

# Fast and Robust Single Particle Reconstruction in 3D Fluorescence Microscopy

Thibaut Eloy , Étienne Baudrier , Marine Laporte , Virginie Hamel , Paul Guichard , and Denis Fortun

**Abstract**—Single particle reconstruction has recently emerged in 3D fluorescence microscopy as a powerful technique to improve the axial resolution and the degree of fluorescent labeling. It is based on the reconstruction of an average volume of a biological particle from the acquisition of multiple views with unknown poses. Current methods are limited either by template bias, restriction to 2D data, high computational cost or lack of robustness to low fluorescent labeling. In this work, we propose a single particle reconstruction method dedicated to convolutional models in 3D fluorescence microscopy that overcomes these issues. We address the joint reconstruction and estimation of the poses of the particles, which translates into a challenging non-convex optimization problem. Our approach is based on a multilevel reformulation of this problem, and the development of efficient optimization techniques at each level. We demonstrate on synthetic data that our method outperforms the standard approaches in terms of resolution and reconstruction error, while achieving a low computational cost. We also perform successful reconstruction on real datasets of centrioles to show the potential of our method in concrete applications.

**Index Terms**—Fluorescence microscopy, image reconstruction, image restoration, inverse problems, machine vision.

## I. INTRODUCTION

**F**LUORESCENCE microscopy has experienced great resolution improvements in recent years thanks to progress in super resolution microscopy [1] and more recently in expansion microscopy [2]. These advances have motivated the use of fluorescence microscopy in structural biology, to decipher the structure of macromolecular assemblies that were previously observable only with electron microscopy [3]. However, fluorescence microscopy is still limited by two issues. Firstly, the resolution of most fluorescence microscopes is strongly anisotropic (the resolution in the axial direction of the microscope is typically

Manuscript received 17 October 2022; revised 24 February 2023; accepted 28 March 2023. Date of publication 19 April 2023; date of current version 30 June 2023. The work of Virginie Hamel and Paul Guichard was supported by Swiss National Science Foundation under Grant 310030\_205087. The work of Paul Guichard was supported by European Research Council under StG Grant 715289. This work was supported in part by French National Research Agency through SP-Fluo Project under Grant ANR-20-CE45-0007, in part by the ANR Spflio A21R340B, and in part by Swiss National Science Foundation under Grant PP00P3\_187198. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Alejandro F. Frangi. (*Corresponding author: Thibaut Eloy*)

Thibaut Eloy, Étienne Baudrier, and Denis Fortun are with the ICube - UMR7357, CNRS, University of Strasbourg, 67081 Strasbourg, France (e-mail: thibaut.elyo@etu.unistra.fr; baudrier@unistra.fr; dfortun@unistra.fr).

Marine Laporte, Virginie Hamel, and Paul Guichard are with the Centriole Lab, Department of Molecular and Cellular Biology, University of Geneva, 1205 Geneva, Switzerland (e-mail: marine.laporte@unige.ch; virginie.hamel@unige.ch; paul.guichard@unige.ch).

Digital Object Identifier 10.1109/TCI.2023.3266996

3 to 5 times lower than in the lateral plane). Secondly, spurious intensity variations along the protein structures can hinder the biological interpretation. This phenomenon is explained by a low and non uniform degree of labeling (DOL), which characterizes the density of fluorescent dyes that attach to the protein of interest and emit fluorescent light.

We address these issues with a single particle reconstruction (SPR) approach. The principle is to combine several acquisitions of identical biological particles observed from different orientations, to reconstruct a single particle model. Thanks to the fusion of complementary information in each view, we aim at reconstructing volumes with isotropic resolution and uniform DOL. This approach has received increasing attention in recent years and offers new perspectives in structural biology to decipher the protein architecture of large macromolecular assemblies [2], [4], [5], [6], [7], [8], [9].

In this work, we focus on a category of fluorescence microscopy techniques that we call *convolutional modalities*, for which the observation model is the convolution with a point spread function (PSF). Confocal and STED (stimulated-emission-depletion) microscopy [10] are the main representatives of this category and have been proven to be well suited to observe important macromolecular assemblies such as the centriole or the nuclear pore, in particular when combined with expansion microscopy [2], [6].

### A. Challenges and Related Works

Multiview reconstruction is a common practice in several microscopy systems for which physical calibration procedures provide an accurate estimation of the pose of each view. (e.g. selective plane illumination microscopy, [11], [12], 3D total-internal reflection fluorescence microscopy [13]). By contrast, the specificity of the multiview reconstruction problem in SPR is the absence of prior knowledge about the poses. This results in a particularly challenging non-convex optimization problem that requires the development of dedicated methods.

SPR has a longstanding and successful history in another modality known as cryo-electron microscopy (cryo-EM), for which a lot of efficient reconstruction methods have been proposed over the past 20 years [14], [15]. However, the imaging model of cryo-EM fundamentally differs from that of convolutional modalities. The cryo-EM images are 2D projections of the 3D densities, whereas convolutional modalities acquire 3D images by convolution of the fluorescent signal with a PSF. Despite this difference, the most common practice for SPR

in fluorescence imaging is to apply cryo-EM reconstruction methods on 2D fluorescence data [4], [5], which ignores the true imaging model of optical microscopes and leads to coarse and suboptimal reconstructions.

Most methods developed specifically for fluorescence microscopy are either restricted to 2D input images [16], [17], [18], or require the use of an initial template or a parametric model [19], [20], [21], which inevitably introduces a bias in the reconstruction. The method we proposed in [22] is the only one, to the best of our knowledge, to take into account an appropriate 3D imaging model. Despite good performance on real data [2], [6], it has some major practical limitations. Firstly, it requires a classification step to reduce the number of particles. Automatic classification often lacks robustness on real data, and manual particle selection is tedious and introduces an arbitrary bias. The computation time of [22] can also be prohibitive for non-symmetric particles, for which the space of the registration parameters cannot be reduced. Finally, it is based on a sequential optimization procedure that considers the views one by one in a predefined order, which leads to reconstruction errors for low fluorescent labeling rates.

The SPR principle has also been applied in single molecule localization microscopy (SMLM), a fluorescence imaging technique that generates data in the form of point clouds. A few studies have been dedicated to SPR for SMLM data from the point of view of point clouds registration [7], [20]. Our reconstruction problem with convolutional modalities is radically different since our data is a collection of volumes, which cannot be processed with the same methods as for point clouds. Moreover, our goal is reconstruction and not only registration.

## B. Contributions

We propose a new method to estimate jointly the poses (rotation and translation) of the particles and the reconstruction in SPR for convolutional modalities in fluorescence imaging. Our approach is based on a reformulation of the original joint optimization problem with respect to (w.r.t.) the volume, the rotations and the translations, into an equivalent hierarchical optimization scheme that decouples the minimization w.r.t. each variable in three nested levels. The volume is estimated at the highest level with stochastic optimization, and the key for the success of our approach is the development of fast approximate solvers for the minimization w.r.t. to the rotations and translations at the lower levels. Thanks to this approach, the computational cost of a reconstruction is reduced to less than 30 minutes, with a reference-free initialization. Moreover, our approach does not require any critical hyper-parameter tuning, which simplifies its usage in practice. We demonstrate on several synthetic datasets that we outperform the standard methods in terms of resolution and visual assessment. Finally, we show the potential of our approach on real data of centrioles acquired with a combination of confocal and expansion microscopy.

Our work builds upon the cryo-EM reconstruction method cryoSPARC described in [15]. We deviate from it on the main following points:

- cryoSPARC cannot be applied to our data since it is designed for a forward model of 2D projection. Instead, we develop a method dedicated to 3D convolutional models.

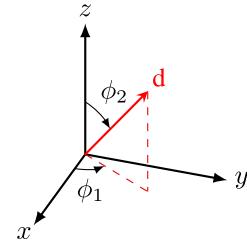


Fig. 1. Representation of a direction  $d$  (point on the unit sphere) with its azimuth  $\phi_1$  and inclination  $\phi_2$ .

- We estimate the poses with a multilevel optimization scheme, whereas the strategy of [15] is to reconstruct the volume without pose estimation by maximizing a marginalized likelihood.
- We estimate the translations at the last level of our hierarchical scheme with a fast phase correlation algorithm. The decoupling of translations and rotations estimation is crucial for computational tractability in our 3D case.

The organization of the manuscript is as follows. In Section II, we present the observation model and formulate the reconstruction as an optimization problem. In Section III, we present our optimization method. In Section IV we present the experimental results obtained both on simulated and real datasets.

## II. PROBLEM FORMULATION

### A. Background on the Representation of Rotations

We use the axis-angle representation  $(d, \psi)$  to represent a 3D rotation, where  $d \in \mathbb{R}^3$  is the rotation axis, and  $\psi \in [0, 2\pi[$  is the rotation angle around this axis. The axis  $d$  is the unit vector:

$$d = \begin{pmatrix} \cos(\phi_1) \sin(\phi_2) \\ \sin(\phi_1) \sin(\phi_2) \\ \cos(\phi_2) \end{pmatrix}, \quad (1)$$

where  $\phi_1 \in [0, 2\pi[$  is the azimuth and  $\phi_2 \in [0, \pi[$  is the inclination (see Fig. 1).

Our method requires a discretization of  $\text{SO}(3)$ . We write it  $\{\theta_{i,j} = (\phi_{1,j}, \phi_{2,j}, \psi_i) \mid (j, i) \in [\![0, M_d - 1]\!] \times [\![0, M_\psi - 1]\!]\}$ , where  $M_d$  and  $M_\psi$  are the discretization sample sizes of  $(\phi_1, \phi_2)$  and  $\psi$ , respectively. For  $\psi$ , we choose a uniform discretization of  $[0, 2\pi[$ :  $\forall i \in [\![0, M_\psi - 1]\!] : \psi_i = 2\pi i / M_\psi$ . For  $(\phi_1, \phi_2)$ , we use the Fibonacci discretization, which gives an almost uniform discretization of the unit sphere. This is formalized as follows

$$\forall j \in [\![0, M_d - 1]\!], \quad \phi_{1,j} \equiv \frac{2\pi j}{G} [2\pi] \quad (2)$$

$$\phi_{2,j} = \arccos \left( 1 - \frac{2j + 1}{M_d} \right), \quad (3)$$

with  $G = \frac{1+\sqrt{5}}{2}$  the golden ratio.

