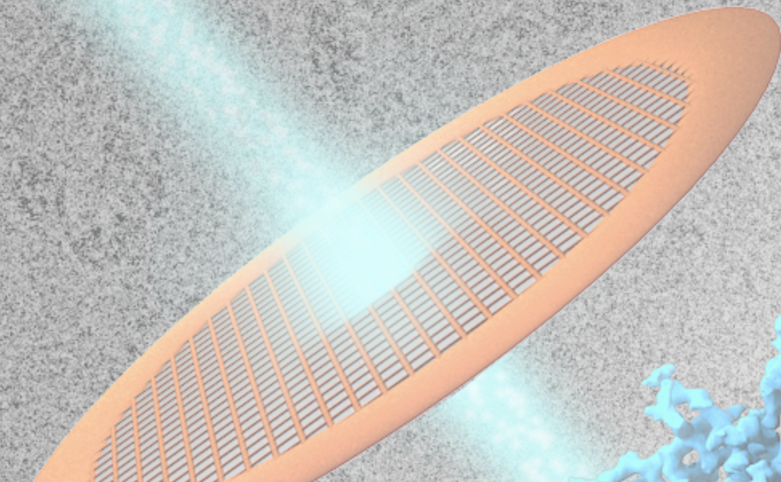
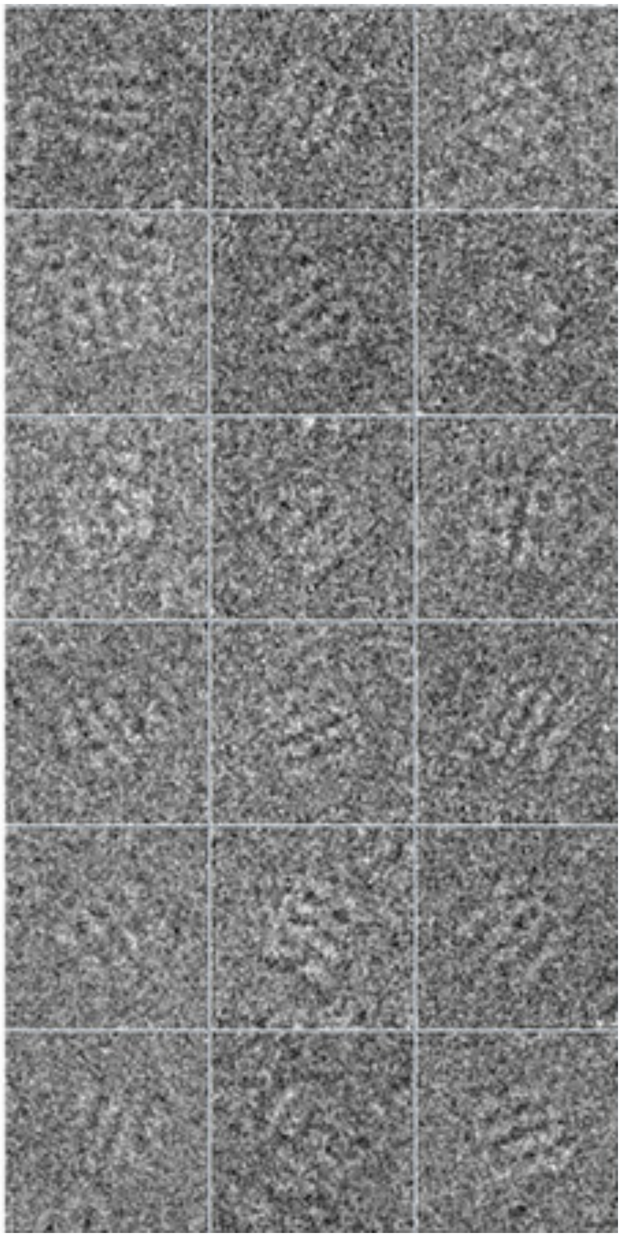


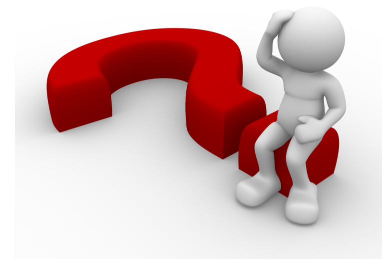
From 2D particles to 3D reconstruction

Hélène Malet

Renafobis 2017

We have collected your nice 2D images. But there are very noisy…



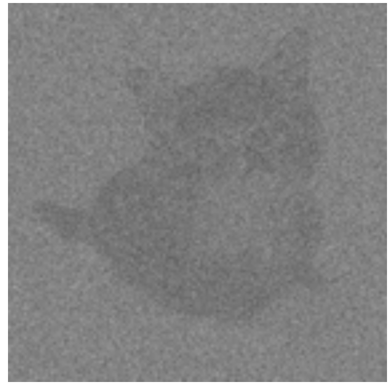


What would you do to improve signal/noise ?

Make groups of similar particles and average them !

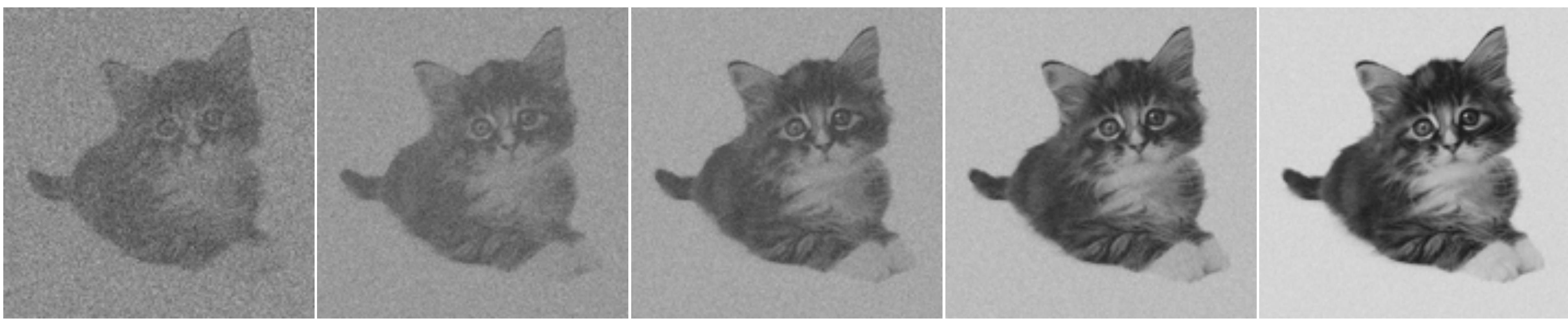
= 2D classification

A single image



Average of several images with the same orientation

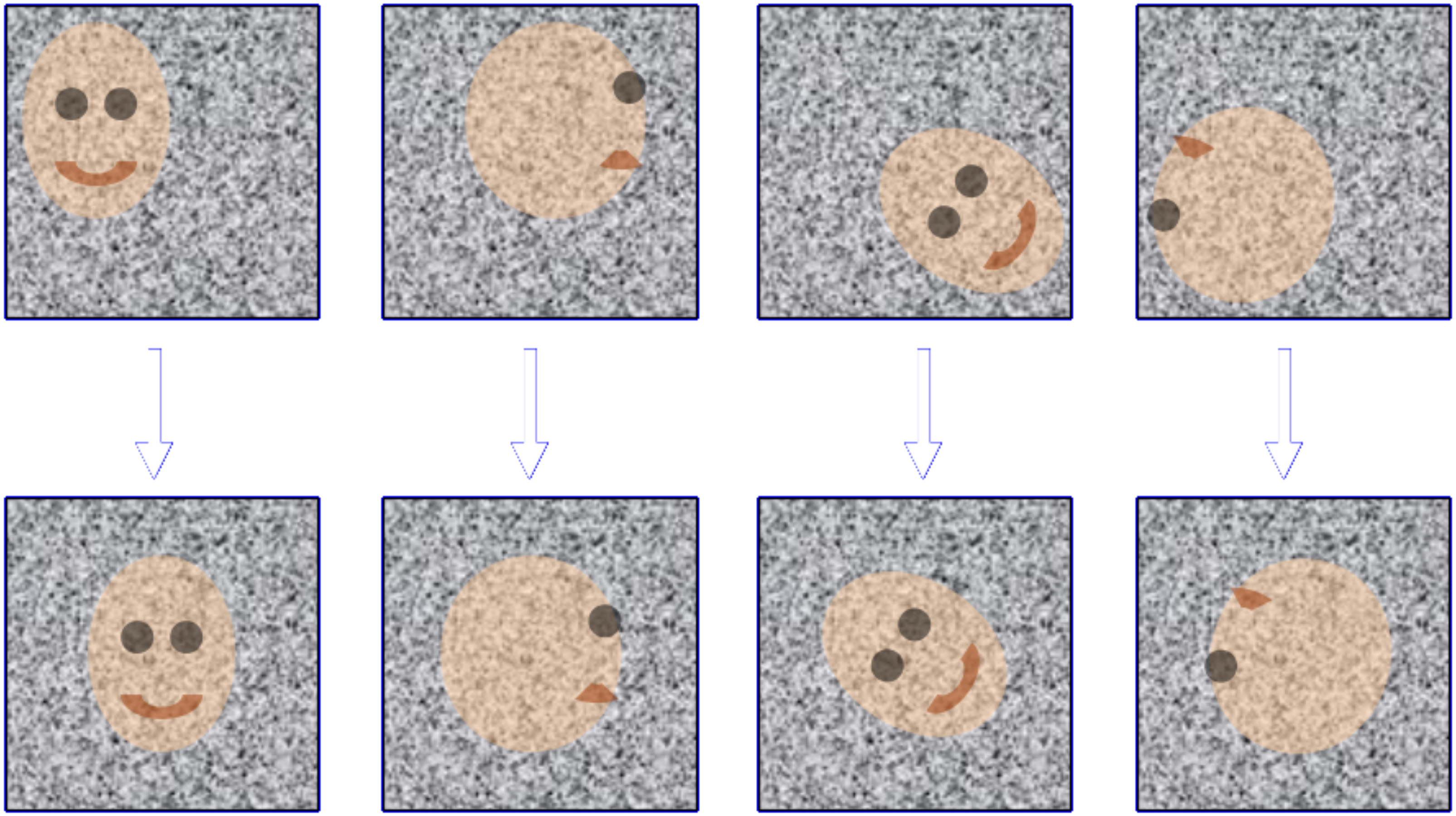
4 16 64 256 1024



The problem: there are several orientations so we need to separate into several class averages

How to create homogeneous 2D class averages?

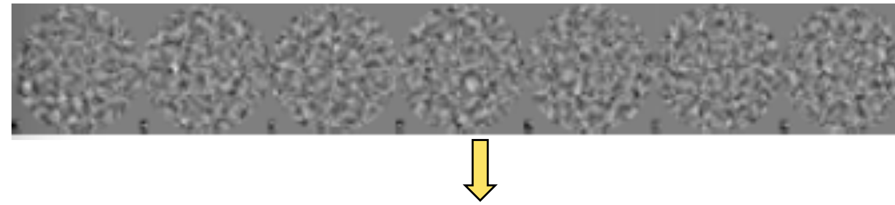
- Translation of images to center them



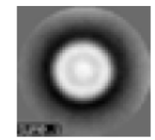
A. Patwardhan

Method used to center images

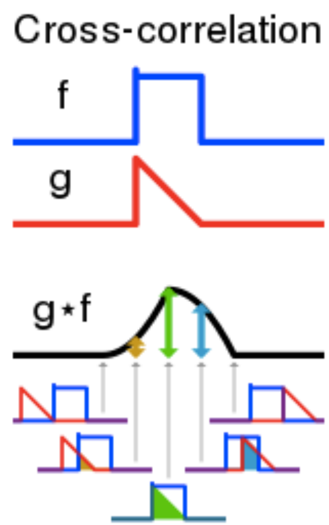
- Average of all images

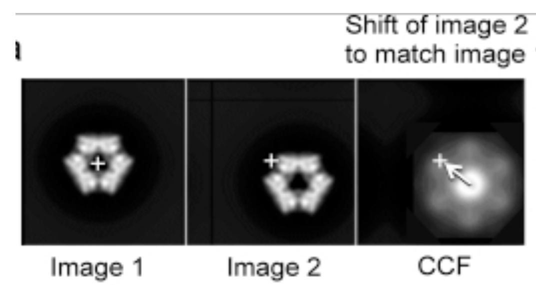


- Rotational average

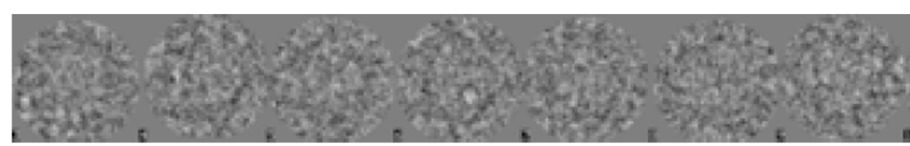


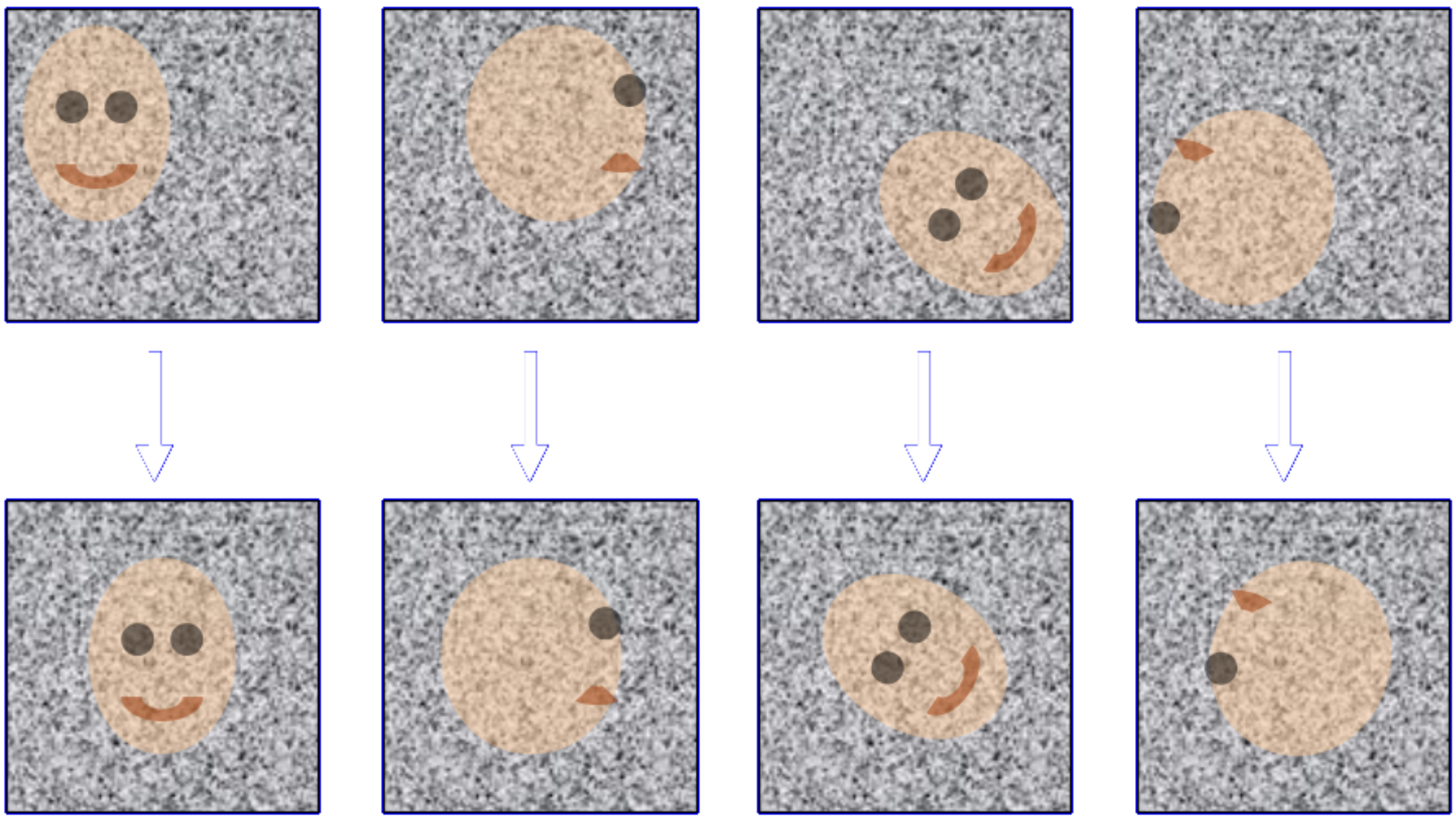
- Cross-correlation between each image and the rotational average





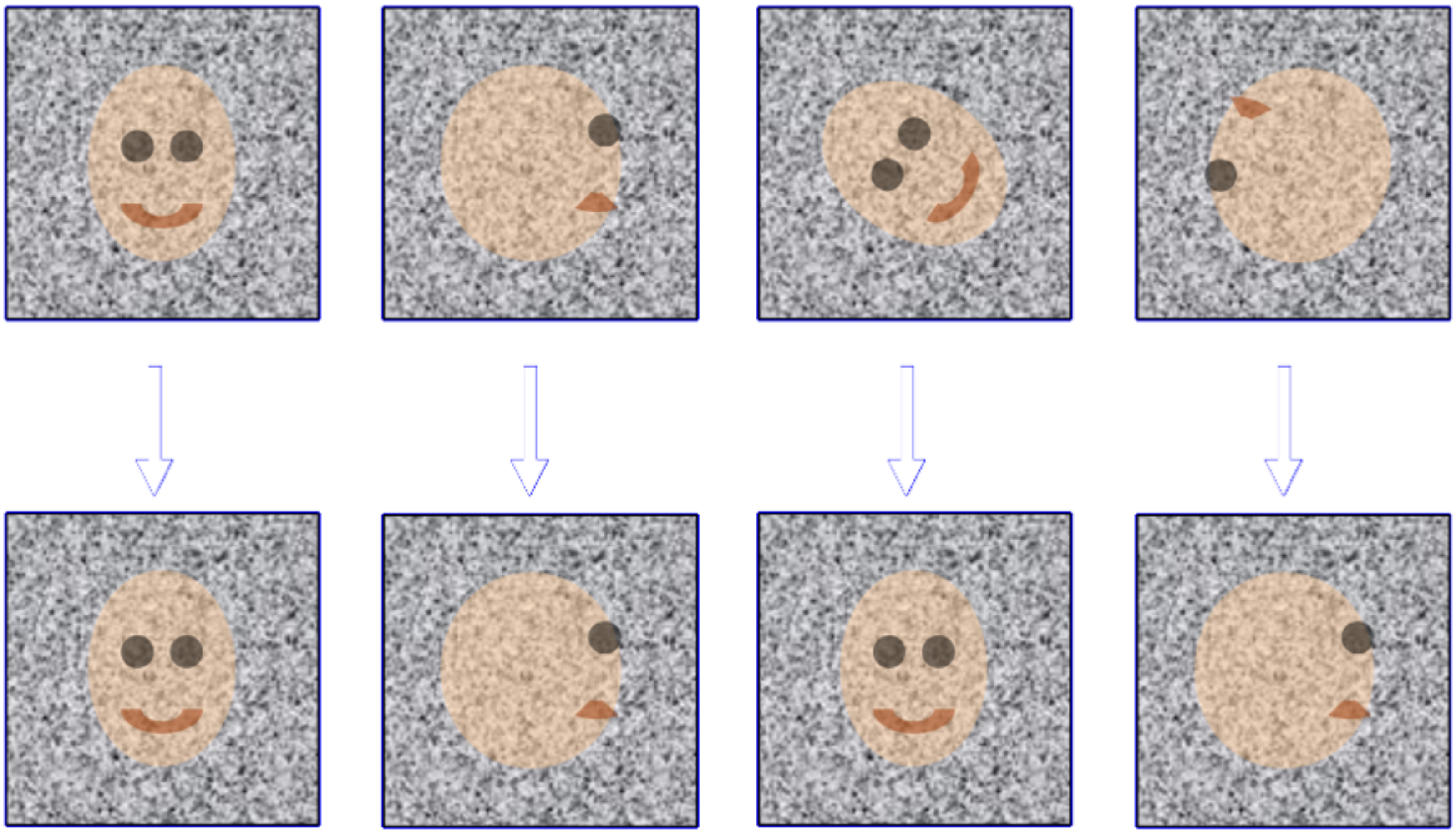
- Translation of each image to center it !





