Notes on Single-Particle Reconstruction in Cryogenic Electron Microscopy

Hemant D. Tagare

Department of Diagnostic Radiology
Department of Biomedical Engineering
Yale University, New Haven, CT 06520.

Version: *01/24/2011*

1 Introduction

This note considers the problem of single-particle reconstruction from an estimation theoretic point of view. The note is meant for beginning graduate students in image processing who are interested in single particle reconstruction, but have little knowledge of molecular biology or experimental electron microscopy. In this note, biological and experimental aspects are discussed in a highly simplified manner, but the description is sufficient for understanding the mathematical formulation of the problem. Readers may consult Frank [1] for details.

Our ultimate goal is to formulate and solve SPR in a maximum-likelihood (ML) framework. Familiarity with basic estimation theory, at the level of Young and Smith [2] is useful.

2 A Guided Tour of Cryo-em and SPR

We begin with a brief guided tour to introduce the reader to the problem and its importance.

2.1 Macromolecules and their structure

A central fact of modern biology is that large polymeric molecules and their assemblies (proteins, DNA etc.) are vital to cell function. These marcomolecules not only make biochemical cell reactions possible, but they also maintain cell structure, cause cell motion, sense and respond to signals in the environment, etc.

Macromolecules are able to do all this because of their three-dimensional structure. Their structure allows macromolecules to dock with other molecules. Docking facilitates reactions between the docked molecules. In addition, structural flexibility allows motion, signal transduction, etc.

X-ray crystallography and Nuclear Magnetic Resonance are classical methods of reconstructing the three-dimensional structure of macromolecules. A more modern method is *cryogenic electron microscopy* (cryo-em), which is described shortly below. In cryo-em, random projections of the macromolecule are obtained and the three-dimensional structure of the macromolecule is reconstructed from the projections.

Biological macromolecules and their assemblies are generically referred to as *particles* in cryo-em. The reconstruction problem is called *single-particle reconstruction* (SPR).

The terms cryo-em and SPR are sometimes used interchangeably. To avoid confusion, we will adopt the following convention. The term cryo-em will refer to the experimental process of obtaining images, while SPR will refer to the algorithmic problem of reconstrucing the three-dimensional structure from the projections.

Summary: Understanding the three-dimensional structure of a macromolecule is key to understanding how it works. Cryo-em is a modern technique for reconstructing the three-dimensional structure of macromolecules.

2.2 Cryo-em

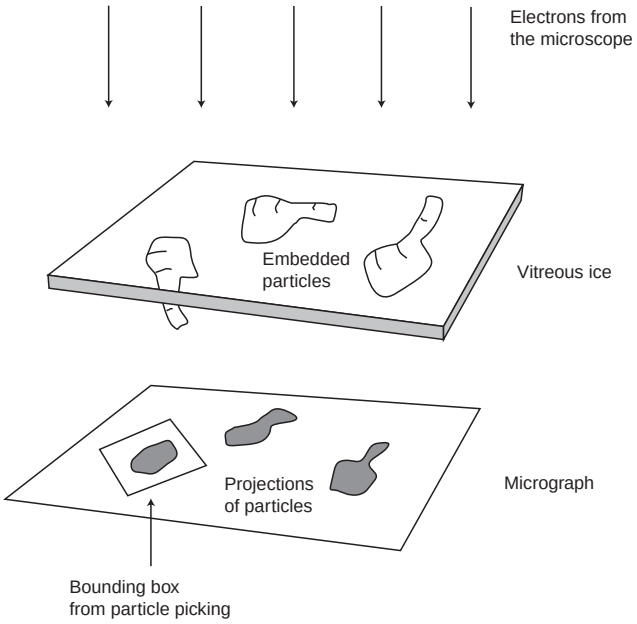

Figure 1: A simplified schematic of a cryo-EM experiment.

In very simple terms, the cryo-em approach is to freeze several identical copies of the particle in a layer of vitreous (non-crystallized) ice, and then to obtain a single transmission electron microscope image of the preparation (figure 1). This image, called a *micrograph*, contains projection images of the particle at random orientations – the orientation being determined by how the particle is frozen in ice. After a micrograph is obtained, each projection of the particle in the micrograph is manually isolated by a bounding box. This is called *particle picking*. We will call the content of an individual bounding box an *image*. The SPR problem is to estimate the 3D structure of the particle using all images obtained from one or more micrographs.

Because particles are embedded at random orientations, an image of the particle is essentially a projection of the particle along a random direction. This observation gives a "particle centric" description of cryo-em that is useful for mathematical formulation. It is illustrated in figure 2. In this description, the particle is in a fixed standard orientation. A randomly oriented orthonormal frame is positioned at the center of the particle.

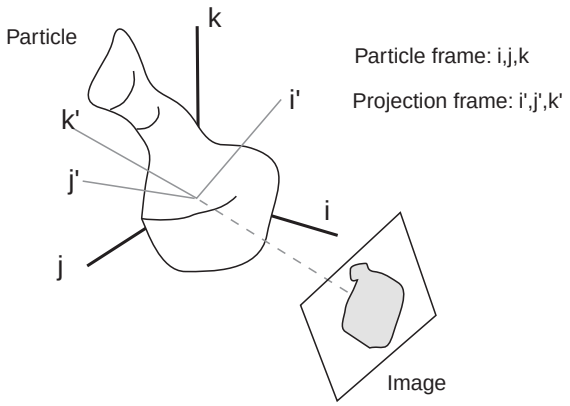

Figure 2: The equivalent cryo-EM experiment: The particle is fixed but projected along the z-axis of a random projection frame.

