



**HAL**  
open science

# A novel fast Fourier transform accelerated off-grid exhaustive search method for cryo-electron microscopy fitting

Alexandre Hoffmann, Valérie Perrier, Sergei Grudin

► **To cite this version:**

Alexandre Hoffmann, Valérie Perrier, Sergei Grudin. A novel fast Fourier transform accelerated off-grid exhaustive search method for cryo-electron microscopy fitting. *Journal of Applied Crystallography*, 2017, 50 (4), pp.1036-1047. 10.1107/S1600576717008172 . hal-01553293

**HAL Id: hal-01553293**

**<https://inria.hal.science/hal-01553293>**

Submitted on 1 Aug 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright



# A novel fast Fourier transform accelerated off-grid exhaustive search method for cryo-electron microscopy fitting

**Alexandre Hoffmann, Valérie Perrier and Sergei Grudinin**

*J. Appl. Cryst.* (2017). **50**, 1036–1047



**IUCr Journals**  
CRYSTALLOGRAPHY JOURNALS ONLINE

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <http://journals.iucr.org/services/authorrights.html>



# A novel fast Fourier transform accelerated off-grid exhaustive search method for cryo-electron microscopy fitting

Alexandre Hoffmann,<sup>a,b,c</sup> Valérie Perrier<sup>a,b</sup> and Sergei Grudin<sup>a,b,c\*</sup>

<sup>a</sup>University of Grenoble Alpes, LJK, F-38000 Grenoble, France, <sup>b</sup>CNRS, LJK, F-38000 Grenoble, France, and <sup>c</sup>Inria, F-38000 Grenoble, France. \*Correspondence e-mail: sergei.grudin@inria.fr

Received 16 December 2016

Accepted 1 June 2017

Edited by Th. Proffen, Oak Ridge National Laboratory, USA

**Keywords:** cryo-EM fitting; fast Fourier transform; FFT; exhaustive search; off-grid search; trust-region problem.

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

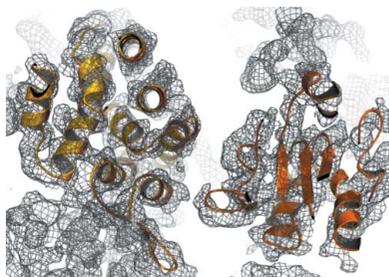
This paper presents a novel fast Fourier transform (FFT)-based exhaustive search method extended to off-grid translational and rotational degrees of freedom. The method combines the advantages of the FFT-based exhaustive search, which samples all the conformations of a system under study on a grid, with a local optimization technique that guarantees to find the nearest optimal off-grid conformation. The method is demonstrated on a fitting problem and can be readily applied to a docking problem. The algorithm first samples a scoring function on a six-dimensional grid of size  $N^6$  using the FFT. This operation has an asymptotic complexity of  $O(N^6 \log N)$ . Then, the method performs the off-grid search using a local quadratic approximation of the cost function and the trust-region optimization algorithm. The computation of the quadratic approximation is also accelerated by FFT at the same additional asymptotic cost of  $O(N^6 \log N)$ . The method is demonstrated by fitting atomic protein models into several simulated and experimental maps from cryo-electron microscopy. The method is available at <https://team.inria.fr/nano-d/software/offgridfit>.

## 1. Introduction

In this section we first briefly review the exhaustive search method. Then, we explain how the fast Fourier transform (FFT)-based algorithm can be used to accelerate sampling of a function written as a convolution or a cross-correlation. Finally, we describe the workflow of a standard FFT-based exhaustive search method.

### 1.1. Historical background

Sampling in low dimensions is a problem common to many application fields, such as image processing, crystallography and structural bioinformatics, for example. Sampling algorithms vary from Monte Carlo simulations to FFT-accelerated exhaustive search. Generally, their performance critically depends on the properties of the sampled function. For experimentally measured noisy data, which arise from scattering or cryo-electron microscopy (cryo-EM) experiments, one should ideally perform a full exhaustive search at least for six rigid-body degrees of freedom (Volkman & Hanein, 1999). Exhaustive search methods are thus widely used. If the optimized function can be decomposed into a correlation between two volumetric grids, then the exhaustive search methods can be accelerated in the Fourier domain using the correlation theorem. More precisely, this theorem states that the cross-correlation function in the direct space can be computed as an inverse Fourier transform of a product of two functions in the reciprocal space.



The first application of the FFT-accelerated search in biology was Crowther's fast rotation function, expressed as a correlation between two Patterson functions (Crowther, 1972). It computes a spherical harmonic expansion of the Patterson functions and uses an FFT in order to calculate the correlation in a very efficient way. This method was further improved by Navaza, as implemented in the *AMoRe* program (Navaza, 1987, 1994). The FFT-accelerated search on a three-dimensional grid was later rediscovered by Katchalski-Katzir and colleagues in application to protein–protein docking (Katchalski-Katzir *et al.*, 1992). Their method uses a three-dimensional Cartesian Fourier expansion to accelerate the search. Later on, this method was also applied to the cryo-EM fitting problem (Rossmann, 2000; Chacón & Wriggers, 2002) and later expressed in spherical coordinates using spherical harmonics for protein–protein docking (Ritchie & Kemp, 2000; Ritchie *et al.*, 2008) and cryo-EM fitting (Kovacs & Wriggers, 2002; Kovacs *et al.*, 2003; Garzon *et al.*, 2007). Recently, we have also combined the two coordinate systems and developed the *HermiteFit* method, which uses the orthogonal Hermite polynomials for the rotational search and the Cartesian Fourier expansion for the translational search (Derevyanko & Grudin, 2014).

### 1.2. FFT-accelerated sampling

As hinted above, the FFT-accelerated sampling is a fast sampling method and can be applied to any function that can be written as a correlation of two other functions. A brute-force sampling of such a cross-correlation function on a regular one-dimensional grid of  $N$  points would cost  $O(N^2)$  operations. However, using the Fourier correlation theorem, the same sampling can be achieved by computing two FFTs and one inverse FFT, which would cost  $O(N \log N)$  operations.

On a three-dimensional  $[N \times N \times N]$  Cartesian grid, the FFT sampling becomes even more efficient and allows one to sample the cross-correlation function in  $O(N^3 \log N)$  operations instead of  $O(N^6)$  operation of the brute-force search. On a spherical grid, the FFT can be used to sample the angular part of the function by taking advantage of the functional form of the spherical harmonics, their orthogonality and the structure of the Wigner D-matrix (Crowther, 1972; Navaza, 1987;

Ritchie & Kemp, 2000; Kovacs & Wriggers, 2002; Kovacs *et al.*, 2003; Ritchie *et al.*, 2008). Fig. 1 schematically represents correlations computed in the two coordinate systems.

We should note that both of these sampling approaches have certain computational issues. For example, in Cartesian coordinates handling the rotation is either computationally costly or inaccurate. On the other hand, in spherical coordinates, first, the sampling becomes coarser when far from the origin (Neveu *et al.*, 2016); second, translation operators are computationally expensive for high expansion orders if computed analytically (Ritchie, 2005).

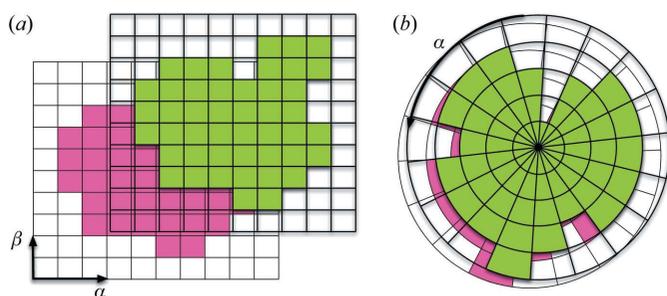
### 1.3. Summary of the standard FFT-based fitting algorithm

The standard FFT-based three-dimensional fitting algorithm operates according to the workflow shown in Fig. 2 (Katchalski-Katzir *et al.*, 1992; Gabb *et al.*, 1997; Chacón & Wriggers, 2002). The input of this algorithm is an atomic structure of a molecule, typically a protein,  $X = X_1, \dots, X_{N_{\text{atoms}}}$ , where  $(X_k \in \mathbb{R}^3)_{k=1 \dots N_{\text{atoms}}}$ , often called the model, determined experimentally by, for example, X-ray crystallography. Another input is an experimental electron-density map (EDM)  $d: \mathbb{R}^3 \mapsto \mathbb{R}$  determined by means of, for example, cryo-EM. Firstly, the algorithm represents the experimental EDM in the Fourier domain using the FFT method. It then rotates the model to a certain orientation  $R_\theta$  and computes the Fourier expansion of the electron density of the rotated model  $d^{\theta, m}: \mathbb{R}^3 \mapsto \mathbb{R}$ , where the superscript  $m$  stands for the model density. Alternatively, some recent approaches (Hoang *et al.*, 2013; van Zundert & Bonvin, 2015) directly rotate one of the two densities  $d$  or  $d^m$  instead of the atomic model, which avoids the computation of  $d^{\theta, m}$  for each orientation. In the present method, we choose to rotate the EDM map. The model electron density is typically computed as a sum of Gaussians centred on heavy (non-H) atoms of the model,