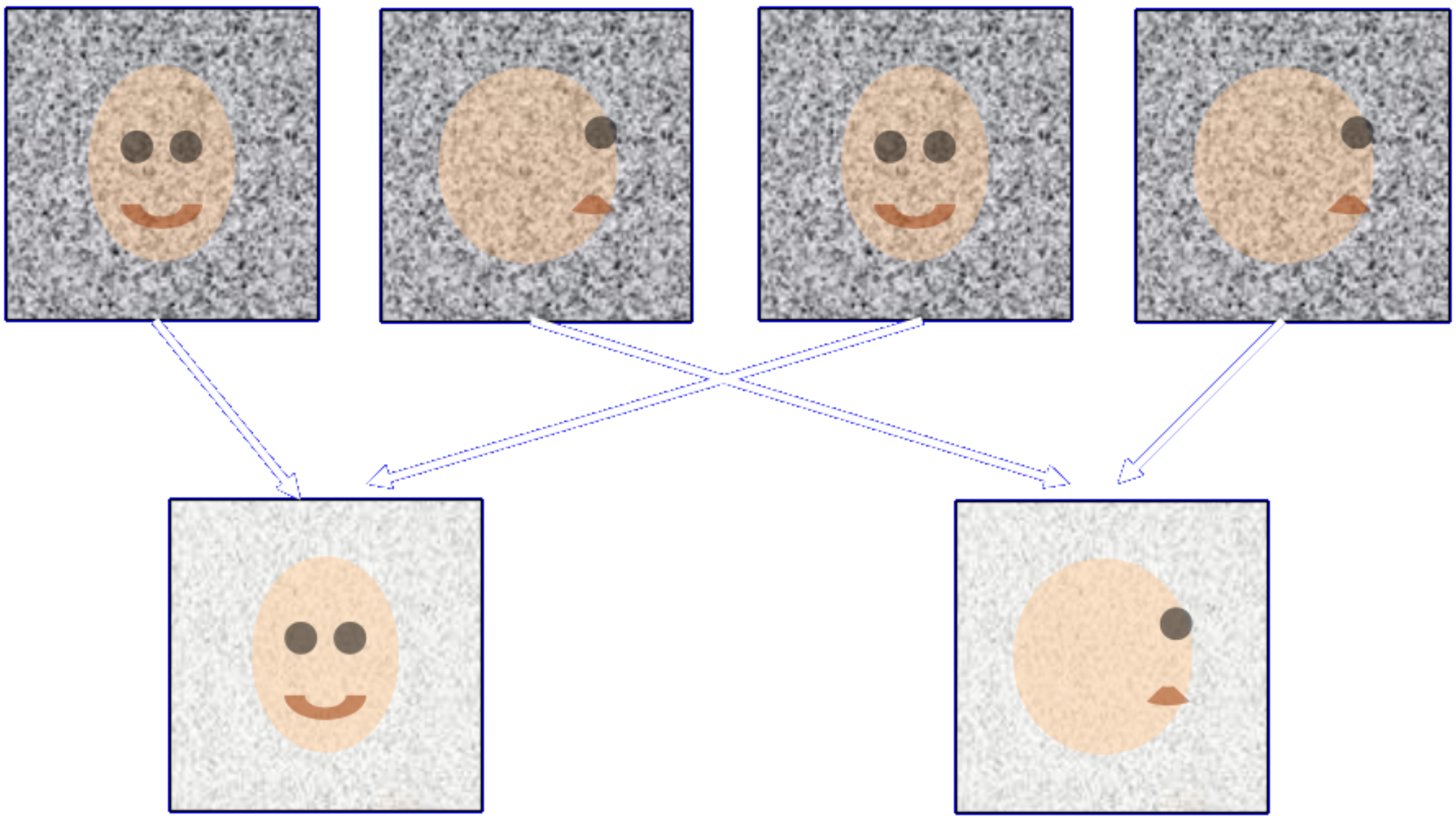
A. Patwardhan

- Rotation (requieres a reference). Again done using cross-correlations.



A. Patwardhan

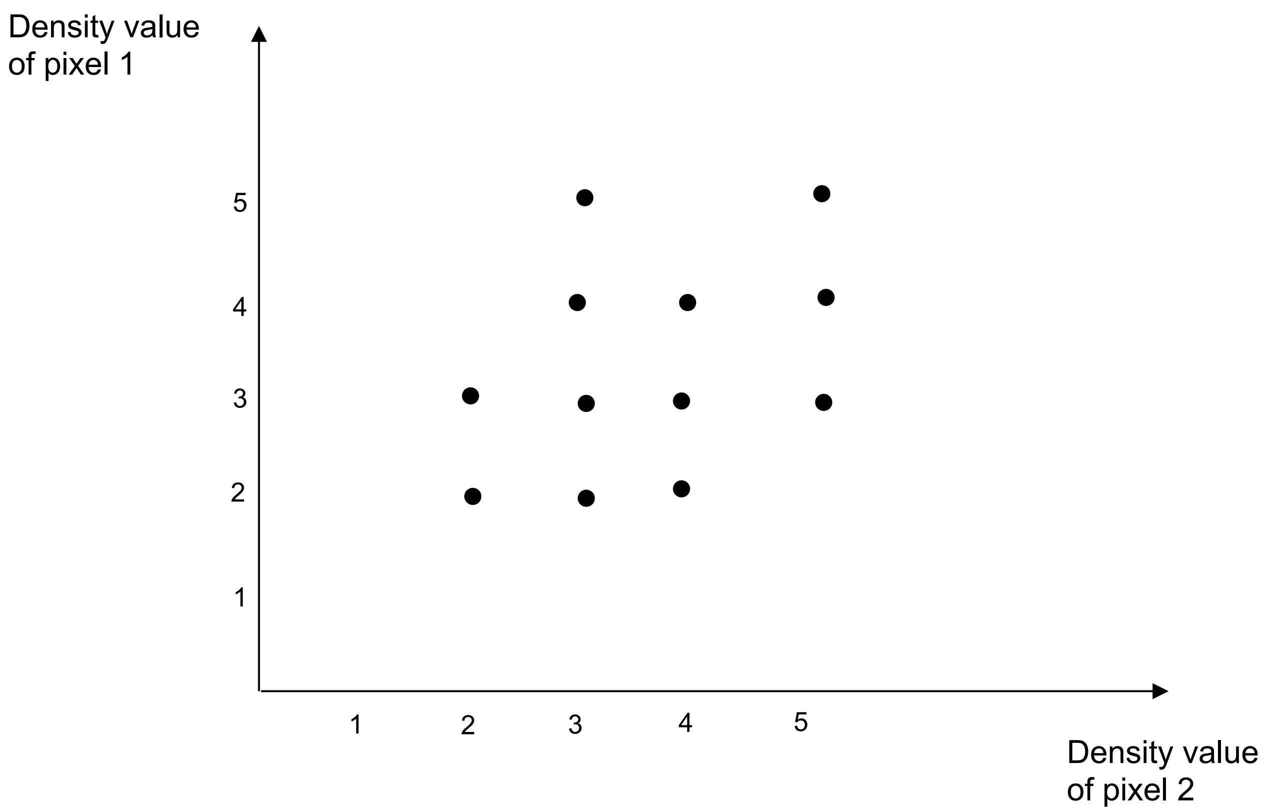
- Centered and rotated images with low signal to noise ratio

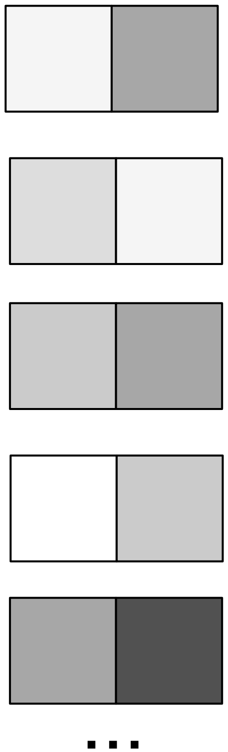


- Class averages with higher signal to noise.

Principle of Multivariate Statistical Analysis to obtain 2D class average

- Let's consider the simplest case: images with only 2 pixels

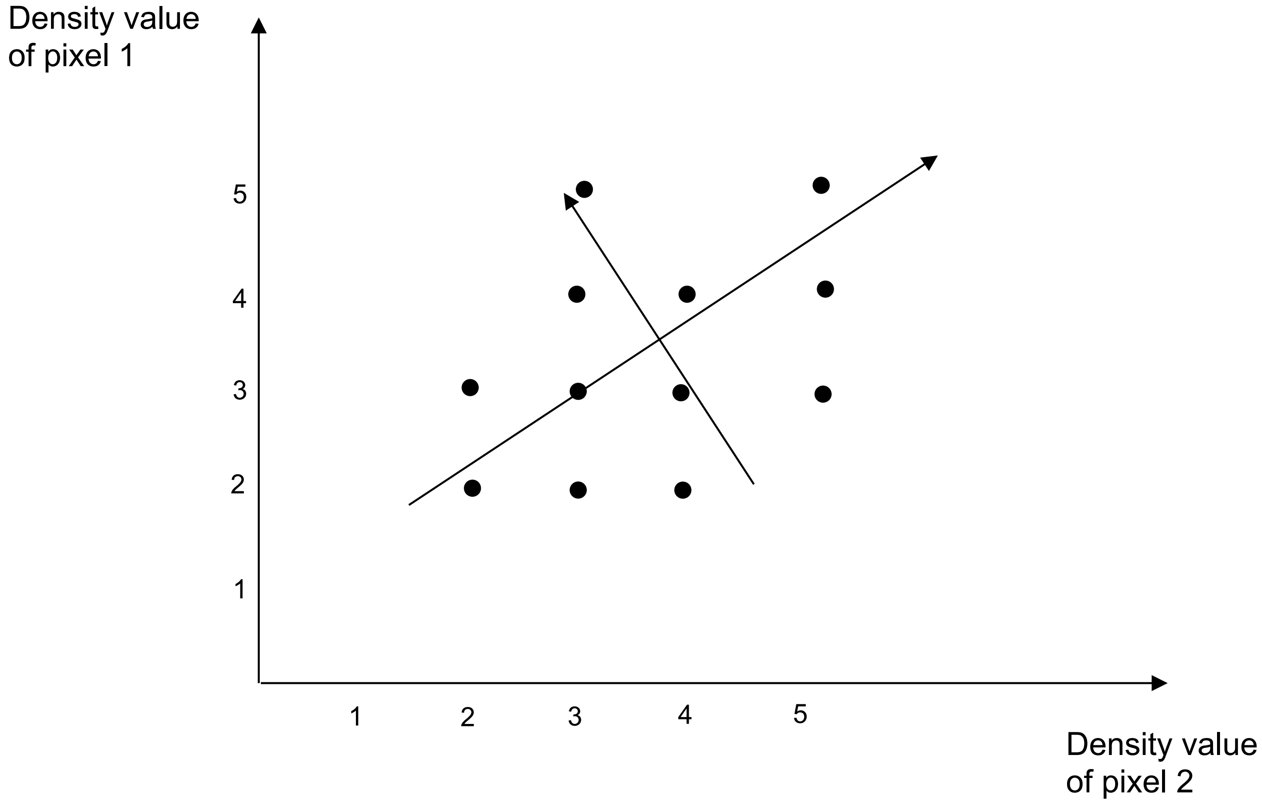


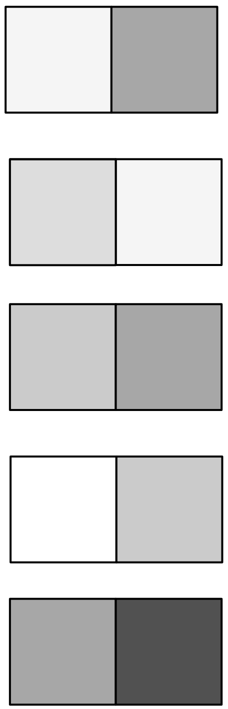


Adapted from B. Klaholz

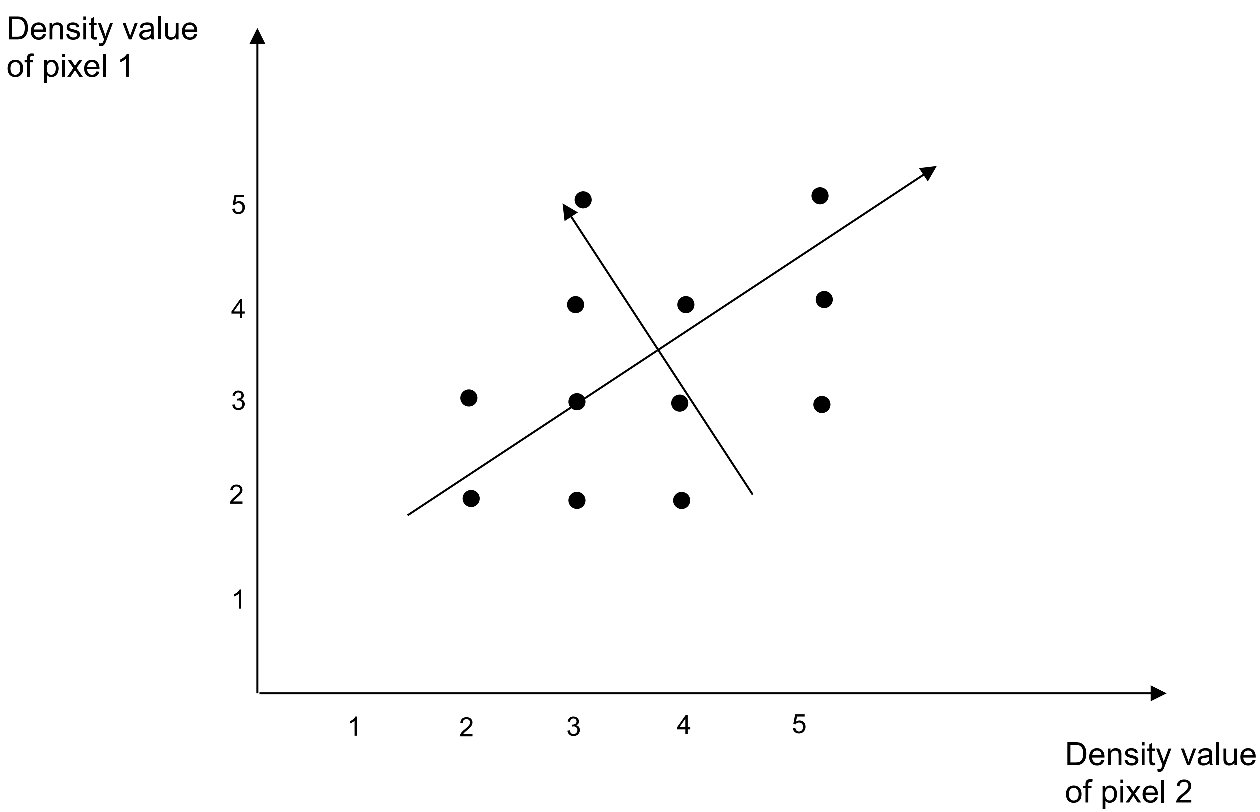
Aim of the MSA:
adapt the coordinate system to the shape of the data cloud

- 1st axis: longest elongation of the data cloud, i.e. highest variance
- 2st (orthogonal) axis: corresponds to the next strongest variance

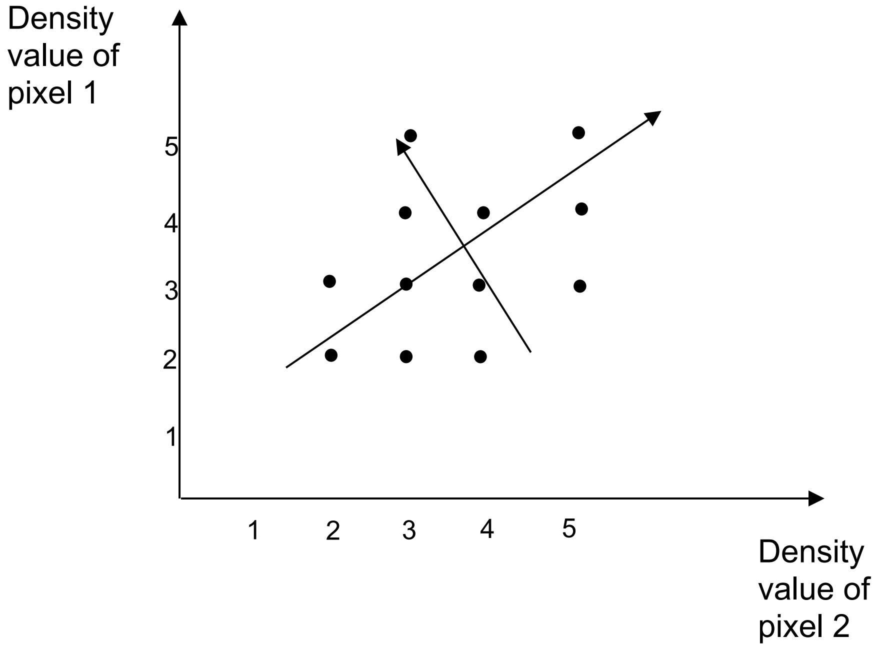




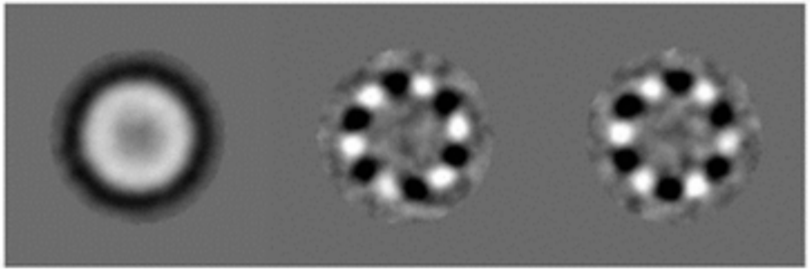
- If we now have images with 100x100 pixels. Instead of having two axes, we will have 100x100 axes….
- But we can still rotate the coordinate systems to make it correspond to the largest elongations of the data cloud.
- Only few axes = eigenvectors will correspond to the main directions of variations. This new coordinate system will be used for an MSA analysis ➡ data reduction



- Eigenimages will be calculated and will correspond:
  - for the 1st one to the total sum
  - for the others: to the higest variances of the dataset

