Expansion-maximization-compression algorithm with spherical harmonics for single particle imaging with x-ray lasers

Julien Flamant*  
Univ. Lille, Centre National de la Recherche Scientifique, Centrale Lille, UMR 9189, CRIStAL Centre de Recherche en Informatique Signal et Automatique de Lille, 59000 Lille, France

Nicolas Le Bihan  
Centre National de la Recherche Scientifique, GIPSA Laboratory, 11 Rue des Mathématiques, Domaine Universitaire, BP 46, 38402 Saint Martin d’Hères Cedex, France

Andrew V. Martin  
ARC Centre of Excellence for Advanced Molecular Imaging, School of Physics, The University of Melbourne, Victoria 3010, Australia

Jonathan H. Manton  
Department of Electrical and Electronic Engineering, The University of Melbourne, Victoria 3010, Australia  
(Received 3 February 2016; published 11 May 2016)

In three-dimensional (3D) single particle imaging with x-ray free-electron lasers, particle orientation is not recorded during measurement but is instead recovered as a necessary step in the reconstruction of a 3D image from the diffraction data. Here we use harmonic analysis on the sphere to cleanly separate the angular and radial degrees of freedom of this problem, providing new opportunities to efficiently use data and computational resources. We develop the expansion-maximization-compression algorithm into a shell-by-shell approach and implement an angular bandwidth limit that can be gradually raised during the reconstruction. We study the minimum number of patterns and minimum rotation sampling required for a desired angular and radial resolution. These extensions provide new avenues to improve computational efficiency and speed of convergence, which are critically important considering the very large datasets expected from experiment.

DOI: 10.1103/PhysRevE.93.053302

I. INTRODUCTION

A. Single particle imaging with x-ray lasers

Three-dimensional (3D) single particle imaging with x-ray free-electron lasers (XFELs) is being actively pursued for its potential biological applications, which notably include determining the structures of single molecules [1]. It is also one of the most challenging goals in x-ray laser science, because each diffraction measurement provides a very weak signal and is destructive. As few as $5 \times 10^{-7}$ photons per Shannon-Nyquist pixel (at high scattering angles) are expected for a 500-kDa protein [2]. To overcome low signal to noise, the experiment involves a large number of diffraction measurements ($10^4$–$10^6$) and each measurement must be of a new copy of the particle, which is assumed to have an identical structure. Such datasets are achievable because XFELs have high repetition rates (>100 Hz [3]) and serial injection technology has been developed to continuously deliver fresh samples into the beam path [4,5]. A consequence of this experimental approach is that particle orientation is not measured and must be determined by analyzing the diffraction data [6]. In practice, a single orientation is not associated with each diffraction pattern, but rather the low signal per pattern is more effectively handled by treating the data globally to directly obtain a 3D diffraction intensity volume [2,7]. This intensity volume is then given as input for a phase retrieval algorithm that produces an image of the particle [8].
such as changing beam conditions or sample heterogeneity. The advantage here is that the generation of the manifold is not biased by human assumptions. The challenge is interpreting the manifold and identifying the correct relationship between the three degrees of freedom on the manifold and the rotation group. One way to explicitly relate the data space to rotations is via mapping geodesics [14]. The first attempt to introduce geometrical constraints in the processing of diffraction patterns was made by Saldin et al. [15], where the three-dimensional intensity function was expanded on the spherical harmonic basis. Their approach was based on the cross-correlation of the diffraction patterns and exploits the orthogonality of spherical harmonics to obtain a decomposition of the cross-correlation function. It is interesting to note that this method has been tested experimentally [16] on large dimers with known structure. For objects of known rotational symmetry, the correlation function can be converted into the 3D single particle Fourier intensity, which can then be converted to an image of the particle via phasing. For some time, it was not known how to extend this method to an asymmetric 3D object, but recently an algorithm has been proposed to phase directly from the correlation function [17]. The correlation function can also be used with molecular replacement techniques [15]. It has the added advantage that multiple particles can be illuminated per measurement, which not many other algorithms can handle.

In [7], Loh and Elser have introduced the expansion-maximization-compression (EMC) algorithm which relies on an expectation-maximization (EM) technique. As a Bayesian method, it aims to use known information about the noise statistics to handle very low signal-to-noise levels. The algorithm iteratively maximizes the likelihood of the reconstructed intensity given the set of diffraction patterns. The algorithm does not assign a single molecular orientation to each pattern, but rather it estimates the probability for a diffraction pattern to be associated with a certain orientation. Recently, this algorithm has proven its feasibility with the reconstruction of Mimivirus from experimental data collected at Linac Coherent Light Source (LCLS) [18]. The EMC algorithm is studied in detail in this work, with the introduction of several improvements. Recently, Waleczak and Grubmiller [19] developed the Bayesian approach further using seed structural models to speed up convergence and to discriminate between different conformations. Finally, we note that Tegze and Bortel [20] have proposed a simplified version of the EMC algorithm, where the diffraction pattern orientation is fully assigned through the intensity best fit, rather than in a probabilistic way. A drawback of the EMC algorithm is that it is computationally expensive to implement on large datasets that are expected in experiment. The EMC algorithm stores a matrix that grows proportionally with the number of diffraction patterns and must be recalculated at every iteration. This is not a trivial step if there are $10^5 - 10^6$ measured diffraction patterns. Hence, there is still scope to develop the EMC algorithm further to enable parallelization and efficient computation. Another related issue is analyzing smaller datasets at lower resolution, which would be very valuable during an experiment to provide feedback on data collection. Algorithms that are readily scaled with the number of patterns and resolution are also valuable.

With these considerations in mind, we study here the application of spherical harmonic analysis to the EMC algorithm. Our goal is to produce an EMC algorithm that can be scaled in terms of angular and radial resolution. This facilitates starting with lower resolution reconstruction with less patterns, higher signal, and less computational time, before proceeding to higher resolutions. By implementing the algorithm shell by shell, we also enable it to be easily parallelized. After reviewing the relevant properties of spherical harmonics in Sec. II, we formulate the EMC algorithm for a single q shell in Sec. III and impose an angular bandwidth limit. Section IV describes how reconstructions on neighboring shells can be aligned via correlations, and finally numerical tests are provided in Sec. V.

While our work improves upon the computational requirements for EMC, it does not match the computational advantages of correlation methods based on spherical harmonics that do not require more memory as the number of patterns increases [15]. As described above, however, correlation-based methods have other issues for particles without rotational symmetry, while EMC has potential advantages for analyzing low signal data, which contributes to its popularity. Explorations of how to retain the advantages of EMC at lower computational expense are thus valuable for the development of XFEL single particle imaging.

II. THE GEOMETRY OF INTENSITY FUNCTIONS

The intensity function of a biomolecule $I(q)$ is linked to $F(q)$, the Fourier transform of its 3D electron density $\rho^{\text{mol}}(r)$, with $r$ and $q$, respectively, being the real and reciprocal space coordinates. During XFEL experiment, each diffraction pattern is obtained from a randomly rotated copy of the biomolecule. A rotation of the molecule in the real space corresponds to the same rotation around the origin of the Ewald sphere in the dual space. The accumulation of thousands of diffraction patterns from rotated molecules can cover a spherical volume in the reciprocal space, in order to reconstruct the 3D intensity of the biomolecule.

