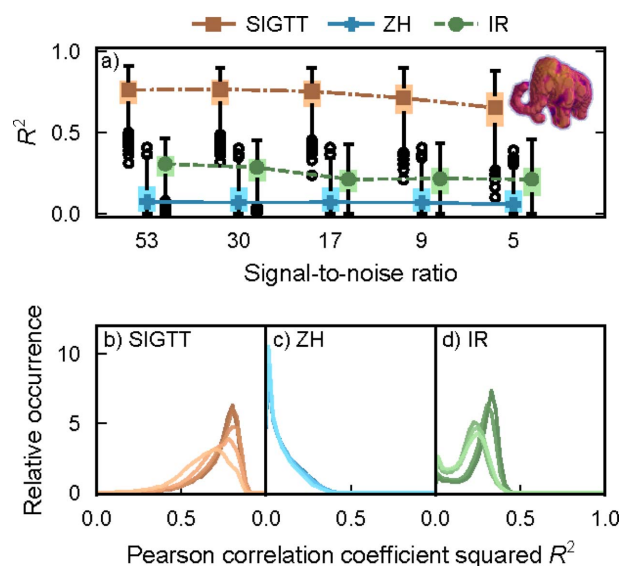


Figure 6

Correlations for sample 'mammoth'. (a) Box plots of R^2 as defined in equation (15), with lines and symbols indicating the respective median of each box plot. Outlier dots each represent 100 voxels. (b)–(d) Correlation coefficient distribution for each of the three methods. The signal-to-noise ratio goes from 53 (darkest lines) down to 5 (lightest lines). The image inset shows a volume render of the simulated sample.

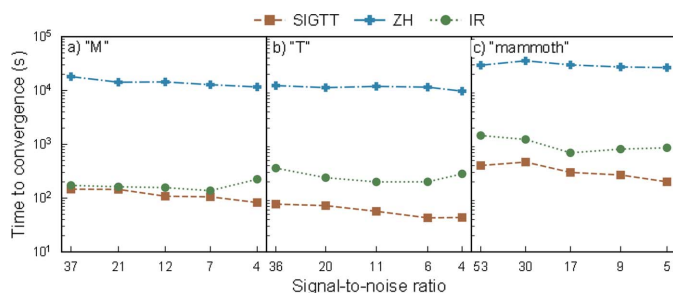


Figure 7

Timing comparison of methods. (a) Timing for sample 'M'. (b) Timing for sample 'T'. (c) Timing for sample 'mammoth'. Note the logarithmic y axis. In all cases, SIGTT is the fastest method, followed by IR, with ZH being the slowest by a considerable margin.

implementation has a number of parameters which affect convergence in different ways, which were adjusted with the aim of obtaining a reconstruction that correlates well with the simulation, rather than optimized for speed. Moreover, SIGTT is implemented in Python, whereas ZH and IR are implemented in Matlab, which means that the conditions for optimizations such as multithreading and efficient memory handling are different (see Section 4.6). It is therefore likely that at least some of the speedup seen when comparing SIGTT with the other methods is due to a more efficient implementation, *i.e.* the particular code used to carry out computations, rather than improvements in the basic method. It is likely that all methods could be sped up considerably by employing a GPU-based projection algorithm (*e.g.* Nikitin, 2023).

2.3. Experimental data

An ensemble of ten reconstructions, each with some initial conditions randomized, were performed using SIGTT, ZH and IR on a sample of trabecular bone. For experimental details, see Liebi *et al.* (2015), sample B. The chosen q range of $0.0379\text{--}0.0758\text{ nm}^{-1}$ for this reconstruction does not contain the collagen meridional peak, and therefore its reciprocal-space map has fiber symmetry from equatorial diffuse mineral scattering. A comparison of a virtual section from each of the three methods can be seen in Fig. 8. Because there is no ground truth with which to compare for experimental data, the ensemble of reconstructions was instead analyzed to investigate the robustness of each method against perturbations in the initial conditions. The spherical harmonic representations of the ten reconstructions were averaged over, voxel-by-voxel, and Fig. 8 shows the results of the averaged reconstruction. The colors of the orientation glyphs indicate the degree to which the anisotropy of the reciprocal-space map changes across every reconstruction; the quantity Q is defined in equation (16). The glyphs are scaled according to the square root of the mean anisotropic power of each reciprocal-space map [the anisotropic power is defined in equation (11)] across the ensemble. The results indicate that SIGTT and IR are robust to perturbations of the initial conditions, but that ZH is not. However, the orientations of the averaged ZH recon-

struction agree well with those of SIGTT and IR. A plausible reason for this difference is that ZH is the only method out of the three which depends on Euler angles. Depending on the initial conditions of the angles, the solution may be confined to local minima, as the symmetries of its reciprocal space can only vary across a limited subspace of the total spherical harmonic coefficient space. While IR performed similarly to SIGTT in the chosen q range, a rank-2 tensor cannot contain more than one local maximum per hemisphere. This poses a problem for the method in the case of the reciprocal-space map of the collagen meridional peak q range of bone, which contains two distinct maxima – an equatorial maximum from diffuse mineral scattering which lies along a great circle, and a meridional maximum from the spacing of the collagen fibril d spacing, which lies on the poles orthogonal to that great circle (*e.g.* Zhou *et al.*, 2016). This symmetry, which requires at least a rank-4 tensor, can be represented by both SIGTT and ZH (Guizar-Sicairos *et al.*, 2020).

3. Discussion

This work has demonstrated the SIGTT method for SAXSTT reconstruction of the reciprocal-space map in samples using a band-limited spherical function expressed in spherical harmonics (see Section 4 for details). In three case studies using simulated data with approximately zonally symmetric reciprocal-space maps, rank-2 tensors and complicated-textured higher-order reciprocal-space maps, the method produces results superior to the approaches of Liebi *et al.* (2018) and Gao *et al.* (2019). In addition, SIGTT is faster than the existing implementations of the methods of Liebi *et al.* (2018) and Gao *et al.* (2019), in many cases considerably so. SIGTT has also been shown to be robust to perturbations in the initial conditions when reconstructing experimental data. The reconstruction of the reciprocal-space map using higher spherical harmonic orders will enable the use of more specific methods of characterization that reveal information about the nature of the sample beyond the main orientation or adher-

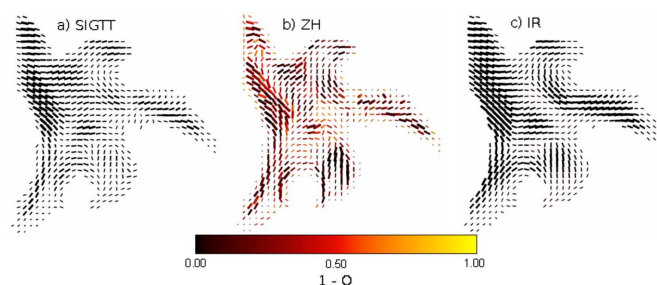


Figure 8

Experimental ensemble reconstructions. (a) Virtual slice from an ensemble of ten reconstructions each with randomized initial conditions of a sample of trabecular bone using SIGTT. (b) Ensemble reconstruction using ZH. (c) Ensemble reconstruction using IR. The glyphs are scaled by the square root of the anisotropic power, defined in equation (11). The quantity Q , defined in equation (16), is a measure of how much the anisotropy of each reciprocal-space map changes across the ensemble of reconstructions. The methods generally agree in terms of the orientation of each reciprocal-space map, but only ZH shows a change in the anisotropy across the ensemble.