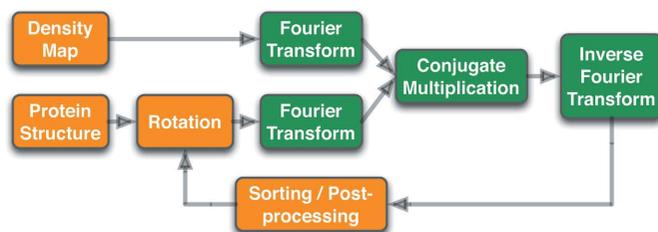
$$d^m(x, X) := (2\pi\sigma^2)^{-3/2} \sum_{k=1}^{N_{\text{atoms}}} \exp[y_k(x, X)], \quad (1)$$

$$y_k(x, X) := -(2\sigma^2)^{-1} \|x - X_k\|_2^2 \quad \forall k = 1 \dots N_{\text{atoms}},$$

where  $\sigma$  is the standard deviation (width) of the Gaussian distribution, which can be also seen as the model resolution. Afterwards, the algorithm exhaustively explores translational degrees of freedom of the rotated model with respect to the EDM. For every translation  $\tau$ , it determines the corresponding



**Figure 1** Comparison of a Cartesian CCF-based sampling with a spherical CCF-based sampling. (a) The green grid is shifted with respect to the pink grid along the  $x$  and  $y$  axes by  $\alpha$  and  $\beta$ , respectively. (b) The green grid is rotated with respect to the pink grid by an angle  $\alpha$ . Note that the spherical grid becomes coarser when far from the origin.



**Figure 2** Flowchart of an FFT-accelerated exhaustive search method.

score, which is usually given by the correlation between the two densities. This procedure is equivalent to computing the correlation of two functions,

$$\text{CCF}(\tau, \theta) = \int_{\mathbb{R}^3} d^\theta(x) d^m(x + \tau, X) dx = [d^\theta \star d^m(\cdot, X)](\tau), \quad (2)$$

where  $d^\theta(x) \equiv d(R_\theta x)$  is the electron density of the rotated EDM and  $d^m(\cdot, X)$  means that  $d^m$  is viewed as a function of the first variable only, with the second variable fixed at  $X$ . To speed up this step, the algorithm computes the values of the Fourier transform of the cross-correlation function (CCF) for all translational degrees of freedom at once using the correlation theorem,

$$\mathcal{F}\{\text{CCF}\}(\xi, \theta) = \mathcal{F}\{d^\theta\}^*(\xi) \mathcal{F}\{d^m(\cdot, X)\}(\xi), \quad (3)$$

where  $\mathcal{F}\{f\}$  denotes the Fourier transform of  $f$  and the star sign denotes the complex conjugation. Finally, the algorithm computes the inverse Fourier transform (IFT) of the correlation, generates a new rotation of the model structure and returns to the second step. The overall cost of this operation is  $O[N^3 \log(N)]$  operations on a  $[N \times N \times N]$  grid instead of  $O(N^6)$  operations of the brute-force computation. This procedure is repeated until all rotational degrees of freedom of the model with respect to the EDM have been explored (see Fig. 2). The solution of the fitting problem is then given by

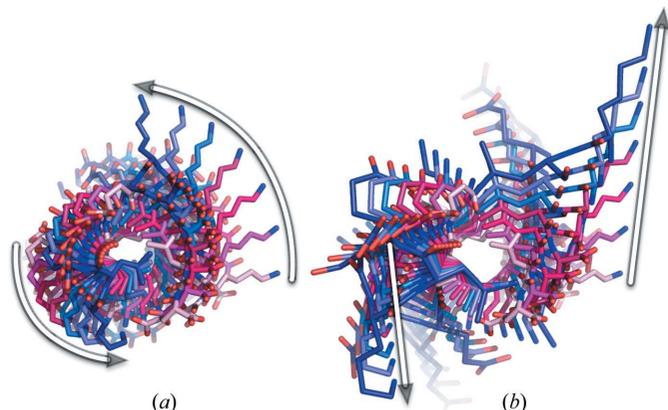
$$(\tau^*, \theta^*) = \text{argmax}_{\tau, \theta} [\text{CCF}(\tau, \theta)]. \quad (4)$$

While in our approach we stick to the cross-correlation scoring function, other functions that can be written as a cross-correlation also exist and may be more accurate. For example, the Laplace-filtered cross-correlation function is a CCF on which the Laplace filter was applied to both  $d$  and  $d^m$ . This scoring function is used by the majority of software packages that perform electron microscopy fitting (Chacón & Wriggers, 2002; Suhre *et al.*, 2006; Siebert & Navaza, 2009; Wriggers, 2010; Derevyanko & Grudin, 2014). The second option is the

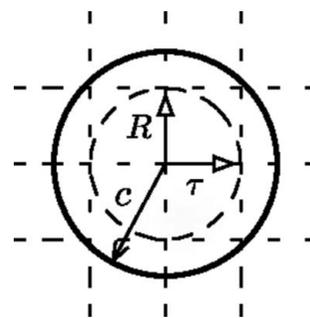
local CCF, which is normalized under the running footprint of a mask of the model (Roseman, 2003; Hoang *et al.*, 2013). The third option is the weighted CCF, which replaces the  $L^2$  scalar product with a weighted scalar product, where the weights describe the depth of a grid point located within the ‘core’ region of the density (Wu *et al.*, 2003). Finally, the recently introduced core-weighted local CCF (van Zundert & Bonvin, 2015) combines a local CCF and a weighted CCF. The local, the weighted and, consequently, the core-weighted local CCFs can all be considered as CCFs with a weighted  $L^2$  scalar product. Below we will describe the ways to adapt our method to these scoring functions.

The main advantage of the FFT-based algorithms is that they allow one to find, on a grid, the global maximum of a CCF with irregular shape. By fitting multiple models into a density map, and by clustering them, or by using a discrete linear programming approach (Amir *et al.*, 2015), it is possible to reconstruct large and complex assemblies. However, as we have mentioned above, the solution is found on a regular grid and a better solution could be found when we make the grid finer. This problem is even more important when dealing with spherical grids, because in this case the distance between the grid points increases with the distance to the grid centre.

To overcome this problem, we propose introducing a set of six additional *off-grid degrees of freedom* (DOF), three for an off-grid translation and three for an infinitesimal off-grid rotation (see Fig. 3). Assuming that the off-grid displacements are small enough, we use a second-order Taylor expansion of our CCF in the neighbourhood of a grid node with respect to these six off-grid DOF. Then, we use an optimization method to find the best off-grid displacement in a certain six-dimensional ball (see Fig. 4). We should mention that local minimization techniques that refine some top predicted model coordinates (which will be referred to as ‘poses’ below) with respect to the EDM have already been widely used (Wriggers & Chacón, 2001; Chacón & Wriggers, 2002; Topf *et al.*, 2008; Siebert & Navaza, 2009). However, these kinds of methods can only be used for a small set of top-ranked solutions, while our technique successfully samples both the gradient and the Hessian over the whole sampling grid, which allows us to refine all poses with a moderate extra cost, which is about 19 times the cost of the in-grid FFT-based fitting method.



**Figure 3**  
Comparison of a real rotation with respect to a linearized twist motion of a coiled coil protein (PDB code 2ch7, Park *et al.*, 2006). Seven snapshots taken at equal intervals are superposed on each other. (a) A real rotation is shown. (b) A linearized rotation is shown. Note the distortion caused by the linearized rotation at large amplitudes.



**Figure 4**  
Schematic representation of the off-grid search domain. Here,  $c$  refers to the RMSD computed for three rotational DOF  $R$  and three translational DOF  $\tau$ . Note that the resultant RMSD is larger than both  $R$  and  $\tau$ .