With the notation previously introduced, the Fourier transform or molecular transform $F$ of the three-dimensional electron density of the molecule reads

$$F(q) \triangleq \int \rho^{\text{mol}}(r) \exp(iq \cdot r) d^3r.$$  \hspace{1cm} (1)

The intensity function $I$ is defined as the square magnitude of the Fourier transform of the electron density, up to a constant normalizing factor $I_0$ depending on the experimental conditions [21]. Thus we have the following:

$$I(q) = |F(q)|^2 I_0.$$  \hspace{1cm} (2)

In what follows, we will be using spherical coordinates for $q$, meaning that its parametrization will be

$$q = (q \sin \theta \cos \phi, q \sin \theta \sin \phi, q \cos \theta)^T,$$  \hspace{1cm} (3)

where $q = ||q||$ is the radial coordinate, and $(\theta, \phi) \in [0, \pi) \times [0, 2\pi)$ are angular coordinates, where the polar angle is relative to the beam axis. In the following we make use of the shorthand notation $q = (q, \Omega)$, where $\Omega$ is a unit vector pointing towards angular coordinates $(\theta, \phi)$.
In order to take advantage of the symmetries of the intensity function $I$, we propose to consider it as a set of concentric shells of increasing radii. On such a shell of radius $q_s$, where $s$ stands for the shell number, we define the intensity $I'(\Omega) = I(q_s, \Omega)$. In the reconstruction process, the three-dimensional intensity function will be recovered by assembling the reconstructed intensities over spherical shells.

The shell-by-shell reconstruction can be made efficient using the fact that each $I'$ can be decomposed on a spherical harmonic basis. Such a spherical basis is presented in Sec. III B. Before introducing this decomposition, we first present the measurement model.

A. Equivalent single shell measurement model

In order to perform shell-by-shell reconstruction of the intensity function, we introduce a measurement model which relates diffraction data to the corresponding part of the intensity on each shell.

In this model, as shown in Fig. 1, the corresponding diffraction patterns on a spherical shell $s$ are monodimensional, i.e., circles on this spherical shell. Formally, we denote by $D^s$ the detector on the shell, that is the reference sampling points. $D^s$ is defined as the set of $N_s$ pixel coordinates on the spherical shell $I'(\Omega)$ and given by

$$ D^s = \{ q_i = (q_s, \Omega_i) | i = 1, 2, \ldots, N_s \}, \quad (4) $$

where $\Omega_i$ is given in angular coordinates by $(\varphi_i, \theta(q_s)) \in [0, 2\pi) \times [0, \pi]$, and $q_s$ is the radius of the spherical shell. The dependence of $\theta$ on the radius $q_s$ is a consequence of the scattering geometry. For completeness we recall that

$$ \theta(q_s) = \frac{\pi}{2} - \arcsin \frac{q_s}{q_{\text{max}}}, \quad (5) $$

where $q_{\text{max}}$ is defined in terms of the wavelength $\lambda$, $q_{\text{max}} = 4\pi/\lambda$.

On the detector, the number of pixels contributing to a shell, i.e., the number of pixels on a circle on the diffraction pattern, increases with the radius $q_s$. The number of points $N_s$ on the detector is scaled such that

$$ N_s = \left\lceil \frac{2\pi q_s}{\Delta q} \right\rceil, \quad \varphi_i = \frac{2\pi}{N_s} i, \quad i \in \{0, 1, \ldots, N - 1\}, \quad (6) $$

where $\lceil \cdot \rceil$ denotes the ceiling function, and $\Delta q$ is the reciprocal space pixel size.

Finally, each measurement $k$ or diffraction pattern on a shell consists of a set of $N_s$ samples $y_{ik}$, $i = 1, \ldots, N_s$, given by a Poisson model [22] (due to low-photon counts expected in single particle experiments):

$$ y_{ik} \sim \text{Pois}(I'(R_k \cdot \Omega_i)), \quad R_k \sim U(\text{SO}(3)), \quad (7) $$

where the random rotation $R_k$ is drawn from the uniform distribution on the rotation group SO(3) and where $\text{Pois}(\alpha)$ is the Poisson distribution of parameter $\alpha$.

B. Harmonic decomposition of the shell intensity

The intensity on a given shell $s$, i.e., $I'(\Omega)$, can be expanded on the spherical harmonics basis (see Appendix A for details) as follows:

$$ I'(\Omega) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} c^{m}_\ell(q_s) Y^m_\ell(\Omega), \quad (8) $$

where the coefficients $c^{m}_\ell(q_s)$ are the spherical harmonic coefficients of the intensity function on the shell $s$. It turns out that the spherical harmonic representation provides an efficient way to deal with the symmetries of the intensity functions. In the following, we exploit the physical properties of the intensity function to simplify further the expansion above.

First, intensity functions are by definition real-valued functions. As a consequence, their spherical harmonic coefficients exhibit the well-known conjugation property of spherical harmonics:

$$ \text{for all } \ell \geq 0, \quad c^{m}_\ell(q_s) = (-1)^m \overline{c^{-m}_\ell}, \quad m \in \{0, 1, \ldots, \ell\}. \quad (9) $$

Moreover, another interesting property is the centrosymmetry of intensity functions, also known as the Friedel property. This symmetry property reads $I'(\Omega) = I'(-\Omega)$ and it is straightforward to see that with this constraint every coefficient of odd degree $\ell$ is set to zero:

$$ \text{for all } p \geq 0, \quad c^{2p+1}_\ell = 0, \quad m \in \{0, 1, \ldots, 2p + 1\}. \quad (10) $$

Finally, the finite size nature of the molecule in real space implies that the spectral representation on each shell is effectively band limited, meaning that the coefficients in the spherical harmonics expansion are nonzero up to a maximum degree $L_s^{\text{lim}}$:

$$ I'(\Omega) = \sum_{\ell=0}^{L_s^{\text{lim}} - 1} c^{m}_\ell(q_s) Y^m_\ell(\Omega), \quad (11) $$

where $L_s^{\text{lim}}$ is the band limit on the shell $s$, and where the sum has been restricted to even values of the degree $\ell$. Since Friedel symmetry imposes condition (10) on the coefficients, the value of $L_s^{\text{lim}}$ is an odd number.

It is interesting to note that rotational symmetries of the particle will further reduce the number of independent coefficients. This makes it easier to implement particle symmetry than for the Cartesian sample of the intensity function used in the original EMC algorithm and may improve performance in the case of highly symmetric particles.

III. SPHERICAL EMC ALGORITHM

We develop in this section an extension of the original EMC algorithm [7] to the spherical harmonic basis [23]. Precisely, we define an update rule in terms of the spherical harmonic coefficients of the shell intensity, and scale efficiently the computation time in terms of the band limit $L$.

In what follows, all quantities with superscript $(n)$ denote the corresponding values at the $n$th iteration of the EM algorithm.