ence to a specific, predetermined symmetry. This representation of the reciprocal-space map in SIGTT makes it possible to study not just individual domains, but also the boundaries and transitions between them, including voxels with more than one orientation, which frequently occur in samples in the fields of biology and materials science (e.g. Georgiadis *et al.*, 2023; Maciel *et al.*, 2018). SIGTT can be applied to more complicated reciprocal-space maps, such as those that occur in SAXS measurements of samples with hexagonal symmetries [as done by Smarsly *et al.* (2005)], or in wide-angle X-ray scattering measurements. Wide-angle X-ray scattering measurements can be used by themselves or as a complement to SAXS measurements; see Mao *et al.* (2018) for an example of using small- and wide-angle X-ray scattering in combination for the study of a polymer under deformation. The extension of SAXSTT methods to also encompass wide-angle scattering is therefore a promising area of further study. Because it is not restricted to lower-order spherical harmonics, SIGTT is able to model reciprocal-space maps which are not well approximated by a rank-2 tensor, a case discussed by Georgiadis *et al.* (2021). For the same reasons, SIGTT should make it easier to reconstruct samples with smaller or less well organized domains [as done by Georgiadis *et al.* (2020)], as these would contain more transitory regions and a greater lack of symmetry in the reciprocal-space map. In this way, the reconstruction of the reciprocal-space map using band-limited spherical functions makes full use of the data obtained from the collection of scanning SAXS data at multiple angles, and opens up many new avenues of analysis. Finally, the anchoring of SIGTT in a framework of integral geometry and linear algebra highlights the potential for algorithms employing alternative schemes for data acquisition, optimization and representation of the reciprocal-space map, e.g. in the vein of the work of Sharma *et al.* (2017) on acquisition schemes for dark-field tomography, and that of Aslan *et al.* (2019) on ptycho-tomographic reconstruction.

4. Methods

4.1. Discretized formalism

The inverse problem to the forward model of equation (5) is to obtain the distributions $a_m^\ell(\mathbf{r})$ for a set of measured data.

In practical terms, the solution to the inverse problem is best discussed using a discrete formalism. Prior to reconstruction, measurements at a particular value of \mathbf{q} are reduced by binning the measured pixel intensities into azimuthal segments, which corresponds to an integral over a great circle segment on the reciprocal-space sphere. Consequently, for the bin i , centered on φ_i of width $\Delta\varphi$, the spherical harmonics in equation (5) are integrated to give the coefficients for these bins,

$$\bar{Y}_m^\ell(\bar{\mathbf{v}}, i) = \frac{1}{\Delta\varphi} \int_{\varphi_i - 0.5\Delta\varphi}^{\varphi_i + 0.5\Delta\varphi} \hat{Y}_m^\ell[C(\tau, v_3, v_4)] d\tau, \quad (6)$$

with v_i as in equation (3), where the integration variable τ replaces the detector angle φ used in equation (5). Since this effectively blurs the reciprocal-space map, it constrains the maximum frequency that can be uniquely represented in the reduced data. It can be concluded that the ℓ_{\max} of the fitted spherical function should follow

$$\ell_{\max} \leq N - 1,$$

where N is the number of azimuthal bins for $\varphi_i \in [0, \pi)$, due to the assumption of Friedel symmetry. This can be shown explicitly by expanding the N reciprocal-space map binned into azimuthal segments as a trigonometric polynomial,

$$\text{RSM}(\varphi_i) = \sum_{m=0,2}^M [C_m \cos(m\varphi_i) + C_{-m} \sin(m\varphi_i)],$$

where m is the frequency of each component of the polynomial, and odd frequencies vanish due to Friedel symmetry. We observe that this has a unique solution only for $M = N - 1$ as a direct consequence of the Nyquist–Shannon sampling theorem (Shannon, 1949). It can be seen that a real trigonometric polynomial is isomorphic to a great circle cut of a spherical harmonic representation by noting that it is the expression for the azimuthal component of the real spherical harmonics (see Note 2 in the supporting information for details). The spherical harmonic rotation theorem implies that any great circle cut of a function expressed in spherical harmonics can be represented in a coordinate system where this great circle cut is the equator of the unit sphere. Since the polar angle is constant at the equator, this means that the great circle cut is determined only by the azimuthal component; thus, for any great circle cut of a spherical harmonic representation, there exists an equivalent trigonometric polynomial. Therefore by setting $\ell_{\max} = N - 1$ for the spherical harmonic representation, a unique solution exists for the great circle cut visible in each measurement. Consequently, the gradient contribution of that measurement becomes unique when solving the system (see Note 5 in the supporting information for details).

The sample volume, spanned by \mathbf{r} in equation (2), is divided into cubic voxels, of the same size as the step size in the scanning SAXS measurement.

To pose our problem in matrix form, we describe our discretized system as a matrix \mathbf{X} of N rows and M columns, where each row corresponds to a voxel, and each column corresponds to a spherical harmonic coefficient. Similarly, the measured data are described by an $I \times J$ matrix which we label \mathbf{D} , with I being the number of scanned points and J the number of detector azimuthal segments across all rotation and tilt configurations. For simplicity of representation, we consider scans and detector azimuthal segments at different rotations and tilts to be distinct, so that \mathbf{D} takes a sparse block-matrix form. The discrete equivalent of the projection operation \mathbb{P} [see equation (4)], considered across all measured projections, is then given by a sparse $I \times N$ matrix \mathbf{P} , describing a mapping between weighted sums of the N voxels to the I scanned points of the sample. Finally, the mapping from the spherical harmonic representation of the reciprocal-

space map to the azimuthal segment-wise detector data is given by an $M \times J$ matrix \mathbf{Y} , consisting of the M coefficients calculated in equation (6) for each of the J detector azimuthal segments.

This gives us the system of equations

$$\mathbf{PXY} = \mathbf{D},$$

and we write the solution as the optimization of the least-squares problem

$$\mathbf{X}_{\text{opt}} = \text{argmin}_{\mathbf{X}} \|\mathbf{PXY} - \mathbf{D}\|^2. \quad (7)$$

See Note 5 in the supporting information for details on how to solve this system with an iterative algorithm.

4.2. Reciprocal-space map evaluations

In tomography, the solution is generally assumed to be a reasonably smooth function, and we impose a regularization term on equation (7) to ensure this (Natterer & Wübbeling, 2001). Since we do not want the evaluation or comparison of reciprocal-space maps [defined by equations (1), (2)] to depend on our choice of coordinate system, it is necessary to use rotational invariants. A general rotational invariant is the canonical inner product of the spherical harmonics, known as the cross-spectrum function,

$$S_{\ell}(g, h) = \sum_{m=-\ell}^{\ell} \mathcal{N}(\ell) g_m^{\ell} h_m^{\ell}, \quad (8)$$

where $\mathcal{N}(\ell)$ is a normalization factor that depends on the choice of spherical harmonic representation, ℓ is the cross-spectrum order, and g and h are two spherical functions (Wieczorek & Meschede, 2018). g_m^{ℓ} and h_m^{ℓ} are the coefficients of the spherical harmonic representation of g and h , given by

$$g_m^{\ell} = \int d\Omega [g(\theta, \phi) \hat{Y}_m^{\ell}(\theta, \phi)].$$

This discussion is therefore applicable to the analysis of spherical functions of any type, but in the particular case of SAXSTT, f and g are reciprocal-space maps as defined by equations (1), (2). To regularize the problem by imposing a smoothness condition, we compute a nearest-neighbor similarity term,

$$(\Lambda g)_{ij} = \lambda \sum_{\ell=0}^{\ell_{\text{max}}} [S_{\ell}(g_i, g_i) + S_{\ell}(g_j, g_j) - 2S_{\ell}(g_i, g_j)],$$

where the set of all (i, j) indicates neighboring pairs of voxels, g_i is the spherical function associated with each voxel and λ is a regularization coefficient.