This is the *projection frame* in figure 2. The particle is projected along the z-axis of the projection frame to form an image. The measured image is this image with a random in-plane shift (because manual placement does not place the center of the bounding box exactly at the center of the projected particle), a random in-plane rotation (because the sides of the bounding box are not oriented along the x-y axis of the projection frame), plus additive noise.

The above description leaves out one important effect, which is due to the electron microscope. The image produced by the microscope is not simply a projection of the particle. It is more accurately modeled as a projection followed by convolution with a filter. The spectral response of the filter is called the *Contrast Transfer Function* (CTF). Figure 3 shows a typical CTF in the Fourier domain. The CTF is real valued, spherically symmetric, and takes positive and negative values.

The CTF has zeros in the Fourier domain and information about the particle is lost at these frequencies. However, the CTF (and its zeros) can be adjusted by changing the *defocus value* of the microscope. To take advantage of this, micrographs of the same particle are obtained at different CTFs whose zeros do not coincide. Thus, some information is available at every frequency.

Figure 4 shows the image formation model taking the CTF and additive noise into account.

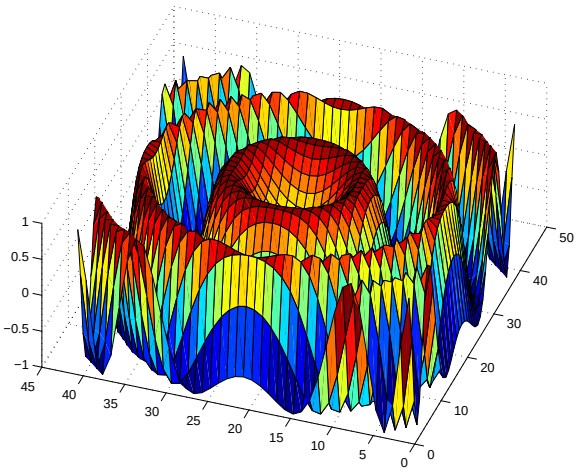

Figure 3: Spectral response of the CTF filter

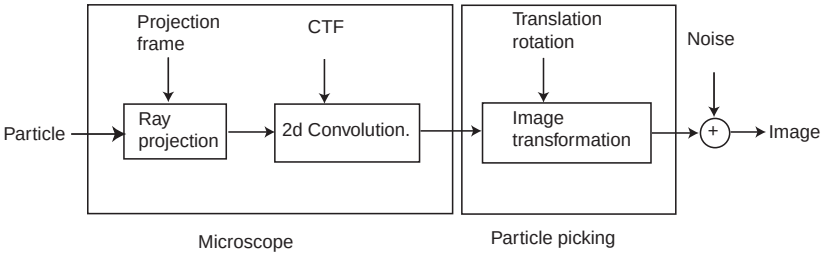

Figure 4: Signal Flow for Cryo-EM.

Summary: In cryo-em, many copies of same particle are frozen in vitreous ice and micrographs are obtained at different CTFs by changing the defocus value of the microscope. Images of individual particles are isolated by particle picking. Each image is a random projection of the particle, filtered by the CTF, further shifted and rotated, with additive noise. Each image has its own projection frame (unknown), CTF (known), image shift and rotation (unknown).

The goal of SPR is to reconstruct the three-dimensional structure of the particle from the images.

2.3 The maximum-likelihood approach to SPR

How should the structure of the particle be reconstructed from its noisy random projections? This problem belongs a category of problems called *estimation problems*. One approach to solving estimation problems is the *maximum-likelihood* approach. To refresh the reader's mind, this section contains a very brief qualitative description of maximum likelihood.

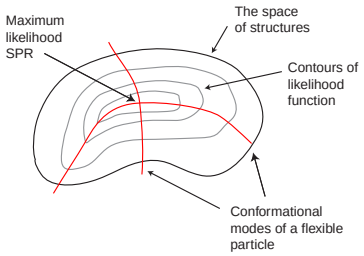

Figure 5: Illustration of the maximum-likelihood approach.

Suppose we have a set (a mathematical space) of possible structures for the particle (figure 5). Each point (element) of this space is a specific rigid structure. Further suppose that we have a set of images from a cryo-em experiment. Using basic probability theory and the image generation model of figure 4, we can calculate the probability that any structure in the space of structures produced the observed experimental images. This calculation reveals a function called the *likelihood function* on the space of structures – by definition the value of the likelihood function at any point in the structure space (i.e. at any structure) is the probability that this structure produced the observed experimental images. Figure 5 illustrates hypothetical contours of the likelihood function for observed images from a cryo-em experiment. The contours are the gray curves in the figure.

From the above discussion, it should be clear that for rigid SPR, the structure with the highest value of the likelihood function is our best estimate of the real structure,

since it is structure that most likely produced the observed images. Finding this maximizing structure is *maximum likelihood* SPR. Using maximum-likelihood involves two steps: first, using the image formation model and probability theory to formulate the likelihood function, and second, finding a procedure for maximizing the likelihood function.

There is one technical point that needs to be mentioned before proceeding. For several reasons, it is usually simpler to maximize the logarithm of the likelihood function rather to than directly maximizing the likelihood function. This does not really change anything - because logarithm is a monotonic function, the structure that maximizes the log-likelihood function also maximizes the likelihood function.