A. A general expectation-maximization framework

We start this section by recalling the framework of EM algorithms. In general, EM algorithms provide an efficient way to deal with missing data [24]. In our spherical setting, we design an EM algorithm which updates the spherical
harmonics coefficients at each iteration; namely, we write the update rule

$$\hat{c}^{(n)} \rightarrow \hat{c}^{(n+1)},$$

where $n$ denotes the iteration index. Note that we distinguish the estimated coefficients $\hat{c}$ from the theoretical ones $c$ to avoid confusion. At each iteration of the EM algorithm, the likelihood of the estimated parameter $\hat{c}$ given the set of measurements $y_1, \ldots, y_K$ is increased. However, EM algorithms do not work directly with the likelihood function, but rather with the intermediate quantity $Q(\hat{c}|\hat{c}^{(n)})$, the value of which is determined in the expectation step (e-step) [25] of the algorithm. For practical purposes, we first introduce the intermediate quantity per pattern $Q_k(\hat{c}|\hat{c}^{(n)})$, which reads

$$Q_k(\hat{c}|\hat{c}^{(n)}) = \int_{SO(3)} \log [p(y_k, R|\hat{c})] p(R|y_k, \hat{c}^{(n)}) d\mu(R),$$

where $d\mu(R)$ denotes the (bi)-invariant Haar measure on $SO(3)$. Since the acquisitions $y_1, \ldots, y_K$ are independent from each other, we can express the total intermediate quantity as the sum

$$(\mathbf{e - step}): \quad Q(\hat{c}|\hat{c}^{(n)}) = \frac{1}{K} \sum_{k=1}^{K} Q_k(\hat{c}|\hat{c}^{(n)}),$$

where the choice of the normalization made here is arbitrary.

The maximization step can now be defined as the maximization of the (total) intermediate quantity, such that it reads

$$(\mathbf{M - step}): \quad \hat{c}^{(n+1)} = \arg \max_{\hat{c}} Q(\hat{c}|\hat{c}^{(n)}).$$

We note here that there are potentially several values of $\hat{c}$ that maximize the intermediate quantity; therefore uniqueness is not guaranteed. This is a general property of expectation-maximization algorithms.

**B. Proposed spherical EMC algorithm**

1. **EM by tomographic grid update**

Considering the EM algorithm expressed directly with the spherical harmonic coefficients is a complicated task, which leads to intractable maximizations procedures. Rather than solving the problem directly in the spherical harmonic domain, we extend the original approach of Loh and Elser [7] to a spherical framework.

Let us introduce a spherical intensity model, denoted by $W$. This spherical intensity model $W$ can be expressed using a spherical harmonic expansion similar to Eq. (8). The
corresponding intermediate quantity per pattern reads
\[ Q_k(W|W^{(n)}) = \int_{\text{SO}(3)} \log \left[ p(y_k, R|W_R) \right] \times p(R|y_k, W^{(n)}_R) \, d\mu(R), \]
where \( W_R \) is a shorthand notation for the set \( \{ W_i = W(R \cdot q_i), \ i = 1, \ldots, N \} \). Recall that the \( q_i \) denote the reciprocal space coordinates of the detector.

Now, let us give details on the different quantities above. The conditional probabilities read
\[ p(y_k, R|W_R) = p(y_k|R, W_R) p(R|W_R), \]
\[ p(R|y_k, W^{(n)}_R) = \frac{p(y_k|R, W^{(n)}_R) p(R|W^{(n)}_R)}{p(y_k|W^{(n)}_R)} \times p(R|W_R). \]

The first equality is simply the definition of the joint probability, whereas the second equality is obtained by applying Bayes’s rule. We note that those quantities involve the prior distribution on the orientations, i.e., Bayes’s rule. We note that those quantities involve the prior ability, whereas the second equality is obtained by applying the first equality is simply the definition of the joint probability of the particle and that the distribution of orientations is uniform, which means with respect to the bi-invariant Haar measure on \( \text{SO}(3) \) that
\[ p(R) = 1, \forall \ R \in \text{SO}(3). \]

Moreover, the Poisson assumption on the photon count and the independence of each pixel yields
\[ p(y_k|R, W_R) = \prod_{i=1}^{N} \left( \frac{(W_{i,R})^{y_{ik}}}{y_{ik}!} \right) \exp(-W_{i,R}), \]
which gives the expression of the intermediate quantity per pattern:
\[ Q_k(W|W^{(n)}) = Z_1 + Z_2^{-1} \int_{\text{SO}(3)} d\mu(R) \times \left[ \sum_{i=1}^{N} y_{ik} \log W_{i,R} - W_{i,R} \right] \times \prod_{i=1}^{N} \left( W_{i,R}^{(n)} \right)^{y_{ik}} \exp(-W_{i,R}^{(n)}), \]
where \( Z_1 \) is a constant only depending on the data \( y_k \) (which is arbitrarily removed hereafter), and \( Z_2 \) is a normalization constant given by
\[ Z_2 = \int_{\text{SO}(3)} d\mu(R) \prod_{i=1}^{N} \left( W_{i,R}^{(n)} \right)^{y_{ik}} \exp(-W_{i,R}^{(n)}). \]

Now, we face the following problem of computing the integrals over the rotation group \( \text{SO}(3) \). Unfortunately there is no exact way to do so due to the lack of quadrature formula for probability functions on the rotation group \( \text{SO}(3) \). Nevertheless, the integrals above can be approximated via a discrete sum over elements of the rotation group. More precisely, if we suppose that we have some sampling set \( \mathcal{X} \subset \text{SO}(3) \) of size \( M \) and the elements of which are indexed by \( j \), the intermediate quantity reads
\[ Q_k(W|W^{(n)}) \simeq \tilde{Z}_2^{-1} \sum_{j=1}^{M} w_j \left[ \sum_{i=1}^{N} y_{ik} \log W_{ij} - W_{ij} \right] \times \prod_{i=1}^{N} \left( W_{ij}^{(n)} \right)^{y_{ik}} \exp(-W_{ij}^{(n)}), \]
where \( w_j \) are equivalent to quadrature weights such that \( \sum_j w_j = 1 \), \( W_{ij} \approx W_{i,R} \), and where \( \tilde{Z}_2 \) is obtained following the same guidelines. Since the model \( W \) is now evaluated at \( M \) rotated versions of the original detector coordinates, we call it a tomographic model. In a similar way, we define the tomographic grid coordinates as \( q_{ij} = \hat{R}_j \cdot q_i \).

In short form, we can rewrite the latter intermediate quantity as a function of the tomographic model:
\[ Q_k(W|W^{(n)}) \simeq \sum_{j=1}^{M} P_{jk} \left( \sum_{i=1}^{N} y_{ik} \log W_{ij} - W_{ij} \right), \]
where \( P_{jk} \) is given by
\[ P_{jk} = \frac{w_j \prod_{i=1}^{N} \left( W_{ij}^{(n)} \right)^{y_{ik}} \exp(-W_{ij}^{(n)})}{\sum_{j=1}^{M} w_j P_{jk}}. \]
Finally, the corresponding expectation step of the EM algorithm is given by
\[ Q(W|W^{(n)}) \simeq \sum_{k=1}^{K} \sum_{j=1}^{M} P_{jk} \left( \sum_{i=1}^{N} y_{ik} \log W_{ij} - W_{ij} \right). \]
procedure, that is,

\[
W_{ij}^{(n+1)} = \frac{\sum_{k=1}^{K} P_{jk} y_{ik}}{\sum_{k=1}^{K} P_{jk}}, \quad i = 1, 2, \ldots, N, \quad j = 1, 2, \ldots, M.
\]  

(27)

We note that in recent papers [8,18], where the EMC algorithm is applied to experimental data, an update rule of the fluence in each diffraction pattern is also derived. Here we make the simplifying assumption that the fluence in each diffraction pattern is constant.

2. Spherical harmonics update through expansion-compression steps

By introducing a tomographic model \( W \), the EM procedure has been made easier. However, our goal is to obtain an algorithm which updates the spherical harmonic coefficients up to some degree \( L - 1 \) for the considered shell, and this is done by adding two extra steps, known as expansion and compression.