In spherical harmonic coefficient space, this reduces to the squared discrete Laplacian operator on our system matrix weighted by λ ,

$$\Lambda X = \lambda (\nabla^2 \mathbf{X})^2. \quad (9)$$

Minimizing this term results in maximizing the covariance between neighboring voxels, since

$$\sum_{\ell=1}^{\infty} S_{\ell}(g, h) = \text{cov}(g, h), \quad (10)$$

with $S_{\ell}(g, h)$ defined in equation (8), and therefore we also have

$$\sum_{\ell=1}^{\infty} S_{\ell}(g, g) = \text{var}(g), \quad (11)$$

where, in particular, we refer in this work to $\text{var}(g)$ as the anisotropic power of the reciprocal-space map represented by g .

Thus, with the addition of the regularization term in equation (9), the solution becomes

$$\mathbf{X}_{\text{opt}} = \text{argmin}_{\mathbf{X}} [\|\mathbf{PXY} - \mathbf{D}\|^2 + \lambda \|\nabla^2 \mathbf{X}\|^2]. \quad (12)$$

We define the isotropic component of a reciprocal-space map as its spherical mean $\bar{\mu}$, and in spherical harmonic representation,

$$\bar{\mu}(g) = g_0^0, \quad (13)$$

where g is the spherical function representing the reciprocal-space map.

We also define the relative anisotropy of a reciprocal-space map as

$$\varsigma(g) = \frac{\sqrt{\text{var}(g)}}{g_0^0} = \frac{\sigma(g)}{\bar{\mu}(g)}, \quad (14)$$

or, in other words, the standard deviation $\sigma(g)$ of the spherical function g normalized by the spherical mean $\bar{\mu}(g)$. We prefer to normalize by the mean rather than by the root-mean-square, because normalization by the root-mean-square would result in an upper bound of 1, which makes the contrast worse for highly anisotropic samples. This is useful to indicate the texture of a sample in many-voxel visualizations, and is used to color the glyphs in Fig. 4.

The relative anisotropy is comparable with the definition of the degree of orientation given by Bunk *et al.* (2009) as the ratio of the first Fourier component of an azimuthally integrated scattering pattern to the zeroth Fourier component (which is equal to the mean). This would correspond to calculating the relative anisotropy as defined by equation (14) using only the $\ell = 2$ components of the spherical harmonic representation of the reciprocal-space map. We include coefficients beyond $\ell = 2$, however, because we are interested in reciprocal-space maps of a more general class of symmetries, as well as figures of merit which are easily extended to other representations of the reciprocal-space map. The primary reason for not using a definition akin to that of the ρ parameter used by Fratzi *et al.* (2004), Pabisch *et al.* (2013), which uses the integrated peak intensity divided by the total intensity (peak intensity plus background intensity), is that defining the background intensity is difficult in a spherical harmonic representation. Moreover, using the standard deviation also incorporates information about the sharpness of the peaks, which is especially useful in cases where the background may be very small. The definition of the relative anisotropy in equation (14) is similar but not identical to the quantity also referred to as the degree of orientation by Liebi *et al.* (2018), which evaluates to the variance of the square root of the reconstructed reciprocal-space map divided by its mean.

4.3. Correlation calculations

Because the variances and covariances of reciprocal-space maps [equations (10), (11)] are rotational invariants, compositions of them are also invariants, and in particular

$$\frac{\text{cov}(g, h)^2}{\text{var}(g)\text{var}(h)} = R^2(g, h), \quad (15)$$

where $\text{var}(g)$ is the variance of the spherical function g , $\text{cov}(g, h)$ is the covariance of two spherical functions, and $R^2(g, h)$ is the squared Pearson correlation coefficient of the two functions g, h . The squared Pearson correlation coefficient is used in the comparison of reconstructed reciprocal-space maps with those in simulated samples. The Pearson correlation coefficient does not map directly to other measures of similarity such as the difference in orientation; it is instead a summary metric of how similar two distributions are, up to a constant offset and a scaling factor. If R^2 is close to 0, then the distributions are very dissimilar, although it is still possible for other measures to show similarity of specific features, such as the orientation. If R^2 is close to 1, then the distributions are very similar and other measures will also show similarity (up to the possible offset of a constant and scaling factor); this generality is the reason why this is our statistic of choice. In the calculation of R^2 shown in Figs. 2, 5 and 6, the calculation is done between the reciprocal-space maps of each voxel in the simulated model (excluding empty voxels) and the reciprocal-space map of the same voxel in the reconstruction. The box plots in Figs. 2(a), 5(a) and 6(a) follow the original definitions of Tukey (1977). They are defined such that the colored rectangles span the interquartile range of the correlation distribution. The black 'whiskers' outside the colored rectangles span the smallest and largest value in the range $[Q_1 - 1.25 \times (Q_3 - Q_1), Q_3 + 1.25 \times (Q_3 - Q_1)]$, where Q_i is the i th quartile of the distribution of R^2 . Values outside the range of the 'whiskers' are represented by small circles, with each circle showing the mean R^2 of 100 reciprocal-space maps. If there is at least one, but fewer than 100 reciprocal-space maps above or below each whisker, a single black circle is shown, representing the mean of R^2 across these reciprocal-space maps. The median, equivalent to Q_2 , is shown by the colored markers in each box plot.

4.4. Simulation framework

For each of the three sample volumes, source points were determined such that the distance between each point was maximized: four source points for 'M', two source points for 'T' and five source points for 'mammoth'. Band-limited spherical functions were constructed such that the spectral power of each order followed power-law decays with respect to ℓ , and assigned to each source point. The interior distance from each source to every other point in the volume was then approximately computed using a combination of a k-d tree and Dijkstra's algorithm. The sources were assigned correlation lengths, with the assumption that for each order of each spherical function in the sample, correlation with the source would decay with distance like a Gaussian distribution with

the correlation length as its standard deviation. The remaining spherical functions were then solved for under several constraints – in all cases, it was assumed that the spherical functions of neighboring voxels in the volume would be correlated with each other, and that the power of each order of the function in each voxel would equal a distance-weighted average of the power of its source. Moreover, it was required that all functions be non-negative. Non-negativity is difficult to enforce perfectly for spherical polynomials, but a dense sampling of each function was performed and the isotropic component was increased to eliminate all the detected negative points.

'M' consists of spherical polynomials up to $\ell_{\max} = 12$ that approximately follow zonal symmetry, *i.e.* there is for each spherical function an axis of orientation, about which they have approximate rotational symmetry. They can thus be well represented in a rotated spherical harmonic expansion with only $m = 0$ coefficients [see Fig. 9(a)]. To enforce zonal symmetry, each spherical function was required to correlate with the ℓ -weighted spherical harmonic Dirac δ function,

$$w(\ell)\delta_\ell^m(\alpha, \beta) = (-1)^{\ell/2}Y_\ell^m(\alpha, \beta),$$

with (α, β) given by the fiber-like orientation vector of the $\ell = 2$ component of the reciprocal-space map (see Note 3 in the supporting information for details on orientation analysis), and $w(\ell)$ being an ℓ -weighting function. In general, the condition of zonal symmetry requires compromise with the demand of continuity [enforced by minimizing the spherical harmonic Laplacian, as in equation (9)]. This is because the different orders of spherical harmonics have differing symmetries with respect to rotations. In effect, this leads to some attenuation of parts of the great circle of intensity. Hence, the reciprocal-space maps of 'M' are said to only approximately follow zonal symmetry. 'T' is entirely represented by symmetric rank-2 tensors, with a distribution required to be continuous [see Fig. 9(b)]. 'Mammoth' is represented by spherical polynomials with $\ell_{\max} = 8$, which in addition to being continuous have a damped $\ell = 2$ component, such that the $\ell = 2$ and $\ell = 4$ components have approximately the same spectral power [see Fig. 9(c)]. This was done to approximate the type of spectral power distributions that