## 2. The method

In this section, we present our method. This section is organized as follows. Firstly, we introduce the new degrees of freedom responsible for the off-grid displacement. Secondly, we describe the quadratic approximation to the CCF and the way to speed up its computation. Finally, we explain how to formulate and solve the optimization problem that maximizes the CCF.

### 2.1. The six off-grid degrees of freedom

As mentioned above, we introduce six additional rigid degrees of freedom  $(\delta\tau, \delta R)$ , where  $\delta\tau \in \mathbb{R}^3$  represents an off-grid translation and  $\delta R \in \mathbb{R}^3$  represents an infinitesimal off-grid rotation. While the effect of  $\delta\tau$  on  $X_k := (x_k, y_k, z_k)$ , the position of the  $k$ th atom, is straightforward, we will describe in more detail the effect of  $\delta R$  on  $X_k$ . More precisely, an infinitesimal rotation is a linearization of a rotation produced by a unit axis  $u \in \mathbb{R}^3$  and an angle  $\theta \in \mathbb{R}$ ,

$$X'_k = RX_k = [\cos\theta I + \sin\theta[u]_{\times} + (1 - \cos\theta)(u \otimes u)]X_k, \quad (5)$$

where  $R$  is an orthogonal rotation matrix,  $I$  is the identity matrix,  $\otimes$  denotes the outer product and  $[u]_{\times}$  is the cross-product matrix,

$$[u]_{\times} = \begin{pmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{pmatrix}. \quad (6)$$

When the angle  $\theta$  is small enough, we can make the following first-order Taylor approximations:

$$\cos\theta \simeq 1, \quad \sin\theta \simeq \theta. \quad (7)$$

Our infinitesimal rotation can thus be approximated as follows:

$$X'_k \simeq (I + \theta[u]_{\times})X_k = X_k + [X_k]_{\times}^T \theta u. \quad (8)$$

Using the notation  $\theta u = \delta R$ , we have defined our linear transformation as a matrix–vector product. Now we define the  $3 \times 6$  off-grid displacement matrix as follows:

$$K_k = (I, [X_k]_{\times}^T), \quad (9)$$

and the new position of  $X_k$  after an off-grid displacement  $(\delta\tau, \delta R)$  is given by

$$X'_k = X_k + K_k \delta = X_k + (I, [X_k]_{\times}^T) \begin{pmatrix} \delta\tau \\ \delta R \end{pmatrix}. \quad (10)$$

Our scoring function can then be rewritten as follows, now also dependent on  $\delta$ :

$$\text{CCF}(\tau, \theta, \delta) = \int_{\mathbb{R}^3} d^\theta(x) d^m(x + \tau, X + K\delta) dx, \quad (11)$$

where  $K$  is a diagonal matrix of size  $3N_{\text{atoms}}$  defined as  $K = \text{diag}(K_1, \dots, K_{N_{\text{atoms}}})$ . Since this transformation is obtained by a linearization of a rotation, we will consider (and impose) it as a small deformation to avoid the distortion effect. Fig. 3 shows a comparison of a linearized twist motion with a real

rotation. We can clearly see a distortion caused by the linearization at large amplitudes.

### 2.2. The quadratic approximation of the CCF

As mentioned above, we consider the off-grid deformation to be small and we also make the assumption that the CCF does not oscillate between two adjacent grid nodes. We then use a second-order Taylor approximation of the CCF with respect to  $\delta \equiv (\delta\tau, \delta R)$ , which reads as

$$\text{CCF}(\tau, \theta, \delta) \simeq \text{CCF}(\tau, \theta, \mathbf{0}) + \langle \nabla_{\delta} \text{CCF}(\tau, \theta, \mathbf{0}), \delta \rangle + \frac{1}{2} \langle \delta, \nabla_{\delta\delta}^2 \text{CCF}(\tau, \theta, \mathbf{0}) \delta \rangle \quad (12)$$

where  $\langle \cdot, \cdot \rangle$  denotes the  $l^2(\mathbb{R}^6)$  dot product, and  $\nabla_{\delta} \text{CCF}(\tau, \theta, \mathbf{0})$  and  $\nabla_{\delta\delta}^2 \text{CCF}(\tau, \theta, \mathbf{0})$  denote, respectively, the gradient and the Hessian of the CCF with respect to  $\delta$  evaluated at  $(\tau, \theta, \mathbf{0})$ . We will now describe both the gradient and the Hessian of the CCF, and explain how to compute these. In this section, we introduce the following abuse of notation:

$$\int_{\mathbb{R}^3} \nabla_y g(x, y) dx := \left[ \int_{\mathbb{R}^3} \frac{\partial g}{\partial y_i}(x, y) dx \right]_{i=1\dots 6}, \quad (13)$$

$$\int_{\mathbb{R}^3} \nabla_{yy}^2 g(x, y) dx := \left[ \int_{\mathbb{R}^3} \frac{\partial^2 g}{\partial y_i \partial y_j}(x, y) dx \right]_{i,j=1\dots 6}.$$

Now, using equation (11) we can write the gradient and the Hessian as follows:

$$\nabla_{\delta} \text{CCF}(\tau, \theta, \mathbf{0}) = \int_{\mathbb{R}^3} d^\theta(x) \nabla_{\delta} d^m(x, X + K\delta)_{\delta=0} dx, \quad (14)$$

$$\nabla_{\delta\delta}^2 \text{CCF}(\tau, \theta, \mathbf{0}) = \int_{\mathbb{R}^3} d^\theta(x) \nabla_{\delta\delta}^2 d^m(x, X + K\delta)_{\delta=0} dx.$$

Then, using equation (1), we get the following identities:

$$\nabla_{\delta} d^m(x, X + K\delta)_{\delta=0} = c \sum_{k=1}^{N_{\text{atoms}}} \exp[y_k(x, X)] \times \nabla_{\delta} y_k(x, X + K\delta)_{\delta=0},$$

$$\nabla_{\delta\delta}^2 d^m(x, X + K\delta)_{\delta=0} = c \sum_{k=1}^{N_{\text{atoms}}} \exp[y_k(x, X)] \times [\nabla_{\delta} y_k(x, X + K\delta)_{\delta=0} \otimes \nabla_{\delta} y_k(x, X + K\delta)_{\delta=0} + \nabla_{\delta\delta}^2 y_k(x, X + K\delta)_{\delta=0}], \quad (15)$$

where  $c = (2\pi\sigma^2)^{-3/2}$ , and the gradient and the Hessian with respect to  $\delta$  of  $y_k$  can be obtained from equation (1). Finally, one finds the following expressions:

$$\nabla_{\delta} y_k(x, X + K\delta)_{\delta=0} := (\sigma^2)^{-1} K_k^T (x - X_k), \quad (16)$$

$$\nabla_{\delta\delta}^2 y_k(x, X + K\delta)_{\delta=0} := -(\sigma^2)^{-1} K_k^T K_k.$$

We note that both the gradient and the Hessian must be computed for each position on the grid. This would asymptotically cost  $O(N^9)$  operations. However, as we can see from equation (14), the gradient and Hessian of the CCF are, respectively, a vector and a matrix of cross-correlations;

therefore, it is possible to compute them in  $O(N^6 \log N)$  operations on a  $[N \times N \times N]$  grid using the Fourier correlation theorem. Therefore, we can pre-compute the gradients and the Hessians on the grid at the same asymptotic cost as that required for the standard FFT-based exhaustive search method. We should mention that overall our method requires about 18 times more computations compared with a standard rigid-body fitting technique. More precisely, it computes six additional cross-correlations per sampled rotation for the gradient term in equation (16) and 12 additional cross-correlations per sampled rotation for the Hessian term in equation (16).

### 2.3. The optimization problem

In this section we will explain how to find the optimal value of  $\delta$ . If we do not provide any constraints on  $\delta$ , our optimization problem becomes

$$\begin{aligned} \max_{\delta} : & \text{CCF}(\tau, \theta, \mathbf{0}) + \langle \nabla_{\delta} \text{CCF}(\tau, \theta, \mathbf{0}) \delta \rangle \\ & + \frac{1}{2} \langle \delta, \nabla_{\delta}^2 \text{CCF}(\tau, \theta, \mathbf{0}) \delta \rangle \end{aligned} \quad (17)$$

for all sampled values of  $\tau$  and  $\theta$ . In order to rewrite it in a standard quadratic optimization fashion, we will introduce the following notations:

$$\begin{aligned} A &= -\frac{1}{2} \nabla_{\delta}^2 \text{CCF}(\tau, \theta, \mathbf{0}), \\ b &= \frac{1}{2} \nabla_{\delta} \text{CCF}(\tau, \theta, \mathbf{0}), \\ x &= \delta. \end{aligned} \quad (18)$$

Note that  $A$  is a symmetric matrix and  $x \in \mathbb{R}^6$  is the optimization variable. Our optimization problem can now be rewritten as follows:

$$\min_x : x^T A x - 2b^T x. \quad (19)$$

This problem has been widely studied. One can distinguish three cases:

- (i) If  $A > 0$  (positive definite), the problem is convex and has a single solution, which is given by  $x = A^{-1}b$ .
- (ii) If  $A \geq 0$  (semi-positive definite) and  $b \in \text{im}(A)$ , the problem has an infinite number of solutions, which are  $x = A^{\dagger}b + q \quad \forall q \in \ker(A)$ , where  $A^{\dagger}$  denotes the pseudo-inverse of the matrix  $A$ .
- (iii) If  $A \geq 0$  and  $b \notin \text{im}(A)$  or  $A \not\geq 0$ , the problem has no exact solutions.

