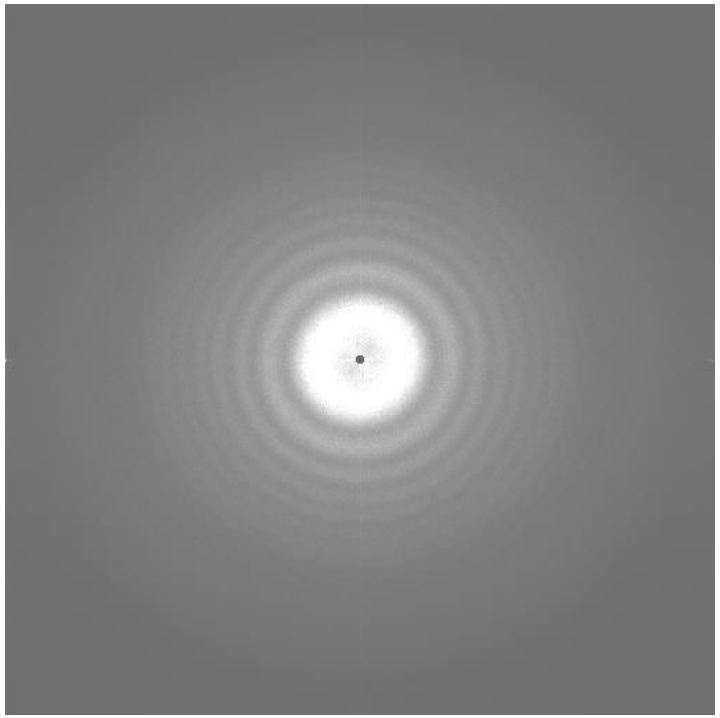
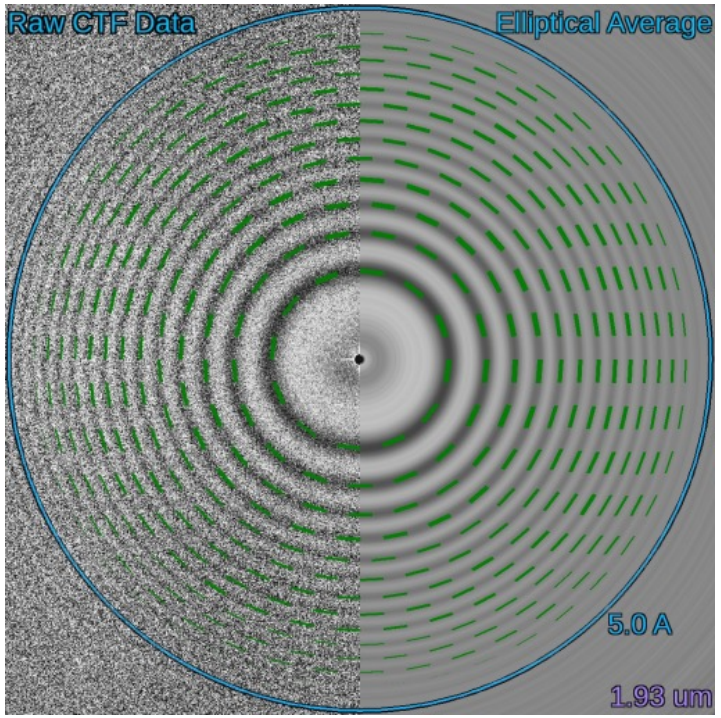
Cameras and Frames

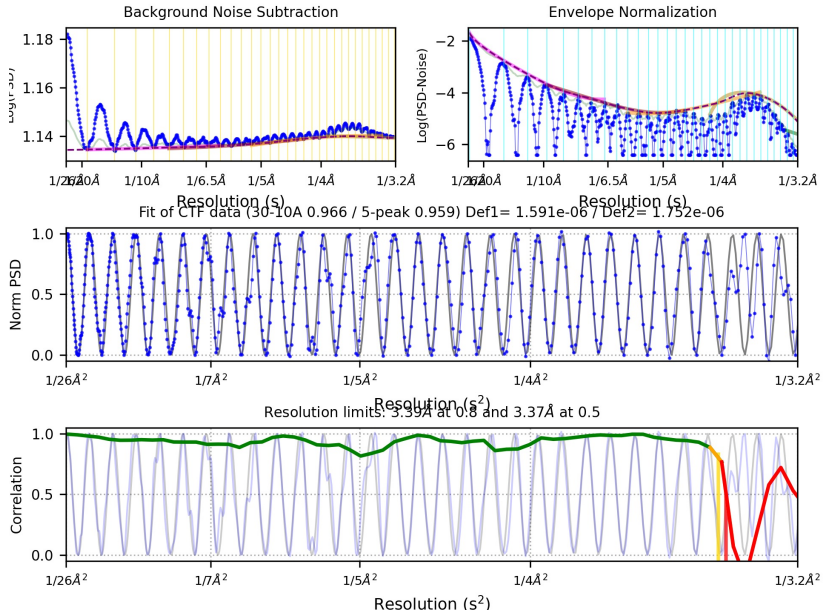
Steps in a single particle reconstruction

- Collect data
- Align frames
- Estimate CTF
- Pick particles
- 2D Classification
- Generate initial model
- Refine data against initial model
- Estimate resolution