### B. Forward Model

Let us denote  $\{y_l\}_{l \in [\![1, N]\!]}$  a set of  $N$  volumes (we call volume a function  $\Omega \rightarrow \mathbb{R}$  where  $\Omega$  is a bounded domain in  $\mathbb{R}^3$ ), where  $y_l$  is a view that contains a single particle, with an orientation

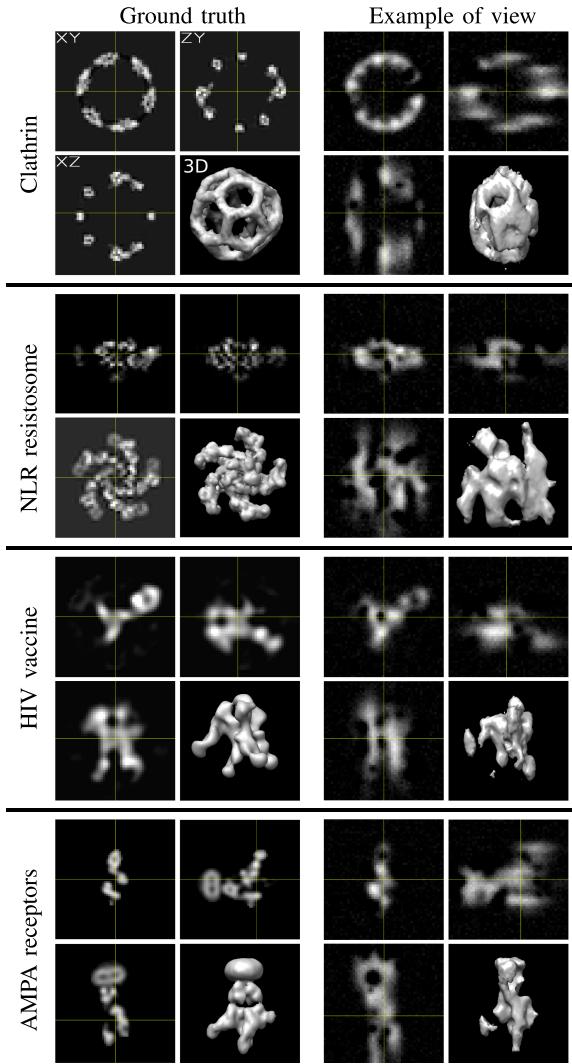


Fig. 2. Examples of simulated data. On the left: ground truth volume; on the right: example of a view simulated from the ground truth. For each volume, we show 3 orthogonal views and a 3D isosurface visualization.

$\theta_l \in \mathbb{R}^3$  and a translation  $t_l \in \mathbb{R}^3$ . From this set of views, the goal is to reconstruct a volume  $f$ , that represents the observed particle. The image acquisition model belongs to convolutional modalities, defined by

$$y_l(x) = (h * \mathcal{T}_{t_l}(\mathcal{R}_{\theta_l}(f)))(x) + \varepsilon(x), \quad (4)$$

where  $h$  is the PSF of the microscope,  $\varepsilon$  is an additive Gaussian noise,  $\mathcal{T}_t$  is the translation operator of vector  $t$  defined by

$$\mathcal{T}_t(f)(x) = f(x - t), \quad (5)$$

and  $\mathcal{R}_\theta$  is the rotation operator of angle  $\theta$  defined by

$$\mathcal{R}_\theta(f)(x) = f(R_\theta^T x), \quad (6)$$

where  $R_\theta \in SO_3(\mathbb{R})$  is the rotation matrix of angle  $\theta$ .

Examples of volumes  $f$  and views  $y_l$  are represented in Fig. 2.

We do not make any assumption on the form of the PSF, though in practice it has an anisotropic shape, more elongated along the Z axis [23], which creates the anisotropy of resolution.

### C. Joint Minimization Problem

We formulate the reconstruction as a maximum likelihood estimation of the volume, the rotation and the translation parameters, which amounts to the following least squares problem:

$$f^*, \Theta^*, T^* = \arg \min_{f, \Theta, T} E(f, \Theta, T), \quad (7)$$

where

$$E(f, \Theta, T) = \sum_{l=1}^N \|y_l - h * \mathcal{T}_{t_l}(\mathcal{R}_{\theta_l}(f))\|_2^2, \quad (8)$$

with

$$\Theta = \{\theta_l\}_{l \in \llbracket 1, N \rrbracket} \text{ and } T = \{t_l\}_{l \in \llbracket 1, N \rrbracket}.$$

We found experimentally that using a regularization in (8) did not improve the results (we considered an  $l_2$  penalization of the Fourier coefficients). While regularization plays an important role in single view deconvolution to prevent noise fitting, the sum over  $N$  views in (8) induces an averaging effect that acts as a smoothness constraint. Therefore, adding a regularization term turned out to be unnecessary in our case

To reduce the computational cost, we formulate the energy (8) in the Fourier domain. This allows us to perform convolutions with simple pointwise multiplication. In addition, rotation and Fourier transform commute, and the translation in the Fourier domain is simply a phase shift. Formally, we have

$$\mathcal{F}(h * \mathcal{T}_{t_l}(\mathcal{R}_{\theta_l}(f))) = \hat{h} \rho_{t_l} \mathcal{R}_{\theta_l}(\hat{f}), \quad (9)$$

where  $\mathcal{F}$  is the Fourier transform,  $\rho_{t_l}$  is a phase factor defined by  $\rho_{t_l} : \omega \mapsto e^{it_l \omega}$ , and we use the notation  $\mathcal{F}(f) = \hat{f}$ . The energy, written in the Fourier domain is obtained with Parseval's theorem:

$$E(f, \Theta, T) = \sum_{l=1}^N \|\hat{y}_l - \hat{h} \rho_{t_l} \mathcal{R}_{\theta_l}(\hat{f})\|_2^2. \quad (10)$$

## III. OPTIMIZATION

The optimization problem (7) is non-convex with many local minima. It is reminiscent of auto-calibration problems in microscopy, where the volume and some parameters of the observation model are estimated jointly. The alternating optimization scheme commonly used for auto-calibration is unable to escape from local minima, and it relies on theoretical observation models (such as PSF models in blind deconvolution) that provide a good initialization of the parameters. In our case, an alternating scheme is destined to fail because we do not have any initial guess of the volume or the poses.

### A. Multilevel Reformulation of the Optimization Problem

To overcome this issue, we reformulate the problem (7) in the multilevel equivalent form

$$\hat{f}^* = \arg \min_{\hat{f}} E(f, \Theta^*, T^*) \quad (11a)$$

$$\text{s.t. } \Theta^* = \arg \min_{\Theta} E(f, \Theta, T^*) \quad (11b)$$

$$\text{s.t. } T^* = \arg \min_T E(f, \Theta, T). \quad (11c)$$

In this scheme, a subproblem at a given level in (11) is nested as a constraint in the subproblem of the upper level. The main interest of this reformulation is to decompose the original intractable problem into three tractable subproblems, similarly with alternating optimization. However, differently from an alternating scheme, the locally optimal points of the joint problem (7) are also optimal points of the multilevel formulation.

The counterpart is that if an iterative method is used to solve a problem at a given level, the lower level problem must be solved at each iteration. Thus, the computational cost of this nested optimization can become prohibitive if the solvers of the subproblems are not efficient enough (for more details on multi-level optimization see [24]). In the following three subsections, we detail our fast and complementary optimization strategies for the subproblems in (11).

### B. Optimization With Respect to the Volume

When the pose parameters  $\Theta$  and  $T$  are known, the minimization problem (11a) can be solved easily in closed form [22]. However, since pose estimation is formulated as constraints (11b) and (11c), we have to adopt an iterative approach to update the poses at each iteration. Moreover the constrained problem is non-convex, and the sum over a potentially large number  $N$  of particles can dramatically increase the computational cost. To overcome these issues, we use a simple stochastic gradient descent strategy. The volume is updated with the following equation:

$$\hat{f}^{(n+1)}(\omega) = \hat{f}^{(n)}(\omega) - \mu \nabla_{\hat{f}} E_{l_n}(f^{(n)}, \theta_{l_n}^*, t_{l_n}^*)(\omega), \quad (12)$$

where  $l_n \in \llbracket 1, N \rrbracket$  is a randomly selected view index at iteration  $n$ ,  $E_l(f, \theta_l, t_l) = \|\hat{y}_l - \hat{h}\rho_{t_l} \mathcal{R}_{\theta_l}(\hat{f})\|_2^2$  is the energy associated to index  $l$  and  $\mu$  is the gradient descent step. The random indices  $l_n$  are imposed to cycle through all the views, in order to divide the optimization procedure in epochs. The computational cost at iteration  $n$  is reduced to the evaluation of the gradient of  $E_{l_n}$  (the derivation of the gradient is detailed in Appendix VII). The robustness of SGD to non-convex problems has been established for similar objective functions in the context of learning applications or for cryo-EM reconstruction [15], although convergence properties can be guaranteed only in the convex case.

### C. Optimization With Respect to the Orientations

The secondary optimization problem (11b), dedicated to the estimation of the orientations, has to be solved at each iteration of the main problem (11a) for the evaluation of the gradient  $\nabla_{\hat{f}} E_{l_n}$ . The  $N$  views are considered independently, so (11b) can be rewritten as:

$$\forall l \in \llbracket 1, N \rrbracket \quad \theta_l^* = \arg \min_{\theta_l} E_l(f, \theta_l, t_l^*). \quad (13)$$

A gradient based approach to solve (13) is very likely to reach a poor local minimum because of the non convexity of the problem, such that it can be used only for refinement. On the

other hand, an exhaustive search over a discretization of  $SO(3)$  would be computationally too expensive.