While the quadratic approximation of the CCF gives us a well known framework for our optimization problem, it is still an approximation. From Taylor's theorem we know that the approximation error grows as  $O(\|\delta\|_2^3)$  with respect to  $\delta$ ; it is therefore important to restrict the norm of our off-grid DOF such that our quadratic approximation does not degenerate.

Since  $\delta$  corresponds to an off-grid atomic displacement, it makes sense to bound its norm by the root-mean-square deviation of atomic positions (RMSD) generated by an in-grid transformation between the adjacent grid points. Given the spacing between two adjacent grid points  $h$ , the total mass of the molecule  $M$ , the molecule's inertia tensor  $I_T$  with the largest eigenvalue  $\lambda_{\max}$  and the angular search step  $\alpha$ , the total

RMSD of an in-grid transformation can be bound by (Popov & Grudinin, 2014)

$$\text{RMSD}^2(h, \alpha) \leq h^2 + \frac{4}{M} \sin^2(\alpha) \lambda_{\max}(I_T). \quad (20)$$

Now, once we know how to properly bound our off-grid deformations, we can re-formulate equation (17) as follows:

$$\begin{aligned} \min_x : & x^T A x - 2b^T x, \\ \text{s.t.} : & \|x\|_2^2 \leq \text{RMSD}^2(h, \alpha). \end{aligned} \quad (21)$$

The constraint in the aforementioned optimization problem can be regarded as a six-dimensional ball constraint. More precisely, three degrees of freedom of the six-dimensional ball correspond to translations and another three degrees of freedom correspond to rotations. Fig. 4 schematically shows the RMSD as a function of these six DOF.

This optimization problem is known as the trust-region sub-problem (TRS) and plays a particularly important role in the trust-region method (Sorensen, 1982; Fortin, 2000; Conn *et al.*, 2000). The TRS has been widely studied and several algorithms have been developed to solve it (Moré & Sorensen, 1983; Golub & Matt, 1991; Gould *et al.*, 1999; Toint *et al.*, 2009). We chose to use the algorithm developed by Moré & Sorensen (1983), because it is used in other optimization libraries such as *Dlib* (King, 2009) and because all the newer algorithms report an advantage only when dealing with the large-scale problems, which is not our case.

The algorithm developed by Moré & Sorensen (1983) first tries to solve the unconstrained problem (17) using a Cholesky factorization. If the Cholesky factorization succeeds and if  $\|A^{-1}b\|_2^2 \leq \text{RMSD}^2(h, \alpha)$ , then the TRS defined in equation (21) is solved; else the algorithm falls back onto a dual algorithm and finds the Lagrange multiplier with Newton's method (Fortin & Wolkowicz, 2004). The dual problem is solved with a modification of Newton's method, which computes the ratio between the function and its derivative with a series of Cholesky factorizations. This costs  $O(P^3)$  operations, with  $P$  being the dimension of the problem. It also allows us to determine if our deflated matrix is positive definite and thus satisfies the Karush–Kuhn–Tucker conditions.

### 2.4. Choosing the parameters

There are two parameters in our method that need to be chosen in advance. The first parameter is the number of sampled rotations or, equivalently, the angular sampling distance  $d\alpha$ . The second parameter is the model resolution, which is defined by the Gaussian  $\sigma$  as can be seen in equation (1). Let us first discuss the rotational sampling. As shown in the supporting information, if we aim to find the same minimum starting from two different points, these must be at a distance linearly proportional to the frequency of the map  $\omega$  and quadratically proportional to the grid spacing  $h$ . Now, assuming that the linear size of the fitted molecule is proportional to the length of the grid with  $N$  points in one direction, we get the following inequality for our angular spacing distance  $d\alpha$ :

$$\tan d\alpha < O\left(\frac{\omega h^2}{Nh}\right) \equiv O\left(\frac{nh}{N^2}\right), \quad (22)$$

where  $n$  is the frequency index. Following this inequality, one should sample  $O(N^2)$  rotations to ensure the convergence of our method for high frequencies and  $O(N)$  rotations for low frequencies. In practice, we use the coarser number of  $O(N)$

sampled rotations, which efficiently reduces the RMSD with our method, as we show below.

Let us now discuss the choice of the model resolution parameter. On the one hand, we would like to have a low-resolution model so that the quadratic approximation remains accurate even for large amplitudes of motion. On the other hand, if the resolution of our model is too low, the CCF scoring function will lose its accuracy (Wriggers & Chacón, 2001; Chacón & Wriggers, 2002; Garzon *et al.*, 2007; Hoang *et al.*, 2013; van Zundert & Bonvin, 2016), which introduces fitting errors (see Figs. 7 and 9 below). Our computational experiments demonstrate that taking the Gaussian  $\sigma$ , which defines the resolution of the model, equal to the spacing of the grid  $h$ ,  $\sigma \simeq h$ , generally ensures good fitting results.

### 3. Numerical results

In this section, we demonstrate the benefit of our method on simulated maps and on an experimental map. Then, we measure the computational speed of our method. We perform all the fitting tests using the GroEL complex. We have chosen this system because it is composed of 14 identical subunits such that in order to reconstruct the whole complex it is critical to sample all the angular intervals. Fig. 5 shows the organization of the GroEL complex.

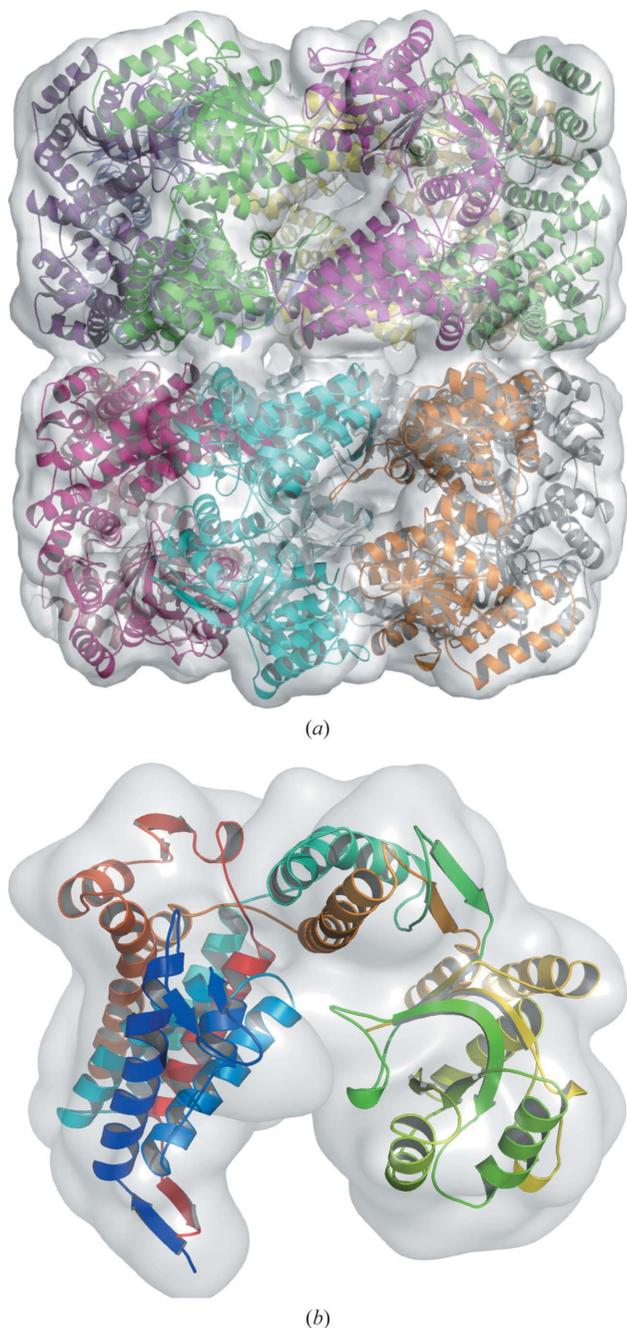
#### 3.1. Fitting into a simulated GroEL complex map

Here we demonstrate the performance of the off-grid exhaustive search on simulated GroEL complex maps of different resolutions. Firstly, we created a symmetric model of the GroEL complex starting with a single chain extracted from the lower ring of the GroEL–GroES crystallographic structure (PDB code 2c7c\_H, where \_H refers to the chain ID of the PDB structure; Ranson *et al.*, 2006). Then, we replicated this chain and rigidly fitted it into 14 individual locations in the GroEL cryo-EM map (code EMD-6422). We subsequently simulated electron-density maps of the obtained assembly model at multiple resolutions of 5, 11, 15 and 21 Å using the *eman2* package (Tang *et al.*, 2007). Finally, we fitted the initial model of a single chain into the simulated density maps. Fig. 5 shows the constructed model and the simulated 11 Å resolution map.