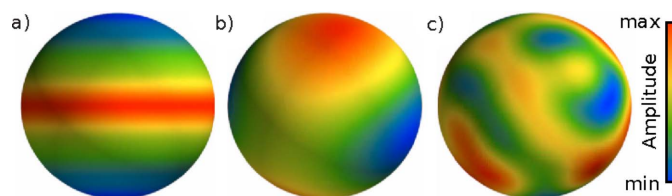


Figure 9
Reciprocal-space map symmetries for simulated samples. (a) Zonally symmetric spherical function with $\ell_{\max} = 12$ with its maximum around a great circle and a minimum at the poles, similar to the reciprocal-space maps of sample 'M'. (b) Spherical function with $\ell_{\max} = 2$, similar to the reciprocal-space maps of sample 'T'. (c) Spherical functions with $\ell_{\max} = 8$, with no particular symmetry constraints, similar to the reciprocal-space maps of sample 'mammoth'. The color of the surface of each spherical function indicates its amplitude on a linear scale, scaled according to the largest and smallest value of each spherical function.

occur in reciprocal-space maps which do not either have equatorial symmetry [with maxima concentrated around a great circle, akin to Fig. 9(a)] or meridional symmetry (with the maxima at the poles), such as in the region of the collagen meridional peak in bone (e.g. Zhou *et al.*, 2016), where the $\ell = 2$ contributions from the meridional peak and the equatorial diffuse mineral scattering have opposing signs and thus tend to cancel out. The reciprocal-space map cuts generated from the projections of each simulation were divided into eight azimuthal segments, accounting for Friedel symmetry, and integrated over, emulating the azimuthal integration approach used for experimental data. The choice of eight bins was made based on previous usage in experimental data, as well as the fact that it will restrict the band-limit of spherical functions that can be precisely retrieved to $\ell_{\max} = 6$, meaning that it will not be possible for SIGTT to exactly solve for the reciprocal-space maps of the samples 'M' ($\ell_{\max} = 12$) and 'mammoth' ($\ell_{\max} = 8$).

4.5. Ensemble reconstructions

For the ensemble reconstructions shown in Fig. 8, each of the methods had their initial conditions randomized. In the case of SIGTT and IR, this consisted of randomizing the coefficients of their solution vectors at values several orders of magnitude below what is expected from their reconstructed value. In the case of ZH, the randomization was only applied to the Euler angles of the orientations of each voxel, which must be initialized before each reconstruction. The Euler angles were randomized such that the orientation vectors of each voxel were uniformly distributed on the unit sphere. The stepwise reconstruction procedure of Liebi *et al.* (2018) was then followed. In order to average over the result of the ensemble, the squared coefficients of the ZH reconstruction, performed with $\ell_{\max} = 6$, were expanded through Driscoll–Healy quadrature in a non-squared spherical harmonic representation up to $\ell_{\max} = 12$ (Driscoll & Healy, 1994). Tests incorporating higher orders and denser grids showed that this approach was accurate to a relative error of approximately 0.1% in the variance of the reciprocal-space map, which was deemed sufficient for the purpose of examining the reconstruction's consistency across the ensemble. To evaluate the robustness of the reconstructions with respect to initial conditions, we defined an anisotropic power quotient for the reciprocal-space map in each voxel,

$$Q = \frac{\text{var}\left(\frac{1}{n} \sum_{i=0}^n g_i\right)}{\frac{1}{n} \sum_{i=0}^n \text{var}(g_i)}, \quad (16)$$

where g_i is the spherical function representing the reciprocal-space map in each voxel for reconstruction run i , and $\text{var}(g_i)$ is the anisotropic power of the reciprocal-space map as defined in equation (11), and n is the total number of reconstructions in the ensemble. We used $n = 10$, as this proved sufficient to illustrate the difference between the methods. If the reciprocal-space maps of every reconstruction in the ensemble are identical, the value of Q will be 1, and it will be in the range [0, 1) if the reciprocal-space maps differ.

4.6. Implementations

SIGTT, and the simulations used in this work, was implemented in Python. The most demanding parts of the code, projections and back-projections, are carried out using *Numba*, part of the software package *Mumott*, whereas other calculations are carried out in *NumPy* and *SciPy* (Harris *et al.*, 2020; Virtanen *et al.*, 2020; Lam *et al.*, 2015). The version of *Mumott* used in this work is available at <https://doi.org/10.5281/zenodo.7798530>. The most recent version of *Mumott* is made available at <https://doi.org/10.5281/zenodo.7919448>. The projection code, written specifically for this work, uses only CPU resources. It performs the John transform by using vectors to trace out the lines of integration, and sampling the voxels that these intersect with, in proportion to the lengths of the intersecting segments. Visualizations of the three-dimensional reconstructions were created using the Python package *Mayavi* (Ramachandran & Varoquaux, 2011). The color maps used throughout this paper were generated with the help of *ColorCET* (<https://colorcet.com>) (Kovesi, 2015). All computations were performed on a workstation using a 12-core, 4.6 GHz AMD Ryzen 9 3900X CPU and 64 GB DDR4 2666 MHz RAM. For the IR and ZH methods, the original code from the *cSAXS* software package written in Matlab was used, with modifications for optimization and termination of each reconstruction upon convergence. The projection code in this package samples voxels using coordinate transforms and bilinear interpolation. It slices the sample along the plane of integration, and samples the four voxels closest to the line of integration, based on the distance in the plane of projection. Because it effectively treats the projection of each voxel as a square at every angle of projection, and does not consider the full three-dimensional distance between voxels, this approach suffers from high-frequency artifacts. However, following testing, this approach was deemed sufficiently accurate for the purpose of comparing ZH and IR with SIGTT.

5. Data availability

The simulated data created for and used in this work are available at <https://doi.org/10.5281/zenodo.7673985>. A notebook demonstrating the analysis and reconstruction using *Mumott* is available at <https://doi.org/10.5281/zenodo.7799517>.

Acknowledgements

The authors extend their gratitude to William Lionheart at the University of Manchester for providing input and feedback on spherical harmonic and tensor field formalism. We acknowledge the Paul Scherrer Institute, Villigen, Switzerland, for provision of synchrotron radiation beamtime at the beamline cSAXS of the SLS. Conceptualization: ML. Methodology: LCN, PE, MG-S, ML. Supervision: PE, ML. Writing – review and editing: LCN, PE, MG-S, ML. Software: LCN, PE. Visualization: LCN. Formal analysis: LCN, ML. Data curation: MG-S. Investigation: MG-S, ML.

Funding information

This work was funded by the Swedish Research Council (VR 2018-041449) and the European Research Council (ERC-2020-StG 949301). Some of the computations in this work were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at NSC, C3SE and PDC partially funded by the Swedish Research Council through grant agreement No. 2018-05973.