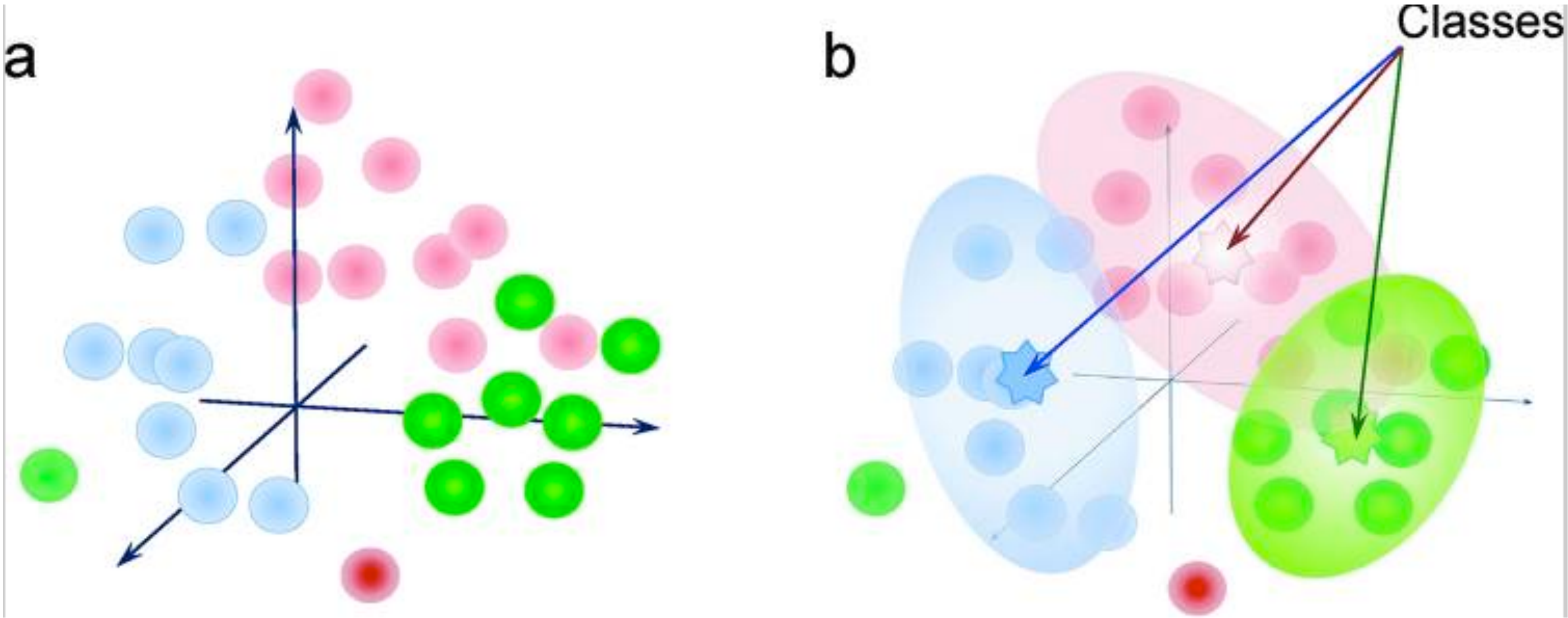


Eigenimages



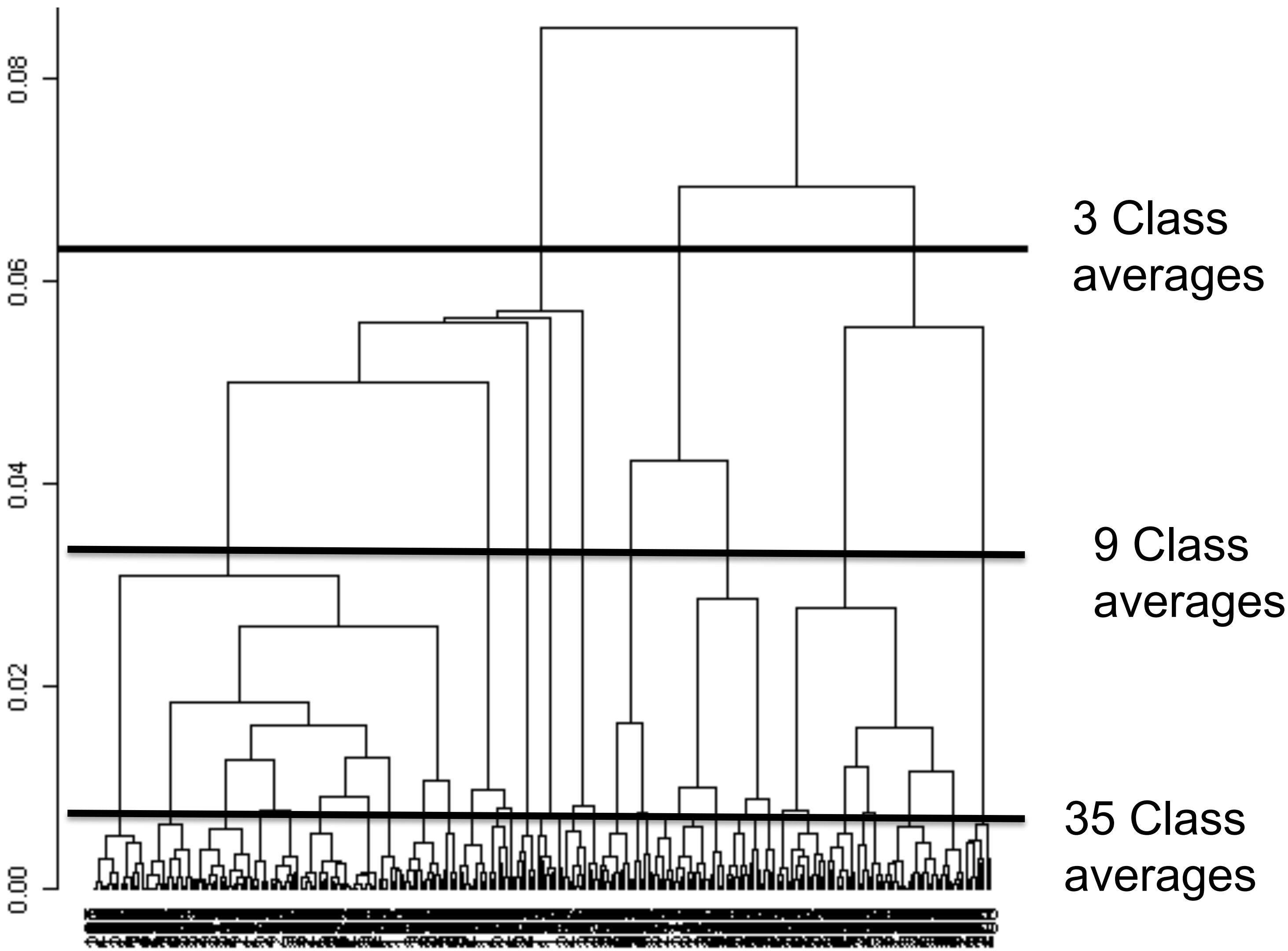
Clustering

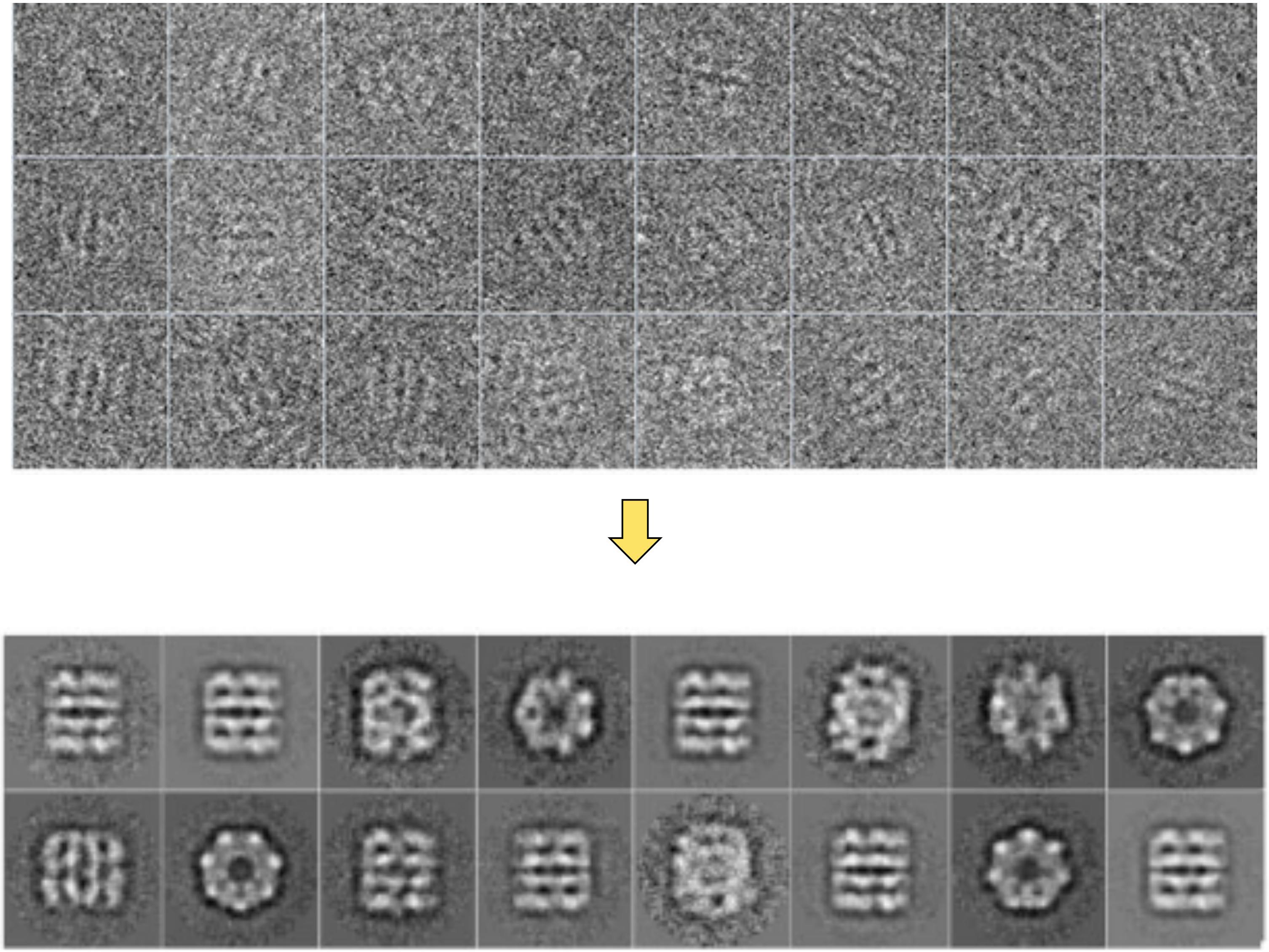
- minimization of the intraclass variance in a cluster (between the members of the cluster)
- maximization of the interclass variance between the centers of mass of the clusters



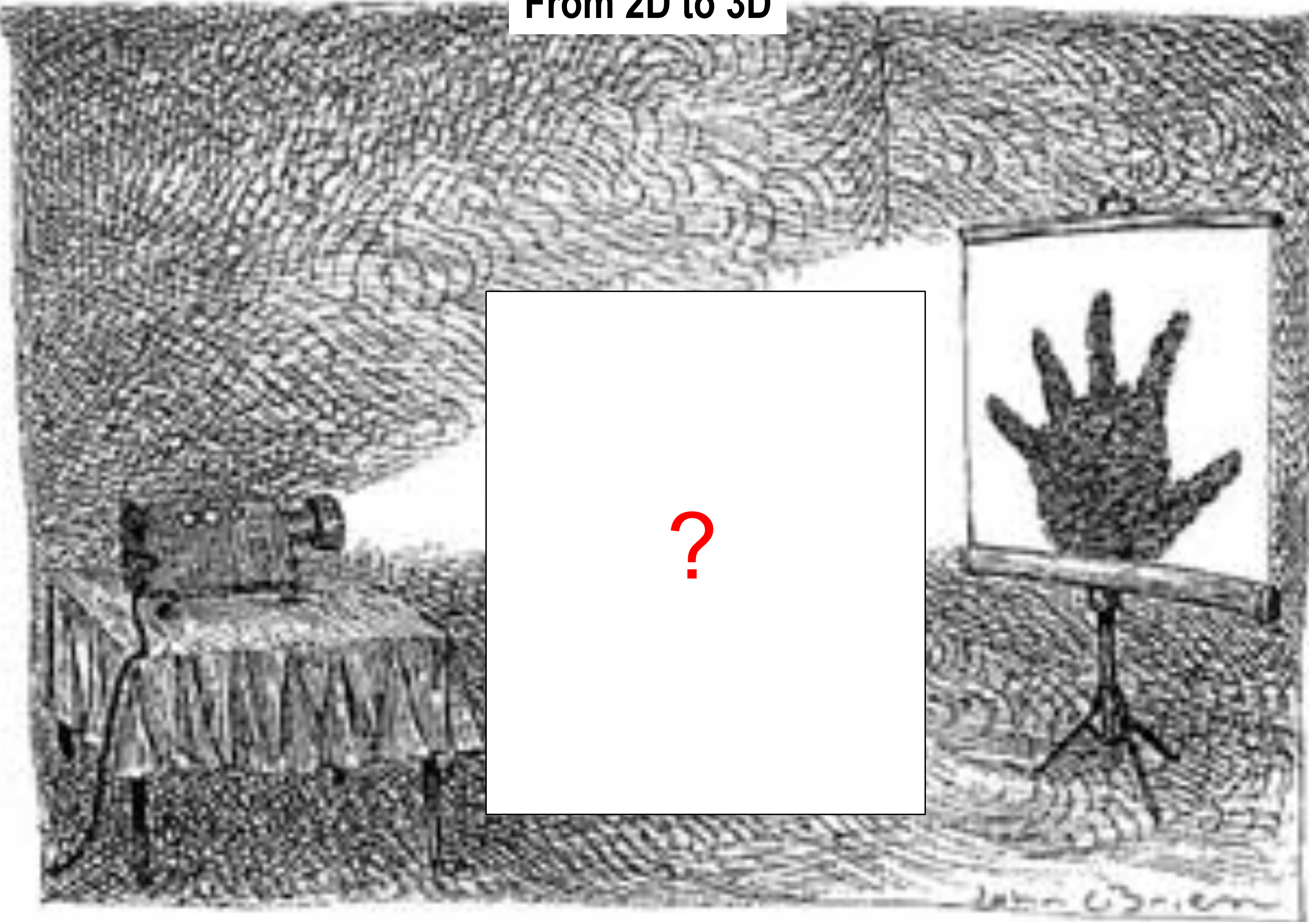
Orlova, Saibil, Chem Rev. 2011

Hierarchical (agglomerative) ascendant classification



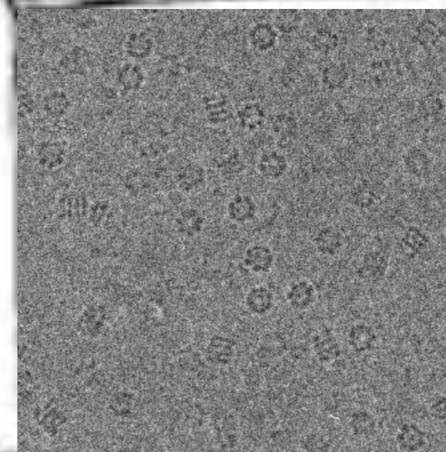


From 2D to 3D

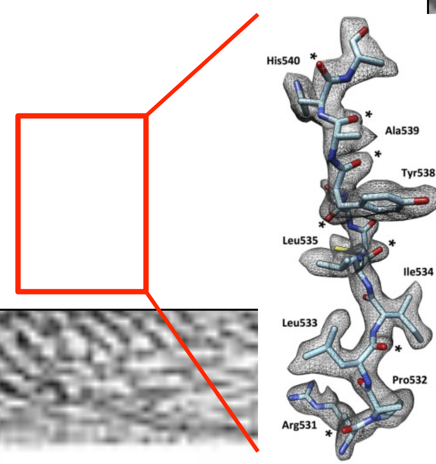


From 2D to 3D

Electron microscope

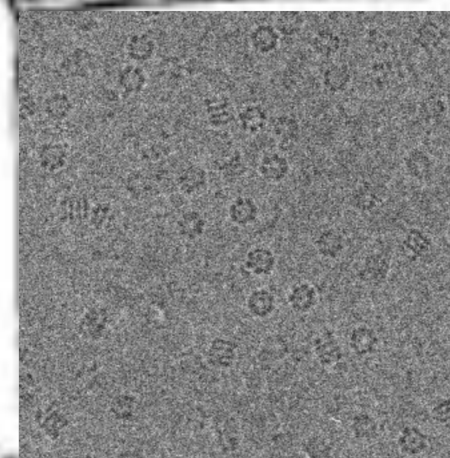


?

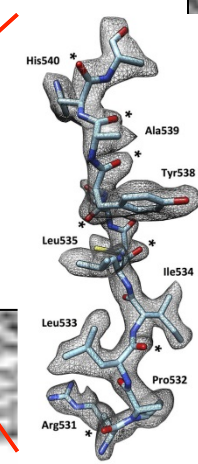


From 2D to 3D

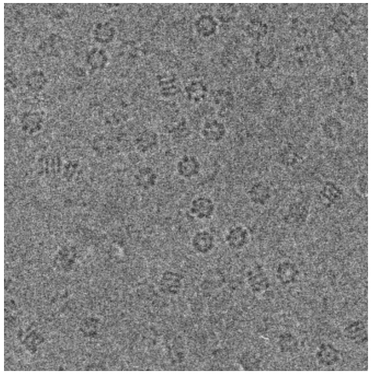
Electron microscope

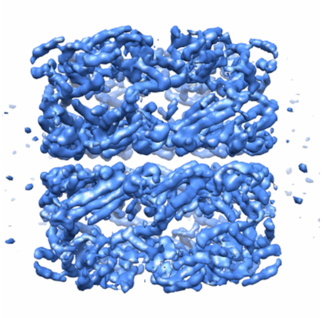


?



From 2D… … To 3D



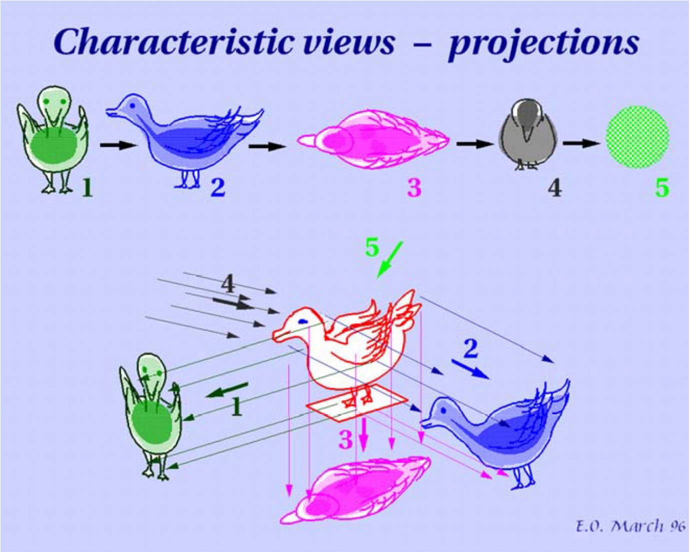


- Step 1: determine orientations

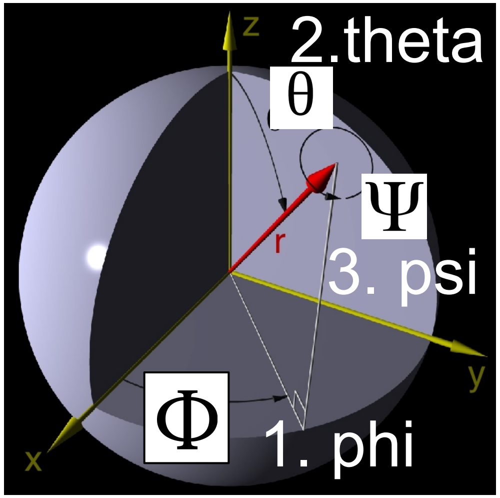
- Step 2: combine these orientations to obtain a 3D reconstruction

- Step 3: refine the structure

Step 1: How to determine orientations?

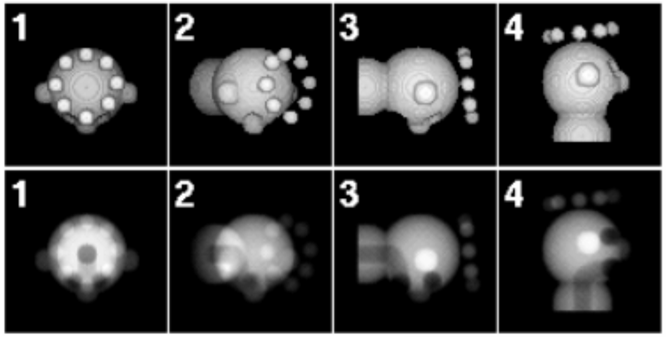


- We need to assign to which view which image corresponds.

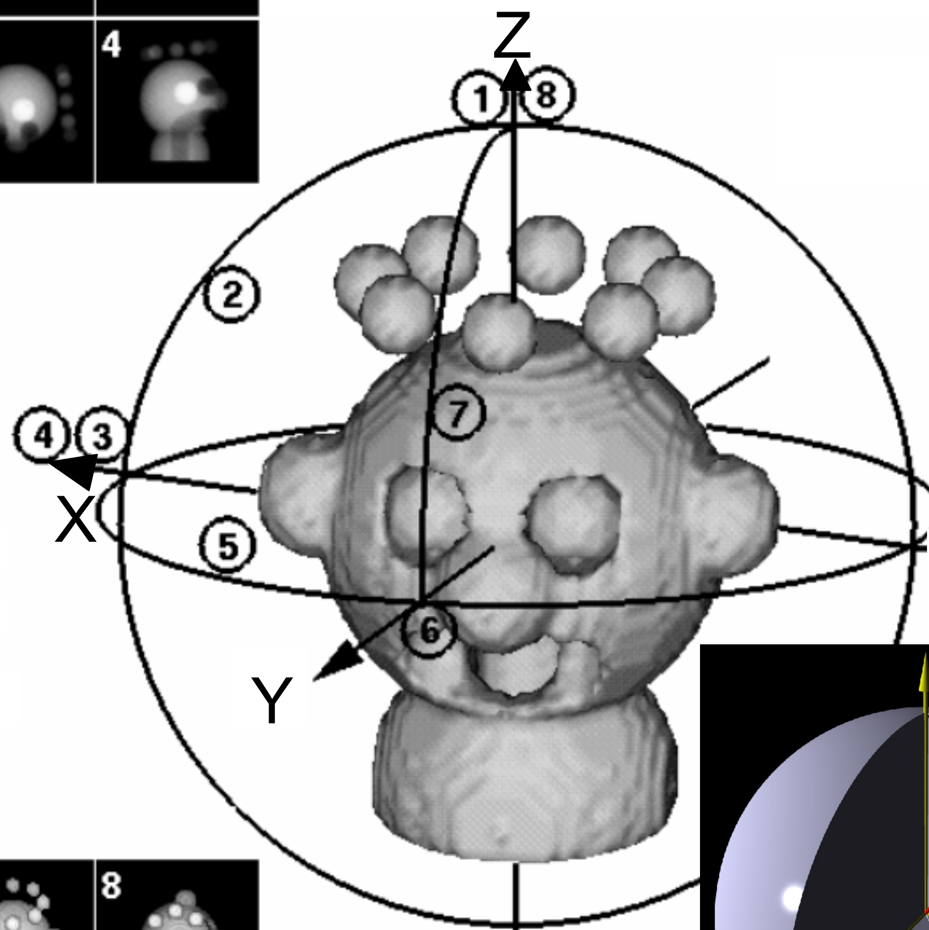


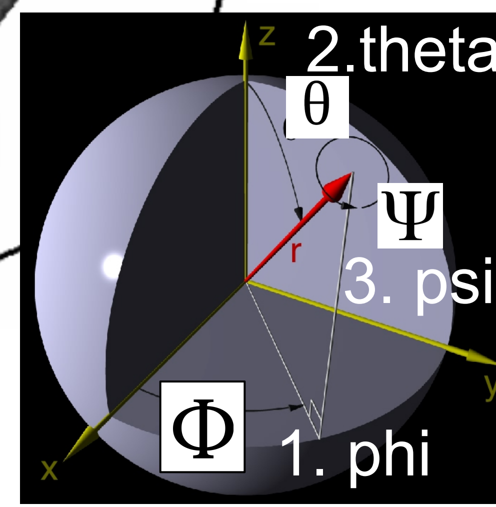
- We need to assign euler angles to images !