Estimating the CTF





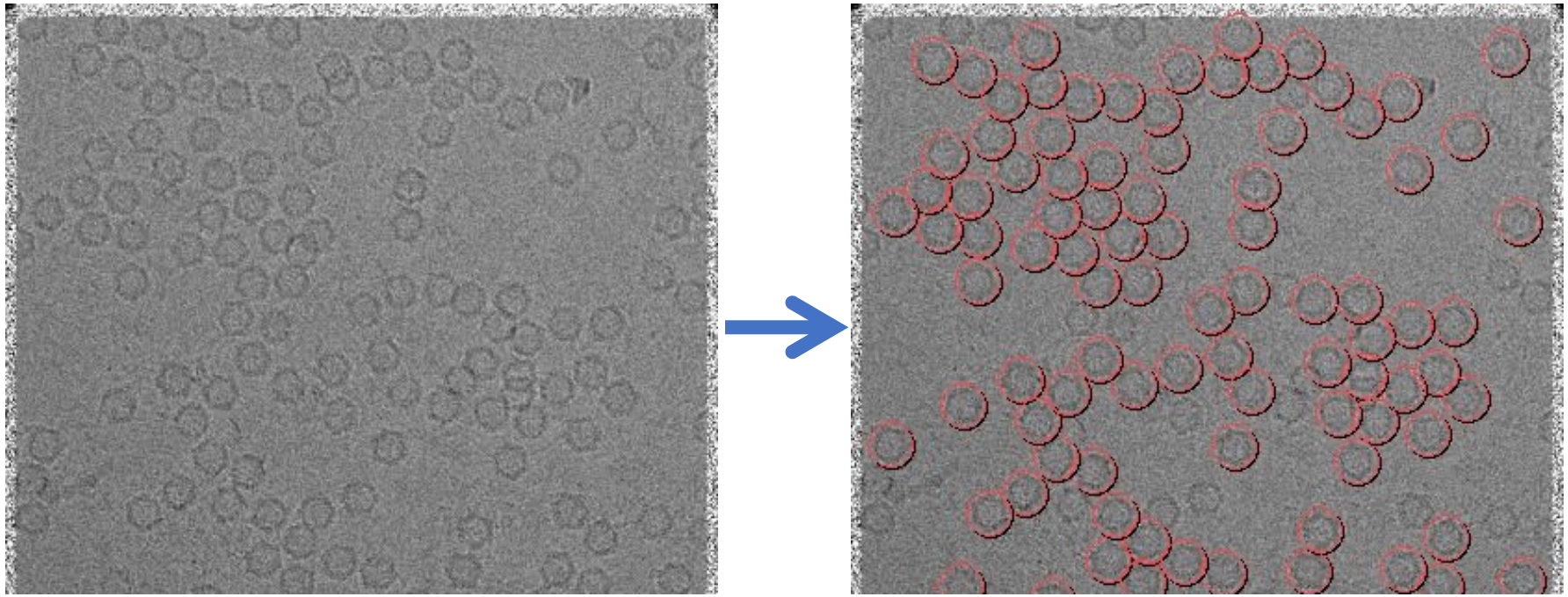


- Many tools exist for automatic CTF estimation
  - Determines defocus and astigmatism for each image
  - Resolution potential present in image can be estimated from agreement between estimate and data

Steps in a single particle reconstruction

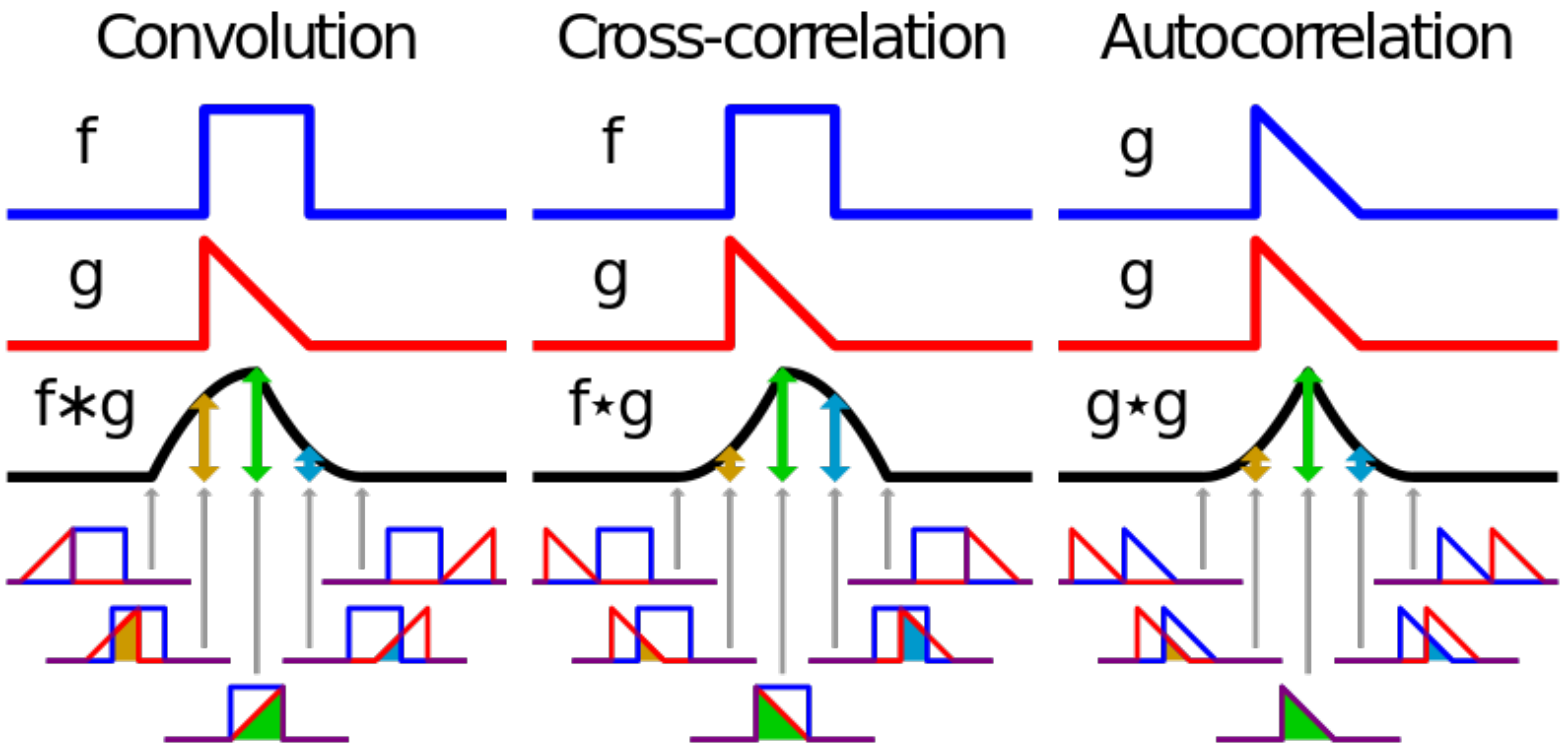
- Collect data
- Align frames
- Estimate CTF
- Pick particles
- 2D Classification
- Generate initial model
- Refine data against initial model
- Estimate resolution

Picking particles



- Mostly automated
  - Many tools exist for template-based picking
  - Many tools for machine learning based picking
  - Some tools for general picking based on size