References

- Ahmadi, A. A., Hall, G., Papachristodoulou, A., Saunderson, J. & Zheng, Y. (2017). *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pp. 453–462.
- Álvarez-Murga, M., Bleuett, P. & Hodeau, J.-L. (2012). *J. Appl. Cryst.* **45**, 1109–1124.
- Aslan, S., Nikitin, V., Ching, D. J., Bicer, T., Leyffer, S. & Gürsoy, D. (2019). *Opt. Express*, **27**, 9128–9143.
- Bunk, O., Bech, M., Jensen, T., Feidenhans'l, R., Binderup, T., Menzel, A. & Pfeiffer, F. (2009). *New J. Phys.* **11**, 123016.
- De Falco, P., Weinkamer, R., Wagermaier, W., Li, C., Snow, T., Terrill, N. J., Gupta, H. S., Goyal, P., Stoll, M., Benner, P. & Fratzl, P. (2021). *J. Appl. Cryst.* **54**, 486–497.
- Driscoll, J. & Healy, D. (1994). *Adv. Appl. Math.* **15**, 202–250.
- Feldkamp, J. M., Kuhlmann, M., Roth, S. V., Timmann, A., Gehrke, R., Shakhverdova, I., Paufler, P., Filatov, S. K., Bubnova, R. S. & Schroer, C. G. (2009). *Phys. Status Solidi A*, **206**, 1723–1726.
- Fratzl, P., Gupta, H., Paschalis, E. & Roschger, P. (2004). *J. Mater. Chem.* **14**, 2115–2123.
- Fratzl, P., Jakob, H. F., Rinnerthaler, S., Roschger, P. & Klaushofer, K. (1997). *J. Appl. Cryst.* **30**, 765–769.
- Fratzl, P., Schreiber, S. & Klaushofer, K. (1996). *Connect. Tissue Res.* **34**, 247–254.
- Gao, Z., Guizar-Sicairos, M., Lutz-Bueno, V., Schröter, A., Liebi, M., Rudin, M. & Georgiadis, M. (2019). *Acta Cryst.* **A75**, 223–238.
- Georgiadis, M., Menzel, M., Reuter, J. A., Born, D. E., Kovacevich, S. R., Alvarez, D., Taghavi, H. M., Schroeter, A., Rudin, M., Gao, Z., Guizar-Sicairos, M., Weiss, T. M., Axer, M., Rajkovic, I. & Zeineh, M. M. (2023). *Acta Biomaterialia*, **164**, 317–331.
- Georgiadis, M., Müller, R. & Schneider, P. (2016). *J. R. Soc. Interface*, **13**, 20160088.
- Georgiadis, M., Schroeter, A., Gao, Z., Guizar-Sicairos, M., Liebi, M., Leuze, C., McNab, J. A., Balolia, A., Veraart, J., Ades-Aron, B., Kim, S.-L., Shepherd, T. M., Lee, C. H., Walczak, P., Chodankar, S., DiGiacomo, P. S., Dávid, G., Augath, M., Zerbi, V., Sommer, S., Rajkovic, I., Weiss, T. M., Bunk, O., Yang, L., Zhang, J., Novikov, D. S., Zeineh, M. M., Fieremans, E. & Rudin, M. (2021). *Nat. Commun.* **12**, 2941.
- Georgiadis, M., Schroeter, A., Gao, Z., Guizar-Sicairos, M., Novikov, D. S., Fieremans, E. & Rudin, M. (2020). *NeuroImage*, **204**, 116214.
- Guizar-Sicairos, M., Georgiadis, M. & Liebi, M. (2020). *J. Synchrotron Rad.* **27**, 779–787.
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C. & Oliphant, T. E. (2020). *Nature*, **585**, 357–362.
- Hochbaum, D. (2007). *Ann. Oper. Res.* **153**, 257–296.
- Jensen, T. H., Bech, M., Bunk, O., Menzel, A., Bouchet, A., Le Duc, G., Feidenhans'l, R. & Pfeiffer, F. (2011). *NeuroImage*, **57**, 124–129.
- Kosmann-Schwarzbach, Y. & Singer, S. F. (2010). *Groups and Symmetries*, edited by Y. Kosmann-Schwarzbach, pp. 93–106. New York: Springer.
- Kovesi, P. (2015). *arXiv:1509.03700*.
- Lam, S. K., Pitrou, A. & Seibert, S. (2015). *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, pp. 1–6. Association for Computing Machinery, New York, USA.
- Lichtenegger, H., Müller, M., Paris, O., Riekel, C. & Fratzl, P. (1999). *J. Appl. Cryst.* **32**, 1127–1133.
- Liebi, M., Georgiadis, M., Kohlbrecher, J., Holler, M., Raabe, J., Usov, I., Menzel, A., Schneider, P., Bunk, O. & Guizar-Sicairos, M. (2018). *Acta Cryst.* **A74**, 12–24.
- Liebi, M., Georgiadis, M., Menzel, A., Schneider, P., Kohlbrecher, J., Bunk, O. & Guizar-Sicairos, M. (2015). *Nature*, **527**, 349–352.
- Liu, Y., Manjubala, I., Roschger, P., Schell, H., Duda, G. N. & Fratzl, P. (2010). *J. Phys. Conf. Ser.* **247**, 012031.
- Maciel, M., Ribeiro, S., Ribeiro, C., Francesko, A., Maceiras, A., Vilas, J. & Lanceros-Méndez, S. (2018). *Composites Part B*, **139**, 146–154.
- Mao, Y., Bucknall, D. G. & Kriegel, R. M. (2018). *Polymer*, **143**, 228–236.
- Natterer, F. & Wübbeling, F. (2001). *Mathematical Methods in Image Reconstruction*. Society for Industrial and Applied Mathematics.
- Nikitin, V. (2023). *J. Synchrotron Rad.* **30**, 179–191.
- Pabisch, S., Wagermaier, W., Zander, T., Li, C. & Fratzl, P. (2013). *Research Methods in Biomineralization Science*, edited by J. J. De Yoeo, Vol. 532 of *Methods in Enzymology*, pp. 391–413. London: Academic Press.
- Paris, O. (2008). *Biointerphases*, **3**, FB16–FB26.
- Ramachandran, P. & Varoquaux, G. (2011). *Comput. Sci. Eng.* **13**, 40–51.
- Schaff, F., Bech, M., Zaslansky, P., Jud, C., Liebi, M., Guizar-Sicairos, M. & Pfeiffer, F. (2015). *Nature*, **527**, 353–356.
- Schroer, C. G., Kuhlmann, M., Roth, S. V., Gehrke, R., Stribeck, N., Almendarez-Camarillo, A. & Lengeler, B. (2006). *Appl. Phys. Lett.* **88**, 164102.
- Seidel, R., Gourrier, A., Kerschnitzki, M., Burghammer, M., Fratzl, P., Gupta, H. S. & Wagermaier, W. (2012). *Bioinspired, Biomimetic and Nanobiomaterials*, **1**, 123–131.
- Shannon, C. (1949). *Proceedings of the IRE*, **37**, 10–21.
- Sharma, Y., Schaff, F., Wiczorek, M., Pfeiffer, F. & Lasser, T. (2017). *Sci. Rep.* **7**, 3195.
- Smarsly, B., Gibaud, A., Ruland, W., Sturmayer, D. & Brinker, C. J. (2005). *Langmuir*, **21**, 3858–3866.
- Stribeck, N., Camarillo, A. A., Nöchel, U., Schroer, C., Kuhlmann, M., Roth, S. V., Gehrke, R. & Bayer, R. K. (2006). *Macromol. Chem. Phys.* **207**, 1139–1149.
- Stribeck, N., Nöchel, U., Funari, S. S. & Schubert, T. (2008). *J. Polym. Sci. B Polym. Phys.* **46**, 721–726.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, I., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., Vijaykumar, A., Bardelli, A. P., Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C. N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D. A., Hagen, D. R., Pasechnik, D. V., Olivetti, E., Nolin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G. A., Ingold, G., Allen, G. E., Lee, G. R., Audren, H., Probst, I., Dietrich, J. P., Silterra, J., Webber, J. T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J. L., de Miranda Cardoso, J. V., Reimer, J., Harrington, J., Rodríguez, J. L. C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmeler, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N. J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P. A., Lee, P., McGibbon, R. T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T. J., Robitaille, T. P., Spura, T., Jones, T. R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U.,

- Halchenko, Y. O. & Vázquez-Baeza, Y. (2020). *Nat. Methods*, **17**, 261–272.
- Wieczorek, M. A. & Meschede, M. (2018). *Geochem. Geophys. Geosyst.* **19**, 2574–2592.
- Wieczorek, M., Schaff, F., Pfeiffer, F. & Lasser, T. (2016). *Phys. Rev. Lett.* **117**, 158101.
- Zhou, H.-W., Burger, C., Wang, H., Hsiao, B. S., Chu, B. & Graham, L. (2016). *Acta Cryst. D* **72**, 986–996.