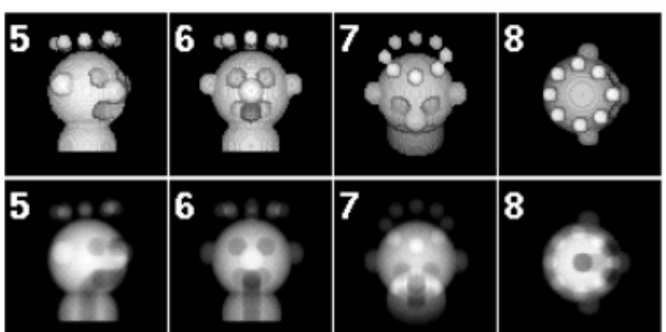
Euler angles



	Phi	Theta	Psi
	φ	θ	ψ
1:			
2:			
3:			
4:			
5:			
6:			
7:			
8:			







From Nicolas Boisset

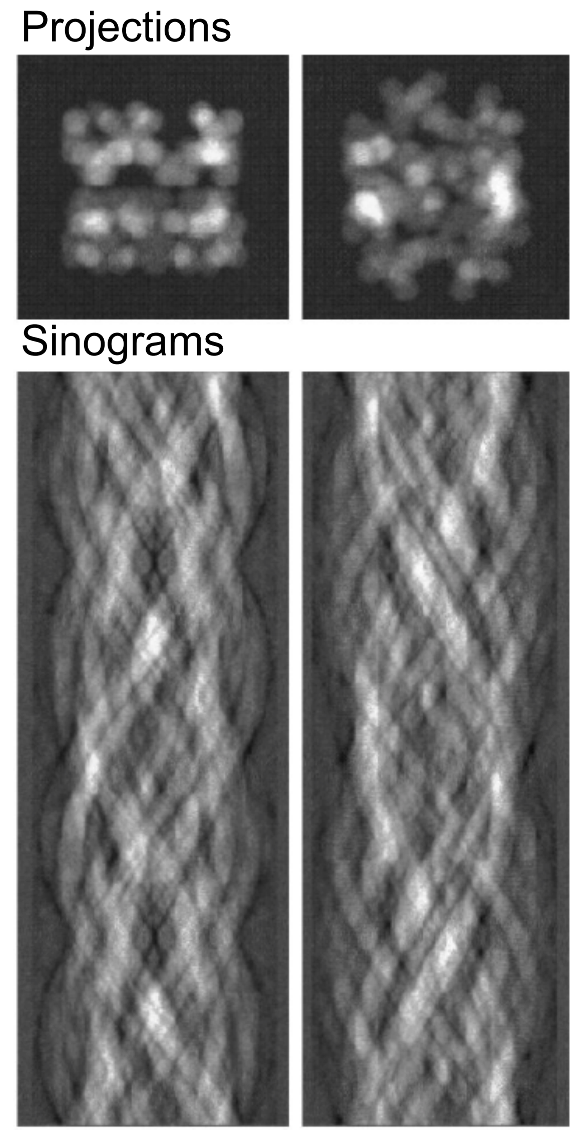
Step 1: How to determine orientations?
1st method: with 2D class averages
Angular reconstitution

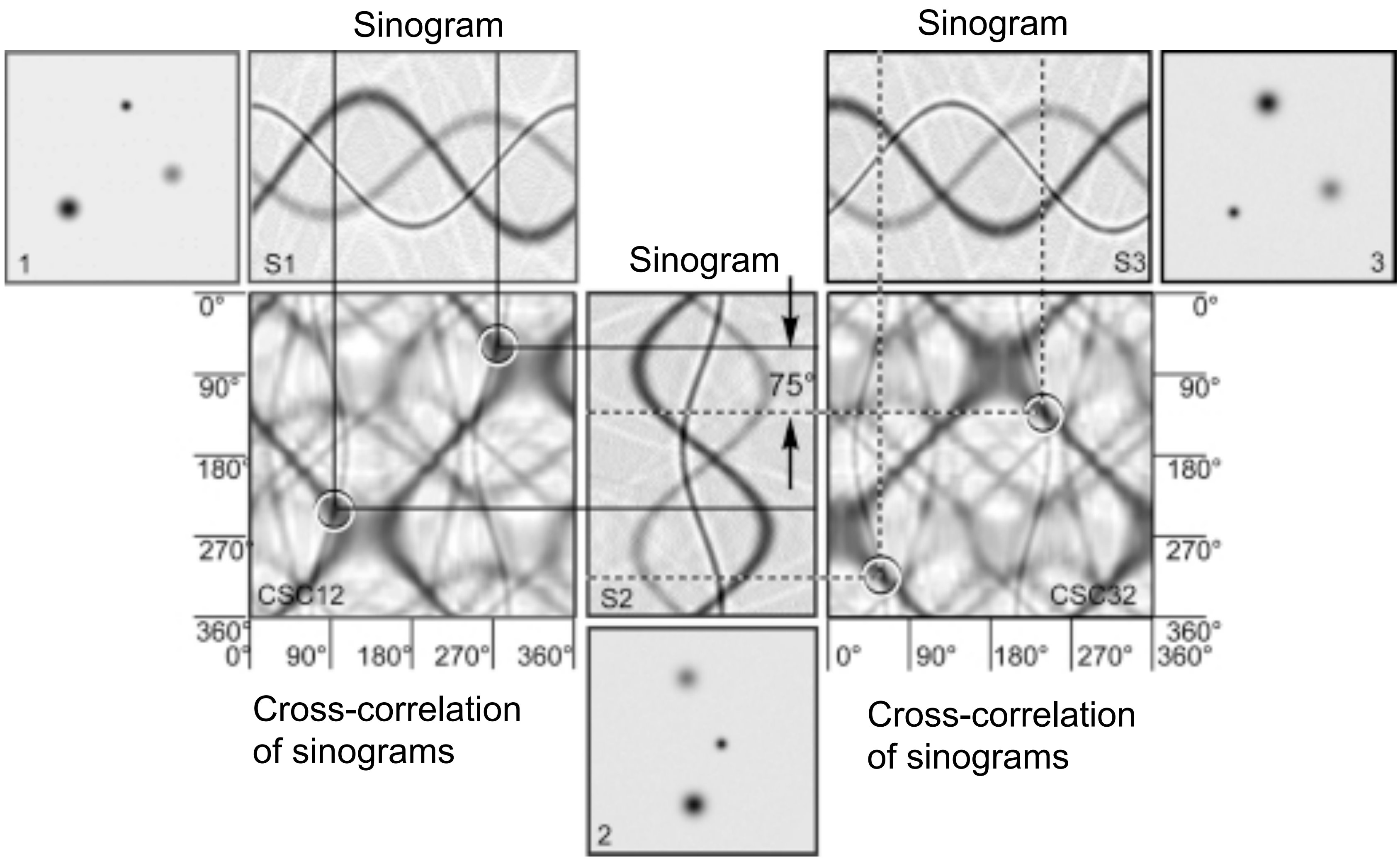
- Determination of the euler angles of the projection: computationally, using the common line theory

"two projections of the same object share a common line"

- Done on 2D class averages

- Necessity to be centered

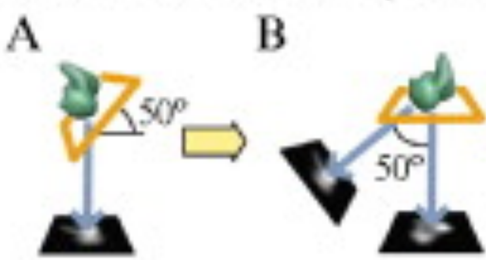


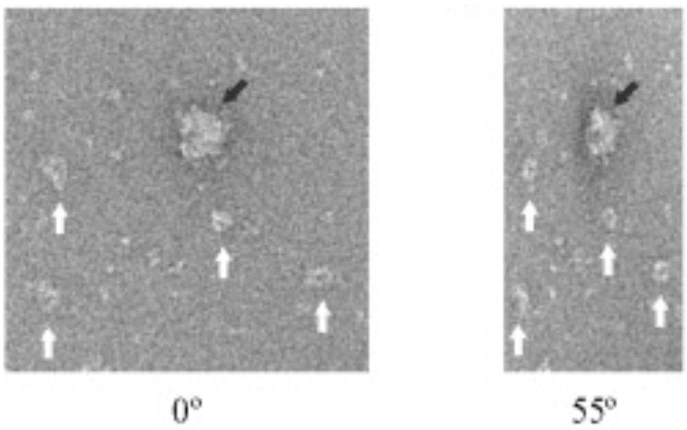


Orlova, Saibil, Chem Rev. 2011

Step 1: How to determine orientations?
2nd method: with data collection of tilt pairs
Random conical tilt

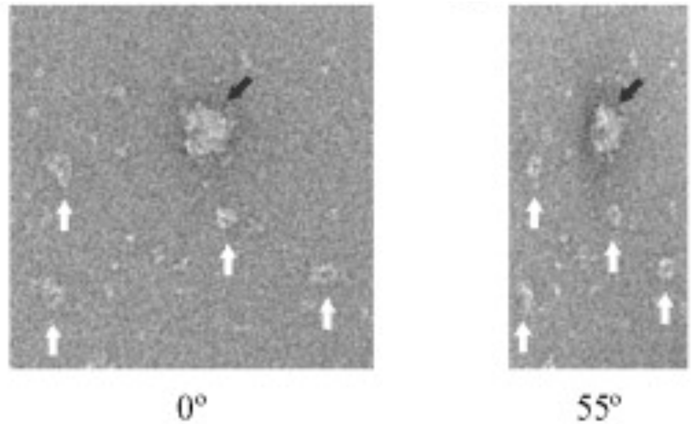
- From projection pairs (usually 0, 60 degree)





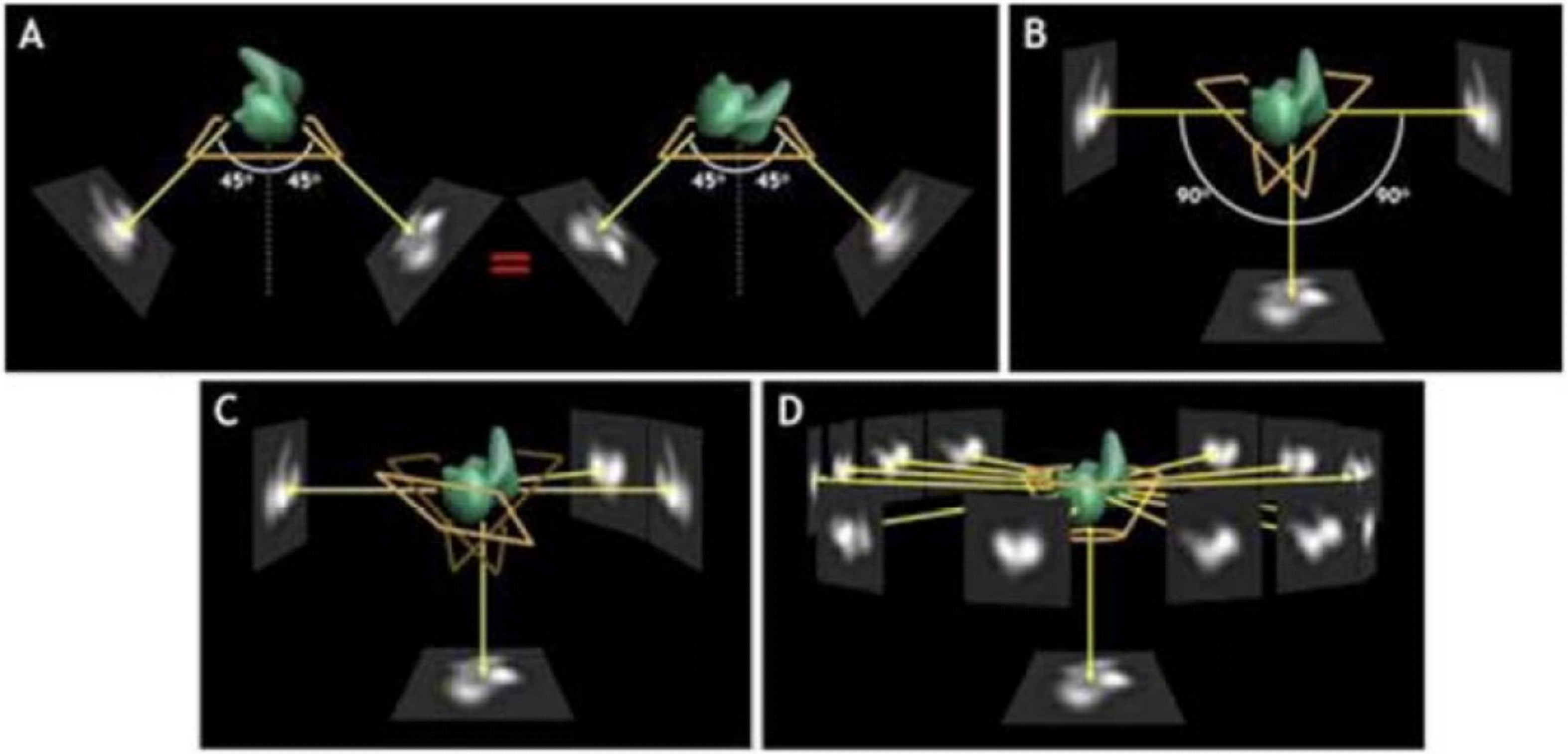
- Possible to assign relative angles between images

From Leschziner AE



From Leschziner AE

- Can also be done with images at -45 and 45 degree
- Orthogonal random conical tilt



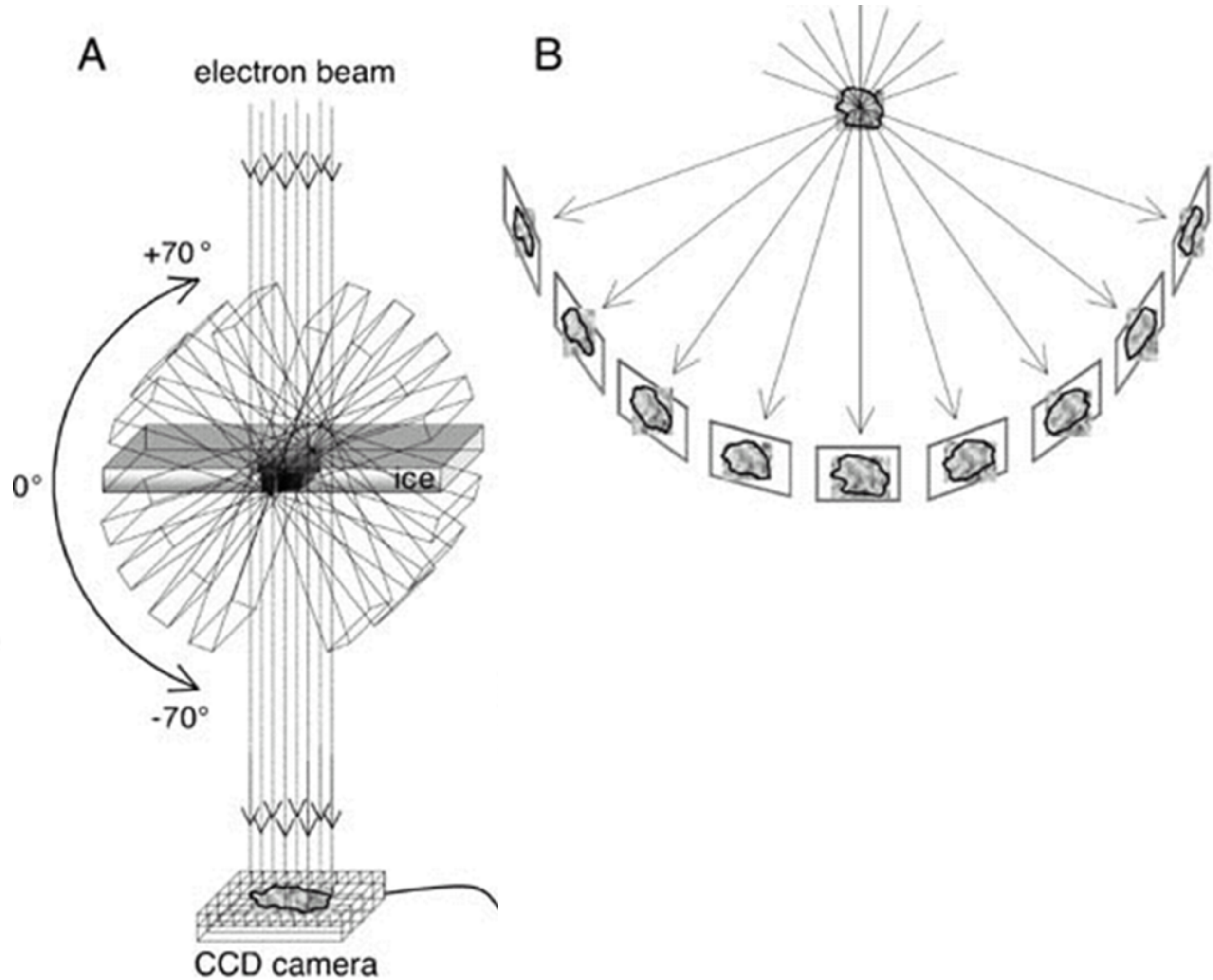
From Leschziner AE

Step 1: How to determine orientations?

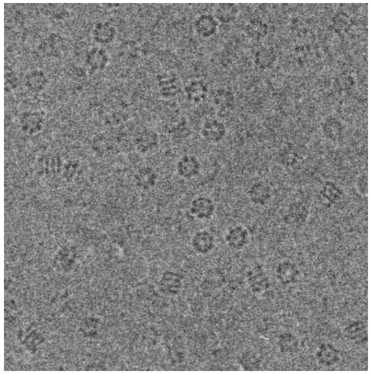
3rd method: tomography

- Advantages:
  - Experimental determination
  - No classification

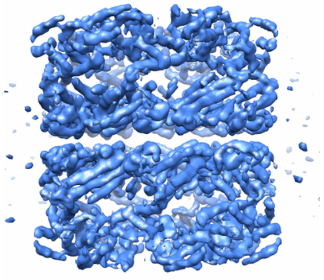
- Disadvantages:
  - Dose issues
  - Artefacts (we'll see later why…)



From 2D…



… To 3D

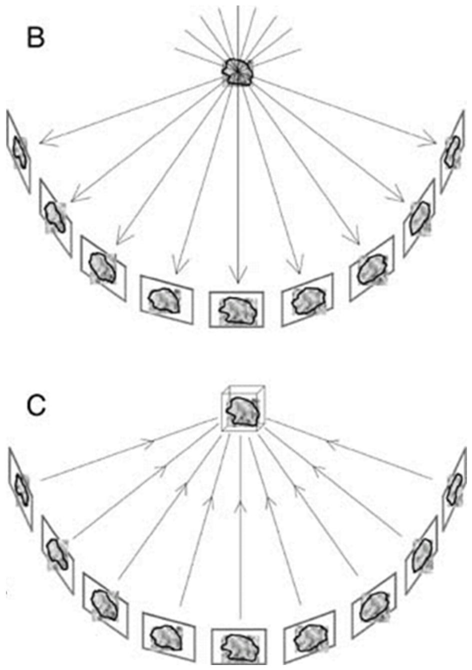


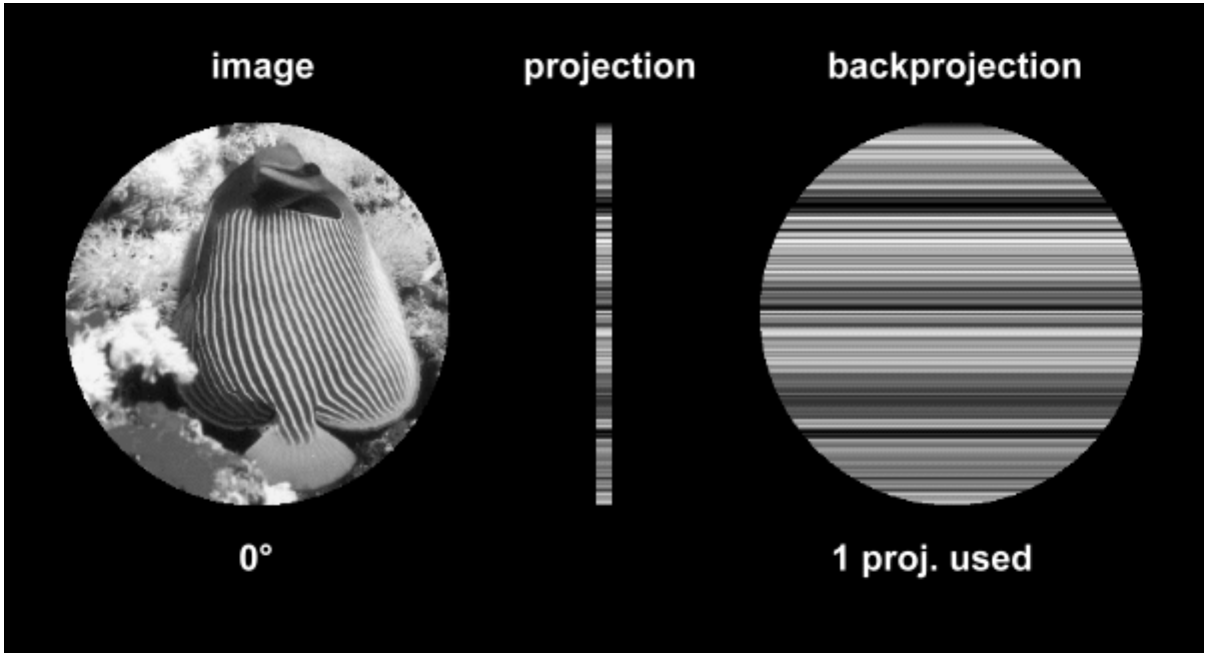
- Step 1: determine orientations

- Step 2: combine these orientations to obtain a 3D reconstruction

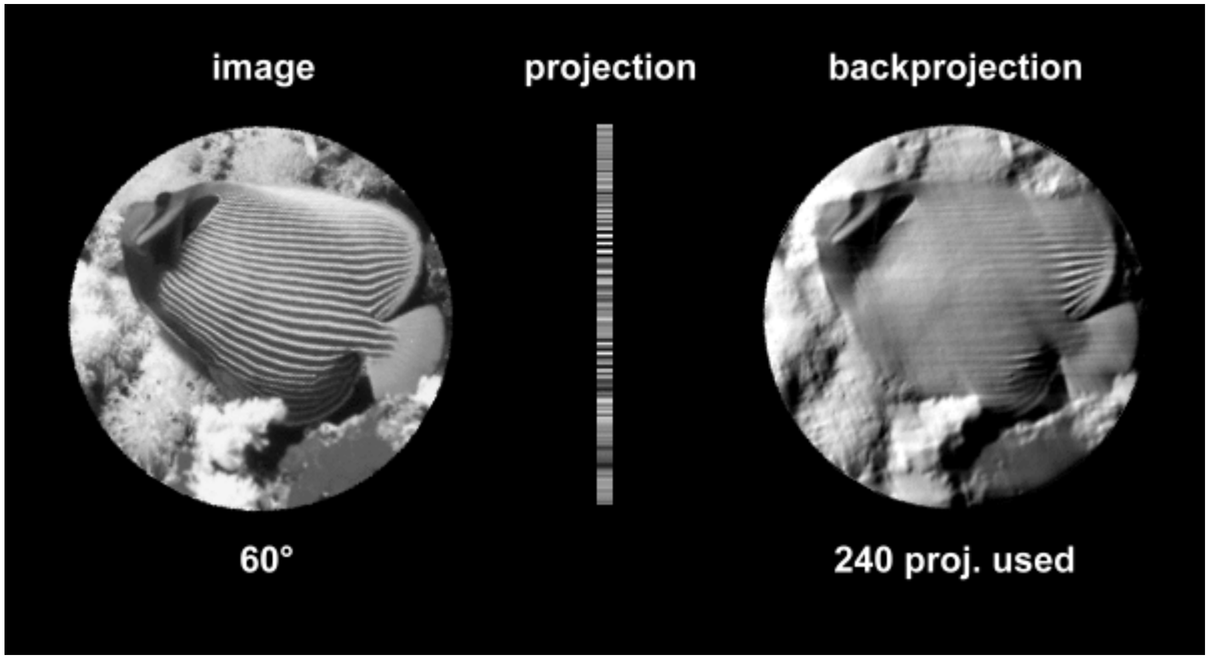
- Step 3: refine the structure

Step 2: How to combine these orientations to obtain a 3D reconstruction?