First, we remark that the tomographic grid coordinates are treated independently: therefore there exist a lot of tuples \((i,j) \neq (i',j')\) such that the coordinates \( q_{ij} \), \( q_{ij}' \) are really close. To enforce the consistency of the estimated model, we introduce a regular spherical grid \( W_{G_L} \) and the corresponding regular model \( W_{G_L} \), the nodes of which are given by \( q_p \). Note that the grid depends on the band limit \( L \), allowing efficient scaling of the grid size. If the grid is well chosen and exhibits nice features, then the spherical harmonics coefficients \( c^L_{ij} \) up to some degree \( L - 1 \) can be obtained using a fast implementation of the spherical harmonic transform.

The above description actually corresponds to the compression step (C-step), which can be schematically given by

\[
(C \rightarrow \text{step}) : W_{ij} \rightarrow W_{G_L} \rightarrow \hat{c},
\]

(28)

whereas the reverse operation is the expansion step (E-step)

\[
(E \rightarrow \text{step}) : \hat{c} \rightarrow W_{G_L} \rightarrow W_{ij}.
\]

(29)

3. Grid \( G \) and implementation of expansion and compression steps

We make use of the material introduced in Appendix A by choosing the HEALPix sampling scheme on the sphere for our grid \( G_L \). The HEALPix grid on the sphere provides equal-area pixelization of the sphere, along with hierarchical resolution and numerous features [26]. The compression step is performed as follows. We first determine the tomographic points \( q_{ij} \) belonging to each HEALPix pixel, and then the intensity value on this pixel is given by inverse distance weighting between the respective \( q_{ij} \) and the pixel center \( q_p \):

\[
W_{G_L}(q_p) = \frac{\sum_{\text{neighbors}} W_{ij} / \| q_{ij} - q_p \|^2}{\sum_{\text{neighbors}} 1 / \| q_{ij} - q_p \|^2}.
\]

(30)

The coefficients \( c^L_{ij}(q_r) \) are then computed up to order \( L - 1 \) as given by the sampling theorem (A7) using a fast spherical harmonic transform (SHT). The Friedel symmetry is restored by canceling the coefficients for odd values of \( \ell \).

The expansion step is done by the successive expansion of the coefficients \( c^L_{ij}(q_r) \) to obtain \( W_{G_L} \), then by computation of the tomographic intensities by interpolation on the sphere.

Precisely the intensity \( W_{G_L} \) is obtained by inverse SHT, again implemented using the HEALPix package routines. The interpolation from the regular grid \( G \) to the tomographic grid is done by bilinear interpolation using the four nearest neighbors on the regular grid.

C. Implementation

Several issues need to be considered when using the proposed spherical EMC algorithm and the adaptive spherical EMC algorithm given in pseudocode in Algorithms 1 and 2, respectively. We discuss thereafter the critical points to control in order to ensure proper behavior of the reconstruction algorithms.

1. Initialization

In general, initialization of EM-like algorithms is critical. Indeed, convergence is often guaranteed only to a local maximum of the likelihood. As a consequence, a bad initialization of the algorithm may lead to convergence to an undesired nonglobal maximum. In order to minimize such an effect, an initialization procedure based on the band limit adaptive scheme proposed throughout this paper can be used. Namely, since the algorithm estimates band limited intensity functions with increasing degree, it seems reasonable to initialize the algorithm with the previous estimated model, at band limit \( L - 2 \). At the first stage of the algorithm, corresponding to band limit \( L = 3 \), this is done the following way. First the tomographic model is initialized with the mean photon count (MPC) of the diffraction patterns on this shell. Note that the MPC is defined here as an average over the detector pixels, not in terms of Shannon-Nyquist pixels. Then, to avoid the algorithm stopping prematurely, we slightly perturbate the MPC. The initialized tomographic model then reads

\[
W_{ij}^{(0)} = \text{MPC} + \varepsilon u_{ij},
\]

(31)

with

\[
\text{MPC} = \frac{1}{N_K} \sum_{i,k} y_{ik},
\]

(32)

and where the \( u_{ij} \)'s are drawn uniformly on the interval \([-\text{MPC}, \text{MPC}]\). Here \( \varepsilon \) is a randomization parameter, \( 0 < \varepsilon < 1 \). Finally, performing one compression step leads to the
estimate of initial spherical harmonic coefficients $\hat{c}^{(0)}$, which completes the initialization step.

2. Convergence assessment

Another important point is the convergence assessment of the algorithm. In the general EM algorithm setting, this is often done by monitoring the likelihood value, the increments of which are given by the so-called fundamental inequality of EM [24,27]:

$$\mathcal{L}(W') - \mathcal{L}(W) \geq \mathcal{Q}(W'|W) - \mathcal{Q}(W|W).$$  \hspace{1cm} (33)

The EM algorithm then stops when likelihood increments become smaller than a certain threshold value. In the EMC setting we added two extra steps—expansion and compression, respectively—to make the expectation-maximization procedure tractable. This yields to a loss of important properties of the EM algorithm, namely, the likelihood $\mathcal{L}(W)$ is not forced to increment at each iteration, and the likelihood value is not directly available.

To obtain the likelihood value, it is thus necessary to proceed as follows. First recall that we have only defined the intermediate quantity $\mathcal{Q}(W|W^{(n)})$ for now. The likelihood is obtained the following way:

$$\mathcal{L}(W) = \mathcal{Q}(W|W^{(n)}) + \mathcal{H}(W|W^{(n)}),$$  \hspace{1cm} (34)

where the entropy $\mathcal{H}(W|W^{(n)})$ is defined like in [27] by

$$\mathcal{H}(W|W^{(n)}) = -\frac{1}{K} \sum_{k=1}^{K} \int_{\text{SO}(3)} \log [p(R|y_k, W_k)] p(R|y_k, W_k) d\mu(R).$$  \hspace{1cm} (35)

We are interested in the likelihood of the model at iteration $n$, denoted by $\mathcal{L}(W^{(n)})$. Using the expression above, it reads $\mathcal{L}(W^{(n)}) = \mathcal{Q}(W^{(n)}|W^{(n)}) + \mathcal{H}(W^{(n)}|W^{(n)})$. Using similar arguments as above, the latter quantity can be rewritten the following way:

$$\mathcal{L}(W^{(n)}) = \frac{1}{K} \sum_{k=1}^{K} \sum_{j=1}^{M} \sum_{i=1}^{N} y_{ij} \log W_{ij} - W_{ij}$$

$$\mathcal{Q}(W^{(n)}|W^{(n)})$$

$$- \frac{1}{K} \sum_{k=1}^{K} \sum_{j=1}^{M} P_{jk} \log \left( \frac{P_{jk}}{W_{j}} \right)$$

$$\mathcal{H}(W^{(n)}|W^{(n)}).$$  \hspace{1cm} (36)

We note that the quantity $-\mathcal{H}(W^{(n)}|W^{(n)})$ is exactly the mutual information quantity proposed by Loh and Elser [see [7], Eq. (19)].

Our stopping criterion is now defined as follows. At each iteration, we compute the relative likelihood variation denoted $\nabla \mathcal{L}$. When the value of $\nabla \mathcal{L}$ goes below a given threshold $\eta$, then the algorithm either stops if the current value of $L$ is equal to $L_{\text{max}}$ or restarts with an increased band limit $L$ and uses the previously estimated spherical harmonic coefficients as a starting point.