To tackle this issue, we perform an exhaustive search on a restricted set of orientations  $\{\theta_{i,j} | (j, i) \in \mathcal{I}^{\psi,l} \times \mathcal{I}^{d,l}\}$ , where  $\mathcal{I}^{\psi,l} \subset \llbracket 0, M_\psi - 1 \rrbracket$  and  $\mathcal{I}^{d,l} \subset \llbracket 0, M_d - 1 \rrbracket$  are subsets of indices in the  $SO(3)$  discretization defined in Section II-A. The sizes  $N_d$  and  $N_\psi$  of  $\mathcal{I}^{l,d}$  and  $\mathcal{I}^{l,\psi}$ , respectively, are chosen such that  $N_d \ll M_d$  and  $N_\psi \ll M_\psi$ . Let us define  $E_{i,j}^l = E_l(f, \theta_{i,j}, t_l^*)$  the energy associated to orientation  $\theta_{i,j}$ . The solution of (13) is approximated by  $\theta_{i^*,j^*}$  with

$$i^*, j^* = \arg \min_{i \in \mathcal{I}^{\psi,l}, j \in \mathcal{I}^{d,l}} E_{i,j}^l.$$

The critical issue is to design a fast method to build subsets  $\mathcal{I}^{\psi,l}$  and  $\mathcal{I}^{d,l}$  that are adapted to the landscape of the energy  $E_{i,j}^l$ . Let us define  $\mathcal{Q}^{l,d}$  and  $\mathcal{Q}^{\psi,d}$  the sampling distributions from which we create  $\mathcal{I}^{\psi,l}$  and  $\mathcal{I}^{d,l}$ . Since we are looking for the maximum likelihood in (13), we can follow the approach of [15] and define  $\mathcal{Q}^{l,d}$  and  $\mathcal{Q}^{\psi,d}$  from the marginal likelihoods w.r.t  $d$  and  $\psi$ , respectively. These marginal likelihoods are defined at each point of the discretization by

$$\forall i \in \llbracket 0, M_d - 1 \rrbracket, \quad \pi_i^{l,\psi} = \sum_{j=0}^{M_\psi-1} p_{i,j}^l, \quad (14)$$

$$\forall j \in \llbracket 0, M_\psi - 1 \rrbracket, \quad \pi_j^{l,d} = \sum_{i=0}^{M_d-1} p_{i,j}^l, \quad (15)$$

where  $p_{i,j}^l \propto \exp(-E_{i,j}^l)$  is the likelihood associated to the orientation  $\theta_{i,j}$ . However the likelihoods are unknown and we need to approximate them. The key observation is that the variation of the volume  $\hat{f}$  between two consecutive iterations of (12) is small. Therefore, the solution of (13) at a given iteration is close to the solution at the previous iteration. Thus, we can reasonably approximate the likelihoods in (14) and (15) by their values at the previous iterations (in what follows,  $p_{i,j}^l$  refers to its value at the previous iteration, without notation change for the sake of readability).

The difficulty is that the likelihood is not evaluated for all the elements of the sums in (14) and (15), but only at the indices  $\mathcal{I}^{l,d}$  and  $\mathcal{I}^{l,\psi}$ . This issue is overcome by using importance sampling [25] to approximate the marginal likelihoods (14) and (15). Since the indices are drawn from the distributions  $\mathcal{Q}^{l,d}$  and  $\mathcal{Q}^{\psi,d}$ , they can be used as importance distributions to create the following approximations of  $\pi^{l,\psi}$  and  $\pi^{l,d}$ :

$$\forall i \in \mathcal{I}^{l,\psi} \quad \tilde{\pi}_i^{l,\psi} = \sum_{j \in \mathcal{I}^{l,\psi}} \frac{p_{i,j}^l}{\mathcal{Q}_j^{l,\psi}}, \quad (16)$$

$$\forall j \in \mathcal{I}^{l,d} \quad \tilde{\pi}_j^{l,d} = \sum_{i \in \mathcal{I}^{l,d}} \frac{p_{i,j}^l}{\mathcal{Q}_i^{l,d}}. \quad (17)$$

To compute the approximated likelihoods  $\tilde{\pi}_i^{l,\psi}$  and  $\tilde{\pi}_j^{l,d}$  associated to all the discretization elements of  $SO(3)$ , a kernel estimation method is used:

$$\forall i \in \llbracket 0, M_\psi - 1 \rrbracket, \quad \tilde{\pi}_i^{l,\psi} = Z_\psi^{-1} \sum_{k \in \mathcal{I}^{l,\psi}} K_{i,k}^\psi \tilde{\pi}_k^{l,\psi}, \quad (18)$$

$$\forall j \in \llbracket 0, M_d - 1 \rrbracket, \tilde{\pi}_j^{l,d} = Z_d^{-1} \sum_{k \in \mathcal{I}^{l,d}} K_{j,k}^d \tilde{\pi}_k^{l,d}, \quad (19)$$

where  $Z_\psi$  and  $Z_d$  are normalization constants given by

$$Z_\psi = \sum_{i=1}^{M_\psi} \sum_{k \in \mathcal{I}^{l,\psi}} K_{i,k}^\psi \tilde{\pi}_k^{l,\psi} \text{ and } Z_d = \sum_{j=1}^{M_d} \sum_{k \in \mathcal{I}^{l,d}} K_{j,k}^d \tilde{\pi}_k^{l,d},$$

and  $K^\psi$  and  $K^d$  are kernels defined by:

$$\forall i, k \in \llbracket 0, M_\psi - 1 \rrbracket, K_{i,k}^\psi = \exp(\beta_\psi \cos(\psi_i - \psi_k)), \quad (20)$$

$$\forall j, k \in \llbracket 0, M_d - 1 \rrbracket, K_{j,k}^d = \exp(\beta_d d_j^T d_k), \quad (21)$$

where  $\beta_\psi$  and  $\beta_d$  are hyperparameters.

In the first iterations of (12), the hypothesis of small variation can be violated until reaching a stable coarse reconstruction. To account for this, the final sampling distributions are defined as linear combinations of the approximated likelihoods with a uniform distribution:

$$\mathcal{Q}^{l,d} = \alpha \mathcal{U}^d + (1 - \alpha) \tilde{\pi}^{l,d}, \quad (22)$$

$$\mathcal{Q}^{l,\psi} = \alpha \mathcal{U}^\psi + (1 - \alpha) \tilde{\pi}^{l,\psi}, \quad (23)$$

where  $\mathcal{U}^\psi = \frac{1}{M_\psi}$  and  $\mathcal{U}^d = \frac{1}{M_d}$  are the uniform components, and  $\alpha$  is the coefficient adjusting the proportion of uniform component. The uniform component makes it possible to explore the entire rotation space at the beginning of the reconstruction process, and the evolution of  $\alpha$  gradually restricts the search space around plausible solutions. It is set to 1 at the beginning of the gradient descent, and is divided by  $\alpha_r$  between two epochs of the SGD (12).

#### D. Optimization With Respect to the Translations

The problem (11c) has to be solved at each evaluation of the energy  $E$  in the minimization of (11b) described in the previous section. It can be written

$$\forall l \in \llbracket 1, N \rrbracket, t_l^* = \arg \min_{t_l} \|\hat{y}_l - \rho_{t_l} \hat{b}_l\|_2^2, \quad (24)$$

where  $\hat{b}_l = \hat{h} \mathcal{R}_{\theta_l}(\hat{f})$  is the rotated and convolved volume in the Fourier domain. The problem amounts to the estimation of a phase shift  $\rho_{t_l}$  of  $\hat{b}_l$  to match the view  $\hat{y}_l$ . It can be straightforwardly reformulated as the maximization of the cross-correlation [26]:

$$t_l^* = \arg \max_{t_l} \int \rho_{t_l}(\omega) \hat{y}_l(\omega) \bar{\hat{b}}_l(\omega) d\omega, \quad (25)$$

where  $\bar{z}$  refers to the conjugate of  $z$ . Since the phase shift corresponds to a translation in the spatial domain, the most efficient way to solve (25) is to find the spatial position of the correlation peak in the inverse Fourier transform of the cross-correlation. In practice we use a standard phase correlation algorithm which considers a normalized cross-correlation instead of a simple cross-correlation in (25), in order to get a sharper and more accurate correlation peak [27]. The computational cost of this step essentially lies in a single inverse Fourier transform. This efficiency is crucial for the computational tractability of the

overall reconstruction algorithm since this step is nested in the two upper problems (11a) and (11b).

The detailed reconstruction steps are regrouped in Algorithm 1.

## IV. EXPERIMENTAL RESULTS

In this section, we present experiments to evaluate the performance of our method. First, we show reconstructions on synthetic data sets and compare our results with other standard approaches. We also analyze the robustness of our method to different imaging conditions and its sensitivity to hyperparameters. Finally, we show results obtained on a real dataset of centrioles.

### A. Simulated Data

We use particle models obtained by reconstruction in cryo-EM and available in the databank EMDataResource.<sup>1</sup> We select the particles AMPA receptor [28], NLR resistosome [29], Clathrin [30], HIV-vaccine [31]. We consider these models as ground truth and simulate the convolutional acquisition process (4) to obtain a set of views, with random orientations. We used a Gaussian PSF  $h$  with covariance  $\Sigma_h$  defined by:

$$\Sigma_h = \begin{pmatrix} \sigma_{xy}^2 & 0 & 0 \\ 0 & \sigma_{xy}^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{pmatrix}, \quad (26)$$

where  $\sigma_z$  and  $\sigma_{xy}$  are the standard deviations along and perpendicular to the microscope axis, respectively. To mimic a realistic anisotropic blur, we choose  $\sigma_z = 5$  and  $\sigma_{xy} = 1.5$  (unless stated otherwise). Note that we choose this particular PSF model because it corresponds to a realistic PSF, but our method can be applied to any PSF. The ground truth volumes are resized to  $50 \times 50 \times 50$  pixels before generating the views. We apply a Gaussian additive noise with a standard deviation of 0.2 (images are scaled from 0 to 1). Intensity variations due to non uniform DOL are simulated by subtracting small Gaussian spots (3D uniform Gaussian functions) at random locations in the image. The standard deviation of the Gaussian spots is uniformly sampled between 2 to 5 percent of the image size, the intensity of the spots is uniformly sampled between 0 and 1 and the number of spots to be removed is set to 120. In Fig. 2, we show the ground truth volumes and examples of simulated views, aligned with the ground truth. To represent them, we use two complementary types of visualizations: three orthogonal views (planes XY, YZ and XZ), and an isosurface 3D visualization (with Chimera<sup>2</sup>). The anisotropy of resolution can be visualized on the intensity maps XZ and YZ.

### B. Metrics

We use three evaluation metrics to evaluate the performance of our method on simulated data. To quantify the visual quality of the reconstruction, we compute the Structural Similarity Index (SSIM) [32] between the ground truth image and the reconstructed volume.