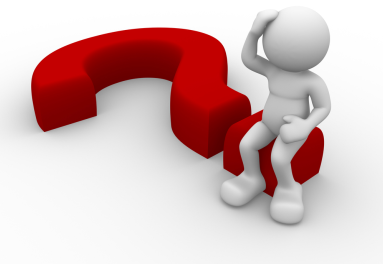


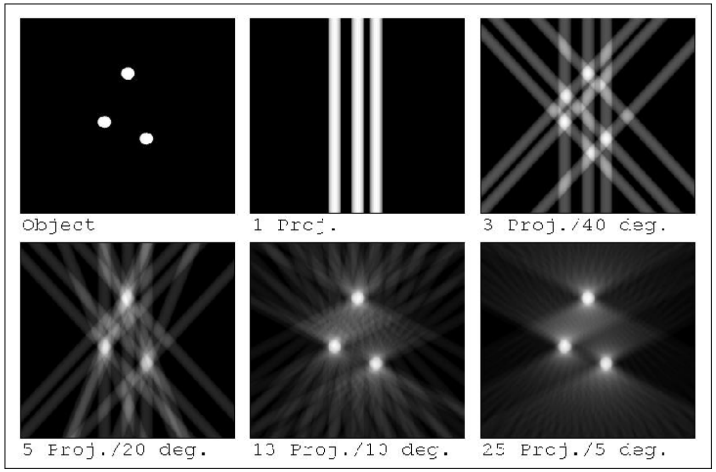
- By backprojection !
- One image : not suffisant. But if we know views from 0 to 60 degrees each ¼ of degree ?



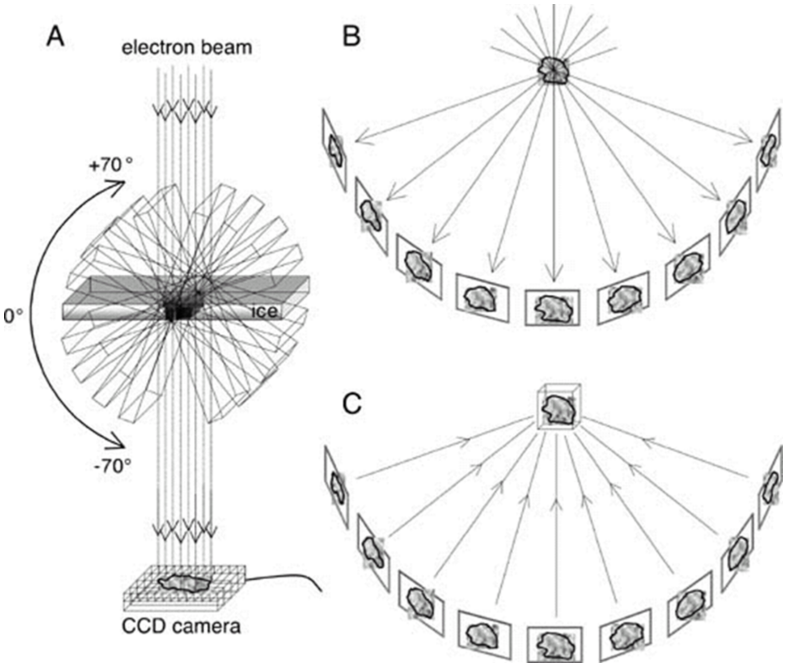
- Much better ! But still distorted. We should go to 90 degree to see no distortion

- Do we encounter such a problem of missing angles?

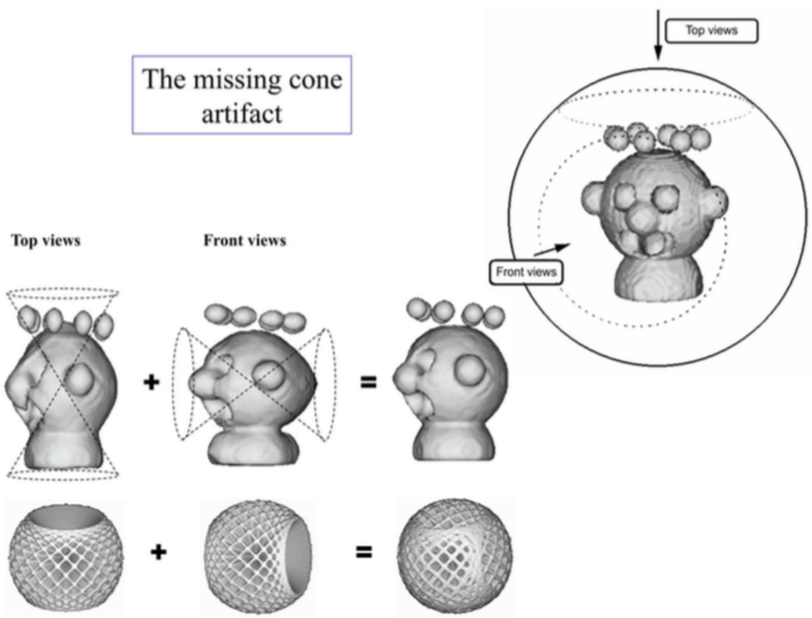




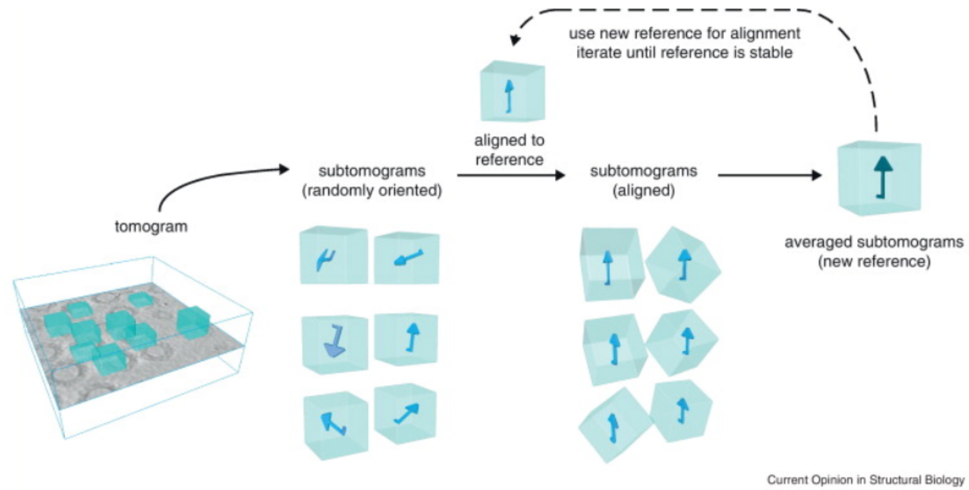
- Yes in random conical tilt and tomography



- Solution: combined different reconstructions with different missing cones

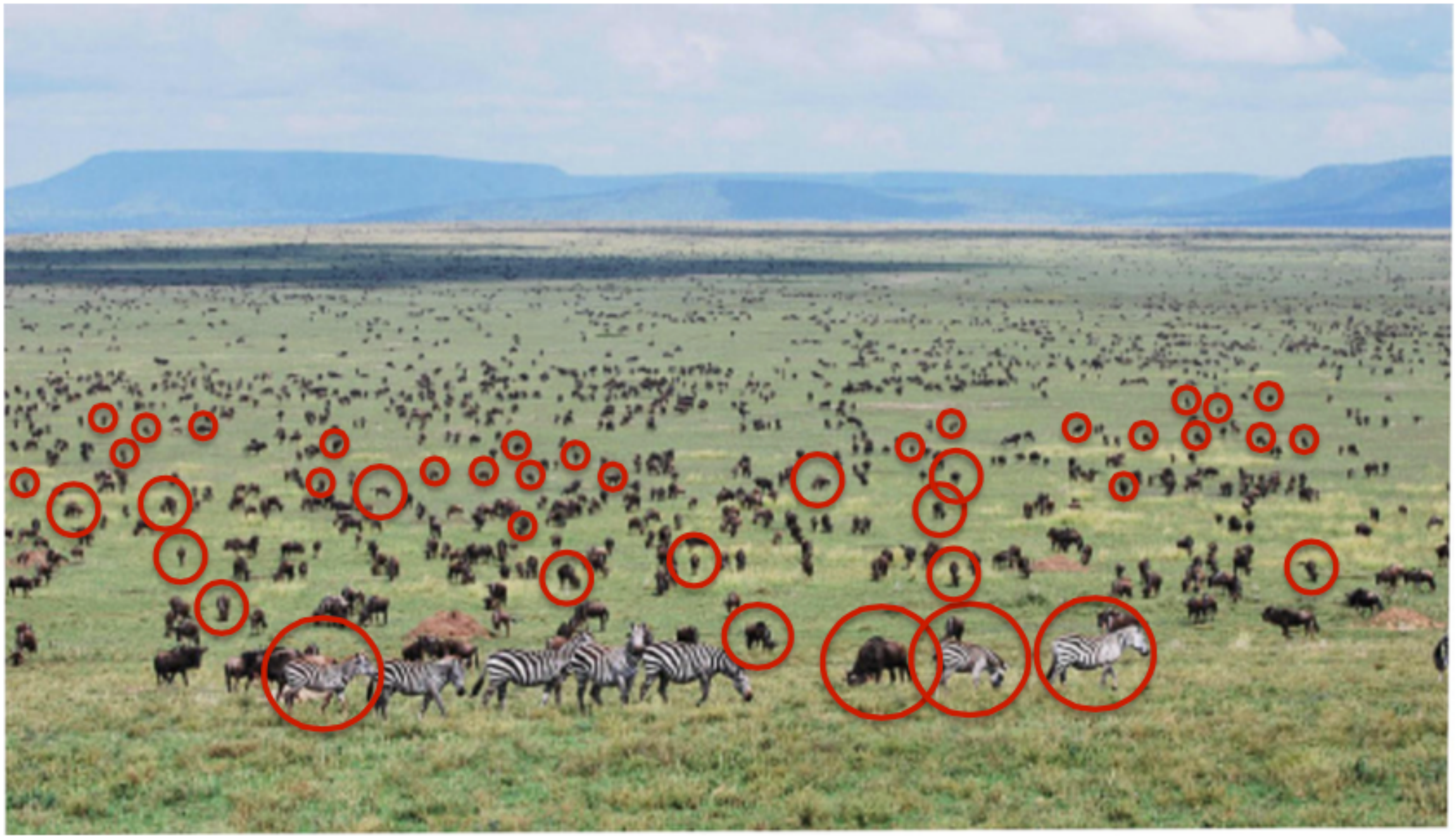


From N. Boisset

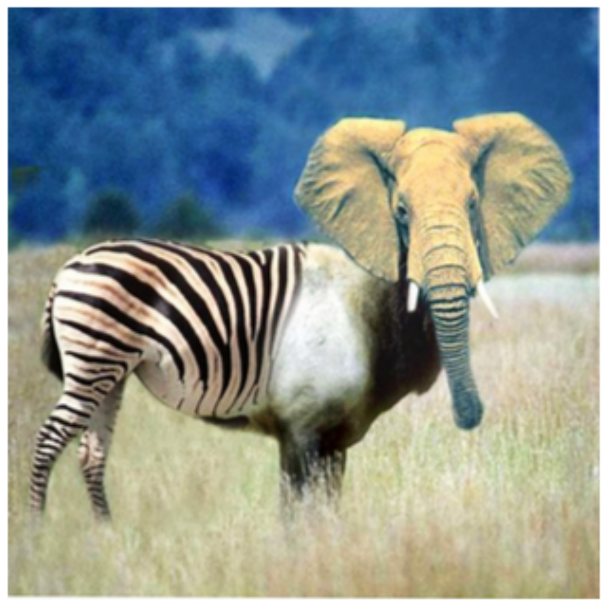


From J. Briggs

Caution: you can always get a 3D volume from your data



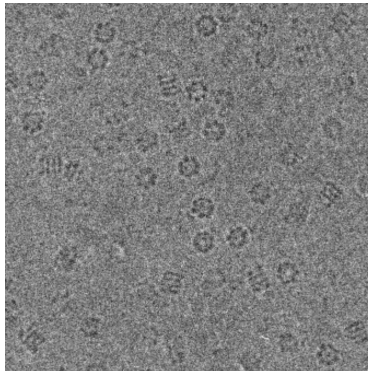
Caution: you can always get a 3D volume from your data

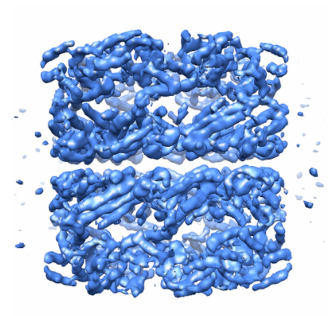


Great ! You got your 1st 3D reconstruction !!

How do you make sure it is correct ?

From 2D… **… To 3D**





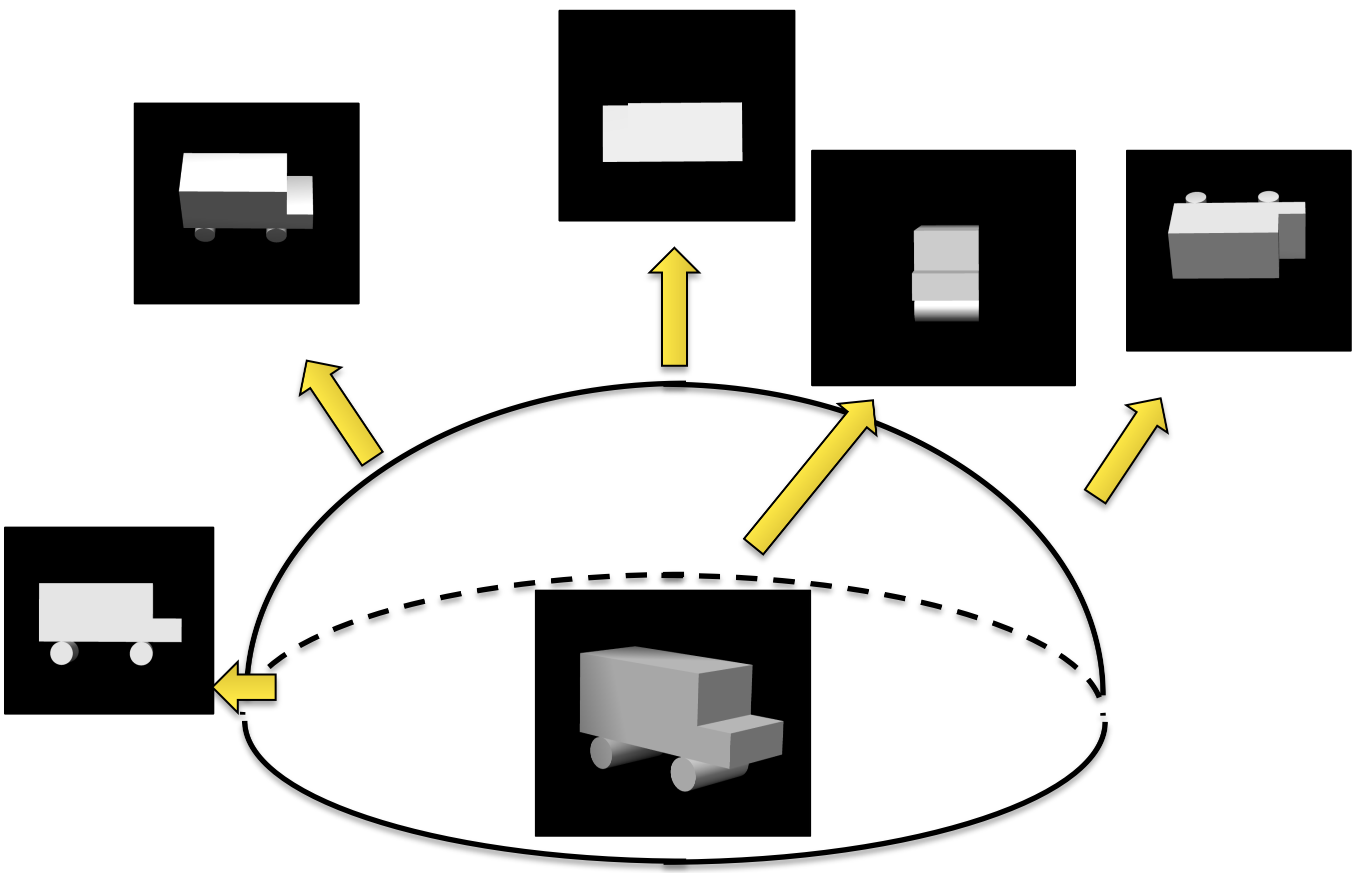
- Step 1: determine orientations

- Step 2: combine these orientations to obtain a 3D reconstruction

- Step 3: refine the structure

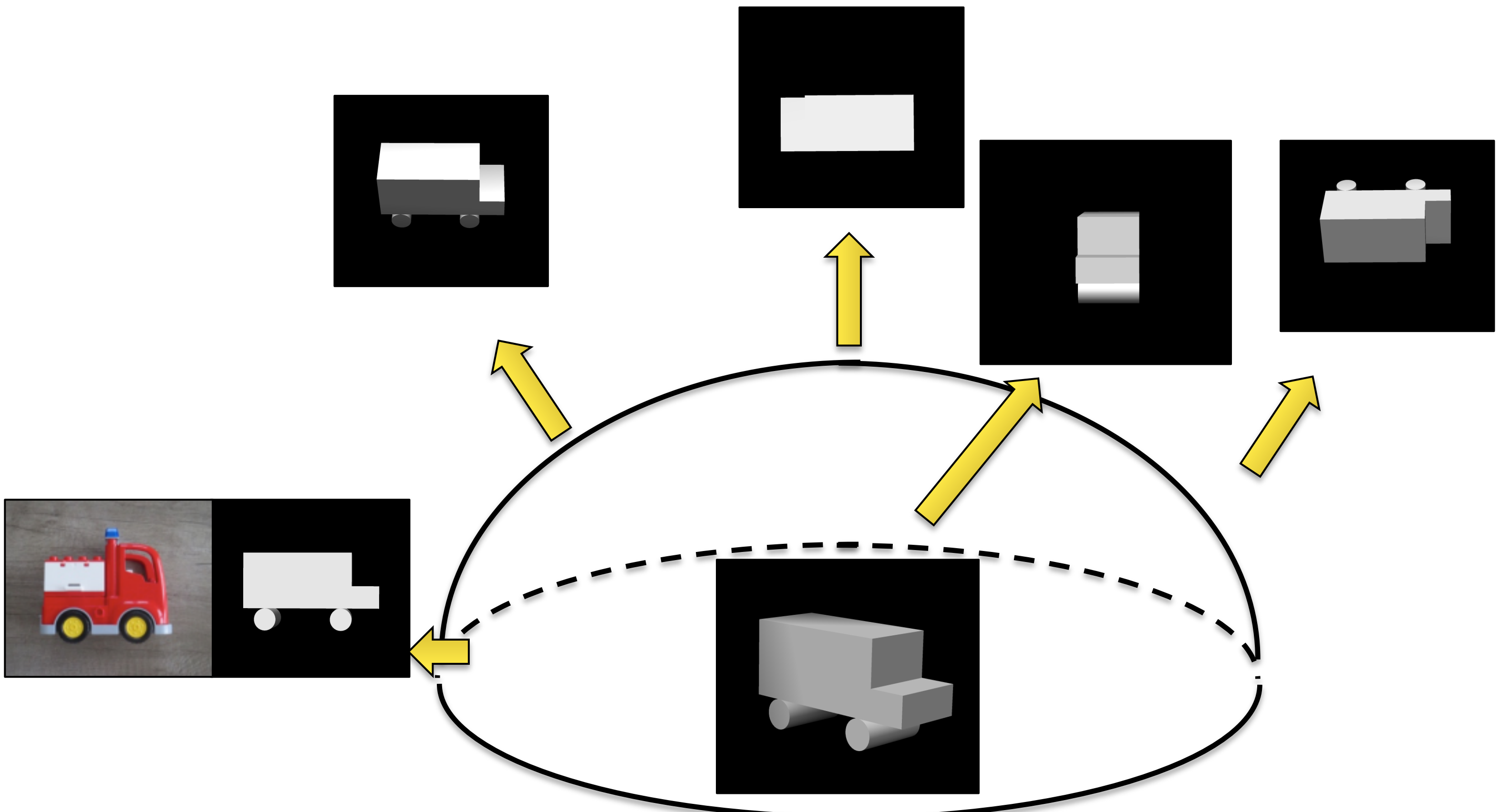
Projections of the initial 3D reconstruction

- Project the initial 3D reconstruction towards all directions (of this asymetric unit)
- Choose even distribution of projections.



