

Resolution Measures in Single Particle Analysis

Houston, March 2005

Pawel A. Penczek

**The University of Texas – Houston Medical School,
Department of Biochemistry and Molecular Biology,
6431 Fannin, MSB6.218, Houston, TX 77030, USA.**

phone: (713) 500-5416

fax: (713) 500-0652

Pawel.A.Penczek@uth.tmc.edu

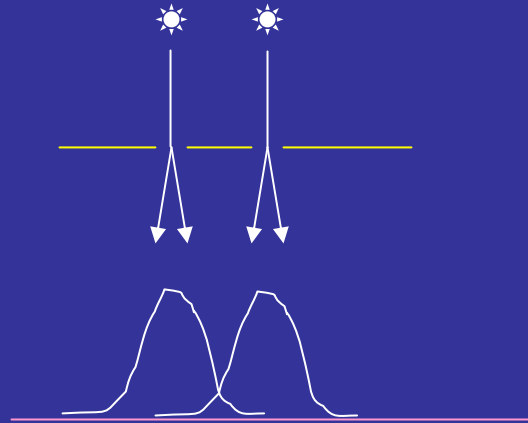


THE UNIVERSITY *of* TEXAS

HEALTH SCIENCE CENTER AT HOUSTON

MEDICAL SCHOOL

Optical resolution



The resolution of a microscope objective is defined as the smallest distance between two points on a specimen that can still be distinguished as two separate entities.

Resolution is a somewhat subjective concept.

The theoretical limit of the resolution is set by the wavelength of the light source:

$$R = \text{const } \lambda$$

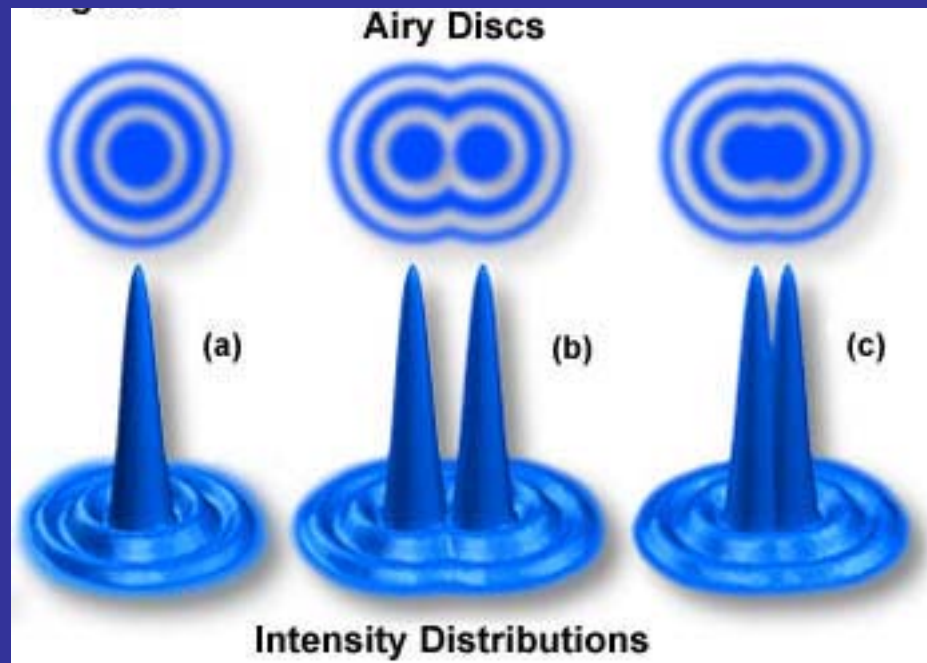
Optical resolution

Hypothetical **Airy disk** (a) consists of a diffraction pattern containing a central maximum (typically termed a zero'th order maximum) surrounded by concentric 1st, 2nd, 3rd, etc., order maxima of sequentially decreasing brightness that make up the intensity distribution.

If the separation between the two disks exceeds their radii (b), they are resolvable.

The limit at which two Airy disks can be resolved into separate entities is often called the **Rayleigh criterion**.