3. Summary

The computational time of the algorithm presented in this paper scales directly with the desired resolution of the desired reconstruction. Actually, the most expensive part is the maximization step, which in our case scales as $O(L^2 N_r K)$. It thus shows a clear dependence on the band limit $L$, and on the number of pixels on the shell. This motivates the shell-by-shell approach: since inner and outer shells show two disjointed sets of parameters (small $L$ and $N_r$ for the inner shell, large $L$ and $N_r$ for the outer shells), it is interesting to gain computation time by parallelization of the shell reconstructions.

Also, the increasing band limit approach is expected to lead to a better conditioning of the algorithm, limiting the number of iterations and improving the quality of the reconstruction.

Finally, we shall point out that the sampling of the rotation group used here is based upon harmonic analysis results on the rotation group. Therefore, one may expect the harmonic approximation of integrals over SO(3) used in the EMC algorithm to perform equally well compared to the 600-cell sampling scheme previously introduced [7].

IV. SPHERICAL SHELL ALIGNMENT

Since we have considered each shell separately (which allows an efficient parallelization of the algorithm and distribution of the data), the shells have to be realigned in order to form a consistent 3D function. Namely, we want to minimize the quadratic error between two successive shells. The problem can be stated as follows: we need to find the rotation $R$ such that

$$\arg \min_{R \in \text{SO}(3)} \|W^{(s)}_G - \Lambda(R)W^{(s+1)}_G\|_2^2,$$  \hspace{1cm} (37)

where we have introduced the rotation operator $\Lambda(R)$ defined by $\Lambda(R)f(\Omega) = f(R^T \Omega)$. As the features shared by the shell intensities evolve with the radius $q_s$, the spherical shell alignment problem between two consecutive shells is only significant if these are sufficiently close to each other. The alignment procedure is indeed expected to fail in the case where $(q_{s+1} - q_s) > \Delta \log 2/2$, that is, the Shannon-Nyquist theorem is no longer verified in the radial direction.

In this paper the shell alignment problem is solved using correlations on the rotation group, using an approach proposed by Kostelec and Rockmore in [28]. For an overview concerning the so-called rotational matching problem, see, for instance, [29] and references therein, especially for the rotational matching problem in biophysics.

The problem as stated in Eq. (37) is equivalent to maximizing the correlation of the two shell intensities:

$$\text{Corr}_{s,s+1}(R) = \int_{\mathbb{S}^2} W^{(s)}_G(\Omega) \Lambda(R) \overline{W^{(s+1)}_G(\Omega)} d\Omega.$$  \hspace{1cm} (38)

We recall that the shell intensities are also represented by their estimated spherical harmonic coefficients (up to a given band limit $L$), respectively, $c_{i,s}(q_s)$ and $c_{i,s+1}(q_{s+1})$. Using Fourier analysis properties on the rotation group, the correlation above
can be expressed the following way:

$$\text{Corr}_{s,t+1}(R) = \sum_{s=0}^{L-1} \sum_{n=-\ell}^{\ell} \sum_{m=-\ell}^{\ell} c_{s}^{m}(q_{s}) \bar{c}_{s+1}^{m}(q_{s+1}) \times D_{\ell}^{m,m}(R),$$

(39)

where we have introduced the Wigner-$D$ functions, $D_{\ell}^{m,m}$, which are detailed in Appendix B. Now, using the same rotation group sampling as in the spherical EMC algorithm, one can evaluate the correlation (39) efficiently. The rotation $R$ that leads to the higher correlation value between the two shells is then used to align the $s+1$ shell with the previous one. Namely, the aligned spherical harmonic coefficients read

$$\tilde{c}_{s}^{m}(q_{s+1}) = \sum_{n=-\ell}^{\ell} D_{s}^{m,n}(R) \bar{c}_{s+1}^{n}(q_{s+1}).$$

(40)

By repetition of the same procedure for successive shells, we are able to reconstruct efficiently the full 3D intensity function. Finally, a few nontrivial points are to be mentioned. First, since the correlation is performed numerically on a discrete rotation sample, it is unlikely that the rotation leading to the higher correlation value will be the exact one. However, by increasing the size of the rotation sampling, the rotation estimation error will be further reduced. Moreover, from a practical perspective, a good approximation of the correlation can be obtained using only spherical harmonics of low degree, even in the case of large band limits $L$ [28]. This could be adequately exploited to obtain low-resolution reconstructions efficiently.

Finally one can note that a method has been proposed in [30] to improve the accuracy of the rotation matching, and could be efficiently applied in our context.

V. NUMERICAL VALIDATION

In this section, we illustrate our approach with simulations. First, we show how the adaptive spherical EMC can effectively reconstruct the shell intensity for different band limits. We illustrate the potential of the approach by considering three typical shells, an inner, middle, and outer shell, respectively. Theoretical intensity functions corresponding to these three shells are presented in Fig. 2.

The behavior of the algorithm with respect to the number of observations is also studied in detail.

Secondly, we investigate the full shell-by-shell reconstruction problem, and analyze the effect of the distance between shells on the resulting resolution. All simulations show results for the Satellite Tobacco Necrosis Virus (STNV, PDB entry: 2BUK), one of the smallest viruses known to date. The reciprocal space size of a pixel is $\Delta q = 0.011 \, \text{Å}^{-1}$. Simulations are performed for a fluence $I_0$ corresponding to $10^{15}$ photons per pulse and 0.1-$\mu$m$^2$ focal spot area. Corresponding mean photon counts per pixel are given in the top of Fig. 2.

We note that the number of photons per pulse used here ($10^{15}$) is about ten times higher than those currently available at XFEL facilities. This choice is motivated here by the fact that it allows us to use fewer diffraction patterns to reconstruct the biomolecule, as the aim of this work is to demonstrate the feasibility of the approach. As it will be seen in the next section, using state of the art photon counts will lead to a higher number of diffraction patterns required for a given band limit $L$.

A. Single shell reconstruction

1. Estimation of the required number of diffraction patterns

In this section we address the important issue of the required number of diffraction patterns in single particle experiments. Namely, we ask the following question: given a specified shell, how many diffraction patterns are needed to reconstruct the intensity defined on this shell, for a given band limit $L$?

This question is in general quite complex, and to allow further investigation we need to make some assumptions. The reconstruction of the intensity on the shell is conditional upon the ability to recover the orientation of each diffraction pattern, yet to have a sufficient signal-to-noise ratio. It is expected that the number of diffraction patterns required will be governed by a combination of these two factors.

When the signal-to-noise ratio is high, as expected for large biomolecules and viruses (and already shown by a recent experiment [18]), the orientation recovery problem drives predominantly the performances of reconstruction algorithms. In the sequel, we perform numerical analysis addressing the question stated previously, taking into account both orientation recovery issue and sufficient signal-to-noise ratio.

Let us consider a shell of radius $q_s$, and that detector coordinates on this shell are given by Eq. (4). We note that the number of pixels on the detector $N_p$ scales with the radius $q_s$. 

FIG. 2. (Top) Mean photon count (MPC) per pixel as a function of the reciprocal space magnitude $q$. The signal becomes weaker as $q$ increases. Vertical lines show three examples of shell intensity distributions, referred to as inner, middle, and outer shell therein. (Bottom) Relative energy per shell distribution in the spherical harmonic domain, for increasing values of $q$. The distribution spreads with increasing values of $q$. 