2.4 Evaluation

Once a structure is obtained from SPR, we have to evaluate its quality. A commonly used measure is the Fourier Shell Correlation (FSC). FSC is explained in further detail in section 6.

3 Mathematical Formulation of SPR

3.1 Notation and Conventions

We now fix some mathematical notation and conventions. We will use them to derive the equation for the image formation process. The notation might seem a little abstract; the point of the notation is to avoid complicated-looking equations.

3.1.1 Spaces

The standard 2D and 3D spaces are $\mathcal{R}^2$ and $\mathcal{R}^3$. Points in these spaces are denoted bold face, $\mathbf{u}, \mathbf{v}$ etc. When we need coordinates, we will write $\mathbf{u} = (u_x, u_y)$ in 2D or $\mathbf{u} = (u_x, u_y, u_z)$ in 3D.

3.1.2 Coordinate vectors

Unit coordinate vectors in $\mathcal{R}^2$ and $\mathcal{R}^3$ are $\mathbf{i},\ \mathbf{j}$ and $\mathbf{i},\ \mathbf{j},\ \mathbf{k}$ respectively.

3.1.3 Composition of functions

The composition of $f : \mathcal{R}^n \rightarrow \mathcal{R}^m$ with $g : \mathcal{R}^m \rightarrow \mathcal{R}^p$ is usually written $g \circ f$. For simplicity we will write it as gf. Potentially this notation can be confusing since it looks like the product of two functions. We will not need the product of two functions often below, and when we do, we will explicitly write it with a "." e.g. $h.g$.

3.1.4 Particle Structure

By a *particle structure* or simply a *structure* we mean a function $S : \mathcal{R}^3 \to \mathcal{R}$. Its value $S(\mathbf{u})$ gives the electron scattering density of the physical particle at the point $\mathbf{u} \in \mathcal{R}^3$. By "reconstructing the particle structure" we mean estimating S from images.

3.1.5 Image

An *image* is a function from $\mathcal{R}^2$ to $\mathcal{R}$. The value of an image I at a point $\mathbf{u} \in \mathcal{R}^2$ is $I(\mathbf{u})$.

3.1.6 Ray projection

A set of unit orthonormal vectors $P = \{\mathbf{i}', \mathbf{j}', \mathbf{k}'\}$ in 3D forms the *projection frame*. The *ray projection operator*, denoted Σ_P, attached to the projection frame P, operates on a structure S to give an image whose x, y coordinates are aligned along $\mathbf{i}', \mathbf{j}'$ and whose value $I(\mathbf{v})$ at the point $\mathbf{v} \in \mathcal{R}^2$ is given by $(\Sigma_P S)(\mathbf{v}) = \int S(v_x\mathbf{i}' + v_y\mathbf{j}' + z\mathbf{k}')dz$.

3.1.7 CTF

The *CTF operator* C applies the contrast transfer function to an image I, i.e it convolves the image I with a circularly symmetric kernel.

3.1.8 Translation and rotation of points and images

The 2D translation operator $t_{\mathbf{v}} : \mathcal{R}^2 \to \mathcal{R}^2$ translates any point $\mathbf{u} \in \mathcal{R}^2$ to $\mathbf{u} + \mathbf{v}$. Similarly the 2D rotation operator $r_\theta : \mathcal{R}^2 \to \mathcal{R}^2$ rotates every point $\mathbf{u} \in \mathcal{R}^2$ by θ in the counter-clockwise direction around the origin.

As defined above, rotation and translation operate on points of $\mathcal{R}^2$. There is a natural extension by which rotation and translation act on functions on $\mathcal{R}^2$. Let $t_{\mathbf{v}}$ be a translation operator and let $I : \mathcal{R}^2 \to \mathcal{R}$ be an image. Then $It_{\mathbf{v}}^{-1}$ is the image I "translated" in the same way as $t_{\mathbf{v}}$ translates points of $\mathcal{R}^2$. In cryo-em it is conventional to talk of "translating the image". To mimic this, we introduce the dual operator $t_{\mathbf{v}}^*$ which acts on I by

$$t_{\mathbf{v}}^* I = It_{\mathbf{v}}^{-1}.$$

$t_{\mathbf{v}}^*$ translates I in the same way that $t_{\mathbf{v}}$ translates points of $\mathcal{R}^2$.

Similarly, r_θ^* is the "dual" operator of rotation defined by

$$r_\theta^* I = Ir_\theta^{-1}.$$

r_θ^* rotates images in the same way the r_θ rotates points of $\mathcal{R}^2$.

Below we will need to jointly apply rotation and the translation operators to an image, and we use the simplified notation $T_{\theta,\mathbf{v}}^* = r_\theta^* t_{\mathbf{v}}^*$ for it. Note that $T_{\theta,\mathbf{v}}^{*-1} = t_{-\mathbf{v}}^* r_{-\theta}^*$.

3.2 Image Formation

As described in section 2 (figure 4) an individual cryo-em image is formed by projecting the structure with an arbitrary projection frame, convolving the resulting image with a CTF, transforming the convolved image by an arbitrary translation and rotation to mimic the manual placement of the bounding box, and adding noise. That is,

$$I = T^*_{\theta,\mathbf{v}} C \Sigma_P S + e, \tag{1}$$

where e is zero mean additive Gaussian noise with a known standard deviation.