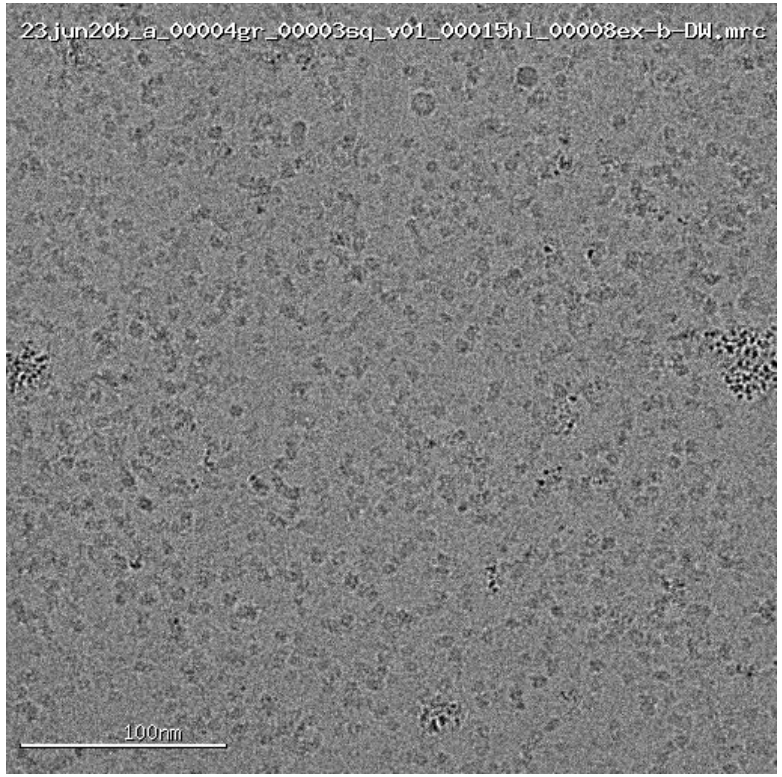
An aside about image alignment

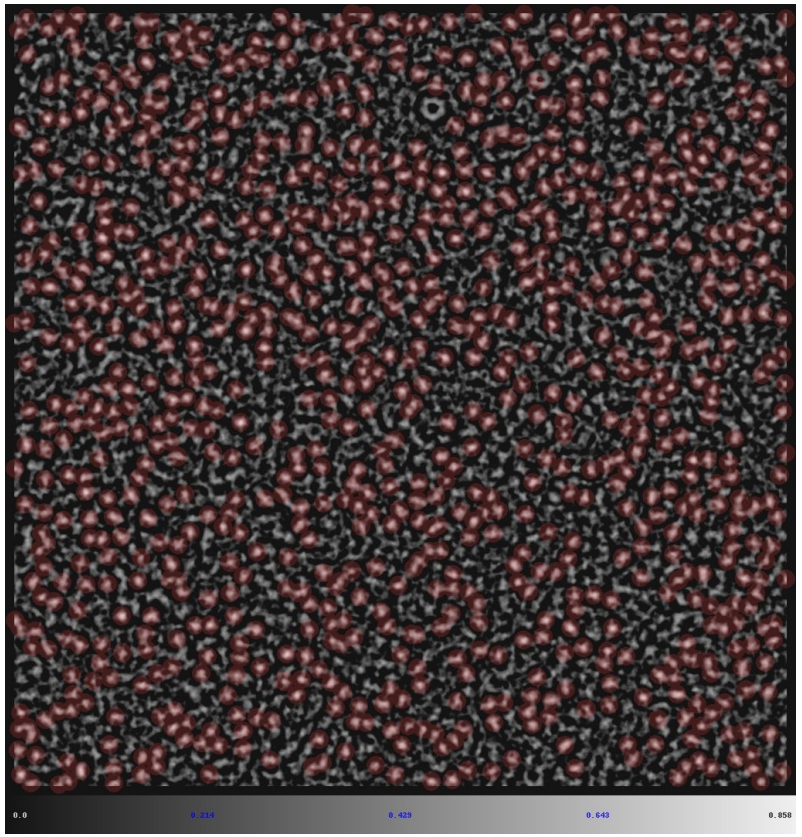


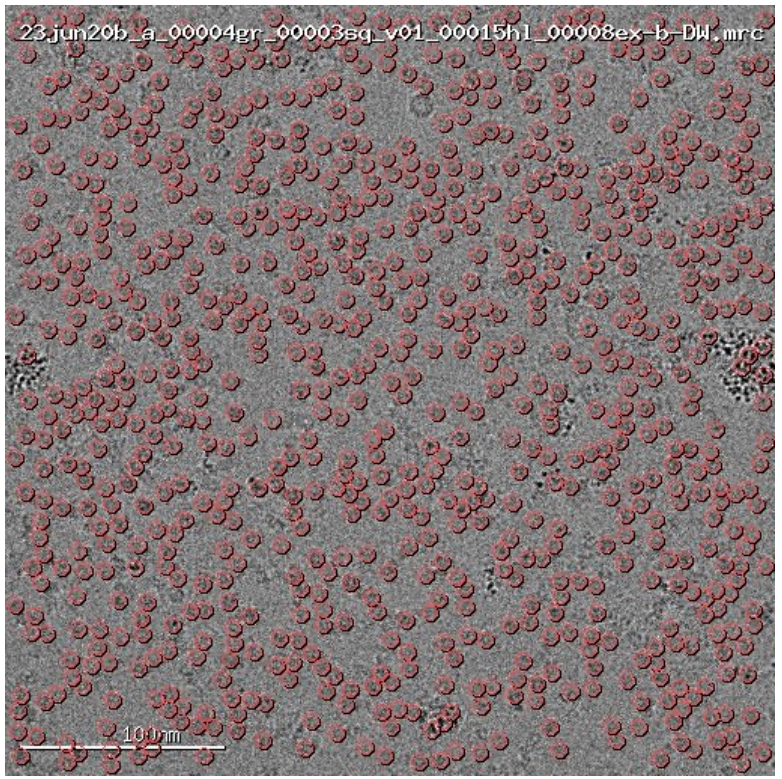
The discrete autocorrelation R at lag l for a discrete signal $y(n)$ is

$$R_{yy}(l) = \sum_{n \in Z} y(n)\,\overline{y}(n-l).$$

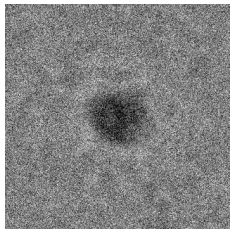
Template based picking







Cross correlated with template



Steps in a single particle reconstruction

- Collect data
- Align frames
- Estimate CTF
- Pick particles
- 2D Classification
- Generate initial model
- Refine data against initial model
- Estimate resolution

Image Alignment in cryo-EM

Importance: Particles in cryoEM have random orientations. Accurate alignment ensures high-resolution 3D reconstruction.