053302-8
FIG. 3. (Top) Required number of patterns with the band limit \(L\) and MPC, where each plot corresponds to different shell radii. In each plot the following MPC values were used: \(\text{MPC} = 10, 1, 0, 0.1, 0.01\). (Bottom) Corresponding required patterns in the case of the STNV virus, using MPC values shown in Fig. 2.

given the pixel spacing \(\Delta q\). Moreover, we fix a mean photon count per pixel on this shell. Now, we define an equal-area grid on the sphere with HEALPix, with a resolution parameter \(n_{\text{side}}\). We draw a uniform random rotation from \(\text{SO}(3)\), and apply it to the detector coordinates. Then, we draw \(N_s\) samples from the Poisson distribution of parameter MPC. Finally, we analyze which HEALPix pixels are visited by the nonzero pixels of this rotated detector and keep track of this visit. The number of runs of this simulation until each pixel has been visited at least once gives us an estimate of the required number of diffraction patterns on this shell, for a given resolution \(n_{\text{side}}\). Repeating this algorithm several times gives a fair estimate of this number.

Here, we only analyzed the number of required patterns in terms of the resolution parameter \(n_{\text{side}}\). Recalling that \(n_{\text{side}}\) is linked to the band limit \(L\) by the approximate sampling theorem (A7), we link each \(n_{\text{side}}\) value with the maximum band limit \(L\) available for this resolution parameter. Extrapolation between missing \(L\) values gives an estimate of the required diffraction patterns for any \(L\) value. Note that with this choice of detector geometry and for a given shell, the maximum value of \(L\) available is given by \(L_{\text{max}} = 1 + [\pi q_s/\Delta q] \). This simply means that the maximum band limit available increases with the shell radius.

The results of the simulation are presented in Fig. 3. In the first row, we analyzed the number of diffraction patterns needed in the case of three typical shells, inner, middle, and outer, and for different MPC on these shells. The MPCs used here are \(10, 1, 0.1, 0.01\). First, one sees that for a given shell and band limit the number of required patterns increases as the MPC decreases, as expected. Similarly the patterns required increase with the band limit \(L\) given a certain shell and MPC.

For a given band limit \(L\) and a fixed MPC, the required number of patterns reduces as the shell radius increases. This is the expected behavior, since at \(L\) fixed the area of each pixel on the sphere increases. Moreover, we note that for each shell the required number of patterns behaves as a power law of \(L\).

At the bottom of Fig. 3, the number of patterns required in the case of the STNV virus is presented. Note that the number of required patterns increases with the shell radii, due to a lower MPC in the outer shells. Black squares show the theoretical band limit of the shell intensity, calculated from Fig. 2 by thresholding the relative energy per degree below some predefined threshold, typically \(10^{-5}\) in our experiments.

2. Multiresolution shell reconstruction

In the sequel, we investigate the behavior of the algorithm in three typical cases of shell intensity reconstructions: inner,
middle, and outer shell cases. The simulations presented here have been performed using 500 diffraction patterns for the inner shell case and 1000 diffraction patterns for the middle and outer shell cases. We make use of the adaptive shell EMC algorithm presented earlier. This algorithm provides an adaptive reconstruction method based on the spherical harmonic decomposition of the shell intensity up to a maximum band limit $L_{\text{max}}$.

The inner shell case is represented in Fig. 4. From Fig. 2, the expected theoretical band limit is $L^{\text{lim}} = 7$. One can notice that the likelihood improvements become rapidly very small, and that the number of iterations for each band limit $L$ is given by the minimum number of iterations imposed (in our case four). This behavior is easily seen from the low band limit distribution of the inner shell. As seen in the reconstruction provided, the main features of the shell intensity are already reconstructed at $L = 3$, as expected.

The middle shell case is represented in Fig. 5. Again, using Fig. 2, we see that the expected theoretical band limit is now $L^{\text{lim}} = 13$. The likelihood is improving over the different band limits, tending to smaller increments as the band limit increases. The reconstruction provided shows how the accumulation of spherical harmonic coefficients of higher degree improves the accuracy of the reconstruction. One can note that low band limit reconstructions exhibit negative values (up to $L = 7$, included); therefore the negative values are thresholded to some small constant (here $10^{-4}$). This stems from the fact that truncation of a spherical harmonic expansion

FIG. 5. Adaptive intensity reconstruction of the middle shell case ($q = 0.21 \text{ Å}^{-1}$). Top: Evolution of the likelihood with the iteration number and corresponding relative likelihood improvements. Vertical lines show changes in the band limit, respectively, $L = 3, 5, 7, 9, 11, 13$. Bottom: Corresponding reconstructions at the end of each fixed band limit and comparison with the theoretical intensity. Color maps were adapted to enhance the contrast of the visualization.

FIG. 6. Adaptive intensity reconstruction of the outer shell case ($q = 0.42 \text{ Å}^{-1}$). Top: Evolution of the likelihood with the iteration number and corresponding relative likelihood improvements. Vertical lines show changes in the band limit, respectively, $L = 3, 5, 7, 9, 11, 13, 15$. Bottom: Corresponding reconstructions at the end of each fixed band limit and comparison with respective truncated theoretical intensities.
of a non-negative function does not preserve the non-negativity of the function.

Finally, the outer shell case is represented in Fig. 6. In this case, the theoretical band limit is $L^{\text{lim}} = 25$, as seen in Fig. 2. Here, the likelihood tends to improve constantly as the band limit increases before stabilizing from $L = 13$. Reconstructions provided show the frequencylike improvement as the band limit $L$ increases. Truncated theoretical intensities for the respective $L$ values are also provided. However, one can notice the difference between the theoretical and reconstructed intensities. The reasons are twofold: first, when estimating low-$L$ intensity functions, there is aliasing of power of higher degree coefficients ($\ell > L$), resulting in a higher overall energy in the reconstruction than in the truncated theoretical intensity. This is well shown by results in the cases $L = 3.5$ and $L = 7$. The second results from an insufficient number of diffraction patterns, since for $L > 7$ the reconstruction barely improves, whereas the truncated theoretical intensity continues to converge to the theoretical intensity.

B. Full shell-by-shell reconstruction

The shell-by-shell approach offers an efficient and faster method to distribute computation of the reconstructions. However, by considering the shells independently the radial information is lost. This radial information has to be retrieved during an alignment process, as described in Sec. IV.

Using 1000 diffraction patterns, we reconstructed each shell intensity independently up to band limit $L = 7$ and maximum frequency $q = 0.42 \text{ \AA}^{-1}$, for a reciprocal space spacing $\Delta q = 0.01 \text{ \AA}^{-1}$ between the shells. The shells were then realigned using the above-mentioned correlation method. Results are presented in Fig. 7.

To evaluate numerically the performance of the method, we introduce an error metric based on the well-know $R$ factor, which is defined such that

$$ R = \frac{\sum_p |I_r(q_p) - I_{\text{th}}(q_p)|}{\sum_p I_r(q_p)}, $$

(41)

where $I_r$ denotes the reconstructed intensity, and $I_{\text{th}}$ is the theoretical intensity. Here $q_p$ denotes the 3D-spherical grid points. To illustrate the reconstruction performances with the band limit $L$, we have computed the $R$ factor for $L = 3.5, 7$ reconstructions. We distinguished two cases, one where the theoretical intensity is truncated up to degree $L$, and the other where the full theoretical intensity is considered. As shown by the results in Table I, the $R$ factor decreases as $L$ increases, meaning that the reconstructed intensity function is improved over $L$. Also one notices that the $R$ factor for the truncated theoretical intensity is smaller than for the full theoretical intensity, which is consistent with the fact that the full intensity function is more complex than the truncated one.

