RELION: In 2D

By Arunabh Ghosh

The objective of RELION

- To objectively determine the structure of the molecule in a Bayesian framework.
- It removes the need to tune arbitrary constants, and shows us how data and prior knowledge should be combined to obtain good resolution estimates.
- It imposes a smoothness in the reconstructed density through a gaussian prior in the Fourier domain.

$$X_{ij} = \mathrm{CTF}_{ij} \sum_{l=1}^{L} \mathbf{P}_{jl}^{\phi} V_l + N_{ij}$$

Linear image formation model

- Almost all existing implementations for cryo-EM structure determination employ the so-called weak phase object approximation, which leads to a linear image formation model in Fourier space.
- X_{ij} is the jth component, with j=1,…,J, of the 2D Fourier transform of the ith experimental image X_i.
- CTF_{ij} is the jth component of the contrast transfer function for the ith image.
- V_l is the lth component, with l=1,…,L, of the 3D Fourier transform of the underlying structure in the data set.
- P^{ϕ} and P^{ϕ^T} are the projection and the back projection operations.
- N_{ij} is the noise in the complex plane.

The iterative algorithm – Maximization step

$$V_l^{(n+1)} = \frac{\sum_{i=1}^{N} \int_{\phi} \Gamma_{i\phi}^{(n)} \sum_{j=1}^{J} \mathbf{P}^{\phi_{ij}^{T}} \frac{\mathrm{CTF}_{ij} X_{ij}}{\sigma_{ij}^{2(n)}} \mathrm{d}\phi}{\sum_{i=1}^{N} \int_{\phi} \Gamma_{i\phi}^{(n)} \sum_{j=1}^{J} \mathbf{P}^{\phi_{ij}^{T}} \frac{\mathrm{CTF}_{ij}^{2}}{\sigma_{ij}^{2(n)}} \mathrm{d}\phi + \frac{1}{\tau_l^{2(n)}}}$$

$$\sigma_{ij}^{2(n+1)} = \frac{1}{2} \int_{\phi} \Gamma_{i\phi}^{(n)} \left| X_{ij} - \mathrm{CTF}_{ij} \sum_{l=1}^{L} \mathbf{P}_{jl}^{\phi} V_l^{(n)} \right|^2 \mathrm{d}\phi$$

$$\tau_l^{2(n+1)} = \frac{1}{2} \left| V_l^{(n+1)} \right|^2$$

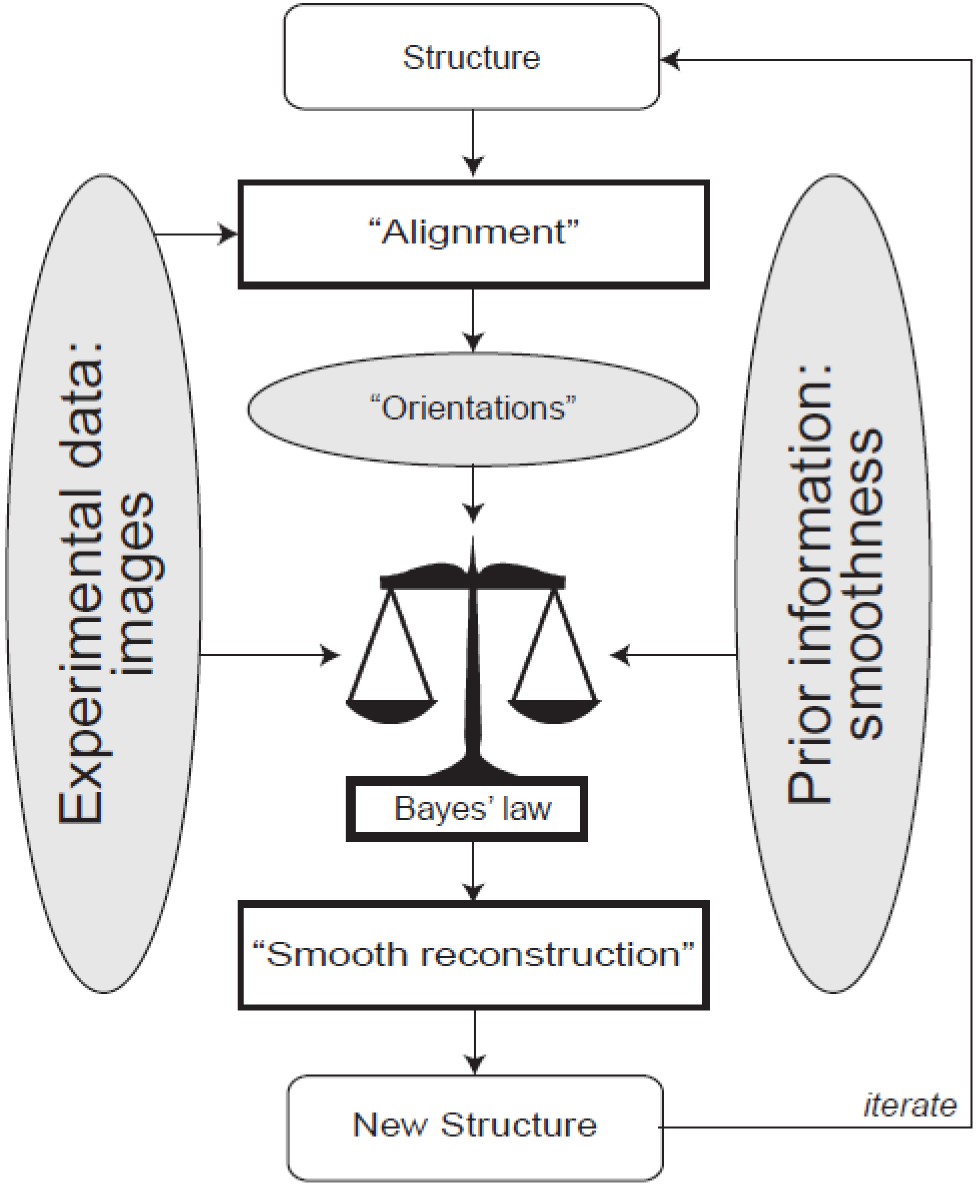
- Optimization by expectation maximization yields the following algorithm. On the side is the maximization step of the iterative algorithm.
- $\mathrm{T}_{i\phi}$ is the probability of projection I having the orientation ϕ. This is calculated in the next step in the so called expectation step.