FOUNDATIONS
ADVANCES

Volume 79 (2023)

Supporting information for article:

Small-angle scattering tensor tomography algorithm for robust reconstruction of complex textures

Leonard C. Nielsen, Paul Erhart, Manuel Guizar-Sicairos and Marianne Liebi

Supplementary Note 1: Coordinate mapping

Supplementary Note 2: Tensors and spherical harmonics

Supplementary Note 3: Orientation analysis

Supplementary Note 4: Additional results

Supplementary Note 5: Solving regularized least-squares system

Supplementary Note 6: Implementation details

Supplementary Note 1. COORDINATE MAPPING

In the ‘‘Formalism’’ section of the paper, the John Transform is described using vectors in the local coordinate system of the sample. This is the most convenient way to parameterize the system, since it is the sample volume that is ultimately being solved for. However, for practical reasons, it is useful to lay out how one can map a coordinate system of a laboratory frame, relative to which the sample is moving. In general, we can specify the relationship between the laboratory system and the sample system by choosing a set of vectors and rotation axes at zero tilt and rotation, and then applying any rotations of the sample to those vectors. Three positioning vectors are needed - one for the direction of the x-ray beam, and two which indicate how the beam moves relative to the sample during scanning. We may call these \mathbf{p} , \mathbf{w}_j and \mathbf{w}_k . The vectors \mathbf{w}_j and \mathbf{w}_k should be shifted to align the system, such that if $\mathbf{w}_j = \mathbf{w}_k = 0$, the beam intersects the center of the sample at all rotations. Additionally, two are required to indicate how angles on the detector map to the reciprocal space sphere, we may call these \mathbf{q}_0 and \mathbf{q}_{90} , defined to indicate where on the reciprocal space sphere detector angle 0 and 90 lie. Two orthogonal rotation axes need to be defined, orthogonal to the projection direction, in order to translate rotations to the sphere of projection, which we call $\hat{\alpha}$ and $\hat{\beta}$. These do not necessarily have to correspond to the axes of real rotation stages. Then, any rotation \mathbf{R} that the sample is subject to must be decomposed into three components, $\mathbf{R}_\alpha(\alpha)$, $\mathbf{R}_\beta(\beta)$, and $\mathbf{R}_p(\gamma)$, for rotations around $\hat{\alpha}$, $\hat{\beta}$ and \mathbf{p} respectively, such that

$$\mathbf{R} = \mathbf{R}_p(\gamma)\mathbf{R}_\beta(\beta)\mathbf{R}_\alpha(\alpha)$$

Then, any one of the vectors \mathbf{a} in the zero-rotation system will become

$$\mathbf{a}' = \mathbf{R}_\alpha^T(\alpha)\mathbf{R}_\beta^T(\beta)\mathbf{R}_p^T(\gamma)\mathbf{a}$$

The John transform parameters j and k are then given by the norms of \mathbf{w}_j and \mathbf{w}_k , whereas α and β are the arguments of the rotation matrices \mathbf{R}_α and \mathbf{R}_β . The rotation matrix \mathbf{R}_p simply specifies a rotation around the axis of projection, and therefore only changes the mapping of j and k , rather than representing a degree of freedom on the sphere of projection. The reciprocal space angles θ and ϕ are given by the azimuthal and polar angle of the reciprocal space vector

$$\mathbf{q}' = \mathbf{R}^T(\cos(\varphi)\mathbf{q}_0 + \sin(\varphi)\mathbf{q}_{90})$$

where φ is the angle on the detector.

The vectors \mathbf{v} and \mathbf{u} in the John transform are given by

$$\begin{aligned}\mathbf{v} &= \mathbf{R}^T(\mathbf{w}_j + \mathbf{w}_k) \\ \mathbf{u} &= \mathbf{R}^T(\mathbf{p})\end{aligned}$$

Using these rules, one can calculate the John transform using rotation information from any laboratory coordinate system.

Supplementary Note 2. TENSORS AND SPHERICAL HARMONICS

The comparison between methods using spherical harmonic representations and tensor representation merits a brief discussion of isomorphisms between these parameterizations. Spherical harmonics are simply harmonic polynomials on the sphere. In this work we use the *real* spherical harmonics, which we may define as

$$\hat{Y}_m^\ell(\theta, \phi) = \begin{cases} L_m^\ell(\theta) \cos(m\phi) & m \geq 0 \\ L_m^\ell(\theta) \sin(|m|\phi) & m < 0 \end{cases}$$

where θ is the polar angle, ϕ is the azimuthal angle, $L_m^\ell(\theta)$ is the associated Legendre polynomial of degree ℓ and order m , with $\ell \geq 0$ and $|m| \leq \ell$. Because of this, they have an intuitive mapping to traceless symmetric Cartesian tensors when representing functions on the unit sphere. In particular, a band-limited spherical function represented in $\ell = n$ maps directly to a traceless rank- n tensor. The rank-0 tensor is simply a constant, so it naturally maps to $\hat{Y}_0^0(\theta, \phi)$, which is a constant term. This mapping naturally leads to restricting the mapping to *traceless* tensors, because the trace of a tensor on the unit sphere is a constant term. Moreover, a spherical harmonic of $\ell = n$ can be represented as a polynomial of Cartesian coordinates of degree n , that is to say

$$P(n, \mathbf{x}) = a_m \hat{Y}_m^n(\mathbf{x}) \\ x_i x_i = 1$$

where Einstein summation over repeated indices is used, and \mathbf{a} is some vector of $2n + 1$ coefficients. In the same way, we have for a traceless rank- n tensor T the mapping

$$P(n, \mathbf{x}) = T^{ijk\dots} x_i x_j x_k \dots,$$

$P(n, \mathbf{x})$ is the value of some polynomial of degree n at \mathbf{x} . Restricting ourselves to the unit sphere, it is then the case that