In an experiment, N images are obtained $I_n,\ n = 1, \ldots, N$. Since the projection frame, CTF etc. can vary from image to image

$$I_n = T^*_{\theta_n,\mathbf{v}_n} C_n \Sigma_{P_n} S + e. \tag{2}$$

To proceed further, assume that the structure and the image are discretized to $V \times V \times V$ voxels and $P \times P$ pixels respectively, and that all operators are replaced by their discrete versions. Then, the likelihood of observing the set of images $\{I_n\}$ is

$$\begin{aligned} & p(\{I_n\} \mid S, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}) \\ = \quad & \frac{1}{(2\pi)^{NP^2/2}\sigma^{NP^2}} \exp\left\{ \sum_n -\frac{\|I_n - T^*_{\theta_n,\mathbf{v}_n} C_n \Sigma_{P_n} S\|^2}{2\sigma^2} \right\}, \end{aligned} \tag{3}$$

where σ is the pixel-wise standard deviation of noise. We assume that σ is known (using regions of the micrograph that do not have particles).

The log-likelihood is

$$\log p(\{I_n\} \mid S, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}) \quad = \quad \sum_n -\frac{\|I_n - T^*_{\theta_n,\mathbf{v}_n} C_n \Sigma_{P_n} S\|^2}{2\sigma^2}, \tag{4}$$

where terms without the conditioning variables are dropped.

4 Maximum-likelihood with Auxiliary Parameters

We are interested in the structure S, but the image likelihood also depends on $\{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}$. These parameters are *auxiliary parameters* for the problem and should be marginalized out for an exact maximum-likelihood reconstruction of the problem. The marginalized log-likelihood does not have a simple closed form solution and the EM algorithm is required to maximize the marginalized log-likelihood with respect to S. We do not pursue this strategy in this note (we have pursued this strategy elsewhere [3]). Instead we formulate the problem as a joint maximum-likelihood estimation of the structure and the auxiliary parameters:

$$\hat{S}, \hat{P_n}, \{\hat{\mathbf{v}_n}\}, \{\hat{\theta_n}\} \quad = \quad \arg \max_{S, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}} \log p(\{I_n\} \mid S, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}) \tag{5}$$

4.1 The objective function J

Maximizing the log-likelihood of equation (4) is equivalent to minimizing a scaled negative version of it. One scaled negative version is the objective function $J(S, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\})$ defined by

$$J(S, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}) = \sum_n \|I_n - T^*_{\theta_n, \mathbf{v}_n} C_n \Sigma_{P_n} S\|^2. \tag{6}$$

This objective function is minimized to obtain estimates of the structure and auxiliary variables.

Before proceeding further, it is useful to rewrite the objective function as follows:

$$\begin{aligned} J(S, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}) &= \sum_n J_n(S, P_n, \mathbf{v}_n, \theta_n), \text{ ,where} \\ J_n(S, P_n, \mathbf{v}_n, \theta_n) &= \|I_n - T^*_{\theta_n, \mathbf{v}_n} C_n \Sigma_{P_n} S\|^2. \end{aligned} \tag{7}$$

Finally, since $T^*_{\theta_n, \mathbf{v}_n}$ and its inverse preserve norm,

$$\begin{aligned} J_n(S, P_n, \mathbf{v}_n, \theta_n) &= \|I_n - T^*_{\theta_n, \mathbf{v}_n} C_n \Sigma_{P_n} S\|^2 \\ &= \|T^{*-1}_{\theta_n, \mathbf{v}_n} (I_n - T^*_{\theta_n, \mathbf{v}_n} C_n \Sigma_{P_n} S)\|^2 \\ &= \|T^{*-1}_{\theta_n, \mathbf{v}_n} I_n - C_n \Sigma_{P_n} S\|^2. \end{aligned} \tag{8}$$

5 The SPR Algorithm

The minimization of J with respect to the structure and the auxiliary variables can be carried out iteratively as described immediately below. It turns out that SPR with different CTFs (unequal C_ns) is a little harder to describe than SPR with equal CTFs. For now, we will assume equal CTFS and replace all C_n with C. We will return to the unequal CTF case in section 5.5.

The minimization algorithm is:

Algorithm A1:

1. *Initialize*: Set the iteration counter $r = 1$, initialize $S^r, \{P_n\}^r, \{\mathbf{v}_n\}^r, \{\theta_n\}^r$.
2. *Minimize w.r.t. auxiliary parameters*:

$$\{P_n\}^{r+1}, \{\mathbf{v}_n\}^{r+1}, \{\theta_n\}^{r+1} = \arg \min_{\{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}} J(S^r, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}). \tag{9}$$

3. *Minimize w.r.t. S*:

$$S^{r+1} = \arg \min_S J(S, \{P_n\}^{r+1}, \{\mathbf{v}_n\}^{r+1}, \{\theta_n\}^{r+1}). \tag{10}$$

4. *Iterate*: If the structure S has not converged, set $r = r + 1$ and go to step 2.

The details of the minimizations in steps 2 and 3 are given immediately below.