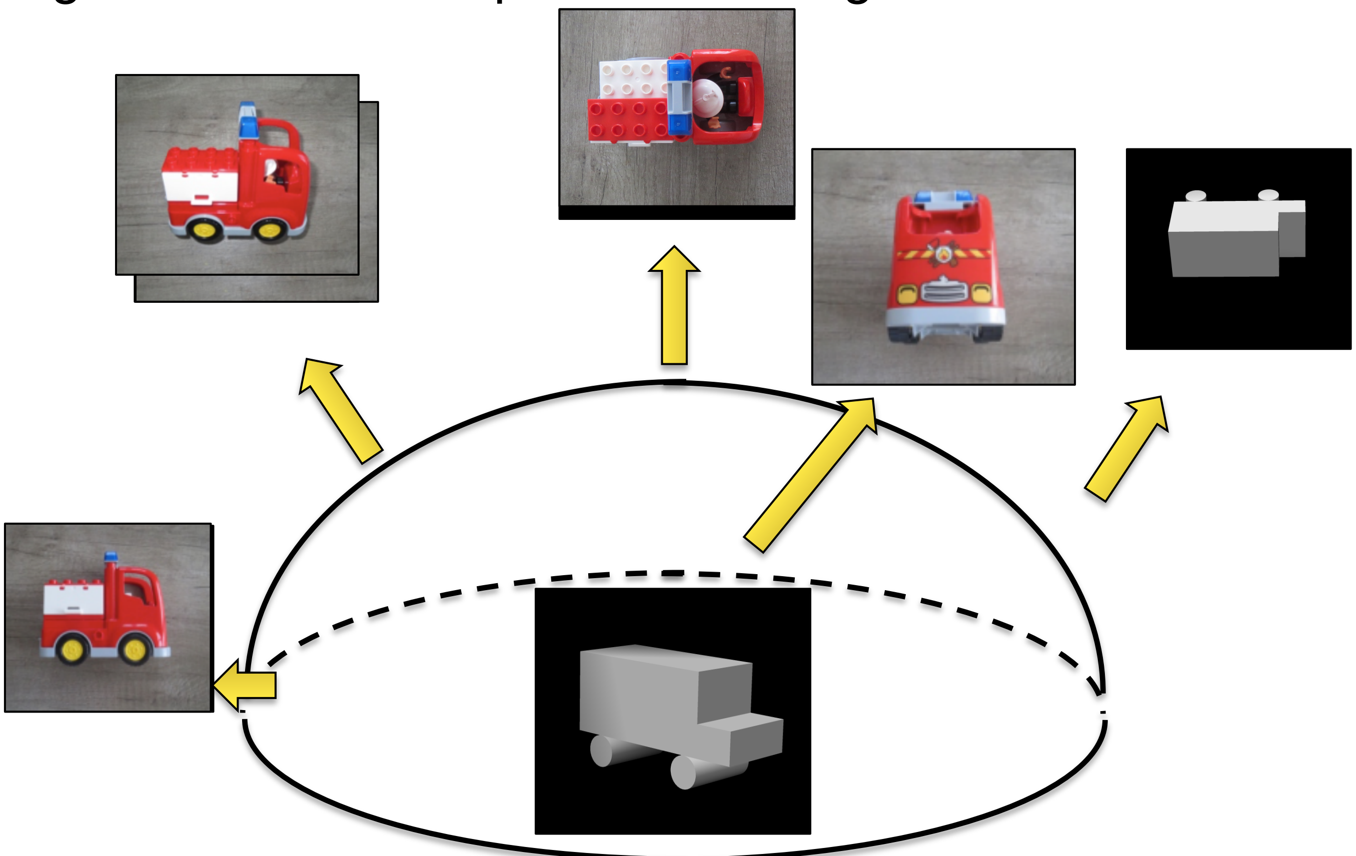
Projection Matching, highest Cross-correlation criteria

- Compare each image to all the projections
- Allow translations and rotations of the image to find the best CC.
- Assign the euler angles of the projection having the highest CC to the experimental image.

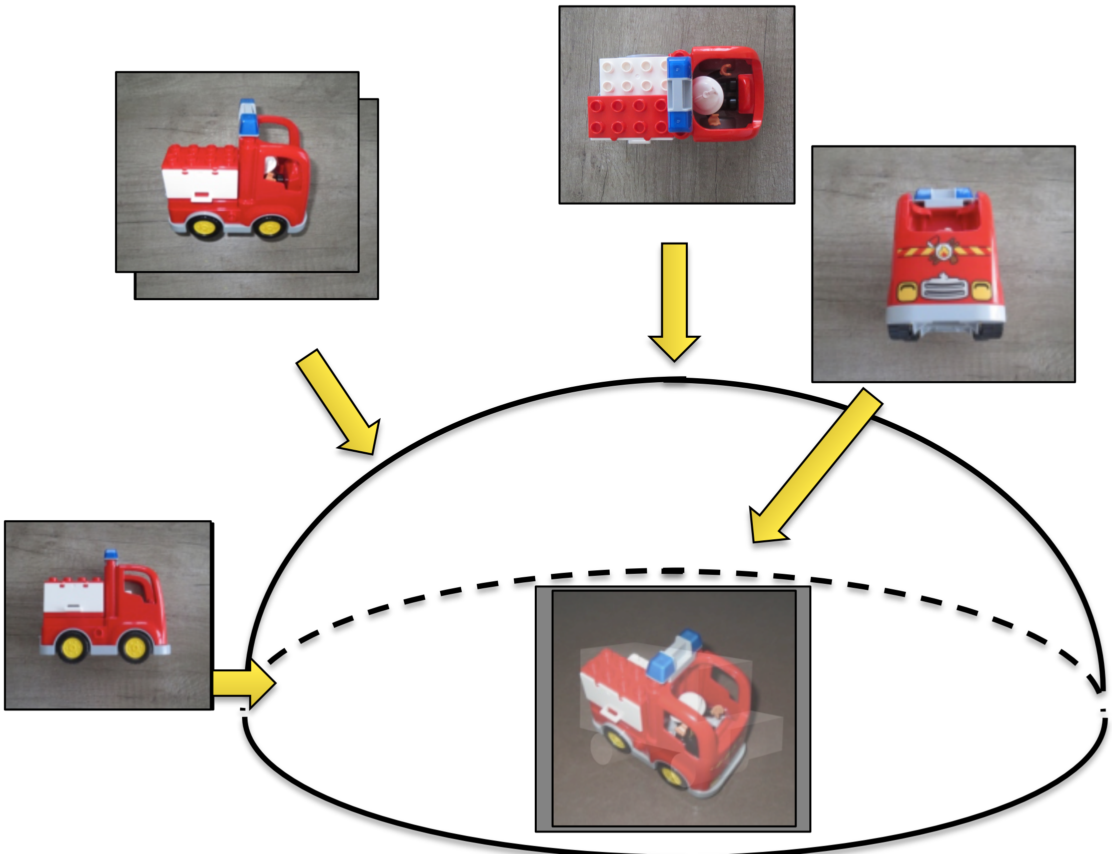


Projection Matching, highest Cross-correlation criteria

- Compare each image to all the projections
- Allow translations and rotations of the image to find the best CC.
- Assign the euler angles of the projection having the highest CC to the experimental image.

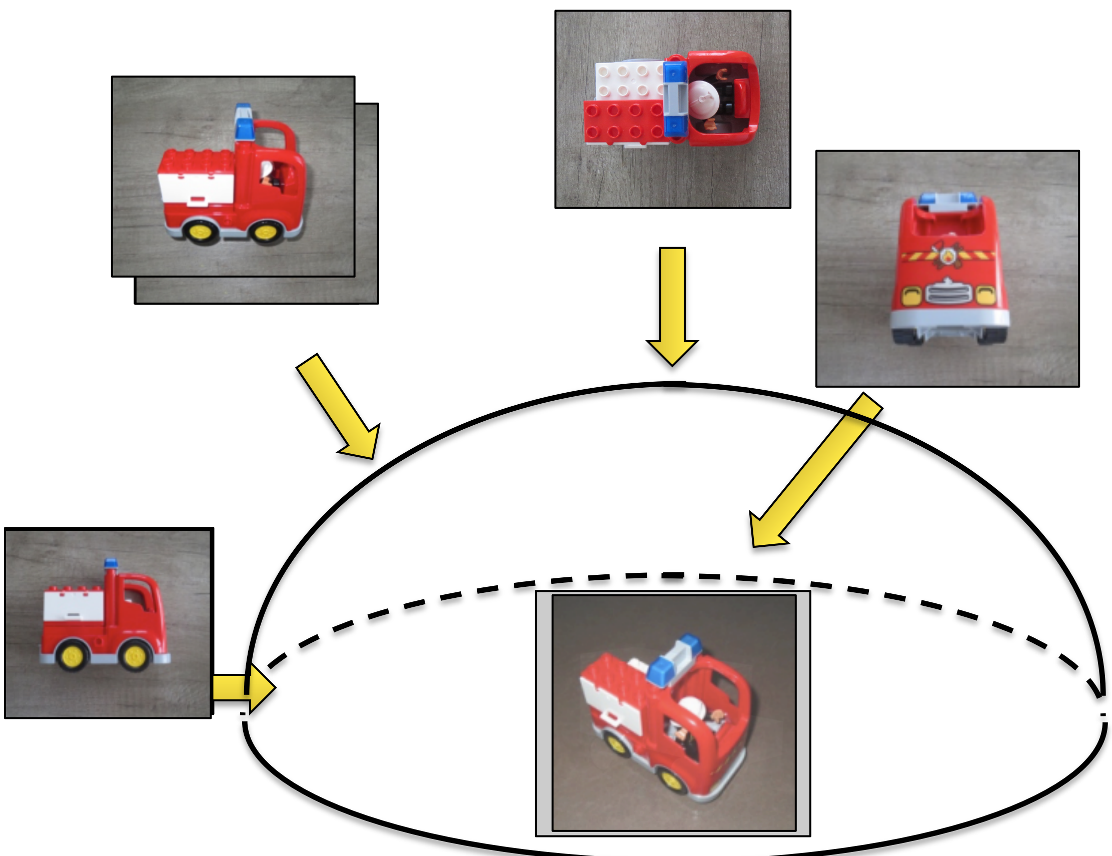


Reconstruction by back projection

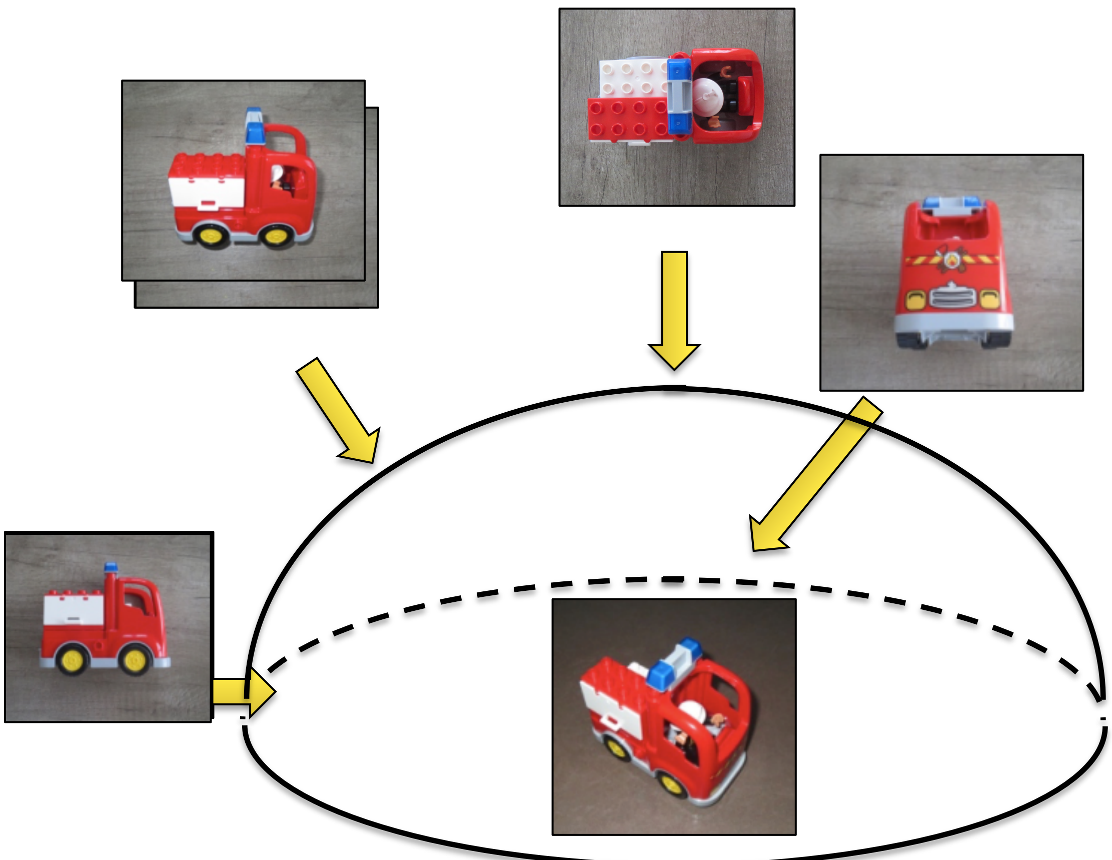


Iterations

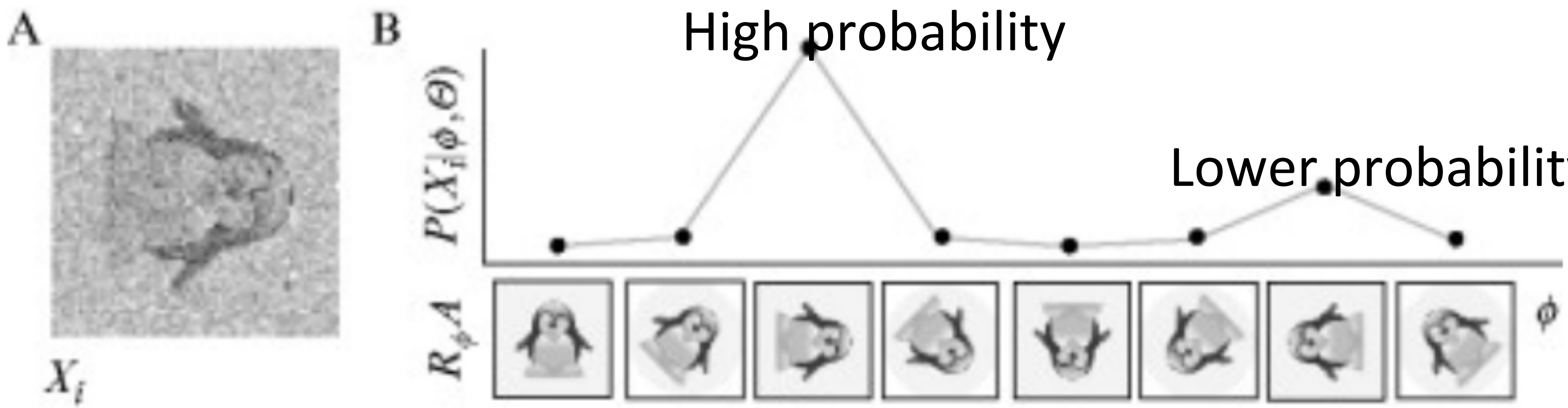
- Projections of the 3D reconstruction
- Projection matching with the highest cross-correlation criterion
- Back-projection



- The 3D reconstruction improves every iteration.
- Its projections are of better quality.
- Angular assignment is more and more precise until convergence.