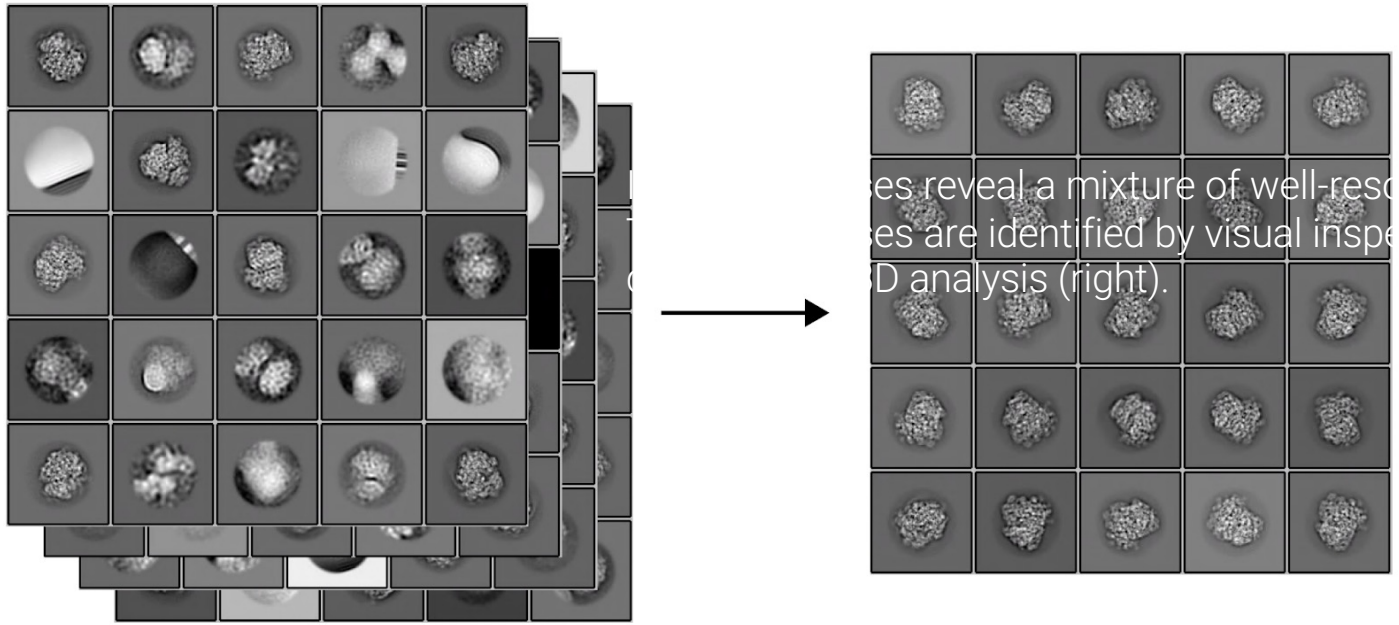
Techniques:

- **Cross-Correlation:** Shifts one image over another to measure similarity.
- **Maximum Likelihood:** Estimates the best orientation and position for particles.
- **Reference-based:** Iteratively aligns using a reference image.

Challenges:

- **Reference Bias:** Wrong references can skew results.
- **Computational Needs:** High-resolution alignment requires powerful computation.
- **SNR Issues:** Noise can hinder alignment.

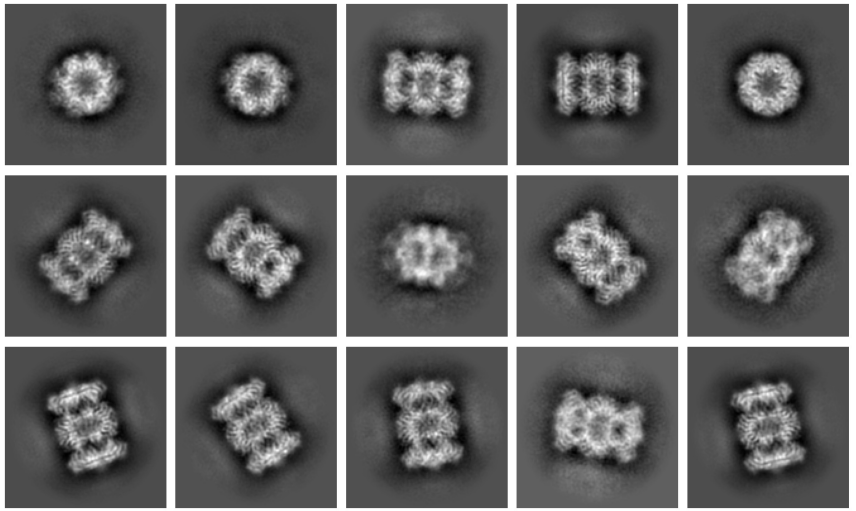
Outcome: Groups similar projections and produces accurate 2D and 3D (later) reconstructions.



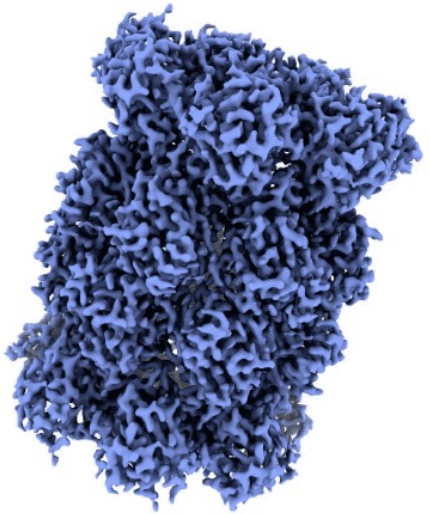
2D classification in cryo-EM

High resolution 2D classes

- Multiple orientations visible
- Well-aligned images
- High-res features visible
- Large number of ptcls per class
- All prerequisites for high-resolution 3D reconstruction