<sup>1</sup>[Online]. Available: <https://www.emdataresource.org/>

<sup>2</sup>[Online]. Available: <https://www.cgl.ucsf.edu/chimera/>

To measure the resolution of the reconstruction, we follow the standard practice in electron microscopy and compute the Fourier Shell Correlation (FSC) [33]. The FSC between the reconstruction  $f^*$  and the ground truth  $f_{gt}$  is a function  $FSC : \mathbb{R} \rightarrow \mathbb{R}$  defined by:

$$FSC(r) = \frac{\sum_{x \in S_r} \mathcal{F}(f^*)(x) \overline{\mathcal{F}(f_{gt})(x)}}{\sqrt{\sum_{x \in S_r} \mathcal{F}(f^*)^2(x) \sum_{x \in S_r} \mathcal{F}(f_{gt})^2(x)}} \quad (27)$$

where  $S_r$  denotes the sphere of radius  $r$ . Thus, the FSC measures the normalized cross-correlation between spherical shells of  $f^*$  and  $f_{gt}$  at different frequency magnitudes in the Fourier space. The resolution of  $f^*$  can be derived from the FSC by determining the frequency limit after which the correlation becomes lower than a cutoff value (we consider a standard cutoff of 0.143). In what follows, we will use the name FSC to refer to the resolution derived from the FSC curve.

In our case, since the resolution of the input data is anisotropic, it is important to assess the ability of our method to make it isotropic. To quantify the degree of isotropy of the resolution, we use a variant of the FSC named conical FSC (cFSC) [34]. The principle of the cFSC is to compute a FSC curve locally for each direction. Thus, we can derive a resolution value for several directions and plot them on a map. Each point on the map is a direction represented by two angles  $\phi_1 \in [0, 2\pi]$  and  $\phi_2 \in [0, \pi]$  (see Fig. 1). In Fig. 4(a) we show the cFSC map between the ground truth object AMPA receptor and a generated view (aligned with the ground truth). This figure highlights the anisotropy of resolution: the resolution in the Z axis ( $\phi_2 = 0$  and  $\phi_2 = 180$ ) is much lower than in other directions. The resolution values are expressed as the inverse of the pixel size.

Since the orientation of the reconstruction does not match the orientation of the ground truth object, the reconstructed volume is registered on the ground truth before computing the FSC, cFSC or the SSIM. The registration is performed with a first step of exhaustive search method maximizing the mutual information. The discretization step is set to 10 degrees. To refine the registration we then use a gradient descent.

### C. Reconstruction Results on Simulated Data

1) *Experimental Setting:* In all our experiments, the volume  $\hat{f}$  is initialized with random values uniformly distributed between 0 and 1. In the default experimental settings, we used 20 views for reconstruction. The standard approach for reference-free particle averaging in fluorescence imaging is the application of cryo-EM reconstruction methods on 2D data. To apply this idea, we create artificial 2D projections from our 3D input volumes, and perform reconstruction with a state-of-the-art cryo-EM reconstruction method based on the RANSAC algorithm [35] (integrated in the Scipion software [36]). We name this method cryo-RANSAC in what follows. We also compare with the results of the SPR method dedicated to 3D convolutional models described in [22]. We name it MP3DR (Multiple Particles 3D Reconstruction) for convenience. This method requires the selection of a small number of particles to

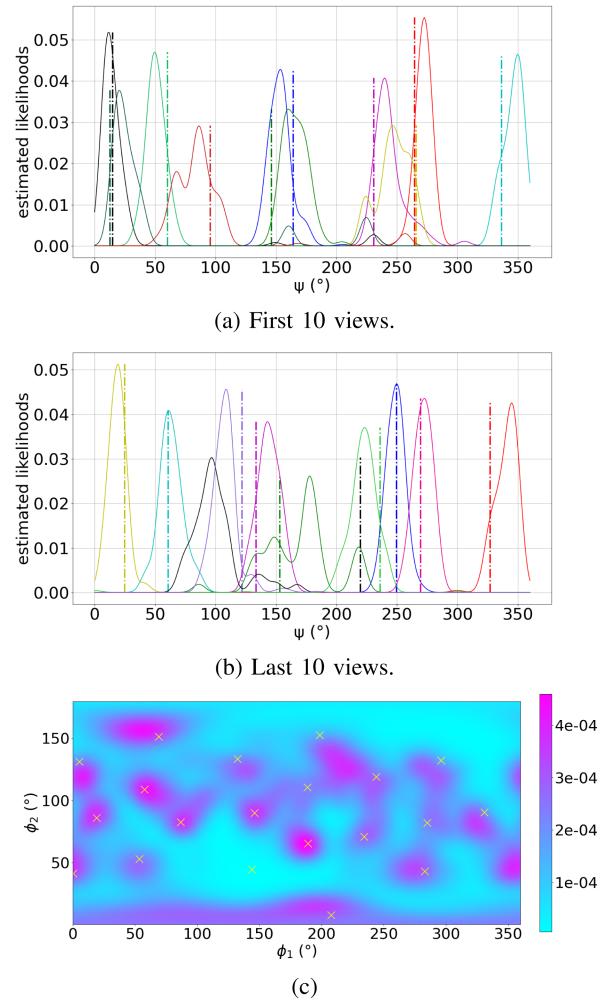


Fig. 3. Approximated likelihoods  $\tilde{\pi}^{l,\psi}$  and  $\tilde{\pi}^{l,d}$  at the last iteration of the reconstruction on data simulated from the AMPA receptors (see Fig. 2). (a) and (b): Likelihoods  $\tilde{\pi}^{l,\psi}$ , where each curve represents a view and the vertical bars represent the true values (the results are split in two plots for a better visualization). (c): Sum of the likelihoods  $\tilde{\pi}^{l,d}$  of all the views. The yellow crosses are the true values of  $d$ .

maintain a tractable computational cost. We selected 5 representative particles manually for each dataset. Finally, in order to evaluate the impact of orientation estimation errors, we show reconstructions results obtained with known poses, using the gradient-based reconstruction method presented in III-B.

2) *Estimation of the Orientations:* In order to evaluate the accuracy of the orientation estimation, we show the approximated marginal likelihoods  $\tilde{\pi}^{l,d}$  and  $\tilde{\pi}^{l,\psi}$  obtained in the last iteration of the reconstruction. We recall that these distributions are used to sample several axes  $d$  and angles  $\psi$ , from which the optimal angle axis-angle representation is estimated. The likelihoods  $\tilde{\pi}^{l,\psi}$  are shown on Fig. 3(a) and (b), where each colored curve represents the likelihood of one view. The vertical bars represent the ground truth values of  $\psi$  for each view. The sum of the likelihoods  $\tilde{\pi}^{l,d}$  of all the views is shown on Fig. 3(c) in a 2D map where each axis represents a dimension of  $d = (\phi_1, \phi_2)$ . The ground truth values of  $d$  are represented with yellow crosses.

TABLE I  
QUANTITATIVE COMPARISON OF THE RESULTS OF OUR METHOD, CRYO-RANSAC, MP3DR AND THE RECONSTRUCTION WITH KNOWN POSES

	Known poses		Our method		cryo-RANSAC		MP3DR	
	SSIM	FSC	SSIM	FSC	SSIM	FSC	SSIM	FSC
AMPA receptors	0.91	0.42	0.82	0.27	0.54	0.05	0.72	0.14
HIV Vaccine	0.93	0.44	0.83	0.28	0.53	0.06	0.72	0.13
clathrine	0.92	0.34	0.82	0.29	0.55	0.11	0.67	0.14
NLR resistosome	0.93	0.35	0.88	0.28	0.53	0.10	0.81	0.17

For both  $\psi$  and  $d$ , we observe that the likelihoods are approximately centered around the true values, which shows that the orientations are well estimated. Note that we aim at a coarse *ab initio* reconstruction that is destined to serve as an initialization for a refinement algorithm [22]. Only one estimated likelihood is not located close to the true value (the black curve on Fig. 3(b)).

3) *Quantitative Reconstruction Results*: Table I reports the SSIM and FSC values obtained on the simulated data with our algorithm, cryo-RANSAC, MP3DR, and the reconstruction with known poses. Since our method, as well as the generation of the views involves randomness, we perform 100 experiments for the 4 particles and report the average.

We found that for all the particles, our algorithm provides significantly better SSIM and FSC than cryo-RANSAC reconstructions. This result illustrates the importance to take into account an appropriate 3D imaging model, whereas working on 2D projections induces an irreversible and critical loss of information. We also consistently outperform MP3DR on the two considered metrics. Note that in addition to this quantitative improvement, one of the main advantages of our method compared to MP3DR is that we do not need to select a reduced number of views, which is subject to classification error or human bias, and we can use directly all the data for the reconstruction. As expected, our method provides slightly lower SSIM and FSC than the reconstruction with known poses. This results shows that the orientation estimation errors analyzed in Section IV-C2 only have a small impact on the final reconstruction.

To quantify the resolution isotropy of the reconstructions, we show in Fig. 4 the cFSC maps associated to a view (aligned with the ground truth image), our reconstruction and cryo-RANSAC reconstructions, for the particle AMPA receptors. We clearly see on Fig. 4(b) that the axial resolution has been improved. Indeed, on the cFSC map associated to the view Fig. 4(a), the cFSC values along the Z axis ( $\phi_2 = 180$  or  $\phi_2 = 0$ ) is almost zero, while it is between 0.35 and 0.45 for our reconstruction. Besides, the resolution along X and Y axis ( $\phi_2 = 90^\circ$ ) of our reconstruction is preserved since we observe similar values to that of the view (between 0.4 and 0.5). Regarding the cryo-EM dedicated algorithm, the resolution along Z axis is also improved, but the resolution along X and Y axis is degraded: the cFSC values are around 0.3, while they are around 0.4 for the view.

4) *Visual Reconstruction*: In Fig. 5, we present visual reconstructions obtained with the different methods. If we compare the intensity map in the plane XZ between a view and our reconstruction, we clearly see that the axial resolution has been enhanced. Visually, we do not see a clear difference between our method and the reconstruction obtained with known poses.