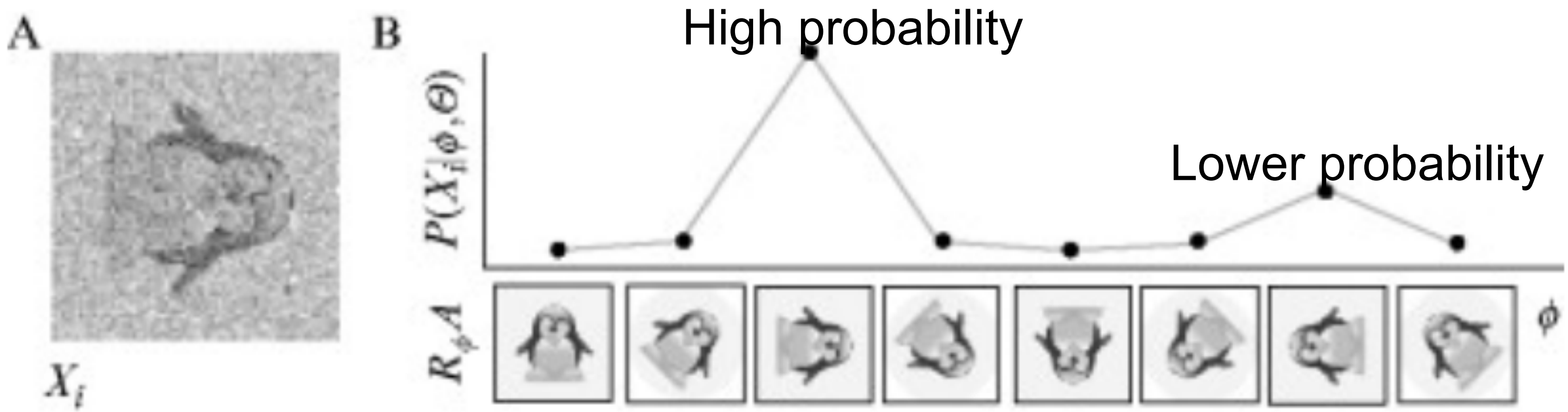


Maximum likelihood methods applied to single-particle reconstruction

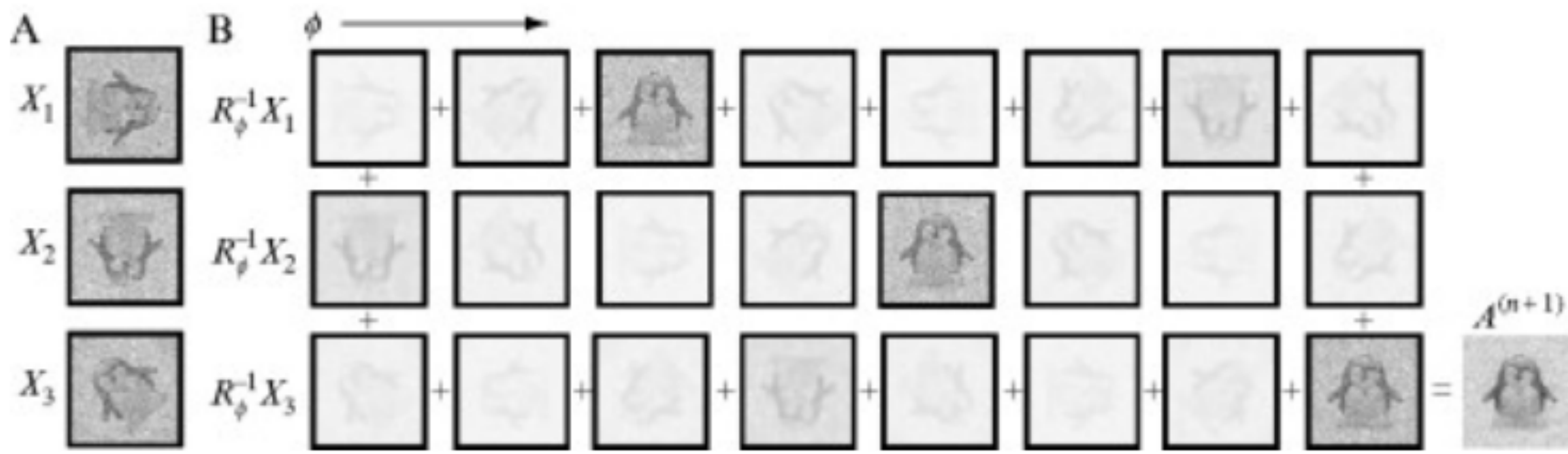


probability distribution function for orientations

Maximum likelihood methods applied to single-particle reconstruction



probability distribution function for orientations



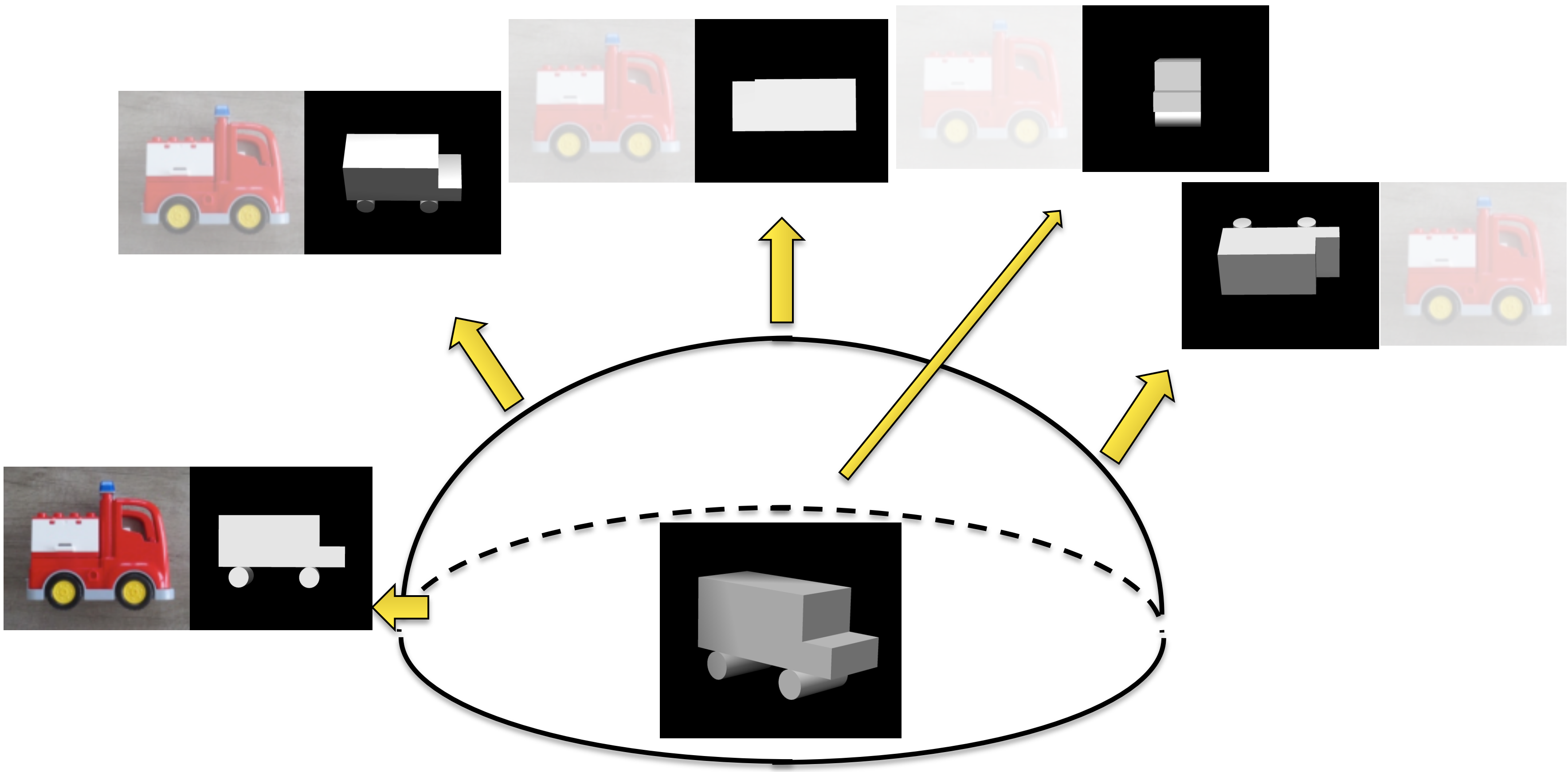
probability-weighted averaging

From Scheres

➔ Cleaner references, higher radius of convergence

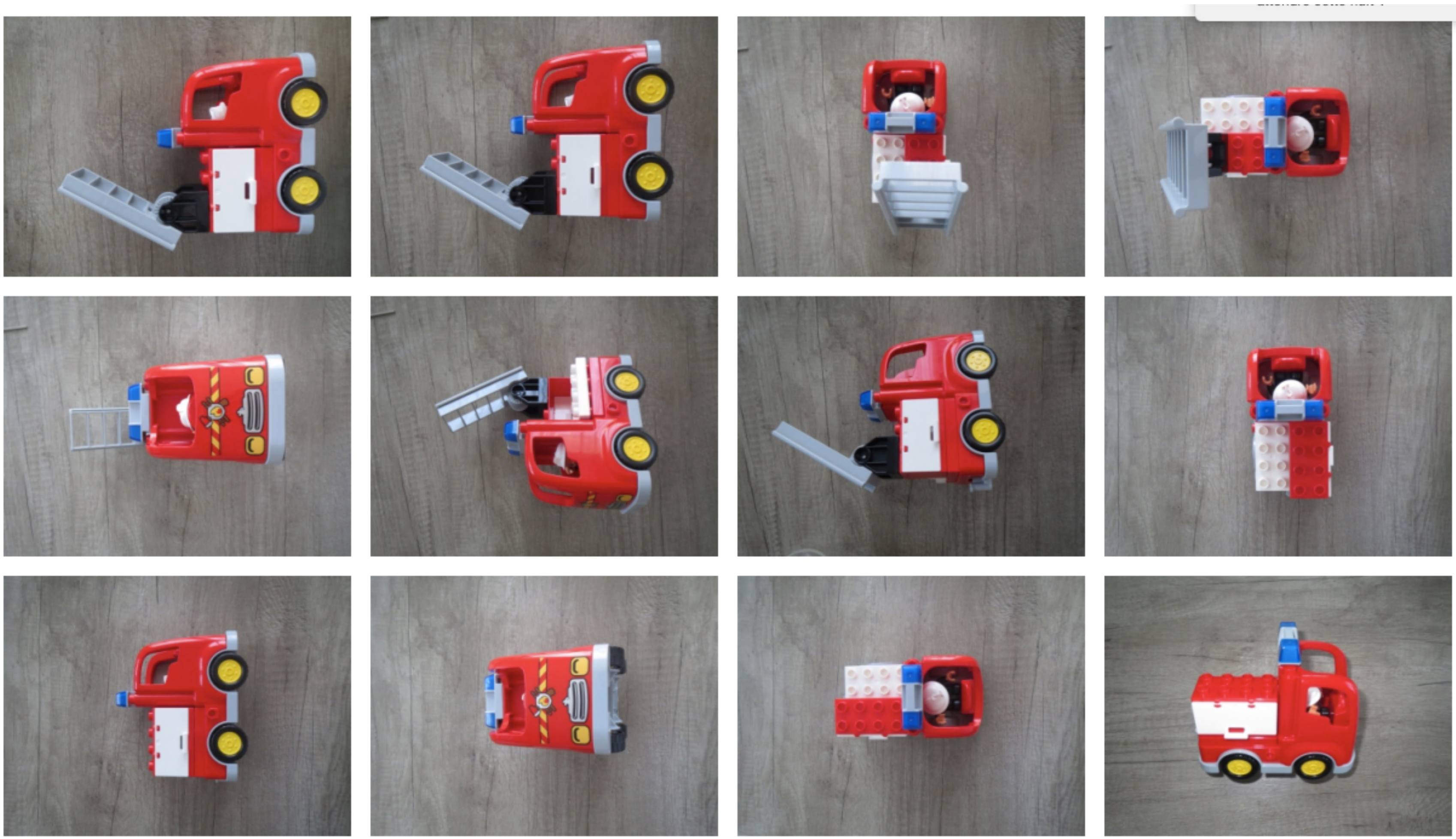
Maximum-likelihood applied to our example

- For each image, a probability is given to every orientation
- More robust as we have images with very low signal to noise

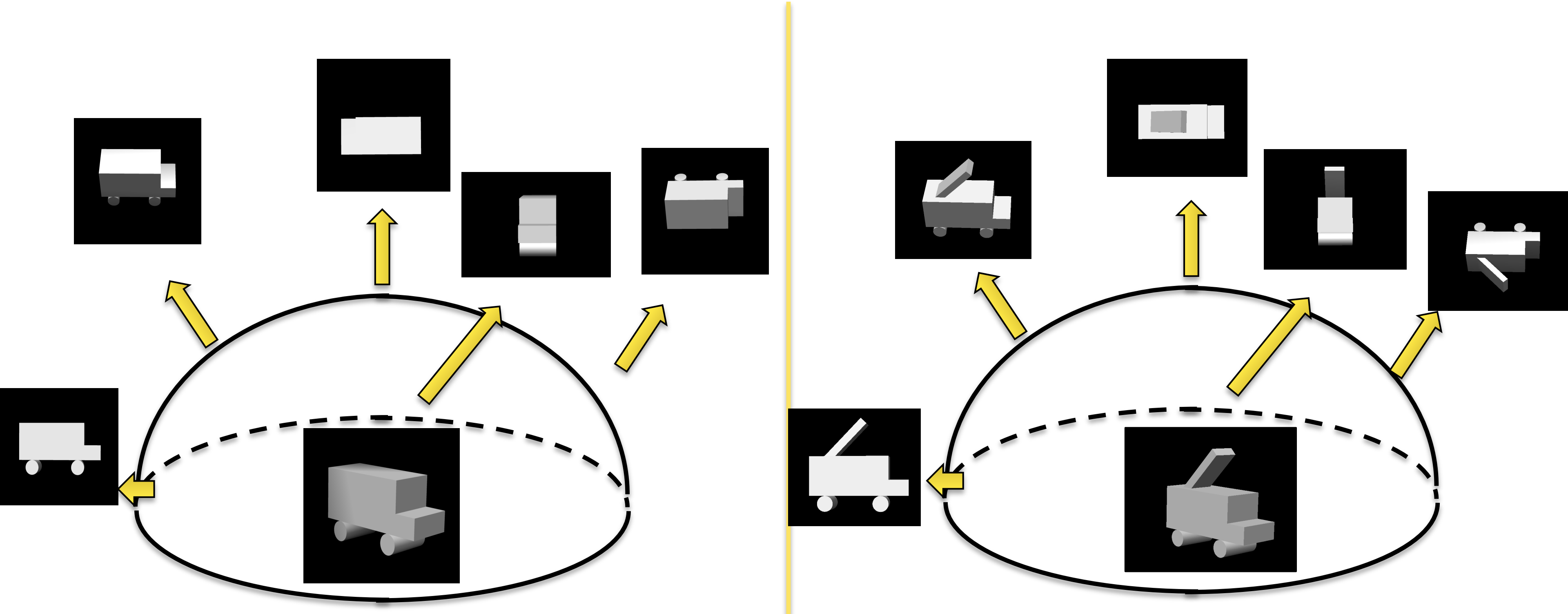


How to deal with structural heterogeneity ?

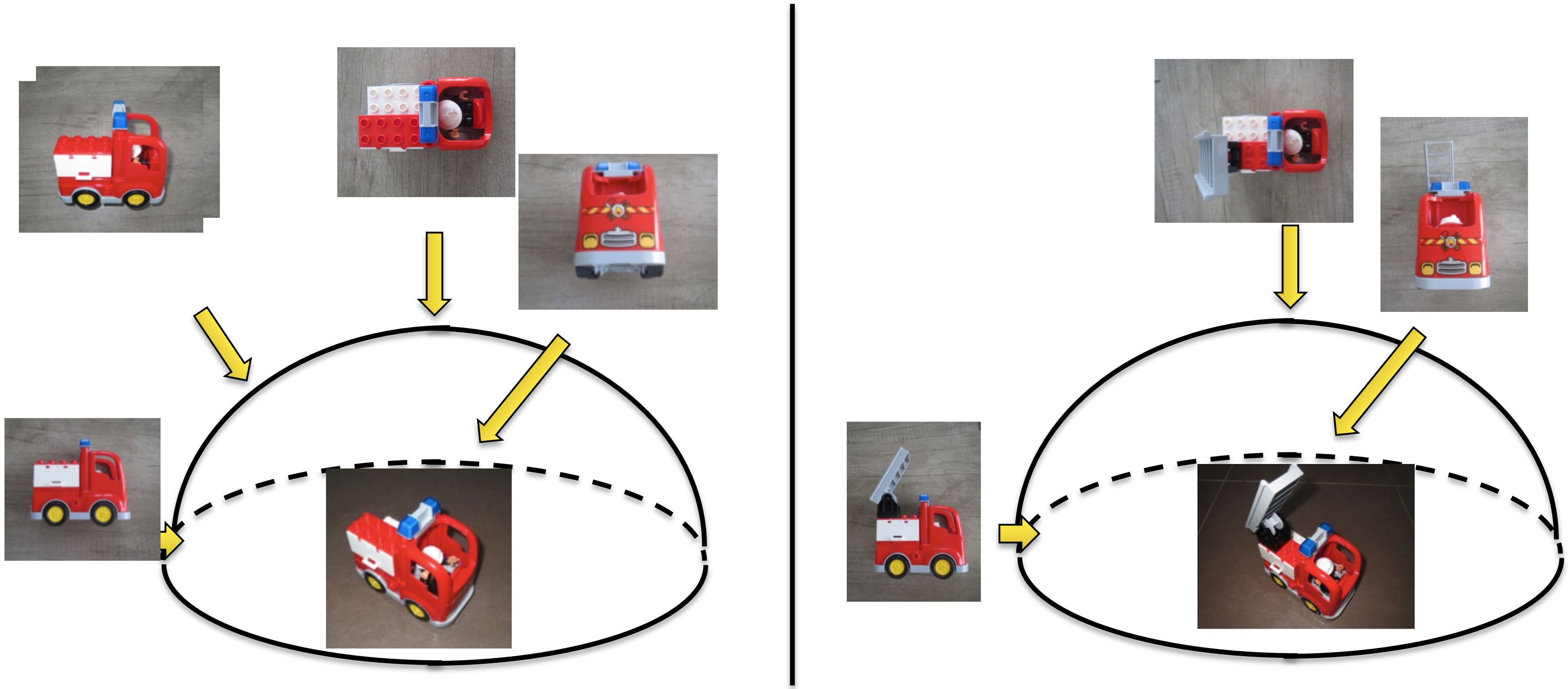
- Most datasets are, to some extent, heterogeneous.
- For example, you can have a dataset with your complex in both active and inactive states.



- Cryo-EM image analysis allows to classify these 3D heterogeneities
- Generation of X initial reconstructions (can be obtained from random subsets of images with assigned euler angles)
- « Competitive » projection matching.
- Can be done with the highest cross-correlation criterion or with maximum likelihood.

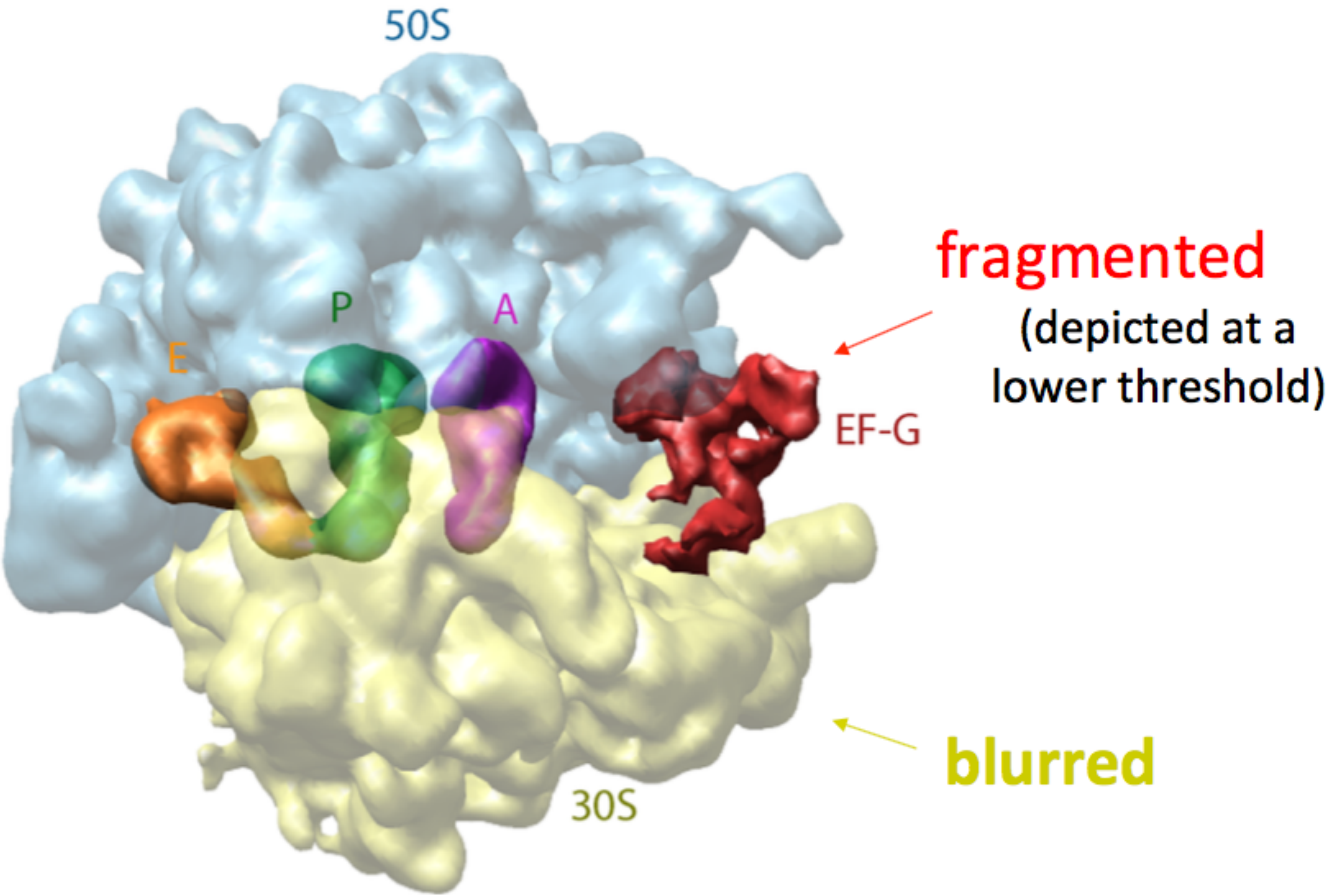


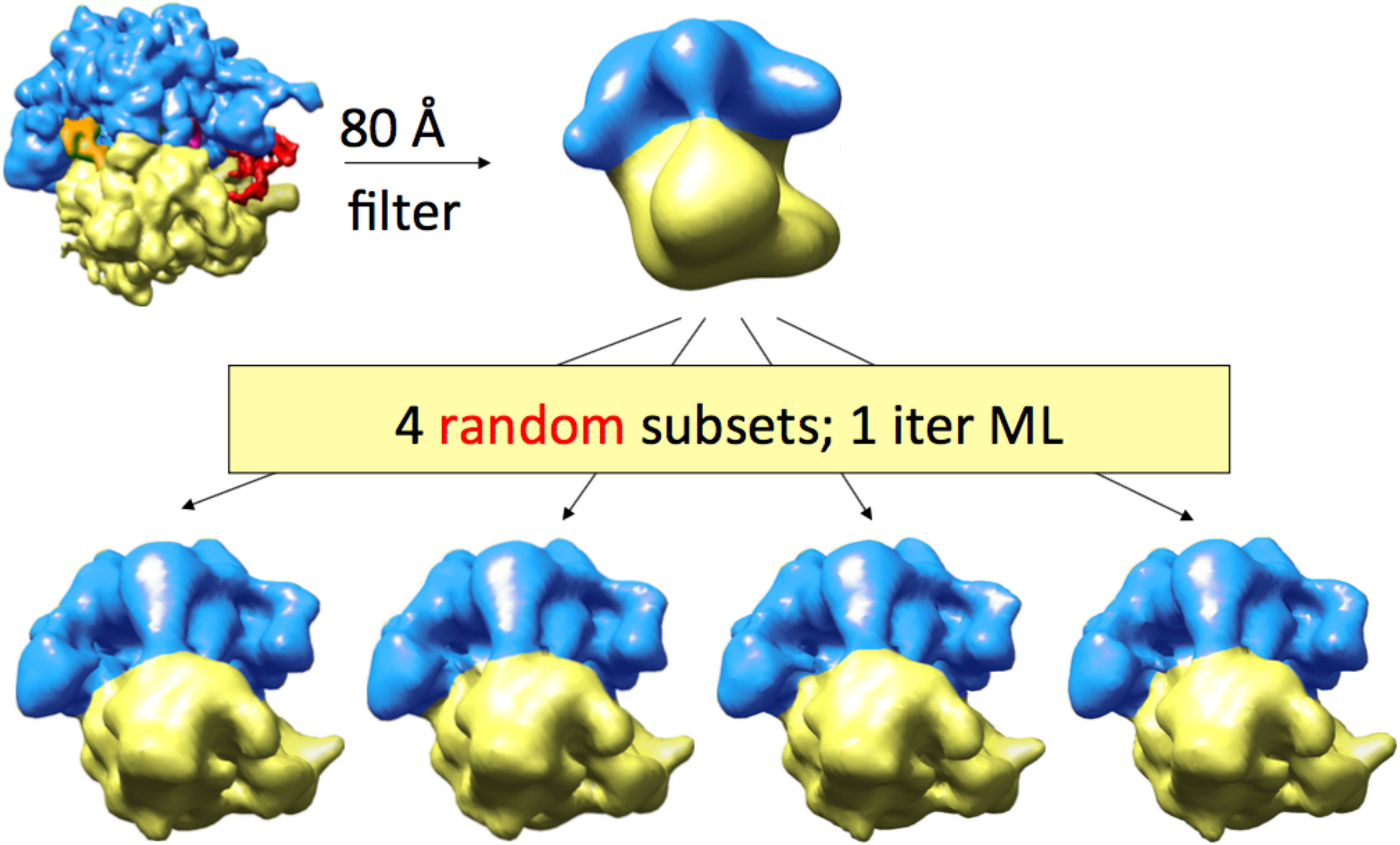
- Allows to obtain several structures out of one dataset !
- As subsets are more homogeneous, they can reach higher resolution !