20S proteasome - 2D class averages



20S proteasome - 3D reconstruction

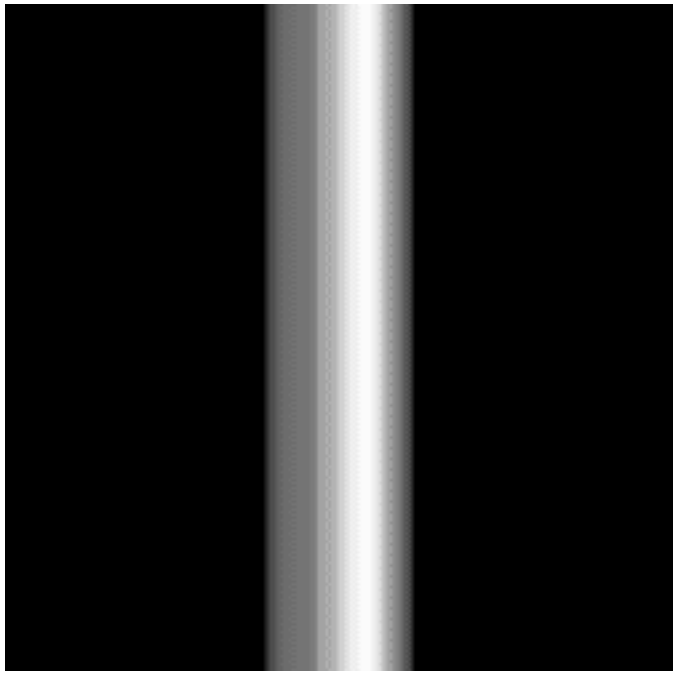
Steps in a single particle reconstruction

- Collect data
- Align frames
- Estimate CTF
- Pick particles
- 2D Classification
- Generate initial model – magic happens (i.e. we’re skipping over this part)
- Refine data against initial model
- Estimate resolution

Basic Principles of 3D Reconstruction

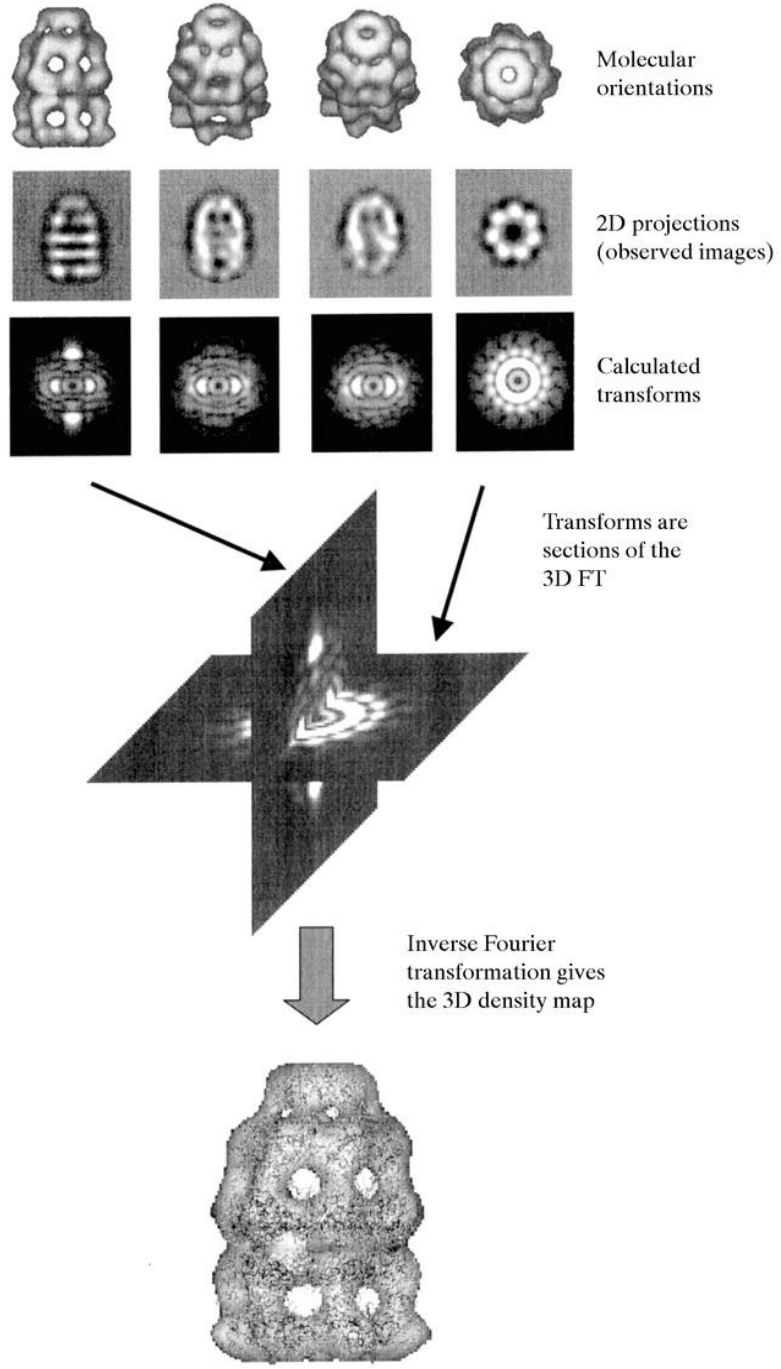
- **Definition:** Transforming 2D particle images from various orientations into a coherent 3D volume, revealing the specimen's structure.
- **Back-projection:** Uses 2D images to estimate density at each point in 3D space.
- **Fourier Space:** CryoEM uses Fourier transforms to convert spatial images into frequency space. This helps in manipulation, alignment, and averaging.
  - **Fourier Inversion:** Converts the frequency information back to real space to produce a 3D structure.
  - **Central Section Theorem:** Each 2D projection provides a slice (or section) of the 3D Fourier transform of the specimen.