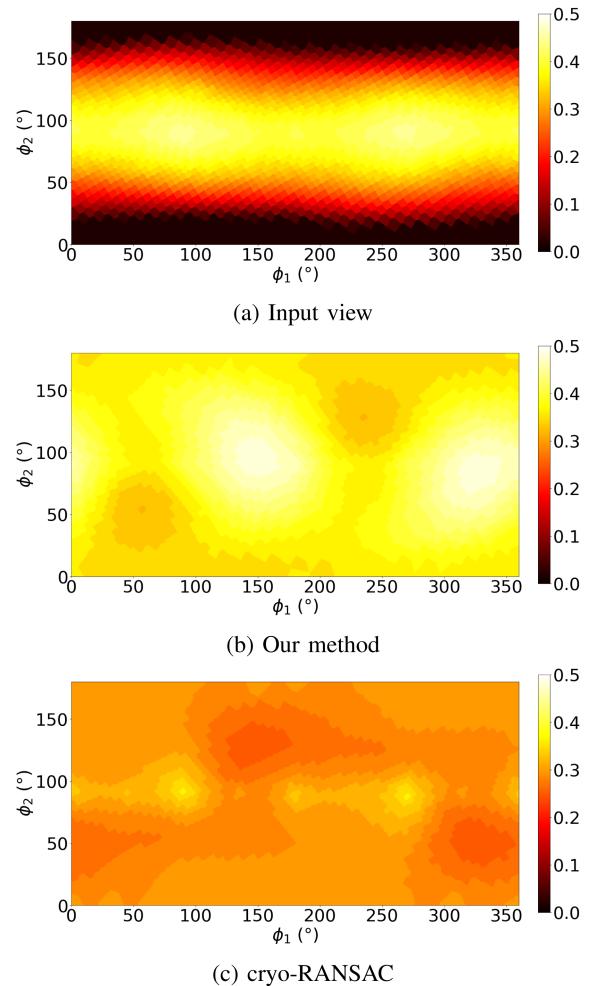


Fig. 4. cFSC maps between the ground truth object (here the AMPA receptor) and a view Fig. 4(a), our reconstruction Fig. 4(b) and cryo-RANSAC reconstruction Fig. 4(c). Each point on the map is a direction, represented by two angles  $\phi_1 \in [0, 360]$  and  $\phi_2 \in [0, 180]$  (see Fig. 1).

Moreover, our reconstructions are visually closer to the ground truth than those of cryo-RANSAC and MP3DR, which confirms the quantitative results of Section IV-C3.

5) *Robustness to Anisotropy*: In order to evaluate the robustness of our method to different levels of anisotropy, we performed reconstructions on data generated with different values of  $\sigma_z$ . The visual results are given on Fig. 6, where we also show the impact of the number of views. When the number of views increases, the reconstruction algorithm is able to combine more complementary information, which improves the quality of the reconstruction. As the anisotropy increases, a single view contains less information and the minimum number of views to

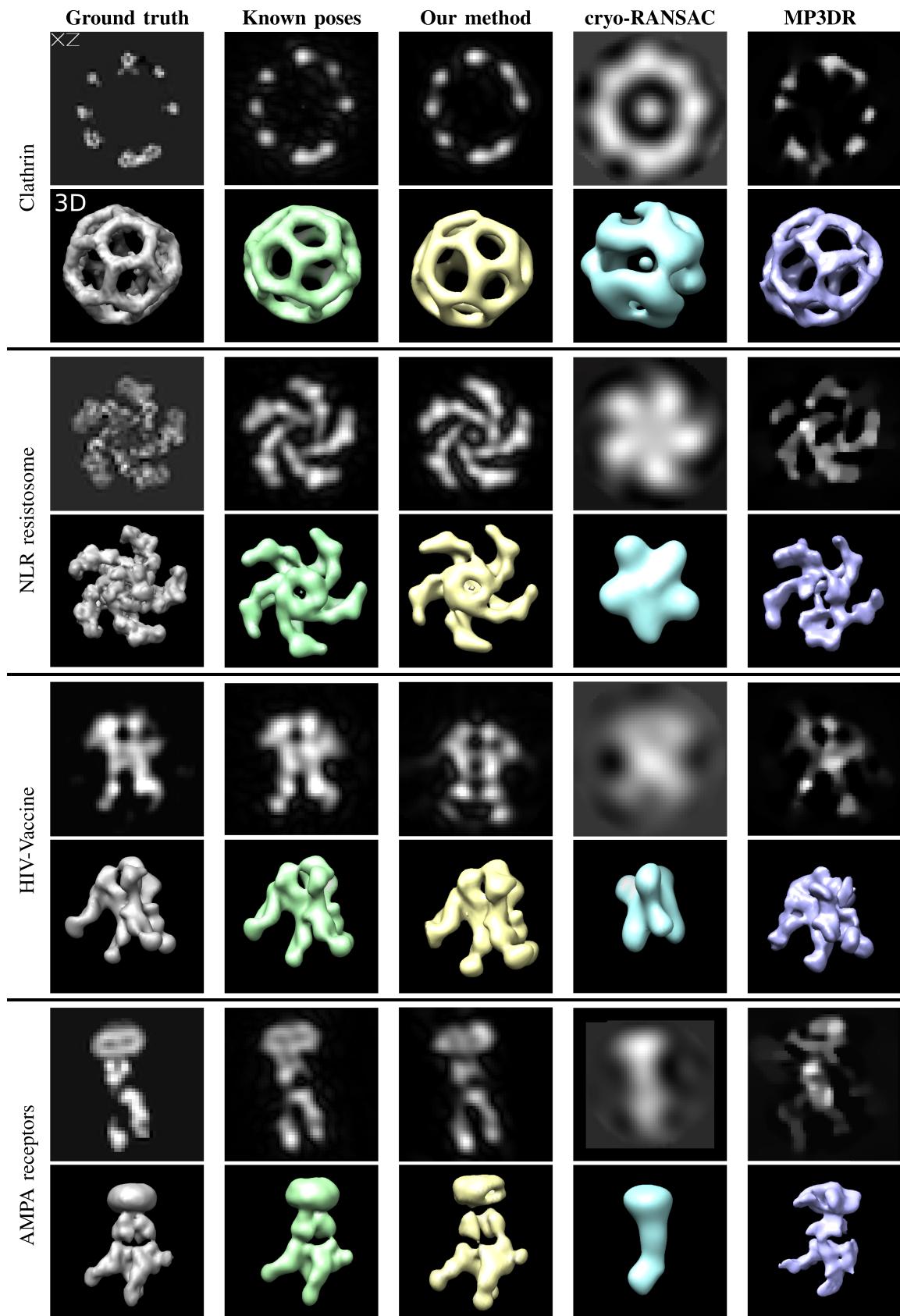


Fig. 5. Reconstruction results obtained on synthetic data. Each volume is represented with an orthogonal view in the XZ plane and a 3D isosurface. From left to right columns: ground truth object, reconstruction with known poses, reconstruction with our method, cryo-RANSAC reconstruction, MP3DR reconstruction. 20 views are used for the reconstruction.

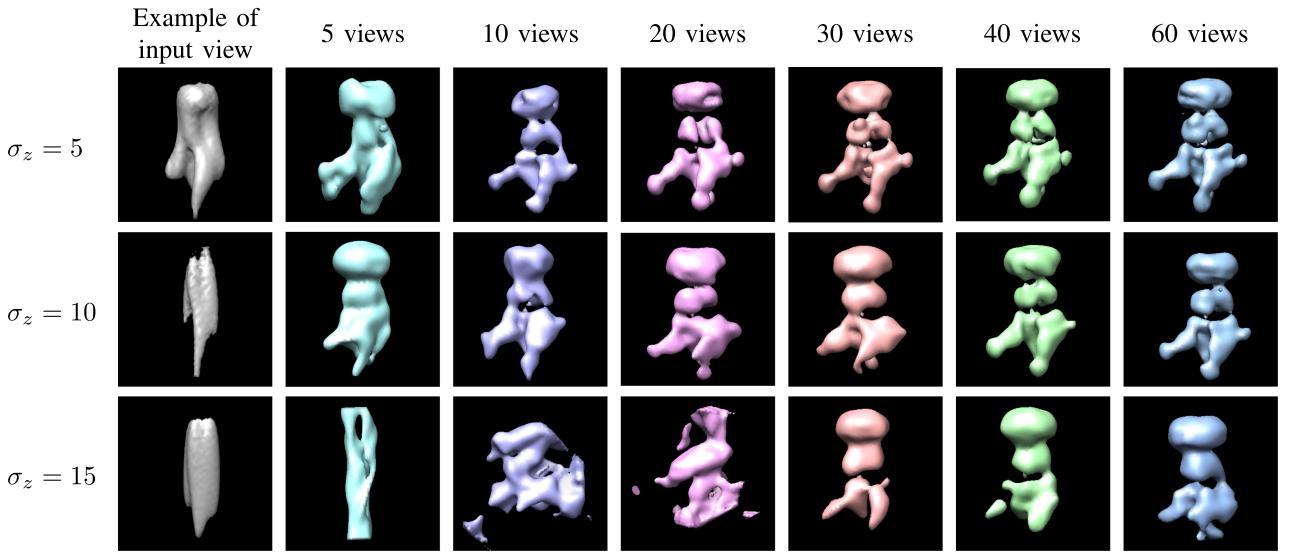


Fig. 6. Reconstructions obtained with different numbers of views and anisotropy levels  $\sigma_z$ .

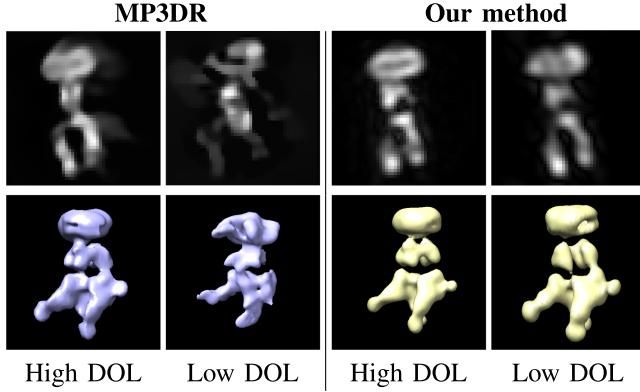


Fig. 7. Comparison of the reconstruction of our method and MP3DR in case of low and high DOL.

obtain a satisfying reconstruction also increases. For  $\sigma_z = 5$ , the minimum number of views is approximately 20, and is about 40 views for  $\sigma_z = 10$ , and for an extremely high anisotropy level of  $\sigma_z = 15$ , more than 60 views would be necessary to obtain a correct reconstruction.