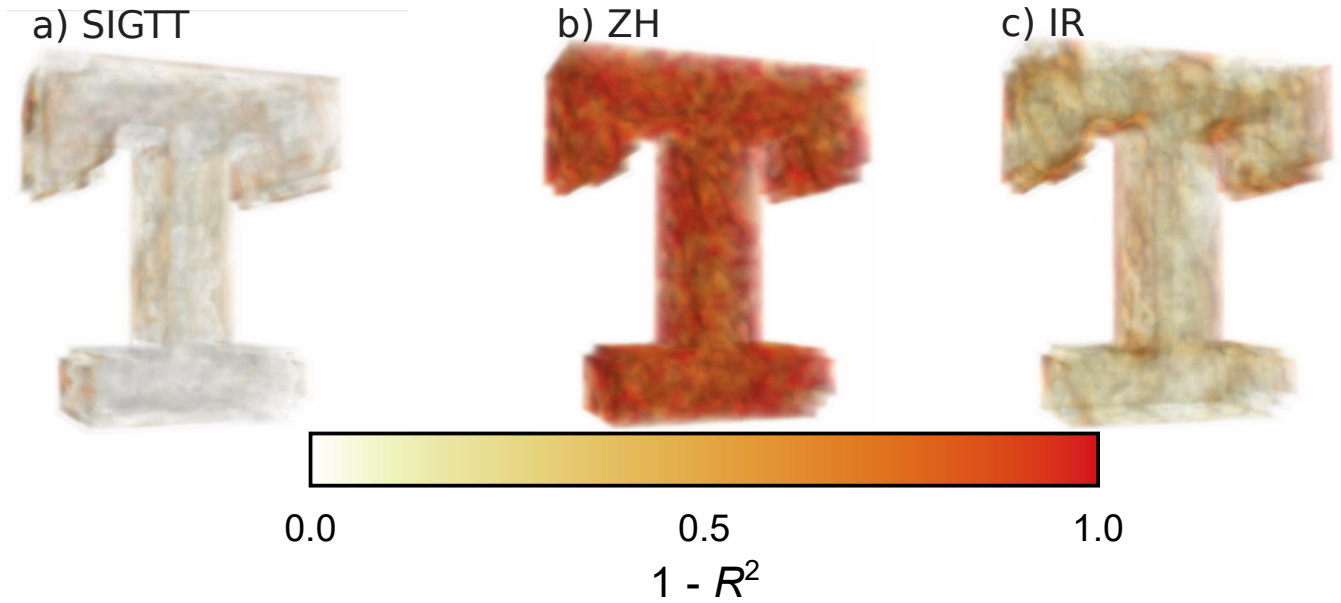
$$a_m \hat{Y}_m^n(\mathbf{x}) = T^{ijk\dots} x_i x_j x_k \dots,$$

By the additivity of polynomials, this mapping can be extended to sums of traceless symmetric tensors of arbitrary orders, and through Gaussian elimination, it is straightforward to calculate a mapping between spherical harmonics and symmetric Cartesian tensors. Tensor and spherical harmonic representations of functions on the unit sphere are thus isomorphic, and their different analytical and algebraic properties render them suitable for different purposes. Spherical harmonics are especially suitable for calculations where rotational invariance is crucial, and for calculating spherical statistics such as the variance or mean over the sphere. On the other hand, it would be easier to perform tensor calculus (such as computing the divergence or curl) on Cartesian tensors.

Supplementary Note 3. ORIENTATION ANALYSIS

Orientation analysis is used in figures 4 and 6 of the main work, as well as in the creation of the simulated sample “M”. The identification of a reciprocal space map’s orientation is done by eigenvector-eigenvalue decomposition of its $\ell = 2$ spherical harmonic coefficients. The $\ell = 2$ coefficients are translated into a symmetric traceless rank-2 tensor by solving for each coefficient in the tensor’s polynomial representation. Specifically, on the unit sphere, the two representations may be expanded as

$$\begin{aligned} \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{xy} & T_{yy} & T_{yz} \\ T_{xz} & T_{yz} & T_{zz} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= x^2 T_{xx} + y^2 T_{yy} + z^2 T_{zz} + 2(xy T_{xy} + xz T_{xz} + yz T_{yz}) \\ \mathcal{N}(2) \begin{bmatrix} a_0^2 & a_1^2 & a_{-1}^2 & a_2^2 & a_{-2}^2 \end{bmatrix} \begin{bmatrix} \frac{2z^2 - y^2 - x^2}{2\sqrt{3}} \\ xz \\ yz \\ xy \\ \frac{x^2 - y^2}{2} \end{bmatrix} &= \mathcal{N}(2) \left(x^2 \left(\frac{a_{-2}^2}{2} - \frac{a_0^2}{2\sqrt{3}} \right) + y^2 \left(-\frac{a_{-2}^2}{2} - \frac{a_0^2}{2\sqrt{3}} \right) + z^2 \frac{a_0^2}{\sqrt{3}} + a_2^2 xy + a_1^2 xz + a_{-1}^2 yz \right) \end{aligned}$$



Supplementary Figure S-1. Errors for “T”. Volume renders of the errors of a) SIGTT, b) SH, and c) IR for sample “T”, defined as $1 - R^2$, where R^2 is given by equation (15) of the main text. Larger errors are rendered with greater opacity and are thus visible even if they are in the interior.

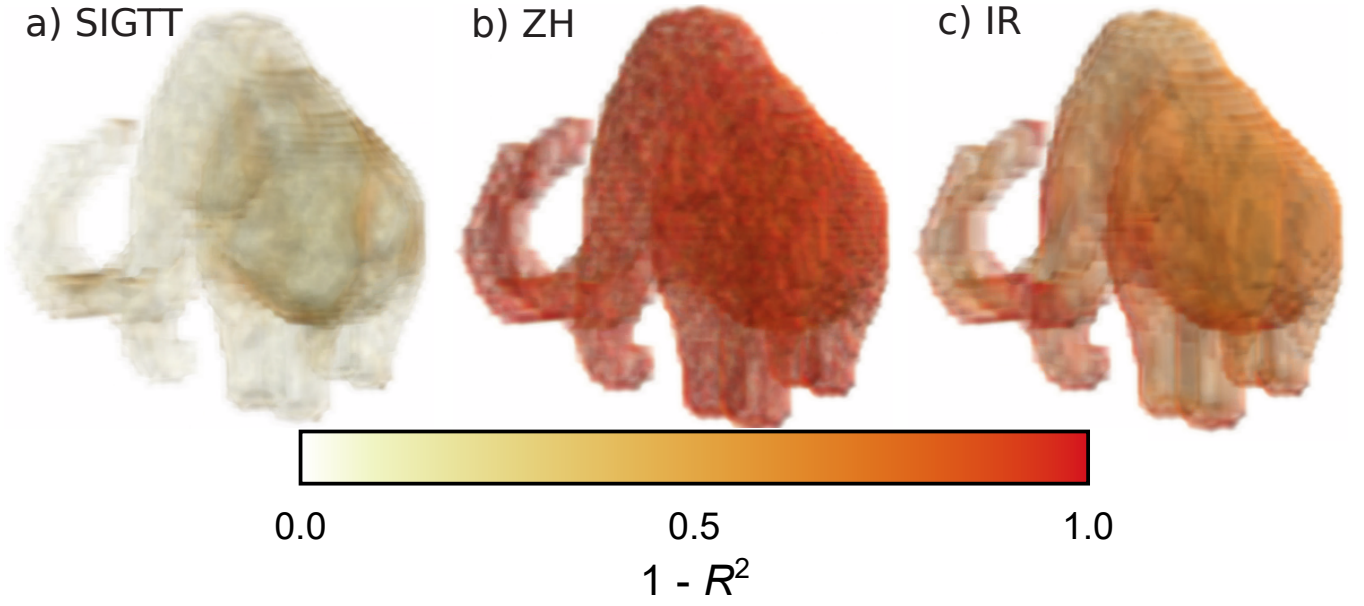
where $\mathcal{N}(2)$ is a normalization factor depending on the spherical harmonic representation, a_m^2 is a spherical harmonic coefficient with $\ell = 2$, T_{ii} are components of a rank-2 tensor, and (x, y, z) are cartesian coordinates. By equating these two systems of equations, we can then identify each T_{ii} with a linear combination of coefficients a_m^2 and thus construct a rank-2 tensor out of spherical harmonic coefficients. Subsequently, we may solve the eigenvalue problem for the rank-2 tensor, to obtain three eigenvalues (which will sum to zero as the tensor will be traceless) and three orthogonal eigenvectors. For a rank-2 tensor with three distinct eigenvalues, the eigenvectors associated with the largest and smallest eigenvalues correspond to the location of the minima and maxima of the spherical function. The eigenvector of the central eigenvalue corresponds to a saddle point. If there is degeneracy in the eigenvalues, but they are nonzero, the unique eigenvalue corresponds to one extremum (minimum if it is negative, maximum if it is positive), whereas the two identical eigenvalues correspond to two orthogonal points on a set of extreme points which lie along a great circle. The orientation analysis then depends on assumptions about the symmetry of the nanostructure associated with the reciprocal space map. If the nanostructure is known or assumed to be scattering from fiber-like structures, orthogonal to the direction of the fibers, the eigenvector associated with the smallest eigenvalue is taken to define the orientation. On the other hand, if the reciprocal space map is known or assumed to be scattering along the orientation, the eigenvector associated with the largest eigenvalue is taken to define the orientation. Nanostructures which contain both of these scattering tendencies, such as bone around the q-range of the collagen peak, cannot in general be robustly analyzed in this manner, because each scattering tendency will tend to cancel out the contribution from the other to the rank-2 tensor component of the reciprocal space map. Such reciprocal space maps require a different treatment, such as correlation analysis with an ensemble of model functions.