$$\Gamma_{i\phi}^{(n)} = \frac{P\left(X_i \mid \phi, \Theta^{(n)}, Y\right) P\left(\phi \mid \Theta^{(n)}, Y\right)}{\int_{\phi'} P\left(X_i \mid \phi', \Theta^{(n)}, Y\right) P\left(\phi' \mid \Theta^{(n)}, Y\right) d\phi'}$$

The iterative algorithm Expectation step

- In the expectation step we calculate the posterior probability of each orientation for each projection.
- Rather than assigning an optimal orientation ϕ_i^* to each image, probability-weighted integrals over all possible orientations are calculated.

Flowchart of structure determination



Problems faced in the algorithm

- Often times, even in cases of little noise and shifts, the probability of even the correct orientation of a projection turns out to be very small. One way to solve it has been to reduce the number of Fourier components to prevent overfitting.
- Implementation of the projection and the back projection operation in the spatial domain and converting it back to the Fourier domain leads to huge discrepancies in the estimates. Effort is being made to keep the entire algorithm in the Fourier domain.
- The model is really sensitive to shifts in the projection, which means while searching for the correct shift the resolution has to be really small which means the program may take a lot of time.

Results achieved by the current algorithm

Parameters of the experiment

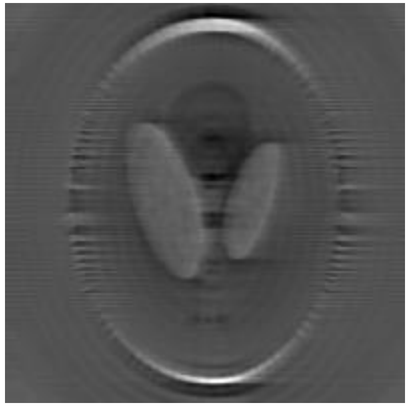
- 5% noise in the projections
- Error in projection angles - ± 5
- Unknown shifts in the projection - 0
- No. of iterations - 10
- Original Image:

Results achieved by the current algorithm

- No. of projections – 30

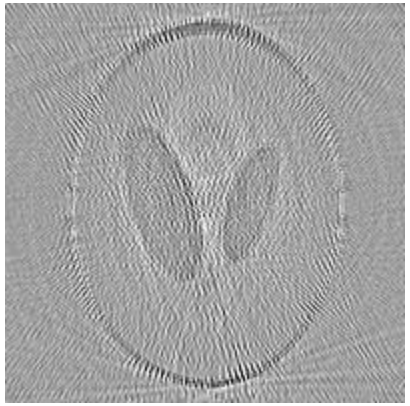
Norm-2 error: 67.35

- No of projections – 180



Norm-2 error: 58.20

- No. of projections – 360



Norm-2 error: 43.37

Further steps to take

- The first step would be to implement the entire algorithm in the Fourier domain without needing to convert back to the spatial domain. Implementing the Fourier slice theorem in the Fourier domain has been very difficult but yesterday I finally was able to accurately reconstruct an image in just the Fourier domain. Using this knowledge I would try to implement this and hopefully obtain better results.
- Further there is another step described in the paper where they have slightly modified the image estimation equation using a **Kaiser–Bessel window** to obtain a smooth reconstruction. I have already implemented it. What's required further is the tuning of some constants to achieve optimum results.