Reconstructing by backprojection

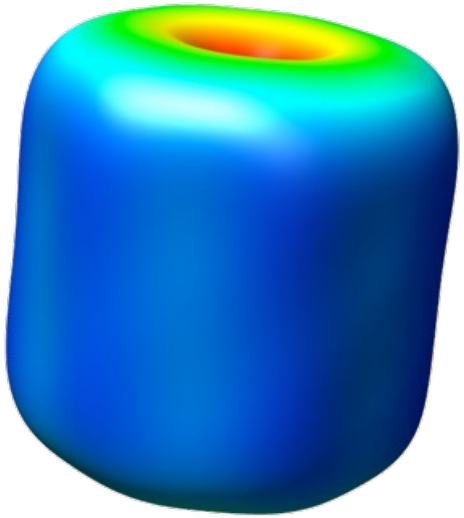


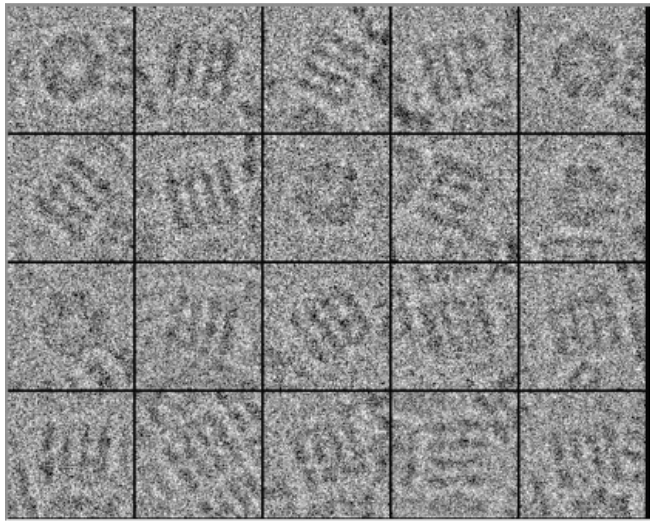
Fourier Inversion and Cryo-EM

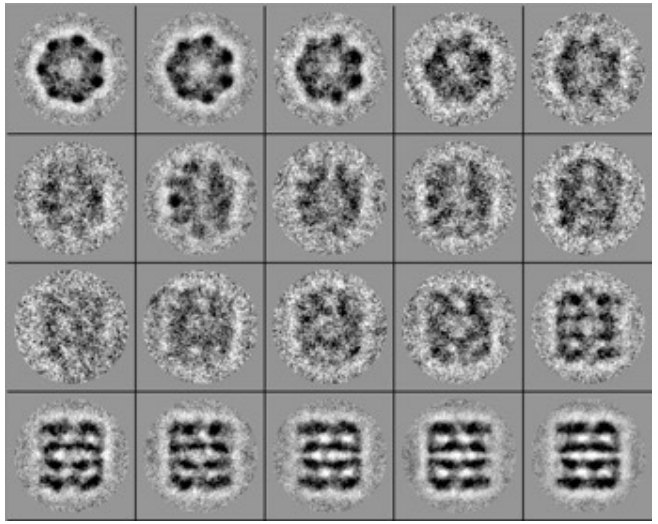
- Transforming data from spatial to frequency domain
- Essential for image processing and 3D reconstruction
- Allows for filtering and resolution enhancement
- Underpins many Cryo-EM algorithms