Supplementary Note 4. ADDITIONAL RESULTS

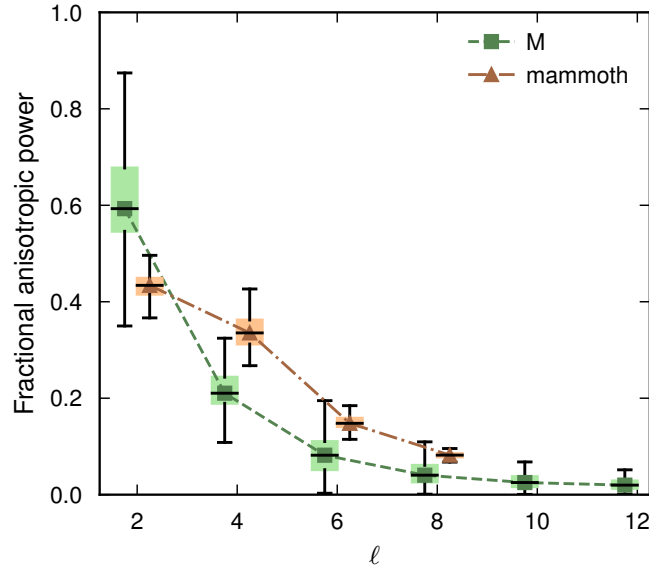
This section shows some additional results for the simulation comparisons to supplement those in the main work.

Volume renders of the errors of each method for “T” can be seen in Supplementary Figure S-1. It is evident from the figure that Spherical Integral Geometric Tensor Tomography (SIGTT) has the smallest errors, and that the large errors are concentrated around the edges. It is followed by IR, which has large regions of small errors but larger errors around the same regions as SIGTT. However, the errors for SH are far greater than for either of the other two methods. This is partly because the squared polynomials of SH cannot exactly represent rank-2 tensors, and partly because sample “T” is not constrained to zonal symmetry.

A similar set of volume renders of errors for each method for “mammoth” can be seen in Supplementary Figure S-2. Again, SIGTT has the smallest errors, with larger errors around some edge regions. In this case, SH and IR are more similar, having large errors throughout, but IR consistently appears to have somewhat smaller errors. This is



Supplementary Figure S-2. Errors for “mammoth”. Volume renders of the errors of a) SIGTT, b) SH, and c) IR for sample “mammoth”, defined as $1 - R^2$, where R^2 is given by equation (15) of the main text. Larger errors are rendered with greater opacity and are thus visible even if they are in the interior.



Supplementary Figure S-3. Fractional anisotropic power. Distribution of the fraction of anisotropic power (equation (S1)) at each ℓ for the samples “M” and “mammoth”. Sample “M” has a distribution that is more dominated by the $\ell = 2$ component, and a steeper decline, whereas for “mammoth”, the $\ell = 2$ and $\ell = 4$ components are closer to each other, and $\ell = 6$ and $\ell = 8$ also occupy a larger fraction than for “M”.

consistent with figure 5 in the main text. where is it clear that IR can achieve a correlation around 0.3, but not more, whereas the correlation for SH is close to 0. These large errors for IR and SH are due to the fact that the “mammoth” sample has a small rank-2 tensor component on the one hand, and on the other hand, no zonal symmetry constraint.

To examine the distribution of anisotropic power in samples, we define the quantity

$$F_\ell(g) = \frac{S_\ell(g)}{\text{var}(g)} \quad (\text{S1})$$

where g is a spherical function, and S_ℓ and $\text{var}(g)$ are defined in Equations (8) and (11) in the main text respectively. This quantity, referred to as the fractional anisotropic power, is then evaluated for every voxel that is part of the sample. In Supplementary Figure **S-3**, we see box plots of the fractional anisotropic power of “M” and “mammoth”. While “M” has a spectrum that goes up to $\ell = 12$, the orders about $\ell = 6$ only contribute a small amount of the overall power.

Supplementary Note 5. SOLVING REGULARIZED LEAST-SQUARES SYSTEM

The system of linear equations to be solved in SIGTT is given as

$$\mathbf{X}_{\text{opt}} = \underset{\mathbf{X}}{\text{argmin}} \left[\|\mathbf{P}\mathbf{X}\mathbf{Y} - \mathbf{D}\|^2 + \lambda \|\nabla^2 \mathbf{X}\|^2 \right].$$

This system is solved through the gradient-based method L-BFGS-B. In detail, a general solution using a quasi-Newton algorithm may be written

$$\begin{aligned} 2r_i &= \|\mathbf{P}\mathbf{X}_i\mathbf{Y} - \mathbf{D}\|^2 + \lambda \|\nabla^2 \mathbf{X}_i\|^2 \\ \nabla r_i &= \mathbf{P}^T (\mathbf{P}\mathbf{X}_i\mathbf{Y} - \mathbf{D})\mathbf{Y}^T + \lambda \nabla^2 \mathbf{X}_i \\ \mathbf{X}_{i+1} &= \mathbf{X}_i - \alpha_i \nabla r_i + \beta_i \mathbf{p}_i \end{aligned}$$

where r_i is the residual for iteration i , \mathbf{X}_i is the estimated solution, α_i and β_i are method-dependent scalars, and \mathbf{p}_i is a method-dependent momentum term. While the multiplication with \mathbf{Y}^T is directly implemented as a matrix multiplication, this is not practical to do for the adjoint projection operation \mathbf{P}^T , due to the prohibitively large size of the matrix. Instead, it is computed using the common iterative tomographic method of back-projection of the residual.

Supplementary Note 6. IMPLEMENTATION DETAILS

For each of the simulated data sets, an overall signal-to-noise ratio (SNR) was estimated. Noise was added to each projected simulated data point per

$$I_{\text{noise}} = \frac{\text{poisson}(I \cdot k)}{k}$$

with k being a noise parameter, and $\text{poisson}(x)$ being a sampling of x with added Poisson noise. The values used for k were $k = 10^t$ with $t \in \{2, 1.5, 1, 0.5, 0\}$. This means that each data point will be a random variable with variance I/k . Thus, the signal-to-noise ratio for a simulated data set with noise parameter k was estimated as

$$\text{SNR} \approx \sqrt{\bar{\mu}(I_i)k}$$

with $\bar{\mu}(I_i)$ being the average intensity of all non-background data points in that data set. This SNR should be understood principally as a relative measure for a particular simulation, not as an absolute measure.