*6) Robustness to Low DOL:* In this section, we focus on the comparison with MP3DR and evaluate the impact of the DOL. Let us recall that we simulate the variation of DOL along the structure by subtracting Gaussian spots at random locations. In Fig. 7, we show reconstructions obtained with our method and MP3DR in two different experimental conditions:

- Low DOL: Default simulation parameters used to simulate the data shown in Fig. 2. It corresponds to a realistic situation.
- High DOL: no Gaussian spot is removed from the ground truth image.

Our method is able to achieve satisfying reconstructions in both cases, whereas MP3DR fails in case of low DOL. This is

explained by the sequential nature of MP3DR: a first reconstruction with two views is performed, and the other views are then introduced one by one. The flaw of this procedure is that in the case of low DOL, the views are too different to achieve a correct pose estimation with only two of them, and it is necessary to combine the information of all the views jointly. Therefore, the initial two-view reconstruction is inaccurate and introduces a detrimental bias that is propagated when the other views are sequentially added. On the contrary, since our approach uses all the views jointly in an unordered way, we are able to use all the information and achieve good reconstruction.

*7) Sensitivity to Hyperparameters:* A low sensitivity to hyperparameters is crucial to avoid tedious and difficult fine tuning on each new dataset. In this section, we show the robustness of our method to the variation of the following hyperparameters:

- The ratio  $1/\alpha_r$  of the geometric progression at each epoch of the proportion  $\alpha$  of uniform component in the distributions  $\mathcal{Q}^{l,\psi}$  and  $\mathcal{Q}^{l,d}$  (22).
- The parameters  $\beta_d$  and  $\beta_\psi$  that determine the width of the kernels used for the estimation of the marginal likelihood distributions (see (20) and (21))
- The number of orientations sampled at each iteration  $N_d$  and  $N_\psi$ .

For each combination of hyperparameters, we performed 10 reconstructions, and report the average SSIM index. The default values are:  $N_\psi = 8$ ,  $N_d = 64$ ,  $\alpha_r = 1.2$ ,  $\beta_\psi = \beta_d = 50$ , and values are taken in a range around them.

Fig. 8 illustrates the influence of these hyperparameters on the SSIM between the reconstruction and the ground truth object (the considered particle is AMPA receptors). We observe in Figs. 8(a) and (b) that the choice of  $\beta_d$  and  $\beta_\psi$  has a low impact on the final reconstruction. In Fig. 8(c), we report the influence of  $\alpha_r$  on the SSIM. We observe that low values of  $\alpha_r$  (from 1.02 to 1.15) provide slightly lower SSIM, and that the SSIM is relatively stable from  $\alpha_r = 1.15$  to  $\alpha_r = 2$ . In Fig. 8(d), we report the influence of the number of sampled

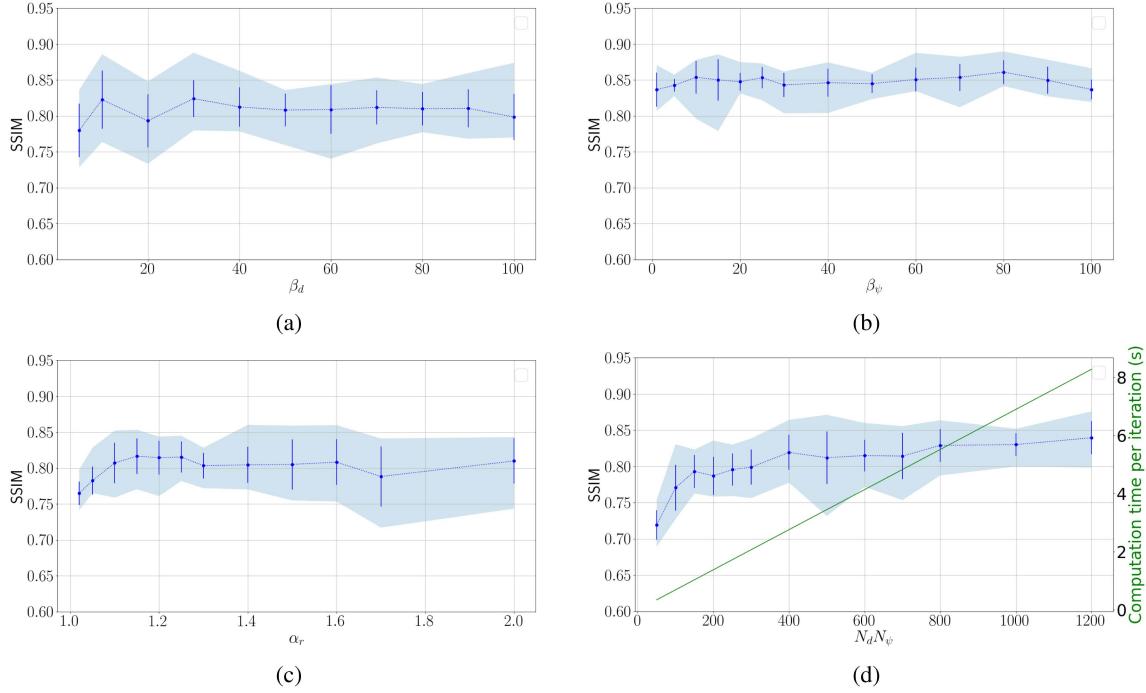


Fig. 8. Influence of the hyperparameters  $\beta_d$ ,  $\beta_\psi$ ,  $\alpha_r$  and  $N_d N_\psi$  on the SSIM and the computation time per iteration. Each point is an average performed on 10 reconstructions on the AMPA receptor. The vertical bars represent the standard deviation on the 10 experiments, and the blue area is the range between the minimum and maximum values.

orientations  $N_d N_\psi$  in terms of both SSIM and computation time per iteration (We set  $N_d = N_\psi^2$ ). We can see that the SSIM increases with the number of sampled orientations. Indeed, if  $N_d N_\psi$  is too small, the algorithm is likely not to explore the region of the search space which contains the optimal solution. On the other hand, the computation time is proportional to  $N_d N_\psi$  (see part IV-D). Then, choosing  $N_d$  and  $N_\psi$  is a compromise between the quality of the reconstruction and the computation time.

#### D. Computation Time

The computation time for one epoch depends on three parameters: the number of views  $N$ , the number of sampled orientations  $N_d N_\psi$  and the number of voxels  $N_v$  in the volume. Each epoch contains as many iterations as the number of views, thus the computation time per epoch is proportional to  $N$ . Most of the computation is devoted to evaluate the energies corresponding to the randomly sampled rotations (lines 9 to 11 of Algorithm 1). Thus, the computation time per iteration is proportional to  $N_d N_\psi$ . Let us give more detail:

- In line 9,  $\hat{f}$  is rotated with orientation  $\theta_{i,j}$ , then a term by term multiplication is performed. The computation time devoted to the multiplication and to a rotation are  $O(N_v)$ .
- In line 10, the phase correlation is computed between  $\hat{h}\mathcal{R}_{\theta_{i,j}}(\hat{f})$  and  $\hat{g}_l$  to estimate the translation vector  $t_l^*(\theta_{i,j})$ . The complexity is dominated by the Fourier transform and is  $O(N_v \log(N_v))$ .
- In line 11, the computation of the energy  $E_{i,j}^l$  given the rotated volume  $\mathcal{R}_{\theta_{i,j}}$  and the translation vector  $t_l^*(\theta_{i,j})$  is  $O(N_v)$ .

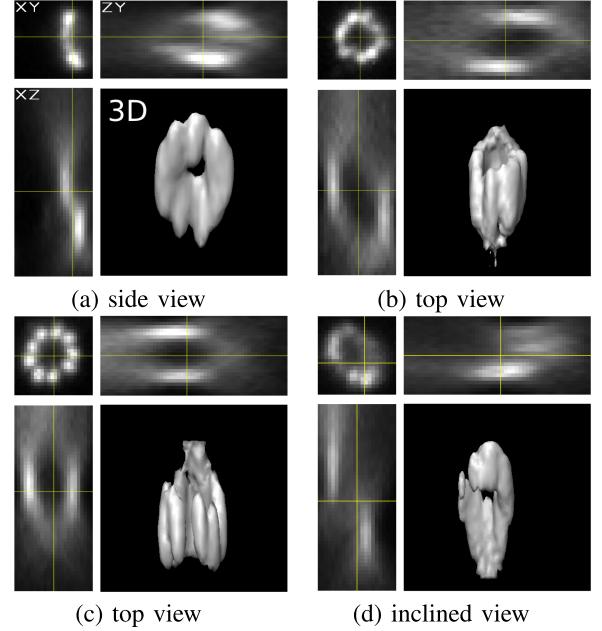


Fig. 9. Examples of views of our real data set.

To speed up the computation, we use a GPU implementation of the rotation and the phase correlation, available in CuPy<sup>3</sup> and cuCIM<sup>4</sup> packages, respectively. The computation time devoted to one epoch  $T_{ep}$  is  $\mathcal{O}(N N_v \log(N_v) N_d N_\psi)$ . With the default parameters  $N = 20$ ,  $N_d = 64$ ,  $N_\psi = 8$  and  $N_v = 50^3$ , a complete reconstruction takes approximately 24 minutes on a Titan