Finally, one can notice that different sources of estimation error arise in the shell-by-shell approach. One is related to the shell intensity estimation problem, where the finite number of diffraction patterns limits the actual accuracy of the reconstruction. The second one is due to the radial decoupling between the shells, which leads to a realignment error. The latter can be minimized by either increasing the size of the grid on which the SO(3) correlation is performed, or using refinement methods such as in [30].

As a consequence, the algorithm presented here may not be as accurate as the original EMC algorithm in the case of low SNR datasets. It is actually expected that the proposed approach is a compromise improved computational performance of a shell-by-shell approach and noise tolerance.

VI. CONCLUSION

Here we have shown that a spherical harmonic basis is a convenient way to design an incremental, parallelizable approach to the reconstruction of 3D intensity functions from.

| $L$ | Truncated | Theoretical |
|-----|-----------|-------------|
| 3   | 0.152     | 0.171       |
| 5   | 0.143     | 0.148       |
| 7   | 0.0979    | 0.0998      |

TABLE I. $R$-factor values computed according to Eq. (41) with respect to the $L$-truncated theoretical intensity or the theoretical intensity, for different $L$ values.
two-dimensional XFEL diffraction patterns. The proposed shell-by-shell algorithm allows control over radial and angular resolution, which can be gradually increased during the reconstruction. As a shell-by-shell approach is intrinsically parallelizable, it is potentially an efficient way to analyze the very large datasets expected in experiment. Using numerical simulations, we have studied how the number of patterns required for convergence depends on the signal and resolution.

ACKNOWLEDGMENTS

N.L.B.’s research was supported by the Research Executive Agency, European Commission, through the International Outgoing Fellowship (Grant No. GeoSToSip 326176) program of the 7th PCRD. A.V.M. was supported by the Australian Research Council’s Discovery Early Career Researcher Award funding scheme (Grant No. DE140100624) and Centre of Excellence programs.

APPENDIX A: FOURIER ANALYSIS ON THE SPHERE

1. Spherical harmonics

The spherical harmonic functions, or in short spherical harmonics, extend Fourier analysis to the two-sphere \( S^2 \). Formally, the spherical harmonic of degree \( \ell \) and order \( m \) is given by

\[
Y_\ell^m(\theta, \phi) = \frac{(2\ell + 1) (\ell - m)!}{4\pi (\ell + m)!} P_\ell^m(\cos \theta) \exp(i m \phi), \tag{A1}
\]

where \( P_\ell^m \) is the associated Legendre polynomial (or generalized Jacobi polynomial). For each integer degree \( \ell \geq 0 \) corresponds \( 2\ell + 1 \) orders \( m \), such that \( |m| \leq \ell \). The spherical harmonic functions form a complete orthonormal basis on the two-sphere \( S^2 \) for square integrable functions. As a consequence any square-integrable complex-valued function \( f : S^2 \to \mathbb{C} \), \( f \in L^2(S^2) \) can be decomposed onto spherical harmonics:

\[
f(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_\ell^m Y_\ell^m(\theta, \phi), \tag{A2}
\]

where we have introduced the spherical harmonic coefficients \( f_\ell^m \), which are given by projection onto the respective spherical harmonic \( Y_\ell^m \) via the canonical inner product on \( L^2(S^2) \):

\[
f_\ell^m \triangleq \left( f, Y_\ell^m \right)_{S^2} = \int_{S^2} f(\theta, \phi) Y_\ell^m(\theta, \phi) \sin \theta \, d\theta d\phi. \tag{A3}
\]

These coefficients are complex-valued, and the properties of the function \( f \) (real-valued symmetries) are encoded in the coefficients. As in classical Fourier analysis, spherical harmonics exhibit the Parseval relation, that is, the energy of \( f \) is preserved in the spherical harmonic domain:

\[
\|f\|_2^2 \triangleq \int_{S^2} |f(\theta, \phi)|^2 \sin \theta \, d\theta d\phi = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \|f_\ell^m\|^2. \tag{A4}
\]

For numerical purposes the infinite series expansion in the spherical harmonic decomposition in Eq. (A2) is not desirable. Rather, we consider band limited functions on the sphere. We say that a function \( f \) is band limited at \( L \) (or equivalently \( L \) band limited) if for all \( \ell \geq L \) the spherical harmonic coefficients are identically zero \( f_\ell^m = 0 \). In the following, we consider \( L \) band limited functions only.

2. Spherical harmonic transforms with HEALPix

The forward and inverse spherical harmonic transforms are given by

\[
f_\ell^m = \int_{S^2} f(\theta, \phi) Y_\ell^m(\theta, \phi) \sin \theta d\theta d\phi \quad \text{(SHT)}, \tag{A5}
\]

\[
f(\theta, \phi) = \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} f_\ell^m Y_\ell^m(\theta, \phi) \quad \text{(inv.SHT)}. \tag{A6}
\]

The computation of the forward SHT requires the evaluation of an integral over the two-sphere \( S^2 \) for each coefficient \( f_\ell^m \). The evaluation of such integrals can be done by conveniently sampling the two-sphere, i.e., distributing nodes on the surface on the sphere in order to obtain a quadrature formula. To this aim, a certain number of sampling schemes have been already proposed [26,31–34], often motivated by the analysis of the cosmic microwave background. In this work, we use the HEALPix sampling scheme [26], which has been designed for high performance, fast, and accurate computation of spherical harmonics on the sphere. The sphere is tessellated into curvilinear equal-area pixels, where the pixel centers are distributed on lines of constant latitude allowing faster computation of spherical harmonic functions due to the separation of angular variables in the spherical harmonics. The HEALPix sampling scheme is hierarchical and provides different levels of resolution through a parameter \( n_{\text{side}} \). The number of pixels \( n_{\text{pix}} \) at resolution \( n_{\text{side}} \) is given by

\[
n_{\text{pix}} = 12 \times n_{\text{side}}^2, \tag{A7}
\]

where the parameter \( n_{\text{side}} \) is a power of 2, due to the construction of the grid [35].

Although the HEALPix grid does not exhibit an exact sampling scheme, a very accurate estimate of the spherical harmonics coefficients in Eq. (A5) is obtained if the following condition is fulfilled:

\[
L \leq 2n_{\text{side}} + 1. \tag{A8}
\]

This last formula stands for an approximate sampling scheme on the grid.