5.1 Minimization w.r.t. auxiliary parameters

The minimization in step 2 is over all auxiliary parameters, with one set of auxiliary parameters $(P_n, \mathbf{v}_n, \theta_n)$ per image. But, as equations (7) and (8) show, the objective function J is separable with respect to these sets of parameters. That is, the minimization with respect to the auxiliary variables for any image can be carried out independent of minimization with respect auxiliary variables for all other images. Hence, for $n = 1, 2, \ldots, N$ independently:

$$P_n^{r+1}, \mathbf{v}_n^{r+1}, \theta_n^{r+1} \quad = \quad \arg \min_{P_n, \mathbf{v}_n, \theta_n} J_n(S^r, P_n, \mathbf{v}_n, \theta_n). \tag{11}$$

The objective functions J_n are highly non-convex in auxiliary parameters and the minimization in equation (11) is non-trivial to carry out. In SPR, one popular strategy for this minimization is to further descretize the problem and then to minimize by exhaustive search.

A simple, general description of this idea is given immediately below. Its application to the minimization in equations (11) is explained after the general description.

Discrete exhaustive search: In general, suppose we want to minimize a function f over a set X. Assume that the set X has infinite elements and that the function f is non-convex and complicated. We choose a finite subset $\tilde{X} \subset X$ and find by exhaustive search the $x \in \tilde{X}$ that has the lowest value of f. We take this x to be an approximation of the minimizer of f. When X is a compact subset of $\mathcal{R}^n$, then $\tilde{X}$ is taken to be the set of vertices of a finite grid in X (figure 6). If the grid spacing is fine enough, then the result of exhaustive search is likely to be a good approximation of the minimizing x.

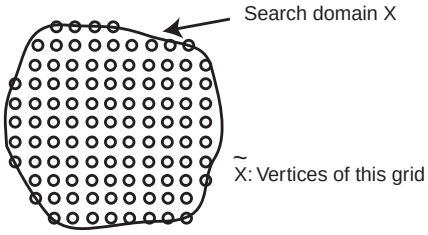

Figure 6: Exhaustive search.

Turning back to the minimization in equation (11), note that the domain of auxiliary variables $\mathbf{v}_n, \theta_n$ is the set of all translations of the image that account for the mis-centering of the bounding box $\times$ the set of all rotations of the image that account for the mis-orientation of the bounding box. Consider these one at a time. Although the bounding box center is not exactly at the center of the projected particle, it is quite close to it. Thus the domain of translations of any image can be considered to be a only a few pixels in the x and y directions. The set of all rotations cannot be similarly

restricted since we have no prior knowledge of the orientation of the projection frame. Thus the domain of rotations is $[0, 2\pi)$. We can discretize the domain of translation and rotation by choosing a rectangular grid for translations (say $\Delta_x, \Delta_y \simeq 1$ pixel), and a rotational grid (say with $\Delta_\theta \simeq 1$ degree). Let $\tilde{D}_T$ be resulting discrete set of translations $\times$ rotations.

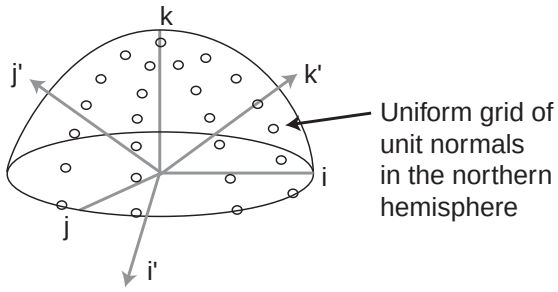

Figure 7: Sampling projection frames.

The domain of the projection frame P_n is the set of unit normals of the north hemisphere (figure 7) for the $\mathbf{k}'$ axis $\times$ the set of rotations of $\mathbf{i}', \mathbf{j}'$ axis around $\mathbf{k}'$. This set can be simplified. Since we are accounting for the $\mathbf{i}', \mathbf{j}'$ rotation in the image rotation, we can simply take the domain of the projection frame to be the set of unit normals of the north hemisphere with a single fixed rotation of the $\mathbf{i}', \mathbf{j}'$ axis. Let this set of discrete projection frames be $\tilde{D}_P$.

The minimization in equation (11) can be carried with the exhaustive search

Exhaustive search:
For $n = 1, 2, \ldots, N$ (i.e for all images),
  For all $P_n \in \tilde{D}_P$, (i.e. for all projection frames),
    For all $\mathbf{v_n}, \theta_n \in \tilde{D}_T$, (i.e. for all image translations and rotations),
      Calculate $\|T^{*-1}_{\theta_n, \mathbf{v}_n} I_n - C\Sigma_{P_n} S\|^2$ and return the minimizing $P_n, \mathbf{v_n}, \theta_n$.

Exhaustive search can be further sped up by noticing that the projection images generated by the $C\Sigma_{P_n} S$ term are identical for all n. These images need be generated only once. The exhaustive search algorithm modified to take this into account is:

Exhaustive search with precomputatiom:
For all $P \in \tilde{D}_P$,
  Generate the projection image $\pi_P = C\Sigma_P S$.
  Let $\Pi = \{\pi_P\}$ be the set of these generated projection images.
  For $n = 1, 2, \ldots, N$
    For all $\mathbf{v_n}, \theta_n \in \tilde{D}_T$ and all $\pi_P \in \Pi$,
      Calculate $\|T^{*-1}_{\theta_n, \mathbf{v}_n} I_n - \pi_P\|^2$ and return the minimizing P as P_n (identified from

the minimizing π_P), and the minimizing $\mathbf{v_n}, \theta_n$.