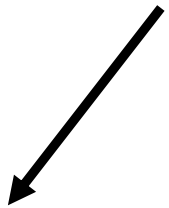


Classification and averaging

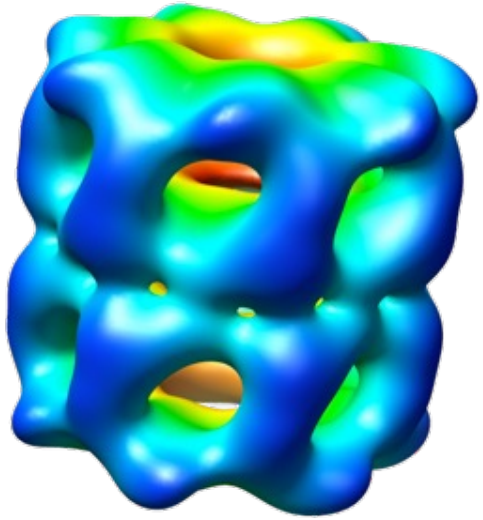


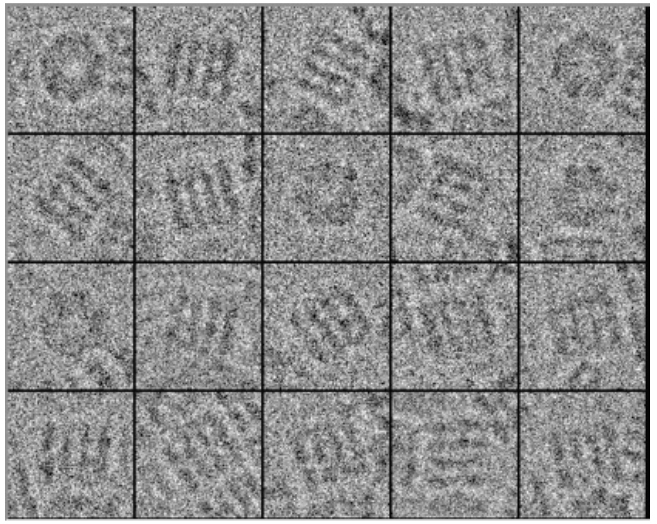


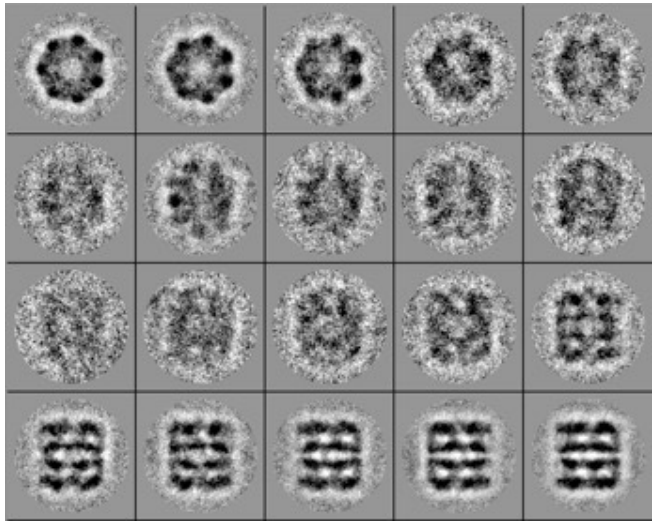




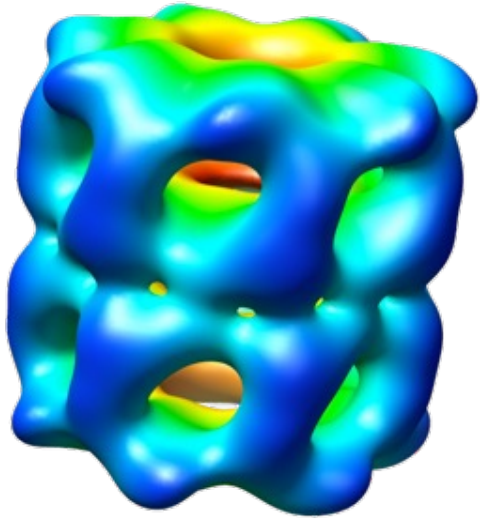
Backprojection yields new 3D model

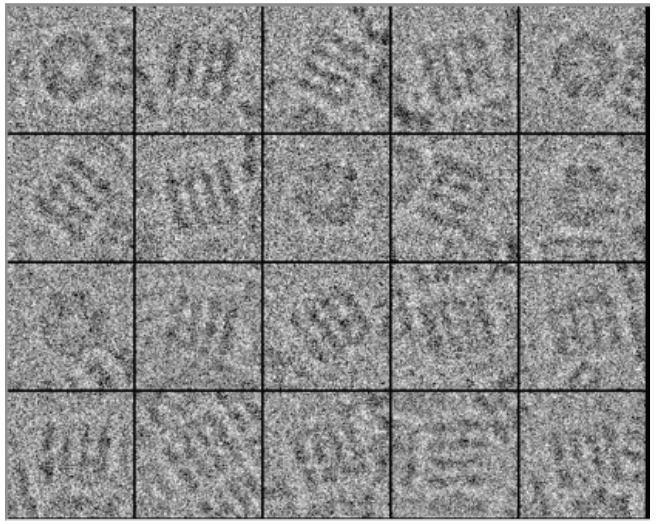


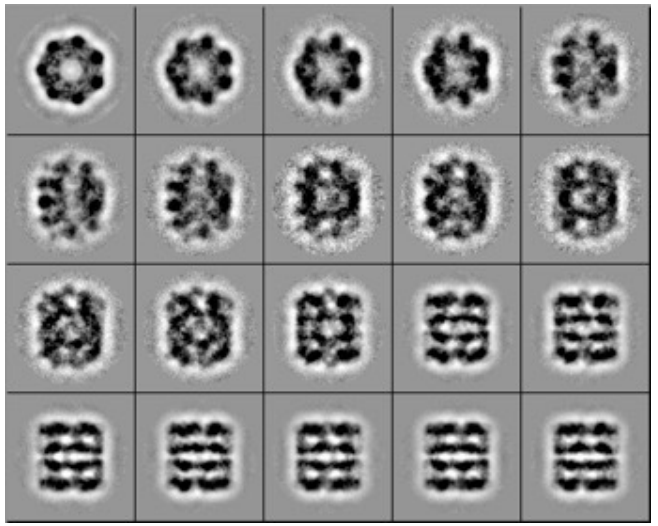




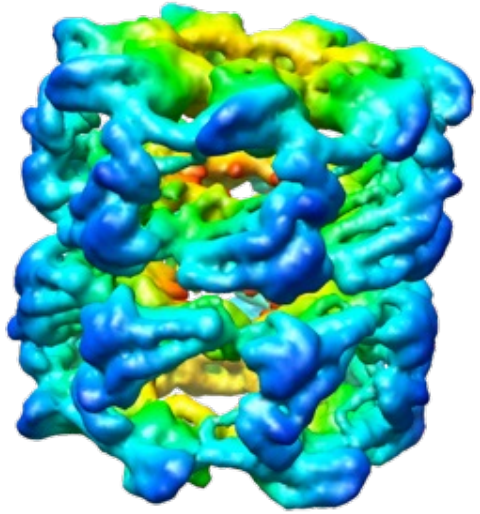
Higher resolution model yields better classification

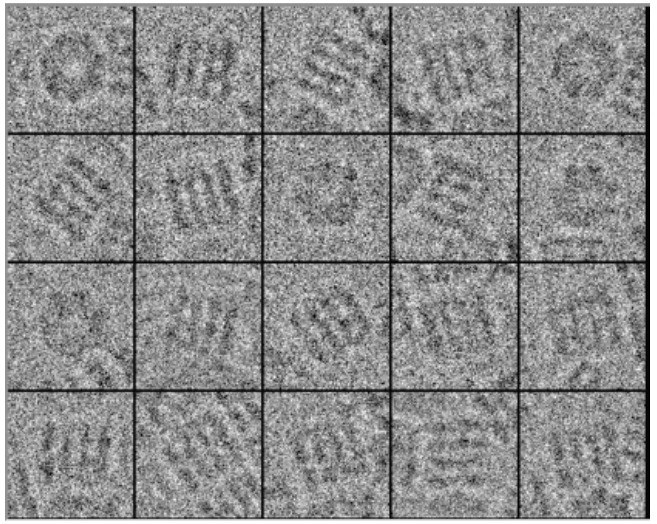


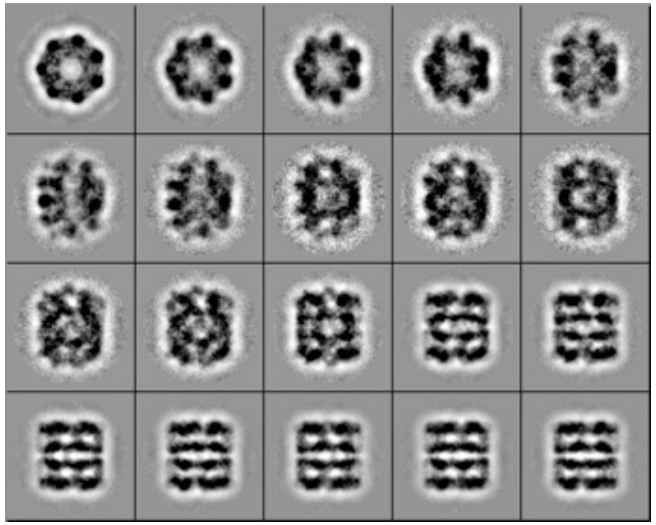




Refinement converges on high resolution reconstruction







Another approach is to use maximum likelihood (ML) methods for alignment and classification

- Statistical approach to optimize alignment and classification
- Improves accuracy and reduces overfitting
- Foundation of many modern cryo-EM software tools
- Has revolutionized high-resolution cryo-EM

Maximum Likelihood in General

Definition: In statistics, maximum likelihood estimation (MLE) is a method used to find the values of parameters that maximize the likelihood of the observed data, given a particular model.

How it works:

- **Likelihood Function:** For a given dataset and a model with certain parameters, the likelihood function calculates the probability of observing that data.
- **Optimization:** MLE aims to find the parameter values that maximize this likelihood function. This means that under the parameters determined by MLE, the observed data is the most probable (or "most likely").

Maximum Likelihood in Cryo-EM

Relevance: In cryoEM, each particle image offers a 2D projection of a 3D structure, but its exact orientation and position in 3D space are unknown. The ML approach estimates these unknown parameters.

How it works in cryoEM:

- **Projection Matching:** For each 2D particle image, the algorithm estimates which 3D orientation and position most likely resulted in that 2D projection.
- **Refinement:** Using these initial estimates, the 3D structure is refined iteratively. Each iteration involves:
  - Generating 2D projections from the current 3D model.
  - Comparing the experimental 2D images to these projections.
  - Adjusting the 3D model to increase the likelihood that it led to the observed 2D images.
- **Regularization:** To avoid overfitting (fitting noise in the data), some form of regularization is applied. This ensures the solution remains physically meaningful.