<sup>3</sup>[Online]. Available: <https://cupy.dev/>

<sup>4</sup>[Online]. Available: <https://pypi.org/project/cucim/>

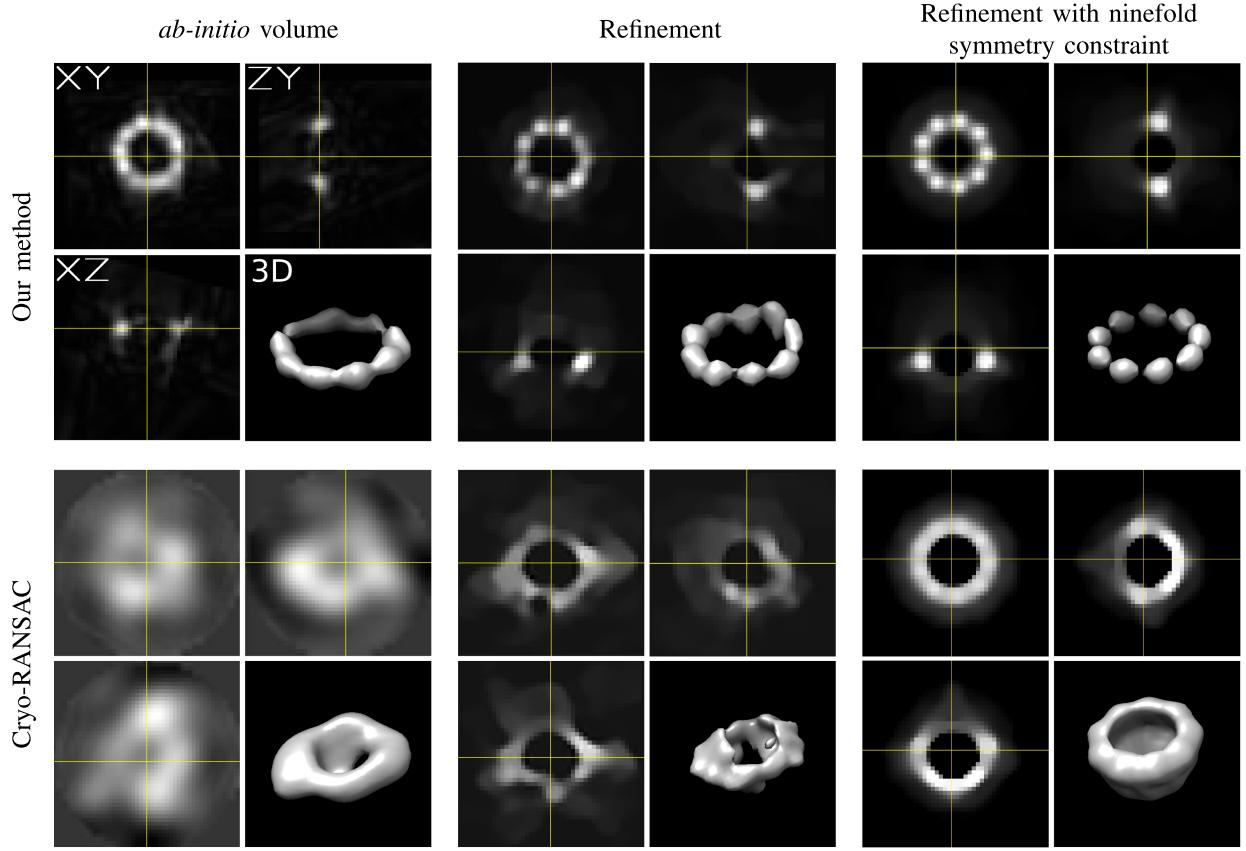


Fig. 10. Reconstruction results obtained on a real dataset of centriole, with our method and cryo-RANSAC.

Xp graphic card. Comparatively, MP3DR takes 90 minutes for a reconstruction with  $N = 5$  and the same  $N_v = 50^3$ .

### E. Experimental Results on Real Data

To test our method on real data, we imaged human centrioles from RPE1 cells, an organelle critical for cell division and templating the cilium (PMID: 19914163). As centriole dimensions, about 450 nm in length and 250 nm in diameter, are below the diffraction limit, we imaged them using ultrastructure expansion microscopy (U-ExM). U-ExM is a super-resolution approach that consists in embedding a biological sample into a swellable polymer that will expand 4 times in pure water, allowing to reach a resolution of 50–70 nm with a conventional widefield microscope [37]. After expansion, the biological sample is immuno-labelled with antibodies against Cep164 (2227-1-AP, Proteintech, 1:500), a protein that composes the distal appendages of the centriole [38], and therefore appears with 9 dots arranged in a ring in super-resolution microscopy [39]. Expanded gels were mounted onto 24 mm coverslips coated with poly-D-lysine (0.1 mg/ml) and imaged with an inverted widefield Leica DM18 microscope. Images were taken with a  $63 \times 1.4$  NA oil immersion objective. 3D stacks were acquired with  $0.21\mu\text{m}$  Z-intervals and a 100 nm XY pixel size. The PSF was experimentally measured with images of beads. We manually selected 47 views in the acquired volumes. The protein Cep164 of the centriole has a ring shape with a ninefold cylindrical

symmetry. In Fig. 9, we show 4 examples of views that include 2 top views of the ring, one side view and one inclined view. The resolution along the Z axis is clearly much lower than the resolution on the XY plane. Besides, the non-uniformity of the distribution of fluorophores is observable by comparing the two top views. The ninefold cylindrical symmetry is also clearly visible on view 9(c).

In Fig. 10, we show the reconstruction obtained with this dataset and the same hyperparameters as those presented in Section IV-C7. We compare our result with that of cryo-RANSAC. In order to get a high-accuracy result, the *ab initio* reconstructions are refined using an alternated optimization scheme described in [22]. In addition to the basic refinement, we also show reconstructions obtained with an additional nine-fold symmetry constraint. For the sake of visualization, the axis of symmetry of the reconstructed centrioles are aligned with the Z-axis in Fig 10. The axis of symmetry is determined by using a PCA on a binarized version of the reconstruction

The reconstructed shape found by our method is a ring that is consistent with the input views. Indeed, the diameter of the ring observable on the XY plane is similar to the diameter of the top views 9(b) and 9(c). Besides, the thickness of the ring, observable on the XZ and ZY cuts is similar to what we observe on the side view Fig. 9(a). The ninefold symmetry is fuzzy in the *ab-initio* reconstruction, but sufficient to allow for a clearly visible symmetry in the refined volume. If we impose the ninefold symmetry in the refinement, the ring has a perfectly

**Algorithm 1:** Reconstruction Algorithm.

---

```

1: Initialize  $\hat{f}$  randomly
2: Initialize probability distributions, with uniform
   distributions  $\mathcal{Q}^{l,\psi} = \mathcal{U}$  and  $\mathcal{Q}^{l,d} = \mathcal{U}$ 
3: while stopping criteria do
4:   Randomly mix the views
5:   for  $l = 1$  to  $N$  do
6:     Randomly draw indices  $\mathcal{I}^{l,\psi}$  and  $\mathcal{I}^{l,d}$  from  $\mathcal{Q}^{l,\psi}$ 
      and  $\mathcal{Q}^{l,d}$ 
7:     for  $(i, j) \in \mathcal{I}^{l,\psi} \times \mathcal{I}^{l,d}$  do
8:        $\theta_{i,j} = (d_j, \psi_i)$ 
9:       Compute  $\hat{h}\mathcal{R}_{\theta_{i,j}}(\hat{f})$ 
10:      Evaluate  $t_l^*(\theta_{i,j})$ , using phase correlation
        between  $\hat{h}\mathcal{R}_{\theta_{i,j}}(\hat{f})$  and  $\hat{y}_l$ .
11:      Compute  $E_{i,j}^l = \|\hat{h}\rho_{t_l^*(\theta_{i,j})}\mathcal{R}_{\theta_{i,j}}(\hat{f}) - \hat{y}_l\|_2^2$ 
12:    end for
13:     $\theta_l^* = (d_{j^*}, \psi_{i^*})$  with
       $i^*, j^* = \arg \min\{E_{i,j}^l \mid i, j \in \mathcal{I}^{l,\psi} \times \mathcal{I}^{l,d}\}$ 
14:     $\forall (i, j) \in \mathcal{I}^{l,\psi} \times \mathcal{I}^{l,d}$ , compute  $p_{i,j}^l = \exp(-E_{i,j}^l)$ .
15:    Compute marginalized likelihoods:
16:     $\forall i \in \mathcal{I}^{l,\psi} \quad \tilde{\pi}_i^{l,\psi} = \sum_{j \in \mathcal{I}^{l,d}} \frac{p_{i,j}^l}{\mathcal{Q}_i^{l,d}}$ 
17:     $\forall j \in \mathcal{I}^{l,d} \quad \tilde{\pi}_j^{l,d} = \sum_{i \in \mathcal{I}^{l,\psi}} \frac{p_{i,j}^l}{\mathcal{Q}_j^{l,d}}$ 
18:    Update probability distributions
19:     $\mathcal{Q}_j^{l,d} = \alpha \mathcal{U}_j^d + (1 - \alpha) Z_d^{-1} \sum_{k \in \mathcal{I}^{l,d}} K_{j,k}^d \tilde{\pi}_k^{l,d}$ 
20:     $\mathcal{Q}_i^{l,\psi} = \alpha \mathcal{U}_i^\psi + (1 - \alpha) Z_\psi^{-1} \sum_{k \in \mathcal{I}^{l,\psi}} K_{i,k}^\psi \tilde{\pi}_k^{l,\psi}$ 
21:     $\alpha = \alpha/\alpha_r$ 
22:     $\hat{f} = \hat{f} - \mu \nabla_{\hat{f}} E_l(f, \theta_l^*, t_l^*)$ 
23:  end for
24: end while
25:  $\hat{f}^* = \hat{f}$ 
26: return  $\mathcal{F}^{-1}(\hat{f}^*), (\theta_l^*)_{l \in [1, N]}, (t_l^*)_{l \in [1, N]}$ 

```

---

regular symmetry. The cryo-RANSAC reconstruction is not able to retrieve the same details, since the ring shape is not correctly reconstructed, and the ninefold symmetry is not visible.

## V. CONCLUSION

We proposed a reference-free SPR approach that jointly estimates the poses and the reconstruction, to address the limitations of resolution anisotropy and non-uniform DOL in fluorescence microscopy. Our approach is based on a reformulation of the original joint optimization problem in a multilevel scheme. We have developed fast approximate solvers at each level to end up with a computationally efficient method able to estimate the particle poses and the reconstruction. We have demonstrated experimentally that our method yields more accurate reconstructions than the standard methods, on both simulated and real data. In addition, our method is fast, it is robust to low DOL, it has a very low sensitivity to hyperparameters, and it does not require any particles classification step. These features are crucial for practical applications of the method on a variety of real data.