APPENDIX B: ROTATION GROUP SAMPLING

1. Definition

The special orthogonal group in three dimensions \( \text{SO}(3) \) denotes the set of real matrices \( R \in \mathbb{R}^{3 \times 3} \) which satisfy

\[
RR^T = I_3 \quad \text{and} \quad \det R = 1, \tag{B1}
\]

where \( R^T \) is the transpose of the matrix \( R \), \( \det R \) is its determinant, and \( I_3 \) is the \( 3 \times 3 \) identity matrix [36]. It is commonly parametrized by Euler angles \((\alpha, \beta, \gamma)\) where
\( \alpha, \gamma \in [0, 2\pi), \) and \( \beta \in [0, \pi) \). In the zyz convention, the matrix \( R(\alpha, \beta, \gamma) \) reads

\[
R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma),
\]

where \( R_z, R_y \) denote rotations around canonical axes \( z \) and \( y \), respectively. These rotations read explicitly

\[
R_z(\alpha) = \begin{pmatrix}
\cos \alpha & -\sin \alpha & 0 \\
\sin \alpha & \cos \alpha & 0 \\
0 & 0 & 1
\end{pmatrix},
\]

(B3)

\[
R_y(\beta) = \begin{pmatrix}
\cos \beta & 0 & \sin \beta \\
0 & 1 & 0 \\
-\sin \beta & 0 & \cos \beta
\end{pmatrix}.
\]

(B4)

2. Harmonic analysis on the rotation group

Harmonic analysis on the rotation group SO(3) is conveyed by the irreducible representations of SO(3). This is a direct consequence of the Peter-Weyl theorem in the SO(3) case [37]. Namely, any square-integrable function \( f : \text{SO}(3) \to \mathbb{C} \) can be decomposed as

\[
f(R) = \sum_{\ell=0}^{\infty} \sum_{m,n=-\ell}^{\ell} f_{\ell,m,n} D_{\ell}^{m,n}(R),
\]

where the \( D_{\ell}^{m,n} \) functions are the Wigner-D functions and \( f_{\ell,m,n} \) are the Fourier coefficients obtained following

\[
f_{\ell,m,n} = \int_{\text{SO}(3)} f(R) \overline{D_{\ell}^{m,n}(R)} \, d\mu(R).
\]

(B6)

The Wigner-D functions are conveniently expressed using the zyz-Euler angle parametrization, that is,

\[
D_{\ell}^{m,n}(\alpha, \beta, \gamma) = e^{-im\alpha} d_{\ell}^{m,n}(\beta) e^{-iny}.
\]

(B7)

where \( d_{\ell}^{m,n}(\beta) \) is a polynomial in \( \cos(\beta/2) \) and \( \sin(\beta/2) \).

The evaluation of the Fourier transform (B6) can be done using fast Fourier transforms on the rotation group, as first developed in [28].

3. Sampling theorem on the rotation group

Let us consider a \( L \) band limited function \( f \) on the rotation group, that is, for all \( \ell \geq L \), \( f_{\ell,m,n} = 0 \). It is possible to compute its Fourier transform exactly using the following equiangular sampling:

\[
\alpha_j = \frac{2\pi j_1}{2L}, \quad \beta_j = \frac{\pi}{4L}(2j_2 + 1), \quad \gamma_j = \frac{2\pi j_3}{2L},
\]

(B8)

where the indices \( j_1, j_2, j_3 \in \{0, 1, \ldots, 2L - 1\} \). Equation (B6) then reads

\[
f_{\ell,m,n} = \sum_j w_j f(R_j) D_{\ell}^{m,n}(R_j),
\]

(B9)

where we have introduced a single index \( j = (j_1, j_2, j_3) \). The weights \( w_j = w_{j_1,j_2,j_3} \) satisfy \( \sum_j w_j = 1 \) and read

\[
w_{j_1,j_2,j_3} = \frac{1}{4L^2} \left( \sum_{k=0}^{L-1} \frac{1}{2k+1} \sin \left[ \frac{\beta_j}{2}(2k+1) \right] \right) \sin \beta_j.
\]

(B10)

This sampling of the rotation group is used throughout this work, at the heart of the EMC algorithm and during the shell alignment process.

Note that in the EMC algorithm the above sampling set is used to approximate the value of the integral of some function. This operation corresponds actually to computing the first Fourier coefficient \( f_{0,0} \). It can be shown that using only half the samples in each angle in the sampling described above is sufficient (aliasing in the higher order coefficients is therefore tolerated). This remark allows us to obtain \( L^3 \) samples instead of \( 8L^3 \), however this was not implemented here since the oversampling of the rotation group allows a better approximation of the intermediate quantity \( Q(W|W^{(0)}) \).
[14] S. Kassemeyer, A. Jafarpour, L. Lomb, J. Steinbrener, A. V. Martin, and I. Schlichting, Phys. Rev. E 88, 042710 (2013).
[15] D. Saldin, V. Shneerson, R. Fung, and A. Ourmazd, J. Phys.: Condens. Matter 21, 134014 (2009).
[16] D. Starodub, A. Aquila, S. Bajt, M. Barthelmess, A. Barty, C. Bostedt, J. D. Bozek, N. Coppola, R. B. Doak, S. W. Epp et al., Nature Communications 3, 1276 (2012).
[17] J. J. Donatelli, P. H. Zwart, and J. A. Sethian, Proc. Natl. Acad. Sci. USA 112, 10286 (2015).
[18] T. Ekeberg, M. Svenda, C. Abergel, F. R. Maia, V. Seltzer, J.-M. Claverie, M. Hantke, O. Jönsson, C. Nettelblad, G. van der Schot et al., Phys. Rev. Lett. 114, 098102 (2015).
[19] M. Walczak and H. Grubmüller, Phys. Rev. E 90, 022714 (2014).
[20] M. Tegze and G. Bortel, J. Struct. Biol. 179, 41 (2012).
[21] H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C. Spence et al., JOSA A 23, 1179 (2006).
[22] V. Elser, IEEE Trans. Inf. Theory 55, 4715 (2009).
[23] A working implementation of the algorithms described in this paper is available at http://github.com/jflamant/sphericalEMC.
[24] A. P. Dempster, N. M. Laird, and D. B. Rubin, Journal of the Royal Statistical Society Series B (Methodological) 39, 1 (1977).
[25] We use the noncapital “e” letter for the term “expectation,” whereas the capital “E” is reserved for “Expansion.”
[26] K. M. Gorski, E. Hivon, A. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, Astrophys. J. 622, 759 (2005).
[27] O. Cappé, E. Moulines, and T. Rydén, Inference in Hidden Markov Models (Springer, New York, 2006).
[28] P. J. Kostelec and D. N. Rockmore, Journal of Fourier Analysis and Applications 14, 145 (2008).
[29] G. S. Chirikjian, P. T. Kim, J.-Y. Koo, and C. H. Lee, International Journal of Computational Intelligence and Applications 04, 401 (2004).
[30] A. Makadia and K. Daniilidis, IEEE Transactions on Pattern Analysis and Machine Intelligence 28, 1170 (2006).
[31] J. D. McEwen and Y. Wiaux, IEEE Transactions on Signal Processing 59, 5876 (2011).
[32] J. R. Driscoll and D. M. Healy, Advances in Applied Mathematics 15, 202 (1994).
[33] A. Doroshkevich, P. Naselsky, O. V. Verkhodanov, D. Novikov, V. Turchaninov, I. Novikov, P. Christensen, and L.-Y. Chiang, Int. J. Mod. Phys. D 14, 275 (2005).
[34] R. G. Crittenden and N. G. Turok, arXiv:astro-ph/9806374 (1998).
[35] This fact arises from the hierarchical decomposition of the grid, where the next resolution level is obtained by dividing each pixel into four equal area pixels.
[36] S. L. Altmann, Rotations, Quaternions, and Double Groups (Clarendon, Oxford, 2005).
[37] A. O. Barut and R. Raczka, Theory of Group Representations and Applications (World Scientific, Singapore, 1986), Vol. 2.