Steps in a single particle reconstruction

- Collect data
- Align frames
- Estimate CTF
- Pick particles
- 2D Classification
- Generate initial model
- Refine data against initial model
- Estimate resolution

FSC (Fourier Shell Correlation)

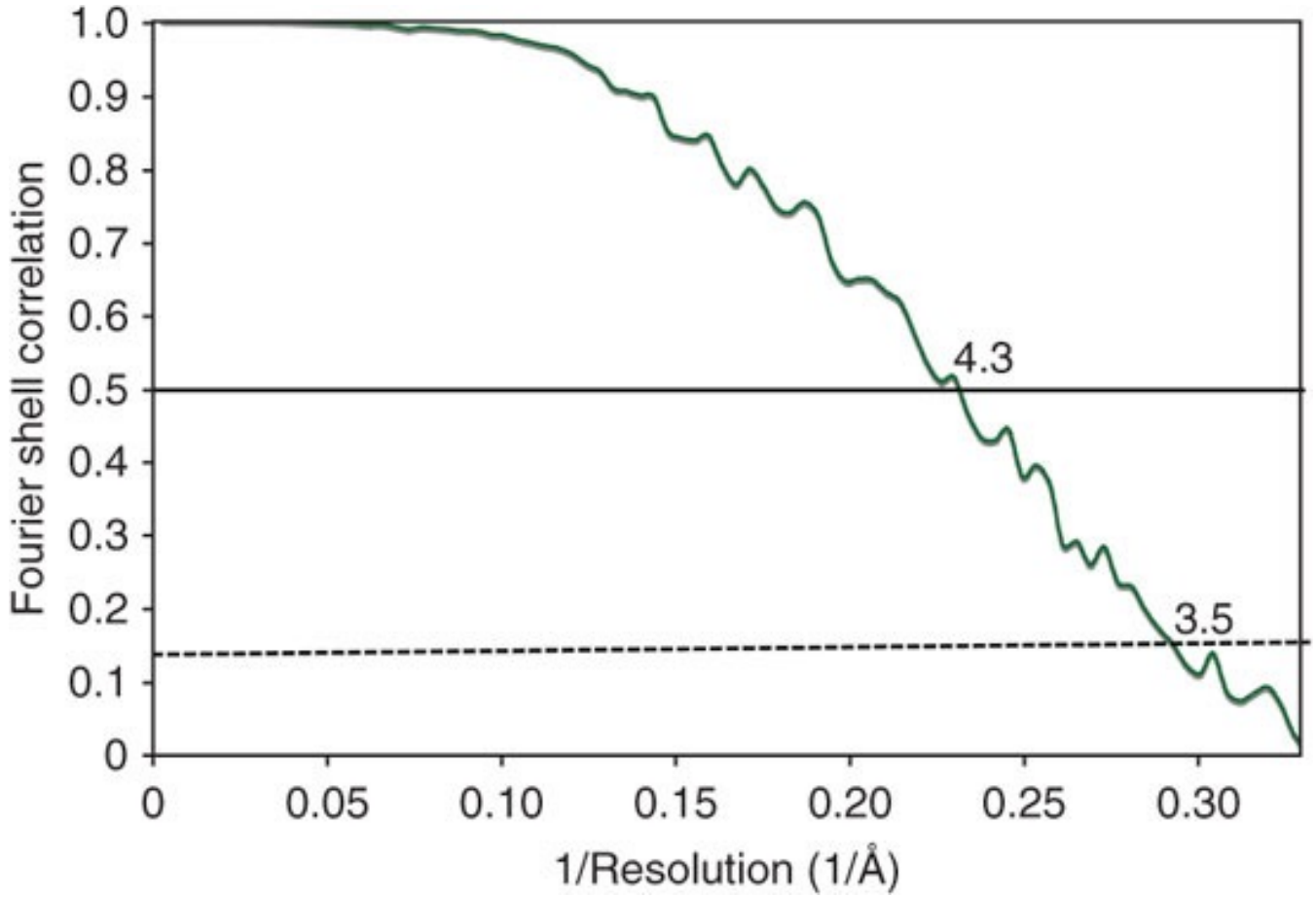
- Quantitative measure of resolution in Cryo-EM
- Compares independent half-maps for consistency
- 0.143 criterion for resolution estimation
- Essential for validating Cryo-EM structures

How to measure resolution?

- In X-ray crystallography, resolution is measured by how far diffraction goes out from origin in diffraction pattern
- No analogous criterion in 3DEM
  - Instead, we use a self-consistency criterion called the Fourier Shell Correlation (FSC)
- With FSC, we split our data into two halves, reconstruct them separately, then compute the correlation between the Fourier transforms of the two reconstructions at discrete resolution shells

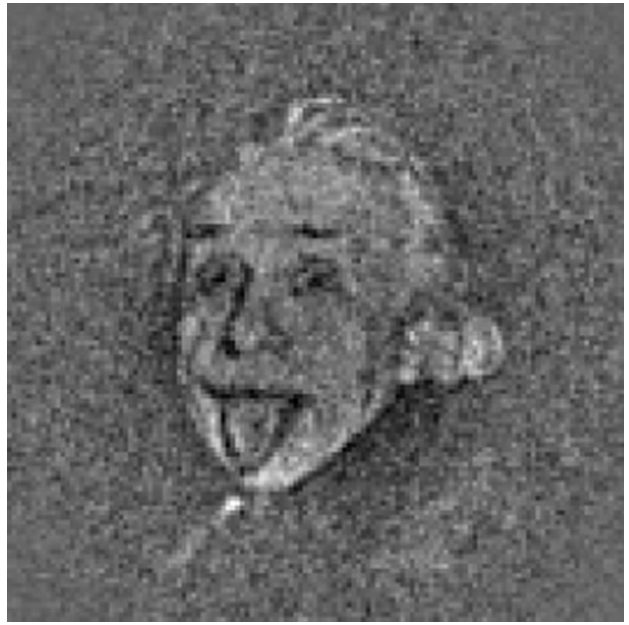
An FSC curve

- Assign resolution at a cutoff correlation value
  - $FSC_{0.143}$ is the standard, but there are other metrics as well



Single particle caveats

- Initial model bias
  - Since the images are so noisy, it is possible to regenerate an image of a reference from aligning pure noise



It is possible to get a “high resolution” structure from a completely incorrect reconstruction

Figure 3