Since all the chains in the GroEL complex are identical with respect to the rotation and the translation operators, in this test we assess our method when reconstructing the whole complex starting from a single chain (PDB code 2c7c\_H) by varying the resolution of the GroEL simulated map, the width of the Gaussian for the model map and the angular sampling distance. We then re-score the obtained solutions using the following symmetry-adapted cost function:

$$J(\text{Solution}) = \min_{j=1\dots 14} \text{RMSD}^2(\text{Solution}, \text{Chain}_j), \quad (23)$$

which computes the RMSD between the solution and the nearest chain of the assembly model. The re-scored solutions are then clustered in order to eliminate redundancy. Finally,



**Figure 5**  
The GroEL complex composed of 14 identical subunits. (a) The symmetric model of the GroEL complex and its simulated electron density at a resolution of 11 Å are shown, where each chain has been coloured differently. (b) A single chain of the GroEL complex (PDB code 2c7c\_H) and its simulated electron density at a resolution of 11 Å are shown.

we compute the mean symmetry-adapted RMSD between the reassembled complex and the original one as follows:

$$\text{RMSD}^2 = \frac{1}{14} \sum_{i=1}^{14} J(\text{Solution}_i), \quad (24)$$

which is a standard quality assessment metric. Fig. 6 shows the RMSD computed with equation (24) as a function of the angular sampling distance for different map resolutions with and without the off-grid optimization. Individual plots for each resolution are given in the supporting information.

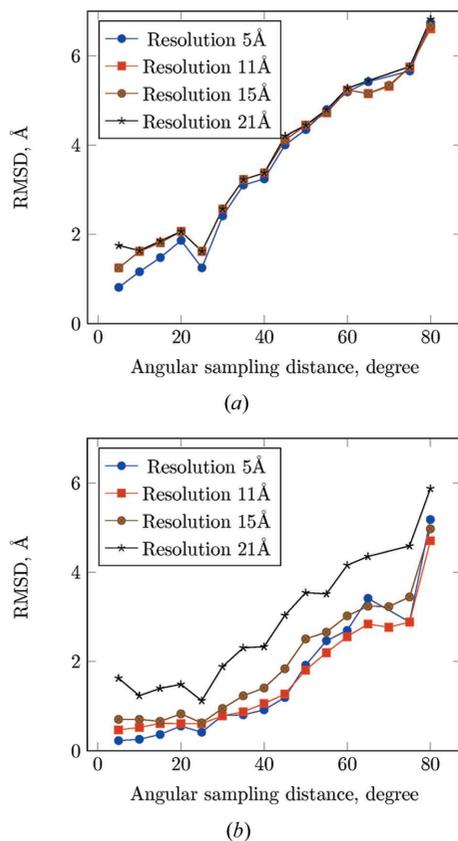
From this figure we can clearly see that our method improves the convergence to the global solution with the angular sampling distance getting smaller. However, one can notice that for very low resolution maps, while the RMSD is effectively reduced, the convergence rate of the refined and unrefined methods is the same (see *e.g.* the 21 Å resolution curves in Fig. 6).

In order to explain this poor convergence rate at low resolutions, let us recall that our CCF is a relatively simplistic scoring function. It has been observed, even with more subtle cost functions, that the success rate of non-Laplace filtered CCFs tends to drop dramatically with the resolution of the

map (Wriggers & Chacón, 2001; Chacón & Wriggers, 2002; Garzon *et al.*, 2007; van Zundert & Bonvin, 2016). It has also been noticed by Chacón & Wriggers that, at low resolution, one may inadvertently displace the model to regions of higher density in the interior of the target map (Chacón & Wriggers, 2002). This means that, even without the quadratic approximation, our cost function has difficulties in discriminating good poses from bad poses for a low-resolution map. This explains why, even with a fine sampling and/or an off-grid search, the method is not able to provide satisfactory results for the low-resolution maps. We should note that while it is in principle simple to adapt our method to a local or core-weighted CCF, which gives better results at low resolutions (van Zundert & Bonvin, 2015, 2016), it would not be easy to use it with a Laplace-filtered CCF. More precisely, regarding the local and core-weighted CCFs, we can either make an assumption that the mask of these scoring functions does not change during the local off-grid deformations, or use a smooth interpolation of the mask and analytically compute its derivatives. The first assumption seems reasonable and thus we will only have to compute additional normalization factors of the in-grid CCF values, using the fast local correlation function method of Roseman (2003), for example. As for the Laplace-filtered CCF, firstly, the quadratic approximation of such a CCF will be different from what we use. Secondly, the quadratic approximation will degenerate faster when far from the expansion centre.

We can also see in Fig. 6(b), especially for the 5 and 11 Å curves, that for the coarser angular sampling the lower resolution gives the best RMSD. We notice that at an angular sampling distance of roughly 45° the two curves cross and the higher resolution starts to give the best RMSD at larger angles. This crossing phenomenon can be explained by recalling that the map resolution is inversely proportional to the maximal frequency in the Fourier domain and that, as we demonstrate in the supporting information, the error reduction of our method is quadratic in terms of the grid spacing and approximately linear with respect to the highest frequency of the map. This explains why, for a coarser angular sampling, our algorithm gives better results with the lower-resolution map, which has a spacing of 3.0 Å. However, for a fine enough angular sampling, the higher-resolution map, which has a spacing of 2.0 Å, gives better fitting results.

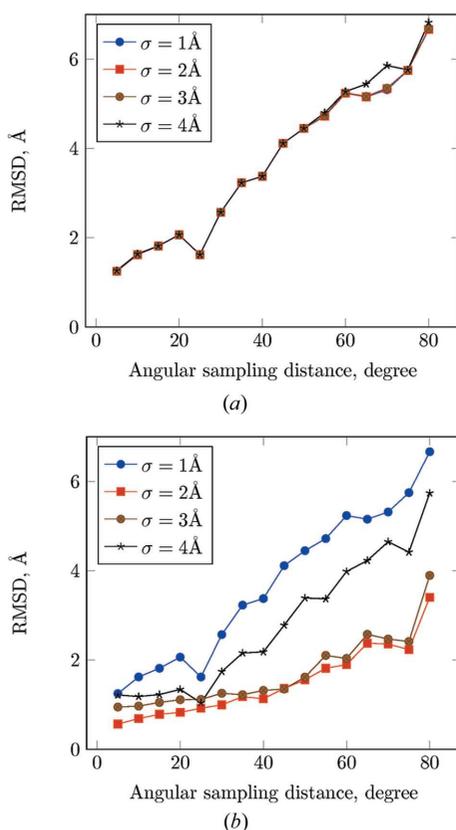
Now, we will demonstrate the effect of the Gaussian  $\sigma$ , which defines the resolution of the model [see equation (1)], on the convergence of our method when fitting a model into the simulated map of 11 Å resolution. Fig. 7 shows the RMSD computed with equation (24) as a function of the angular sampling distance for different values of Gaussian  $\sigma$  with and without the off-grid optimization. Individual plots at different  $\sigma$  values are given in the supporting information. Interestingly, from this figure we can see that for  $\sigma = 1$  Å the off-grid optimization does not improve the results at all. This happens because of the non-concavity of our maximization problem [TRS in equation (21)]. In order to explain this effect, let us first recall that the FFT algorithm approximates the integral in equation (11) by a finite sum, as shown in equation (15). It is



**Figure 6**  
Symmetry-adapted RMSD as a function of the angular sampling distance for fitting a subunit of the GroEL complex to simulated EDMs of different resolutions. (a) Fitting without off-grid optimization. (b) Fitting with off-grid optimization. The grid spacing of the 5 Å resolution map is 2.0 Å and the grid spacing for the 11, 15 and 21 Å resolution maps is 3.0 Å. The resolution of the model is fixed and equals  $\sigma = 2.5$  Å.

thus critical to have a sufficient number of sampling points for the Gaussians and their derivatives. We should note that the integration accuracy would depend on the chosen value of Gaussian  $\sigma$ . For example, the width of  $\sigma = 1 \text{ \AA}$  is definitely too low for the grid spacing of  $3 \text{ \AA}$ , which causes an inaccurate integration of the Hessian in equation (15). These inaccuracies can then lead to a positive semi-definite Hessian, thus making the TRS problem non-concave. Consequently, when the problem is not concave, an optimal solution can be found on the boundary of the search domain (Sorensen, 1982; Fortin, 2000). However, in such a case, the problem is ill posed for at least one degree of freedom (off-grid rotation or translation), and our experience shows that one should keep the original pose in order to avoid an overshoot problem.