In plain words, this algorithm corresponds to first calculating the projected images π using the projection frames in $\tilde{D}_P$. Next the n^{th} data image is compared with each projection using all transformations in $\tilde{D}_T$, and the projection+transformation pair which gives the least mismatch is retained. Figure 8 illustrates this.

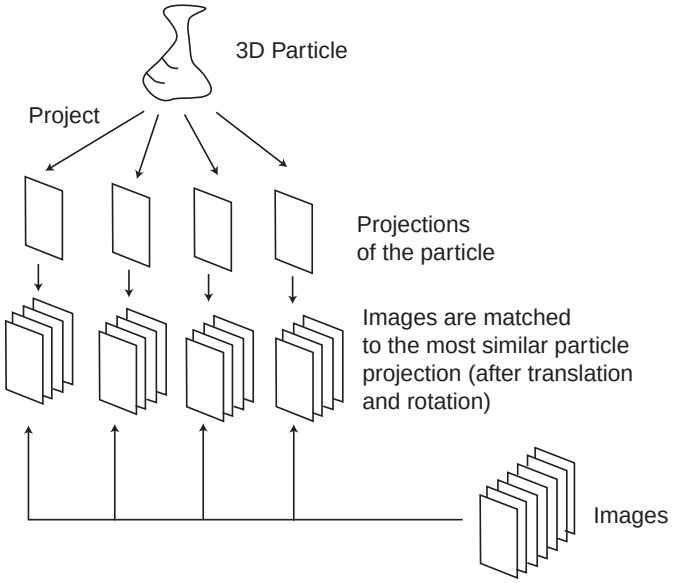

Figure 8: Exhaustive search with precomputation.

5.2 Minimization w.r.t. S

The objective function J of equation (7) is quadratic in S (for fixed values of auxiliary parameters, the ray projection, CTF, image translation and rotation are linear operators). In principle, a closed form solution is available for the minimizing S, but the matrices involved are very large and numerically impossible to work with. Iterative numerical minimization is a natural alternative. The conjugate gradient method is especially attractive because the objective function is quadratic in S. In practice, we find that a few iterations of conjugate gradient are sufficient to find a very good approximation to the minimizing S.

Just as with the minimization with respect to auxiliary variables, it is also possible to speedup the minimization with respect to S. To understand how, notice that at the end of the auxiliary variable minimization step, for each image we have a projection frame P_n^{r+1} and a pair of transformation parameters $\mathbf{v}_n^{r+1}, \theta_n^{r+1}$. Looking at this slightly differently, with every projection frame $P \in \tilde{D}_P$ we can associate a set of

images for which $P_n^{r+1} = P$ (the set of images whose best match projection frame is P). Let $\mathcal{I}_P$ be this set. It should be clear that $\mathcal{I}_{P_1} \neq \mathcal{I}_{P_2}$ for $P_1 \neq P_2$ and that $\cup_P \mathcal{I}_P$ is the set of all images (since every image matches to some P). Thus,

$$\begin{aligned} J(S, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}) &= \sum_n \|T^{*-1}_{\theta_n, \mathbf{v}_n} I_n - C\Sigma_{P_n} S\|^2 \\ &= \sum_{P \in \tilde{D}_P} \sum_{I_n \in \mathcal{I}_P} \|T^{*-1}_{\theta_n, \mathbf{v}_n} I_n - C\Sigma_P S\|^2 \\ &= \sum_{P \in \tilde{D}_P} \left\{ \sum_{I_n \in \mathcal{I}_P} \|T^{*-1}_{\theta_n, \mathbf{v}_n} I_n - \mu_P\|^2 + \|\mu_P - C\Sigma_P S\|^2 \right\}, \\ &= \sum_{P \in \tilde{D}_P} \sum_{I_n \in \mathcal{I}_P} \|T^{*-1}_{\theta_n, \mathbf{v}_n} I_n - \mu_P\|^2 + \sum_{P \in \tilde{D}_P} \|\mathcal{I}_P\| \times \|\mu_P - C\Sigma_P S\|^2, \end{aligned} \tag{12}$$

where,

$$\mu_P = \frac{1}{\|\mathcal{I}_P\|} \sum_{n \text{ such that } I_n \in \mathcal{I}_P} T^{*-1}_{\theta_n, \mathbf{v}_n} I_n$$

is the mean of all images that "match" to the projection frame P, and $\|\mathcal{I}_P\|$ is the number of images in $\mathcal{I}_P$. In SPR, the image μ_P is called the *class mean*.

The first term on the right hand side of equation (12) can be dropped since it is independent of S. Thus, the following form of the objective function can be used in the minimization w.r.t. S:

$$J(S, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}) = \sum_{P \in \tilde{D}_P} \|\mathcal{I}_P\| \times \|\mu_P - C\Sigma_P S\|^2. \tag{13}$$

5.3 SPR for single CTF

Putting all this together, the SPR reconstruction algorithm for a single CTF is:

Algorithm A2:

1. *Initialize*: Set the iteration counter $r = 1$, initialize $S^r, \{P_n\}^r, \{\mathbf{v}_n\}^r, \{\theta_n\}^r$.
2. *Minimize w.r.t. auxiliary parameters* (Project and align):
   (a) Generate all projection images $\pi_P = C\Sigma_P S^r$, for $P \in \tilde{D}_P$.
   (b) For each data image, use exhaustive search to find the projection image and transformation parameters that minimize $\|T^{*-1}_{\theta_n, \mathbf{v}_n} I_n - \pi_P\|^2$. Set P_n^{r+1} to the projection frame of the minimizing projection image and $\mathbf{v}_n^{r+1}, \theta_n^{r+1}$ to the minimizing transformation parameters.
3. *Minimize w.r.t. S* (Reconstruct):

(a) Calculate class means μ_P for all $P \in \tilde{D}_P$.