## APPENDIX

Here we derive the computation of the gradient of the energy  $E_l$  with respect to the volume  $\hat{f}$  used in the SGD update (12). We have

$$E_l(f, \theta_l, t_l) = \|\hat{y}_l - \hat{h}\rho_{t_l}\mathcal{R}_{\theta_l}(\hat{f})\|^2 \quad (28)$$

$$= \int [\hat{y}_l(\omega) - \hat{h}(\omega)\rho_{t_l}(\omega)\hat{f}(R_{\theta_l}^T\omega)]^2 d\omega. \quad (29)$$

Let us consider a particular frequency  $\omega_0$  and compute the derivative  $\frac{\partial E_l}{\partial \hat{f}}(\omega_0)$ . In the integral (29), only the term that verifies  $R_{\theta_l}^T\omega = \omega_0$  depends on  $\hat{f}(\omega_0)$ :

$$\frac{\partial \hat{f}(R_{\theta_l}^T\omega)}{\partial \hat{f}}(\omega_0) = \begin{cases} 1 & \text{if } R_{\theta_l}^T\omega = \omega_0 \\ 0 & \text{otherwise} \end{cases}. \quad (30)$$

Then we have

$$\frac{\partial E_l}{\partial \hat{f}}(\omega_0) = 2 \left[ \rho_{R_{\theta_l} t_l}(\omega_0) \hat{h}(R_{\theta_l} \omega_0) \hat{f}(\omega_0) - \hat{y}_l(R_{\theta_l} \omega_0) \right] \quad (31)$$

$$\begin{aligned} & \rho_{R_{\theta_l} t_l}(\omega_0) \hat{h}(R_{\theta_l} \omega_0) \\ &= 2 \left[ \rho_{R_{\theta_l} t_l}(\omega_0) \mathcal{R}_{-\theta_l}(\hat{h})(\omega_0) \hat{f}(\omega_0) - \mathcal{R}_{-\theta_l}(\hat{y}_l)(\omega_0) \right] \\ & \rho_{R_{\theta_l} t_l}(\omega_0) \mathcal{R}_{-\theta_l}(\hat{h})(\omega_0). \end{aligned} \quad (32)$$

Finally, the gradient can be written

$$\begin{aligned} & \nabla_{\hat{f}} E_l(f, \theta_l^*, t_l^*) \\ &= 2\rho_{R_{\theta_l} t_l} \mathcal{R}_{-\theta_l}(\hat{h}) \left[ \rho_{R_{\theta_l} t_l} \mathcal{R}_{-\theta_l}(\hat{h}) \hat{f} - \mathcal{R}_{-\theta_l}(\hat{y}_l) \right]. \end{aligned} \quad (33)$$

## ACKNOWLEDGMENT

The authors would like to acknowledge the High Performance Computing Center of the University of Strasbourg for supporting this work by providing scientific support and access to computing resources. Part of the computing resources were funded by the Equipex Equip@Meso project (Programme Investissements d’Avenir) and the CPER Alsacalcul/Big Data.

## REFERENCES

- [1] L. Schermelleh et al., “Super-resolution microscopy demystified,” *Nature Cell Biol.*, vol. 21, no. 1, pp. 72–84, 2019.
- [2] D. Gambarotto et al., “Imaging cellular ultrastructures using expansion microscopy (U-ExM),” *Nature Methods*, vol. 16, no. 1, pp. 71–74, 2019.
- [3] F. Chen, P. W. Tillberg, and E. S. Boyden, “Expansion microscopy,” *Science*, vol. 347, no. 6221, pp. 543–548, 2015.
- [4] D. Salas et al., “Angular reconstitution-based 3D reconstructions of nanomolecular structures from superresolution light-microscopy images,” *Proc. Nat. Acad. Sci.*, vol. 114, no. 35, pp. 9273–9278, 2017.
- [5] C. Sieben, N. Banerle, K. M. Douglass, P. Gönczy, and S. Manley, “Multicolor single-particle reconstruction of protein complexes,” *Nature Methods*, vol. 15, no. 10, pp. 777–780, 2018.
- [6] D. Mahecic et al., “Homogeneous multifocal excitation for high-throughput super-resolution imaging,” *Nature Methods*, vol. 17, no. 7, pp. 726–733, 2020.
- [7] H. Heydarian et al., “3D particle averaging and detection of macromolecular symmetry in localization microscopy,” *Nature Commun.*, vol. 12, no. 1, pp. 1–9, 2021.
- [8] A. Mendes, H. S. Heil, S. Coelho, C. Leterrier, and R. Henriques, “Mapping molecular complexes with super-resolution microscopy and single-particle analysis,” *Open Biol.*, vol. 12, no. 7, 2022, Art. no. 220079.

- [9] S. Liu, P. Hoess, and J. Ries, "Super-resolution microscopy for structural cell biology," *Annu. Rev. Biophys.*, vol. 51, pp. 301–326, 2022.
- [10] L. Schermelleh, R. Heintzmann, and H. Leonhardt, "A guide to super-resolution fluorescence microscopy," *J. Cell Biol.*, vol. 190, no. 2, pp. 165–175, 2010.
- [11] M. Temerinac-Ott, O. Ronneberger, P. Ochs, W. Driever, T. Brox, and H. Burkhardt, "Multiview deblurring for 3-D images from light-sheet-based fluorescence microscopy," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1863–1873, Apr. 2012.
- [12] S. Preibisch et al., "Efficient Bayesian-based multiview deconvolution," *Nature Methods*, vol. 11, no. 6, pp. 645–648, 2014.
- [13] J. Boulanger et al., "Fast high-resolution 3D total internal reflection fluorescence microscopy by incidence angle scanning and azimuthal averaging," *Proc. Nat. Acad. Sci.*, vol. 111, no. 48, pp. 17164–17169, 2014.
- [14] J. M. Carazo et al., "Three-dimensional reconstruction methods in single particle analysis from transmission electron microscopy data," *Arch. Biochem. Biophys.*, vol. 581, pp. 39–48, 2015.
- [15] A. Punjani, M. A. Brubaker, and D. J. Fleet, "Building proteins in a day: Efficient 3D molecular structure estimation with electron cryomicroscopy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 706–718, Apr. 2017.
- [16] A. Löschberger et al., "Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution," *J. Cell Sci.*, vol. 125, no. 3, pp. 570–575, 2012.
- [17] A. Szymborska, A. De Marco, N. Daigle, V. C. Cordes, J. A. Briggs, and J. Ellenberg, "Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging," *Science*, vol. 341, no. 6146, pp. 655–658, 2013.
- [18] B. Blundell, C. Sieben, S. Manley, E. Rosten, Q. Ch'ng, and S. Cox, "3D structure from 2D microscopy images using deep learning," *Front. Bioinf.*, vol. 1, 2021, Art. no. 740342.
- [19] V. Mennella et al., "Subdiffraction-resolution fluorescence microscopy reveals a domain of the centrosome critical for pericentriolar material organization," *Nature Cell Biol.*, vol. 14, no. 11, pp. 1159–1168, 2012.
- [20] J. Broeken et al., "Resolution improvement by 3D particle averaging in localization microscopy," *Methods Appl. Fluorescence*, vol. 3, no. 1, 2015, Art. no. 014003.
- [21] X. Shi, G. Garcia III, Y. Wang, J. F. Reiter, and B. Huang, "Deformed alignment of super-resolution images for semi-flexible structures," *PLoS One*, vol. 14, no. 3, 2019, Art. no. e0212735.
- [22] D. Fortun et al., "Reconstruction from multiple particles for 3D isotropic resolution in fluorescence microscopy," *IEEE Trans. Med. Imag.*, vol. 37, no. 5, pp. 1235–1246, May 2018.
- [23] H. Kirchner, F. Aguet, D. Sage, and M. Unser, "3-D PSF fitting for fluorescence microscopy: Implementation and localization application," *J. Microsc.*, vol. 249, no. 1, pp. 13–25, 2013.
- [24] R. Sato, M. Tanaka, and A. Takeda, "A gradient method for multilevel optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2021, pp. 7522–7533.
- [25] S. T. Tokdar and R. E. Kass, "Importance sampling: A review," *Wiley Interdiscipl. Rev.: Comput. Statist.*, vol. 2, no. 1, pp. 54–60, 2010.
- [26] J. R. Fienup, "Invariant error metrics for image reconstruction," *Appl. Opt.*, vol. 36, no. 32, pp. 8352–8357, 1997.
- [27] X. Tong et al., "Image registration with fourier-based image correlation: A comprehensive review of developments and applications," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 10, pp. 4062–4081, Oct. 2019.
- [28] Y. Zhao, S. Chen, A. C. Swensen, W.-J. Qian, and E. Gouaux, "Architecture and subunit arrangement of native AMPA receptors elucidated by cryo-EM," *Science*, vol. 364, no. 6438, pp. 355–362, 2019.
- [29] J. Wang et al., "Reconstitution and structure of a plant NLR resistosome conferring immunity," *Science*, vol. 364, no. 6435, 2019, Art. no. eaav5870.
- [30] K. L. Morris et al., "Cryo-EM of multiple cage architectures reveals a universal mode of clathrin self-assembly," *Nature Struct. Mol. Biol.*, vol. 26, no. 10, pp. 890–898, 2019.
- [31] K. M. Cirelli et al., "Slow delivery immunization enhances HIV neutralizing antibody and germinal center responses via modulation of immunodominance," *Cell*, vol. 177, no. 5, pp. 1153–1171, 2019.
- [32] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [33] M. Van Heel and M. Schatz, "Fourier shell correlation threshold criteria," *J. Struct. Biol.*, vol. 151, no. 3, pp. 250–262, 2005.
- [34] F. G. F. Christoph, A. Diebold, A. J. Koster, and R. I. Koning, "Conical fourier shell correlation applied to electron tomograms," *J. Struct. Biol.*, vol. 190, no. 2, pp. 215–223, 2015.
- [35] J. Vargas, A.-L. Alvarez-Cabrera, R. Marabini, J. M. Carazo, and C. O. S. Sorzano, "Efficient initial volume determination from electron microscopy images of single particles," *Bioinformatics*, vol. 30, no. 20, pp. 2891–2898, 2014.
- [36] J. M. D. la Rosa-Trevin et al., "Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy," *J. Struct. Biol.*, vol. 195, no. 1, pp. 93–99, 2016.
- [37] V. Hamel and P. Guichard, "Improving the resolution of fluorescence nanoscopy using post-expansion labeling microscopy," *Methods Cell Biol.*, vol. 161, pp. 297–315, 2021.
- [38] M. LeGuenne, N. Klena, G. Aeschlimann, V. Hamel, and P. Guichard, "Overview of the centriole architecture," *Curr. Opin. Struct. Biol.*, vol. 66, pp. 58–65, 2021.
- [39] L. Lau, Y. L. Lee, S. J. Sahl, T. Stearns, and W. Moerner, "Sted microscopy with optimized labeling density reveals 9-fold arrangement of a centriole protein," *Biophysical J.*, vol. 102, no. 12, pp. 2926–2935, 2012.