For  $\sigma = 4 \text{ \AA}$  we can see that, as in the previous experiment at very low map resolutions, while the RMSD is effectively reduced, the convergence rates of the refined and unrefined methods are nearly the same. This can be explained by the fact that, as we previously mentioned, our cost function has difficulties in discriminating poses for a low-resolution map. We should mention that, to our knowledge, the effect of a low-resolution model on the success rate of the fitting algorithm has not been thoroughly studied, and it is reasonable to assume that a low-resolution model will lead to the same accuracy issues as a low-resolution experimental map. For

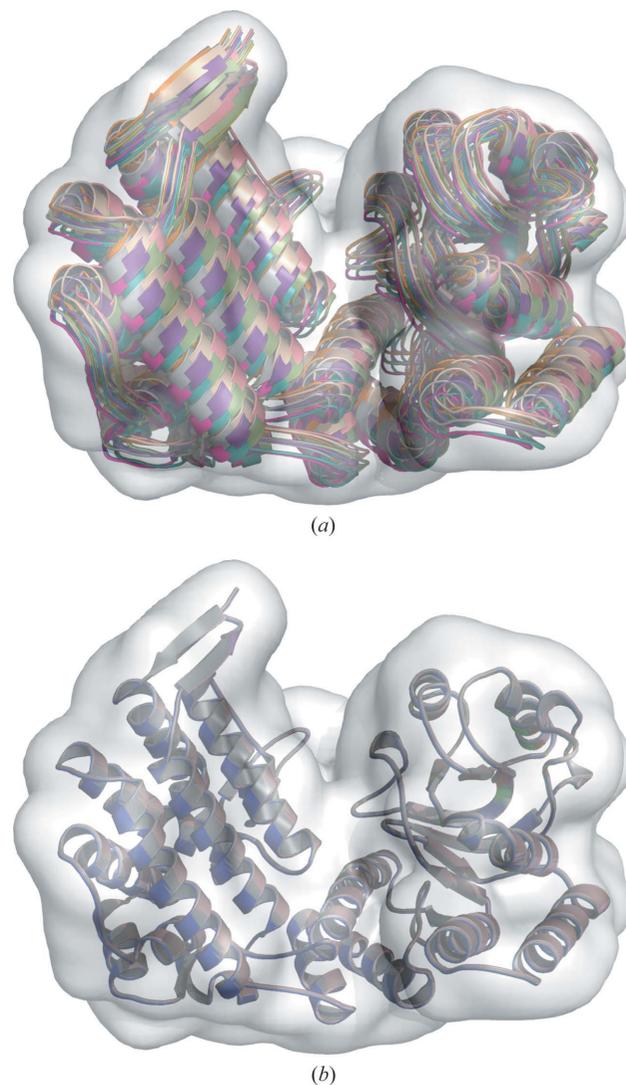


**Figure 7**  
Symmetry-adapted RMSD as a function of the angular sampling distance for fitting subunits of the GroEL complex of different resolutions to a EDM simulated at  $11 \text{ \AA}$  with a grid spacing of  $3.0 \text{ \AA}$ . (a) Fitting without off-grid optimization. (b) Fitting with off-grid optimization.

other values of Gaussian  $\sigma$ , one can see that, when the optimization problem is well posed, our method improves the speed of convergence of our method with respect to the angular sampling distance.

### 3.2. Translational degrees of freedom

In this section we demonstrate the performance of our method when mainly optimizing translational degrees of freedom. To do so, we fitted a single chain of the crystallographic GroEL structure (PDB code *2c7c\_H*) into its simulated map using the correctly rotated model. Firstly, we simulated the maps of 11 different sizes (number of voxels) at a fixed resolution of  $11 \text{ \AA}$  using the *eman2* package (Tang *et al.*, 2007) starting from the atomic model of a GroEL chain (see Fig. 5). Then we fitted the arbitrarily translated model, which had the correct rotation, into the artificial maps without any additional sampling of other rotations. We generated 20 best



**Figure 8**  
The spread of 20 best solutions provided by our fitting algorithm during the translational search. (a) In-grid solutions before the off-grid search. (b) Solutions after the off-grid search. See main text for details.

optimized and non-optimized poses and we then measured the maximum distance between the centres of mass (COM) of two solutions as follows:

$$\text{RMSD}_{\max} := \max_{i,j}(\text{COM}_i, \text{COM}_j), \quad (25)$$

which gives us a certain measure of the maximum spread of solutions generated by our algorithm. We can expect to find the 20 best solutions within this distance. Fig. 8 shows the spread of the 20 best solutions provided by the algorithm with a Gaussian  $\sigma$  equal to 3 Å and a grid with 179 712 voxels. We see that after the off-grid search all the solutions converge to the same off-grid solution, which, according to Fig. 9, has a lower RMSD than the in-grid solutions. More precisely, Fig. 9 shows the RMSD computed with equation (25) as a function of the grid spacing, for different map sizes with the resolution of 11 Å. We can see that for large values of the grid spacing and small values of the model Gaussian  $\sigma$ , the poses have certain difficulties in converging to a single solution. However, we also see that, for small values of  $\sigma = 1$  Å and  $\sigma = 2$  Å, once the grid is sufficiently fine, the poses start to converge to a single solution. We believe that, with such values of  $\sigma$  and a very coarse grid, some Gaussians are not well sampled, which

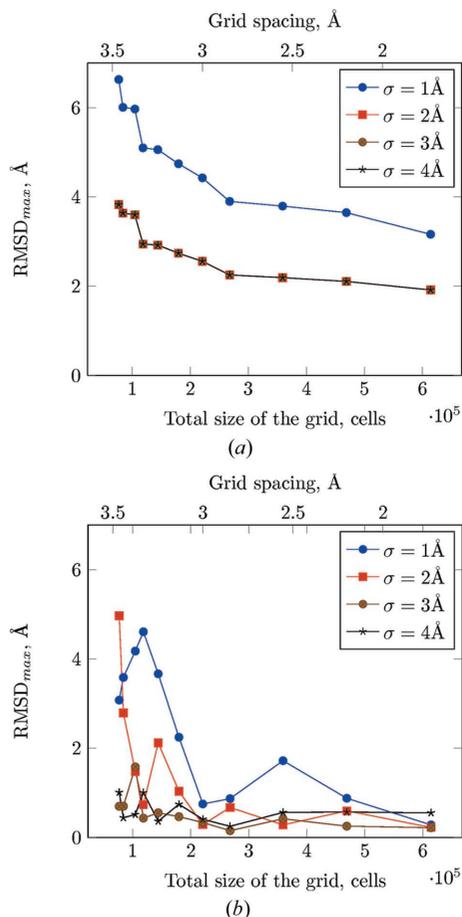
can in turn introduce errors in the gradient and the Hessian of equation (17), and thus lead to a poor fitting quality. At large values of  $\sigma$  equal to 3 and 4 Å, the poses converge to a single solution much better independently of the grid size. This again indicates that the method gives good results as long as the Gaussian width of the model is suited to the spacing of the grid.

It is interesting to note that we do not have the same low-resolution problems as we had in the rotational search examples (see Fig. 9). This observation confirms the results reported by van Zundert & Bonvin, who showed that, for translational degrees of freedom, a relatively simplistic CCF has a good success rate, even for a low-resolution model map (van Zundert & Bonvin, 2016).