(b) Using conjugate gradient minimize J of equation (13) with respect to S. Set S^{r+1} to the minimizing S.

4. *Iterate*: If the structure S has not converged, set $r = r + 1$ and go to step 2.

The algorithm is illustrated in figure 9.
This finishes the description of the rigid particle SPR algorithm with a single CTF.

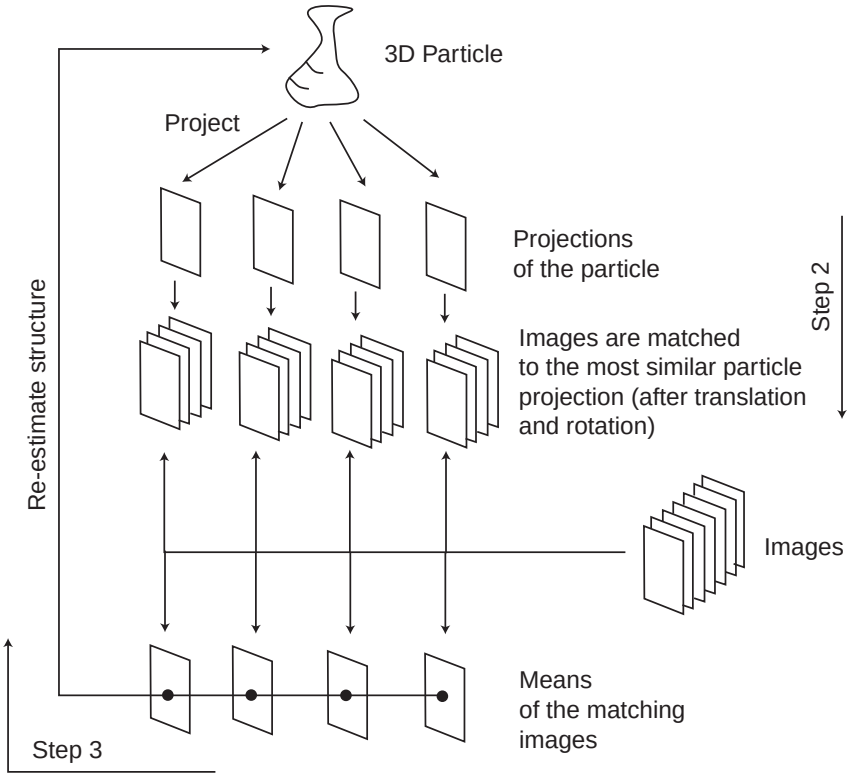

Figure 9: Rigid particle SPR algorithm.

5.4 Comments

Some comments on the above development before we proceed to the multi-CTF case.

1. This algorithm is similar to many reported SPR algorithms in the literature. The difference lies in the fact that this algorithm uses a single objective function whose value is reduced in every iteration, whereas other SPR algorithms do not have a single objective function. Thus the above algorithm is guaranteed to converge.

2. The reconstruction stage of our algorithm falls under the general rubric of *algebraic reconstruction techniques* (ART) of tomography. Many reported SPR algorithms do not use ART; instead they adopt Fourier domain techniques. The comparison of ART to Fourier domain techniques (in the context of SPR) remains to be done.

5.5 Multi-CTF SPR

SPR with mulitple CTFs is a straight-forward extension of Algorithm A2. A clear understanding of what "multiple CTFs" means is useful to understand the extension. In a typical cryo-em procedure, particles are embedded in many different vitreous ice layers. Each layer is imaged with a known, but possibly different, microscope defocus value. Thus each micrograph has its own CTF, and all images from this micrograph inherit that same CTF. Below, we will let $\mathcal{C}$ denote the set of all CTF's present in the cryo-em procedure. The image I_n has the CTF $C_n \in \mathcal{C}$.

Because the CTF is no longer fixed, the first change in the algorithm occurs in the generation of projection images in step 2 of the algorithm. We now have to generate projection images for different CTFs as well. Thus step 2 of the algorithm becomes:

2. *Minimize w.r.t. auxiliary parameters* (Project and align):
   (a) Generate all projection images $\pi_P = C\Sigma_P S^r$, for $(C, P) \in \mathcal{C} \times \tilde{D}_P$.
   (b) For each data image, use exhaustive search to find the projection image and transformation parameters that minimize $\|T^{*-1}_{\theta_n,\mathbf{v}_n} I_n - \pi_P\|^2$. Set P_n^{r+1} to the projection frame of the minimizing projection image and $\mathbf{v}_n^{r+1}, \theta_n^{r+1}$ to the minimizing transformation parameters.

In figure 8 this corresponds to generating more projections – each projection for one combination of CTF and a projection direction. The matching part of the algorithm stays the same.