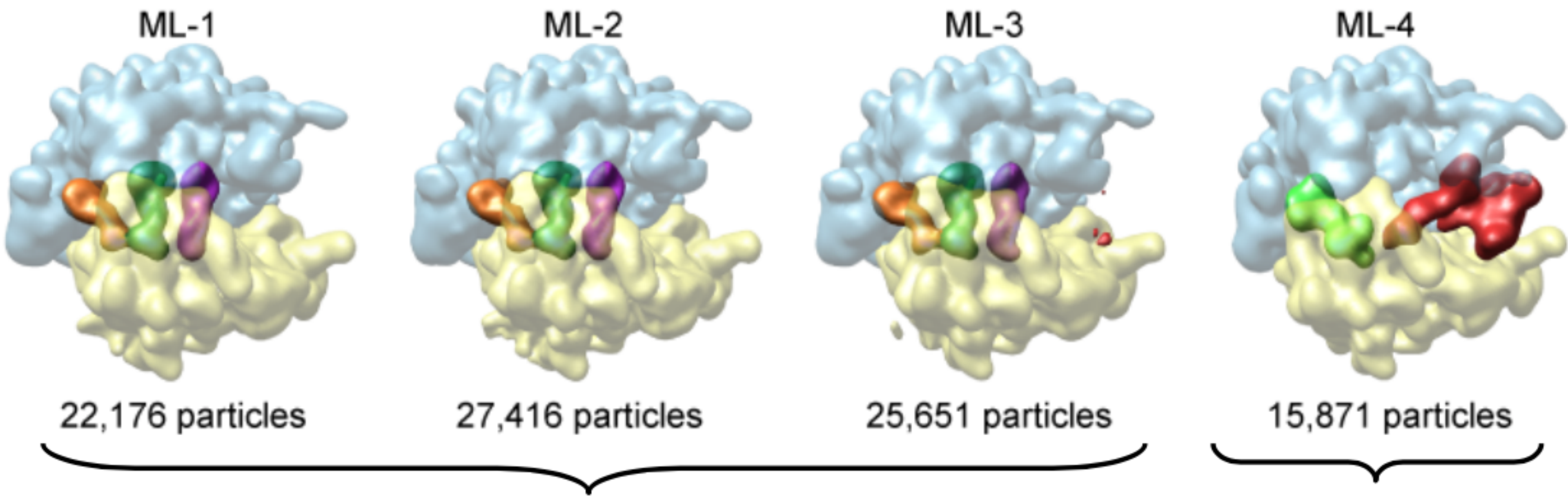


Methods to classify 3D heterogeneity

- « Competitive projection matching » using the highest cross-correlation criteria
- « Competitive projection matching » using maximum likelihood
- 3D MSA


50S
P
A
E
EF-G
fragmented
(depicted at a
lower threshold)
blurred
30S


80 Å
filter
4 random subsets; 1 iter ML

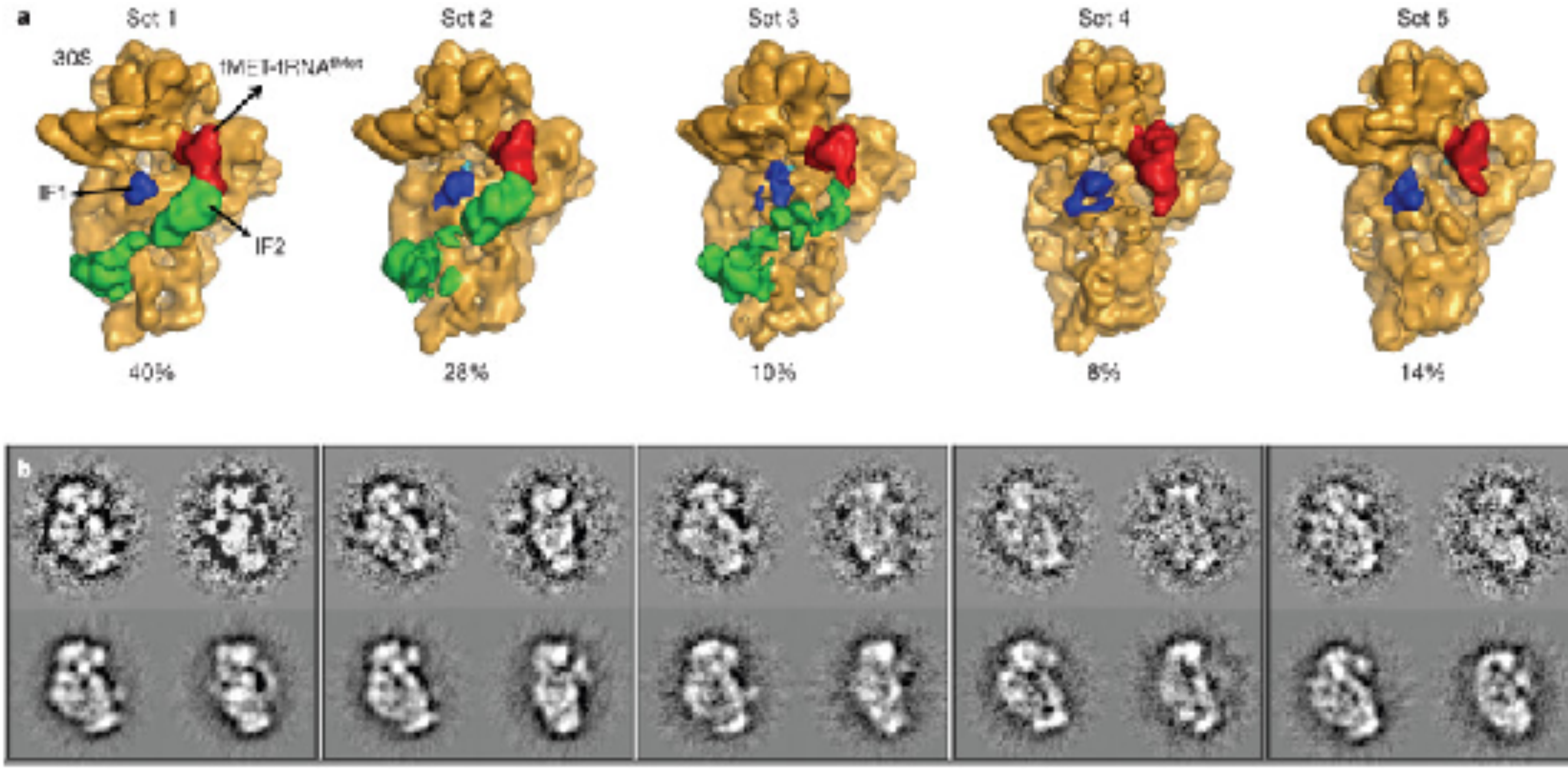


no ratcheting; no EF-G; 3 tRNAs
differences: overall rotations

ratcheting,
EF-G, 1 tRNA

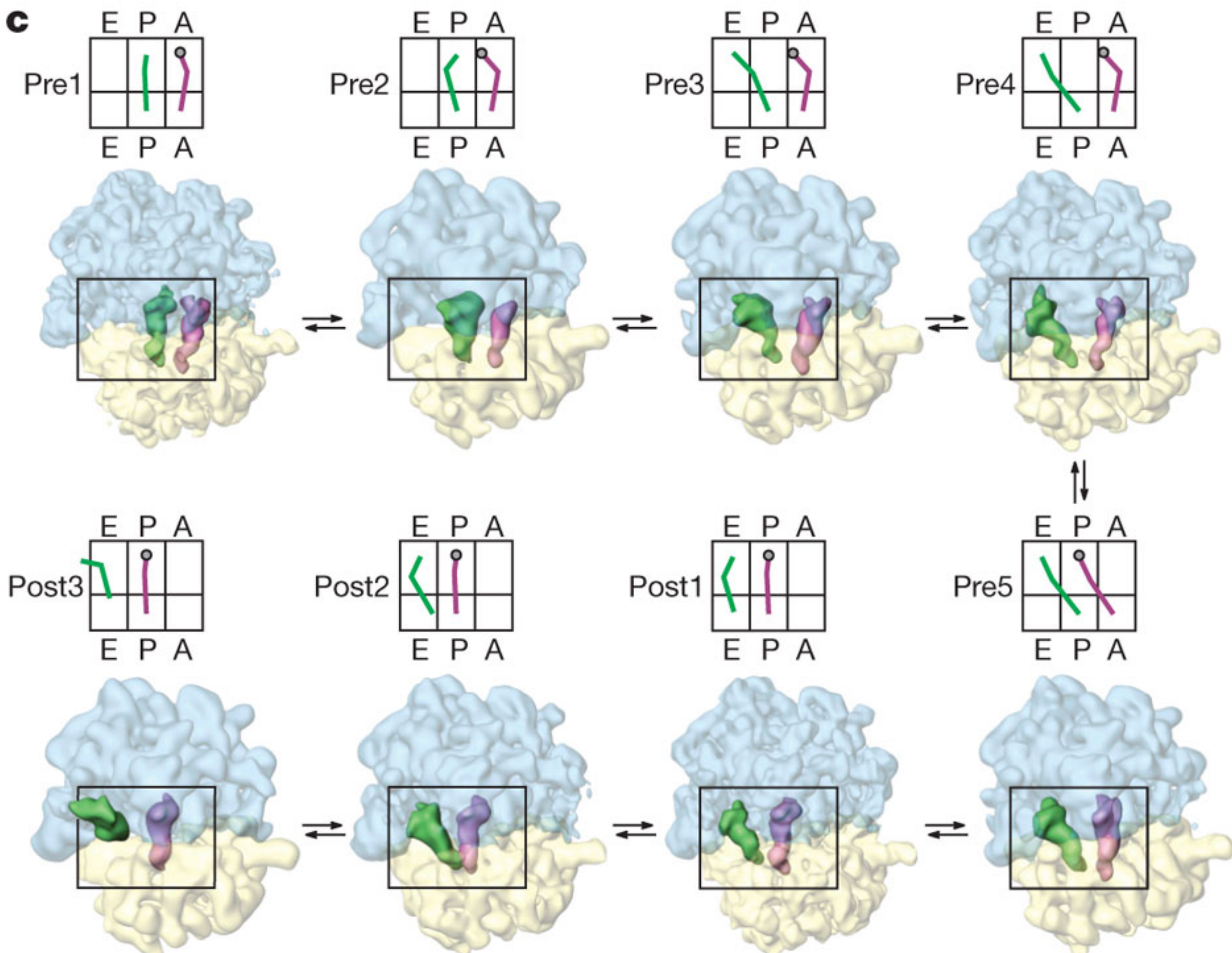
(Results coincided with a supervised classification)

Scheres et al (2007) Nat. Meth.



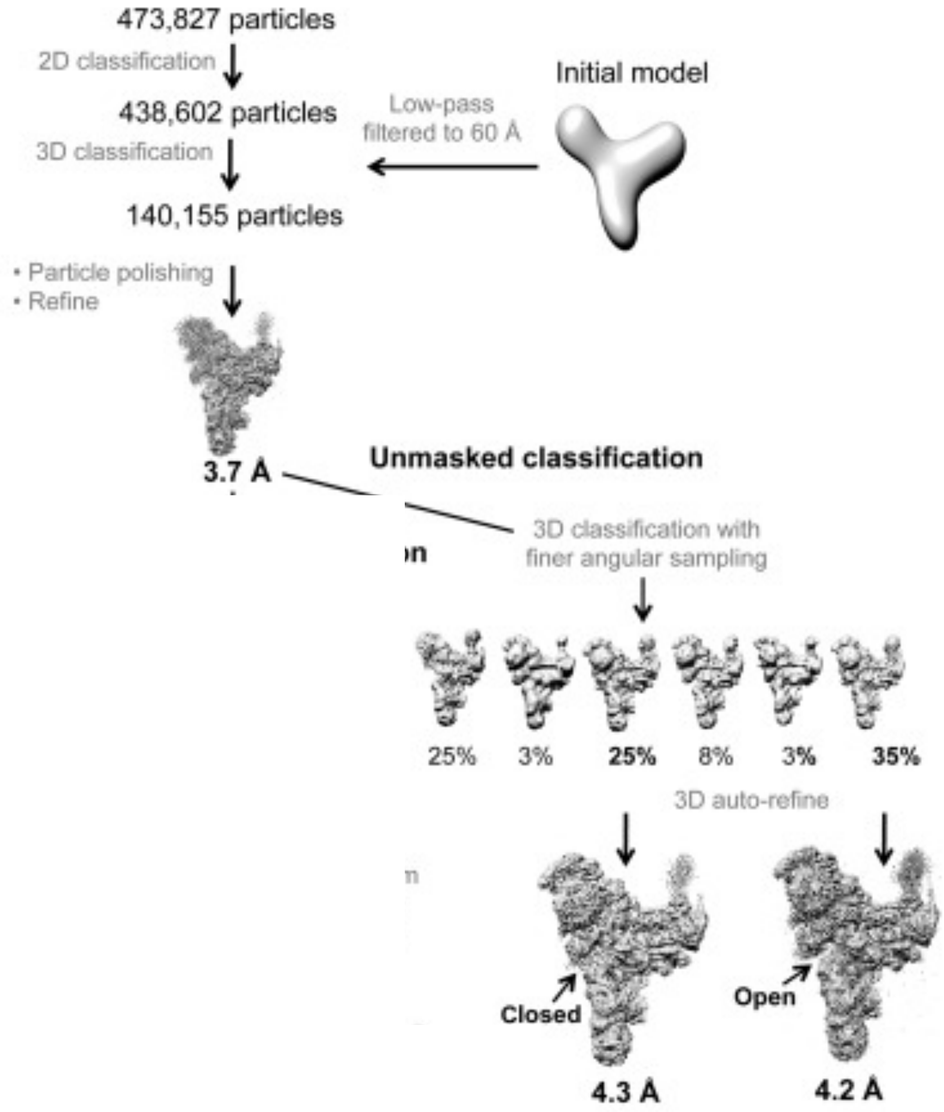
Simonetti et al, Nature 2008

Separation of time-resolved states



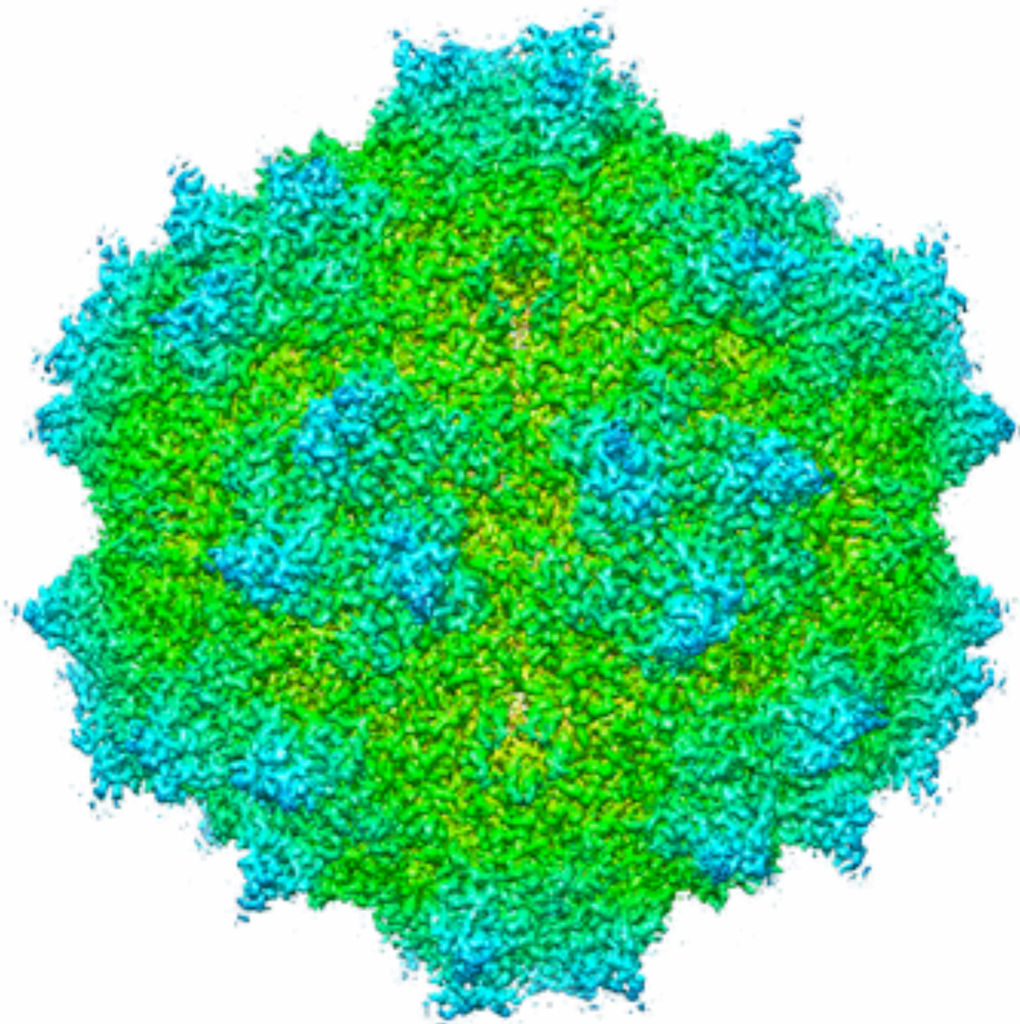
Fischer et al, Nature 2010

Look how great EM image analysis can become !!

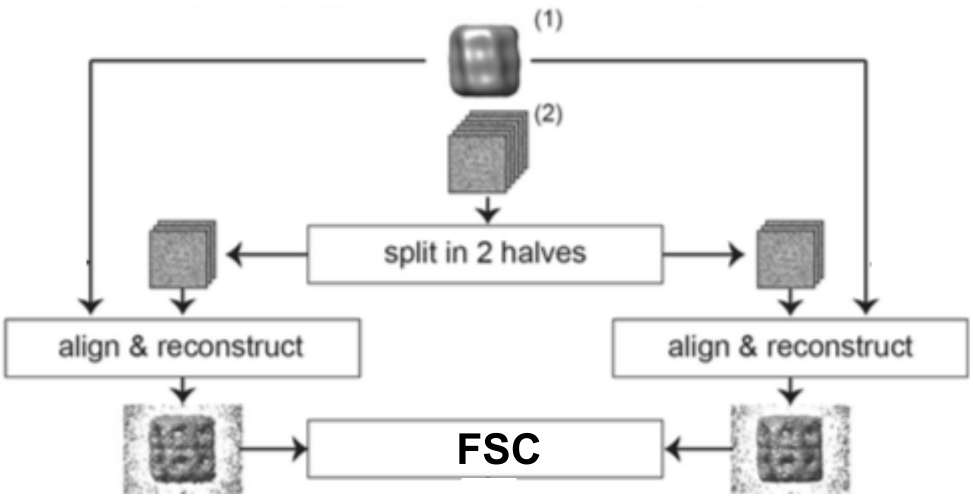


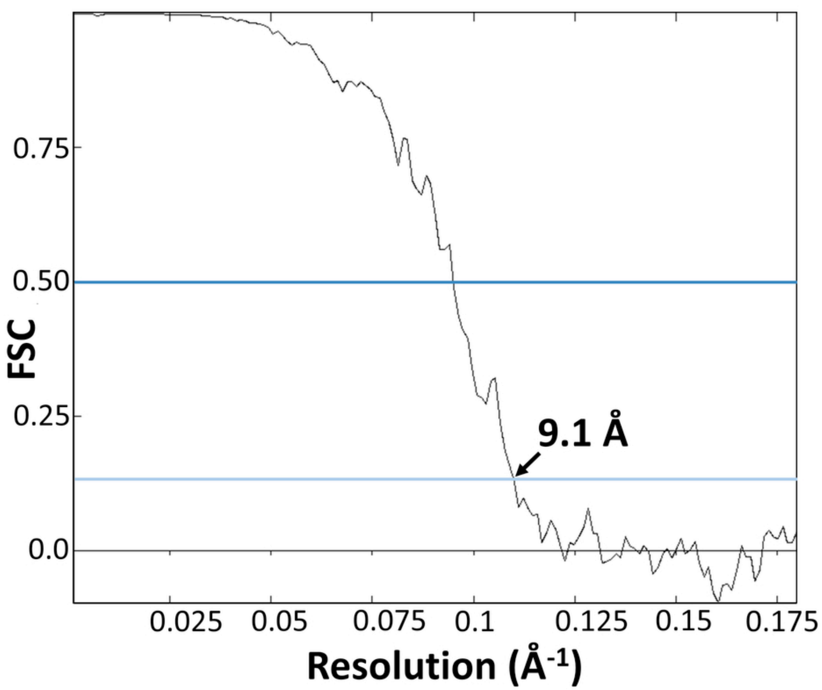
Nguyen et al, Nature, 2016

Congratulations ! You got your 1st refined EM map !



How to calculate the resolution?





Dataset separation in two halves

↓

One reconstruction from each half

↓

Comparison in Fourier space for each Fourier shell

Several criteria:
0.5
0.143
3 sigma
½ bit

Harauz and van Heel (1986)
Rosenthal et al., 2003

Resolution is not the same everywhere, local resolution can be calculated