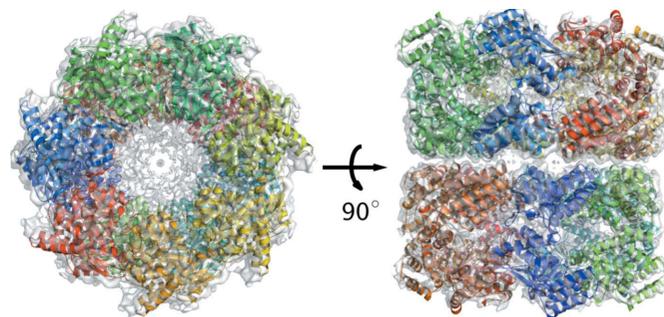
### 3.3. Application to an experimental density map

In this section, we apply our method to an experimental density map. Thanks to recent improvements in cryo-EM instrumentation (Bai *et al.*, 2015), it has become possible to reconstruct cryo-EM maps with resolutions that rival those of X-ray crystallography (Callaway, 2015; Bai *et al.*, 2015; Cheng, 2015). This drastic improvement in resolution created a need for automatic fitting of algorithms with a high precision. We tested our method on a GroEL complex map (code EMD-6422) with  $240 \times 240 \times 240$  voxels and a resolution of 4.1 Å. This map, shown in Fig. 10, was also provided as an assessment exercise in the 2015–2016 EMDDataBank Validation Challenge.

Firstly, we tested the off-grid search when reconstructing the whole GroEL complex by fitting a single chain from the the lower ring of the GroEL–GroES complex (PDB code 2c7c\_H) with the model Gaussian  $\sigma$  of 2 Å into the map. Fig. 10 shows the complex reconstructed by our algorithm with approximately 1000 sampled rotations and 13 824 000 sampled translations. We can see that at the very top and at the very bottom of the complex the individual chains do not fit perfectly. We thus performed a similar experiment with approximately 100 000 rotations without obtaining visually better results. Moreover, we observed that, while the off-grid search improved the fitting of the central part of the complex,



**Figure 9**  $\text{RMSD}_{\max}$  as a function of the grid size for different model resolutions. The resolution of the density map is 11 Å. (a) Fitting without off-grid optimization. (b) Fitting with off-grid optimization.

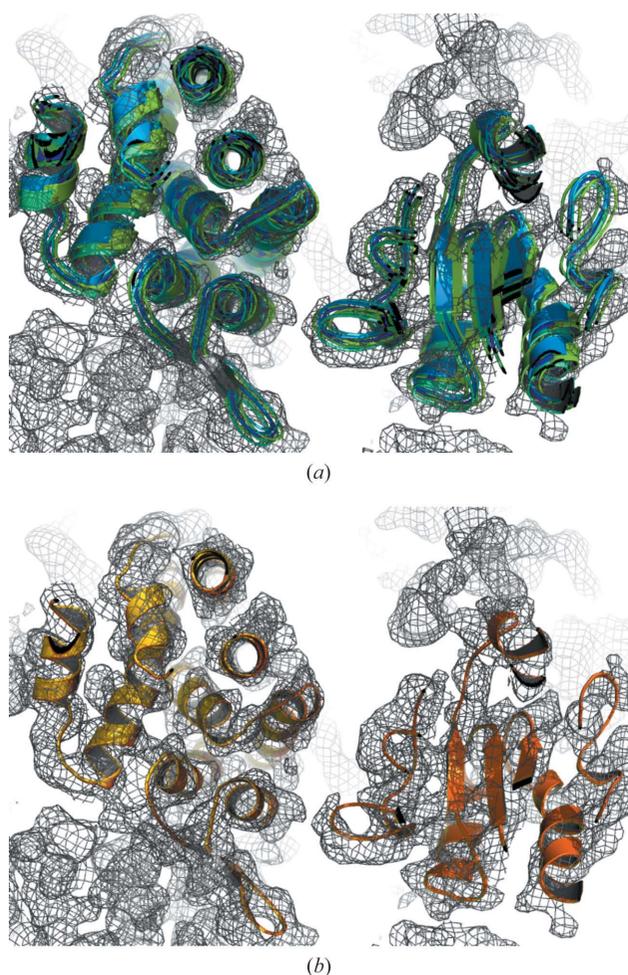


**Figure 10** The experimental GroEL electron-density map (code EMD-6422) with the GroEL structure reconstructed with our algorithm. Left: top view. Right: side view. In total, 13 824 000 translations and approximately 1000 rotations were sampled in order to perform the reconstruction. Each chain of the reconstructed structure is coloured differently.

some inaccuracies can still be observed at the extremes. This can be explained by higher values of the electron density in the centre of the map, which is not currently taken into account by our method.

Secondly, we specifically tested the off-grid translational search by fitting a single correctly rotated chain into the map of the complex. Fig. 11 shows the top five solutions obtained by our algorithm when exhaustively sampling 13 824 000 translations. We see that after the off-grid search all the solutions converge to the same off-grid solution, which visually fits better into the map than the in-grid solutions.

These experiments demonstrate that the additional off-grid rigid search improves the quality of fitting. However, it is still not perfect and our experiments show that for further improvements we need to allow for additional degrees of freedom, such as twisting of secondary-structure elements with respect to each other *etc.* Indeed, in most cases, initial models are determined by X-ray crystallography and thus



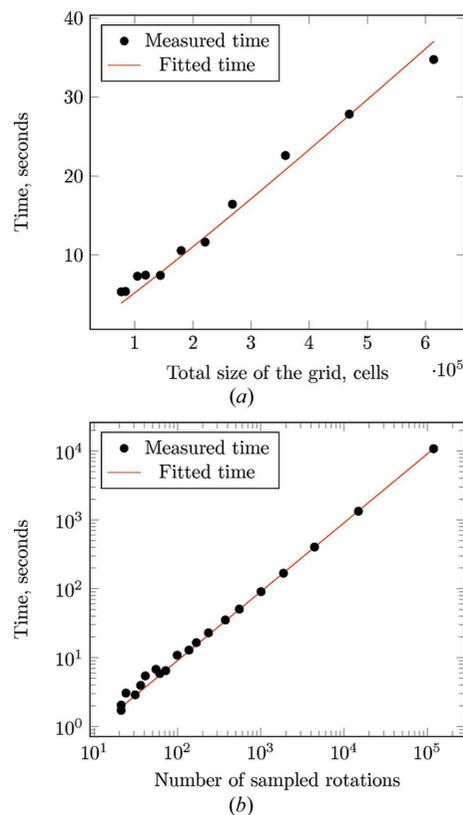
**Figure 11**  
Solutions provided by our fitting algorithm when fitting to the density of a single chain in the map (code EMD-6422) of size  $240 \times 240 \times 240$  voxels and resolution of 4.1 Å. A total of 13 824 000 translations were scanned in order to perform the reconstruction. The top five solutions are shown. (a) Top five in-grid solutions before the off-grid search. (b) Top five solutions after the off-grid search.

might not rigidly fit into a density determined by experiments in solution, such as cryo-EM.

### 3.4. Computational speed

In this section, we demonstrate the running time of our algorithm as a function of both the total size of the grid and the number of sampled rotations. We ran the tests on a Linux machine with 32 GB DDR3 RAM and Intel(R) Core(TM) i7-4800MQ CPU @ 2.70 GHz. For all the tests we used simulated maps of the GroEL structure. The model  $\sigma$  in all the experiments was set to 3 Å. Let us first recall that the overall complexity of our algorithm is  $O[MN \log(N)]$  operations, where  $M$  is the total number of rotations and where  $N$  is the total size of the grid. Below, we will experimentally verify the complexity of our method and measure the wall-clock time for real-size problems.

In order to study the running time as a function of the total size of the electron-density grid, we used 11 maps of a single chain of the GroEL structure (PDB code 2c7c\_H) simulated at a resolution of 11 Å, which range from 76 800 to 614 016 voxels. For this test we sampled only five rotations and measured the time taken by our algorithm to solve the fitting problem. We then fitted the measured time with the



**Figure 12**  
Measured and fitted runtimes of the algorithm. (a) The runtime as a function of the grid size fitted with  $O(N \log N)$ , the theoretical computational complexity, is shown. Only five rotations are sampled here. (b) The runtime as a function of the number of sampled rotations is shown in log–log scale. Here, for each rotation we sampled 163 296 translations. A constant time of 21 s is subtracted from the results in the second test, which is taken by the initialization and the output.

theoretical expectation of  $O[N \log(N)]$ . Fig. 12(a) shows the measured and the fitted times. We can see that the measured computational time fits well to the theoretical  $O[N \log(N)]$  expectation and that our algorithm solves the problem in less than a minute, provided that only five rotations are sampled.