Step 3 of algorithm A2 has to be modified too. Similar to the modification of step 2, the image classes $\mathcal{I}_P$ have to refined so that there is an image class for each combination of CTF and projection direction. Let $\mathcal{I}_{CP}$ be the set of images that match the projection image for $(C, P) \in \mathcal{C} \times \tilde{D}_P$ in step 2. Then the class mean for these images is:

$$\mu_{C,P} = \frac{1}{\|\mathcal{I}_{C,P}\|} \sum_{n \text{ such that } I_n \in \mathcal{I}_{C,P}} T^{*-1}_{\theta_n,\mathbf{v}_n} I_n.$$

Moreover, the objective function in equation (13) modifies to

$$J(S, \{P_n\}, \{\mathbf{v}_n\}, \{\theta_n\}) \quad = \quad \sum_{(C,P) \in \mathcal{C} \times \tilde{D}_P} \|\mathcal{I}_{C,P}\| \times \|\mu_{C,P} - C\Sigma_P S\|^2. \quad (14)$$

Step 3 of the algorithm modifies to

3. *Minimize w.r.t.* S (Reconstruct):
(a) Calculate class means $\mu_{C,P}$ for all $(C, P) \in \mathcal{C} \times \tilde{D}_P$.
(b) Using conjugate gradient minimize J of equation (14) with respect to S.
Set S^{r+1} to the minimizing S.

6 Evaluating SPRs

As mentioned before, the quality of an SPR reconstruction is measured by Fourier Shell Correlation.

6.1 Fourier Shell Correlation

Fourier Shell Correlation (FSC) gives a measure of signal-to-noise ratio at different spatial frequecies in the reconstruction [1]. The data images are randomly split into two sets and a reconstruction is obtained from each set. Suppose that $F_1(\omega)$ and $F_2(\omega)$ are the 3D discrete Fourier Transforms of the two reconstructions, where ω is the 3D frequency. Suppose that the 3D Fourier domain is partitioned into spherical bins of thickness Δr. Let the bins be denoted Ω_i $i, = 1, \ldots, K$. Then, the FSC in the i^{th} bin is

$$FSC(i) = \frac{\text{Re}\{\sum_{\omega \in \Omega_i} F_1(\omega) F_2^*(\omega)\}}{\sqrt{\sum_{\omega \in \Omega_i} F_1(\omega) F_1^*} \sqrt{\sum_{\omega \in \Omega_i} F_2(\omega) F_2^*}}. \tag{15}$$

The FSC is usually plotted as a function of the center frequency of the bin Ω_i. In the presence of noise in the reconstruction, the FSC begins with a high value of 1 and rolls off down to 0. The frequency at which the FSC reaches 0.5 is usually taken to be the resolution of the reconstruction.

7 Conclusions

This note presented the single particle reconstruction (SPR) problem in cryogenic electron microscopy in a maximum-likelihood framework. CTF and multi-CTF problems were addressed.

References

[1] Frank J. *Three Dimensional Electron Microscopy of Macromolecular Assemblies*, Oxford University Press, 2005.

[2] Young G. A., Smith R. L., *Essentials of Statistical Inference*, Cambridge University Press, 2010.

[3] Tagare H. D., Barthel A., Sigworth F. J., "An adaptive Expectation-Maximization algorithm with GPU implementation for electron cryomicroscopy", Journal of Structural Biology, 171, pp. 256-265, 2010.

[4] Bernardo J. M., Smith A. F., *Bayesian Theory*, John Wiley & Sons, 2000.

[5] Tierney L., Kadane J. B., "Accurate approximations for posterior moments and marginal densities", J Amer Stat Assoc, 81, 82-86, 1986.

[6] Ludtke, S.J., P.R. Baldwin, and W. Chiu, "EMAN: semiautomated software for high-resolution single-particle reconstructions", J Struct Biol 128: 82-97, 1999.

[7] Grigorieff, N., "FREALIGN: high-resolution refinement of single particle structures", J Struct Biol 157: 117-25, 2007.

[8] Burnham K. P., Anderson D., *Model selection and multi-model inference*, Springer, 2003.

[9] Claesken G., Hjort N. L., *Model selection and model averaging*, Cambridge University Press, 2008.

[10] Rissanen J., *Information and complexity in statistical modeling*, Springer, 2007.

[11] Carthew, R.W., and E.J. Sontheimer, "Origins and Mechanisms of miRNAs and siRNAs". Cell 136: 642-55, 2009.

[12] Bernstein, E., A.A. Caudy, S.M. Hammond, and G.J. Hannon, "Role for a bidentate ribonuclease in the initiation step of RNA interference", Nature 409: 363-366, 2001.

[13] MacRae, I.J., K.H. Zhou, F. Li, A. Repic, A.N. Brooks, W.Z. Cande, P.D. Adams, and J.A. Doudna, "Structural basis for double-stranded RNA processing by dicer", Science 311: 195-198, 2006.

[14] Bertsekas D., *Non-linear programming*, Athena Scientific, 1999.

[15] Penczek P. A., Yang C., Frank J., Spahn C. M. T., "Estimation of variance in single-particle reconstruction using the bootstrap technique", J Struct Biol, 154: 168-183, 2006.

[16] Penczek P. A., Frank J., Spahn C. M. T., "A method of focused classification based on bootstrap 3D variance analysis, and its application to EF-G-dependent translocation", J Struct Biol, 154: 184-194, 2006.

[17] Zhang W., Kimmel M., Spahn C. M. T., Penczek P. A., "Heterogeneity of large macromolecular complexes revealed by 3D cryo-EM variance analysis", Structure, 16: 1770-1776, 2008.

[18] Borg I., Groenen P., "Modern multidimensional scaling: theory and applications", Springer-Verlag, 2005.