To study the running time as a function of the total number of sampled rotations, we used a map of the GroEL assembly simulated at a resolution of 11 Å, which contains 163 296 voxels. We then fitted a single chain of the GroEL structure into the map and measured the time taken by our algorithm to solve the fitting problem. Fig. 12(b) demonstrates the measured computational time in log–log scale. Again, one can see that the measured time fits well to the theoretical linear expectation. We can see from this figure that, even for rather fine sampling of order  $10^5$  translations and  $10^5$  rotations, our algorithm exhaustively solves the off-grid fitting problem in less than 3 h. To conclude this section, we mention that we have computationally verified the expected complexity of our algorithm and that the pre-factor of our algorithm is not prohibitive, such that the method can be used with realistic large data sets.

#### 4. Conclusions and perspectives

In this paper, we have presented a new method that performs an exhaustive off-grid search in six rigid-body degrees of freedom. The method is applicable to any cost function that can be expressed as a convolution or correlation of two other functions defined on a three-dimensional grid. This method uses an FFT-based technique to quickly sample a quadratic approximation of the cost function and then it uses a well known optimization algorithm to find the optimal off-grid solution.

We have verified the efficiency of the method on simulated maps, and shown that the algorithm reduces the RMSD between the original model and the fitted model compared with the in-grid solutions. We have also tested our method on an experimental density map and demonstrated its efficiency over the in-grid search algorithm. Nonetheless, the method can still be improved in several ways. First, as we have already hinted above, in order to further improve the fitting quality, we should try to add some flexibility to the model. This approach has been widely used in cryo-EM fitting applications (Tama *et al.*, 2004; Suhre *et al.*, 2006; Topf *et al.*, 2008; Rusu *et al.*, 2008; Tan *et al.*, 2008; Wriggers, 2012; López-Blanco & Chacón, 2013). The second improvement we can think of is running several optimization steps while keeping the fast sampling of the off-grid degrees of freedom, which, coupled with a good optimization algorithm, could help us to reduce the RMSD even further and, hopefully, independently of the grid size and the model resolution. It is also worth mentioning that a similar method could be used to sample the off-grid rigid-body DOF in spherical coordinates. The idea would be to apply the Taylor expansion of the Gaussian functions decomposed in a spherical harmonics basis coupled with a radial basis. However, owing to the complexity of both the radial and angular basis

functions, we expect the method to be less stable compared with the presented one.

#### Acknowledgements

The authors acknowledge support from the Inria-associated team PPI-3D and the Agence Nationale de la Recherche, grant No. ANR-11-MONU-006-01.

#### References

- Amir, N., Cohen, D. & Wolfson, H. J. (2015). *Bioinformatics*, **31**, 2801–2807.
- Bai, X., McMullan, G. & Scheres, S. H. (2015). *Trends Biochem. Sci.* **40**, 49–57.
- Callaway, E. (2015). *Nature*, **525**, 172–174.
- Chacón, P. & Wriggers, W. (2002). *J. Mol. Biol.* **317**, 375–384.
- Cheng, Y. (2015). *Cell*, **161**, 450–457.
- Conn, A. R., Gould, N. I. M. & Toint, P. L. (2000). *Trust Region Methods*. Namur: Society for Industrial and Applied Mathematics.
- Crowther, R. (1972). *Comput. Phys. Commun.* **3**, 137–138.
- Derevyanko, G. & Grudin, S. (2014). *Acta Cryst.* **D70**, 2069–2084.
- Fortin, C. (2000). PhD thesis, University of Waterloo, Canada.
- Fortin, C. & Wolkowicz, H. (2004). *Optim. Methods Softw.* **19**, 41–67.
- Gabb, H. A., Jackson, R. M. & Sternberg, M. J. (1997). *J. Mol. Biol.* **272**, 106–120.
- Garzon, J. I., Kovacs, J., Abagyan, R. & Chacón, P. (2007). *Bioinformatics*, **23**, 427–433.
- Golub, G. H. & von Matt, U. (1991). *Numer. Math.* **59**, 561–580.
- Gould, N. I. M., Lucidi, S., Roma, M. & Toint, P. L. (1999). *SIAM J. Optim.* **9**, 504–525.
- Hoang, T. V., Cavin, X. & Ritchie, D. W. (2013). *J. Struct. Biol.* **184**, 348–354.
- Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A. A., Aflalo, C. & Vakser, I. A. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 2195–2199.
- King, D. E. (2009). *J. Mach. Learn. Res.* **10**, 1755–1758.
- Kovacs, J. A., Chacón, P., Cong, Y., Metwally, E. & Wriggers, W. (2003). *Acta Cryst.* **D59**, 1371–1376.
- Kovacs, J. A. & Wriggers, W. (2002). *Acta Cryst.* **D58**, 1282–1286.
- López-Blanco, J. R. & Chacón, P. (2013). *J. Struct. Biol.* **184**, 261–270.
- Moré, J. J. & Sorensen, D. C. (1983). *SIAM J. Sci. Comput.* **4**, 553–572.
- Navaza, J. (1987). *Acta Cryst.* **A43**, 645–653.
- Navaza, J. (1994). *Acta Cryst.* **A50**, 157–163.
- Neveu, E., Ritchie, D. W., Popov, P. & Grudin, S. (2016). *Bioinformatics*, **32**, i693–i701.
- Park, S. Y., Borbat, P. P., Gonzalez-Bonet, G., Bhatnagar, J., Pollard, A. M., Freed, J. H., Bilwes, A. M. & Crane, B. R. (2006). *Nat. Struct. Mol. Biol.* **13**, 400.
- Popov, P. & Grudin, S. (2014). *J. Comput. Chem.* **35**, 950–956.
- Ranson, N. A., Clare, D. K., Farr, G. W., Houldershaw, D., Horwich, A. L. & Saibil, H. R. (2006). *Nat. Struct. Mol. Biol.* **13**, 147–152.
- Ritchie, D. W. (2005). *J. Appl. Cryst.* **38**, 808–818.
- Ritchie, D. W. & Kemp, G. J. (2000). *Proteins Struct. Funct. Bioinf.* **39**, 178–194.
- Ritchie, D. W., Kozakov, D. & Vajda, S. (2008). *Bioinformatics*, **24**, 1865–1873.
- Roseman, A. M. (2003). *Ultramicroscopy*, **94**, 225–236.
- Rossmann, M. G. (2000). *Acta Cryst.* **D56**, 1341–1349.
- Rusu, M., Birmanns, S. & Wriggers, W. (2008). *Bioinformatics*, **24**, 2460–2466.
- Siebert, X. & Navaza, J. (2009). *Acta Cryst.* **D65**, 651–658.
- Sorensen, D. C. (1982). *SIAM J. Numer. Anal.* **19**, 409–426.
- Suhre, K., Navaza, J. & Sanejouand, Y.-H. (2006). *Acta Cryst.* **D62**, 1098–1100.
- Tama, F., Miyashita, O. & Brooks, C. L. III (2004). *J. Struct. Biol.* **147**, 315–326.

- Tan, R. K.-Z., Devkota, B. & Harvey, S. C. (2008). *J. Struct. Biol.* **163**, 163–174.
- Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I. & Ludtke, S. J. (2007). *J. Struct. Biol.* **157**, 38–46.
- Toint, P. L., Tomanos, D. & Weber-Mendonca, M. (2009). *Optim. Method. Softw.* **24**, 299–311.
- Topf, M., Lasker, K., Webb, B., Wolfson, H., Chiu, W. & Sali, A. (2008). *Structure*, **16**, 295–307.
- Volkman, N. & Hanein, D. (1999). *J. Struct. Biol.* **125**, 176–184.
- Wriggers, W. (2010). *Biophys. Rev.* **2**, 21–27.
- Wriggers, W. (2012). *Acta Cryst.* **D68**, 344–351.
- Wriggers, W. & Chacón, P. (2001). *Structure*, **9**, 779–788.
- Wu, X., Milne, J. L., Borgnia, M. J., Rostapshov, A. V., Subramaniam, S. & Brooks, B. R. (2003). *J. Struct. Biol.* **141**, 63–76.
- Zundert, G. C. van & Bonvin, A. M. (2015). *AIMS Biophysics*, **2**, 73–87.
- Zundert, G. van & Bonvin, A. (2016). *J. Struct. Biol.* **195**, 252–258.