When the center-to-center distance between the zero'th order maxima is less than the width of these maxima, the two disks are not individually resolvable by the Rayleigh criterion (c).



Resolution-limiting factors in electron microscopy

- The wavelength of the electrons
(depends on the voltage: 100kV - 0.037 Å; 300kV – 0.020Å)
- The quality of the electron optics
(astigmatism, envelope functions)
- The underfocus setting.
The resolution of the TEM is often defined as the first zero in the contrast transfer function (PCTF) at Scherzer (or optimum) defocus.
- Signal-to-Noise Ratio (SNR) level in the data
- Accuracy of the alignment

The concept of optical resolution is not applicable to electron microscopy and single particle reconstruction

- In single particle reconstruction, there is no “external” standard by which the resolution of the results could be evaluated.
- Therefore, the resolution measures in EM have to estimate “internal consistency” of the results.
- Unless an external standard is provided, objective estimation of the resolution in EM is not possible.

FRC - Fourier Ring Correlation

Saxton W.O. and W. Baumeister.

The correlation averaging of a regularly arranged bacterial cell envelope protein.

J. Microsc., **127**, 127-138 (1982).

FSC – Fourier Shell Correlation (3-D)

DPR – Differential Phase Residual

Frank J., A. Verschoor, M. Boublik.

Computer averaging of electron micrographs of 40S ribosomal subunits.

Science, **214**, 1353-1355 (1981).

2-D & 3-D

SSNR – Spectral Signal-to-Noise Ratio

Unser M., L.B. Trus, A.C. Steven.

A new resolution criterion based on spectral signal-to-noise ratios.

Ultramicroscopy, **23**, 39-52 (1987).

Penczek, P. A.

Three-dimensional Spectral Signal-to-Noise Ratio for a class of reconstruction algorithms.

J. Struct. Biol., **138**, 34-46 (2002)

Q-factor

van Heel M. and J. Hollenberg.

The stretching of distorted images of two-dimensional crystals.

In: Proceedings in Life Science: Electron Microscopy at Molecular Dimensions (Ed.: W. Baumeister).

Springer Verlag, Berlin (1980).

only 2-D

Fourier Ring Correlation

$$FSC(R) = \frac{\sum_{n \in R} F_n G_n^*}{\left\{ \left(\sum_{n \in R} |F_n|^2 \right) \left(\sum_{n \in R} |G_n|^2 \right) \right\}^{1/2}}$$

A. either:

1. Split (randomly) the data set of available images into halves;
2. Perform the alignment of each data set “independently”;

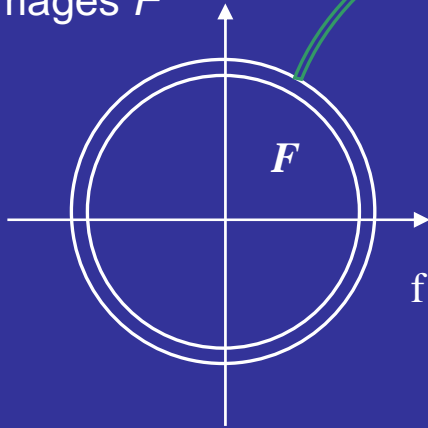
B. or:

1. Perform the alignment of the whole data set;
2. Split (randomly) the aligned data set into halves;
3. Calculate two averages (3-D reconstructions);
4. Compare the averages in Fourier space by calculating the FRC.

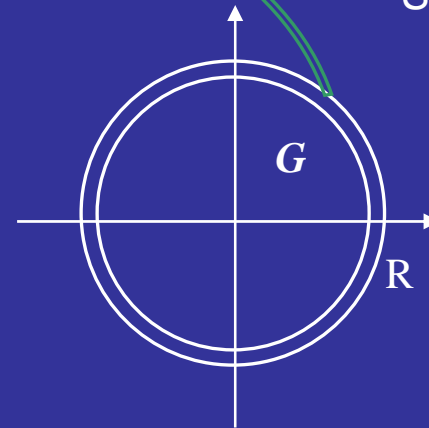
WARNINGS - method B valid *only* if the noise component in the data is independent (not aligned)
- the two sets in method A might not be as independent as one assumes.

Fourier Shell Correlation FSC

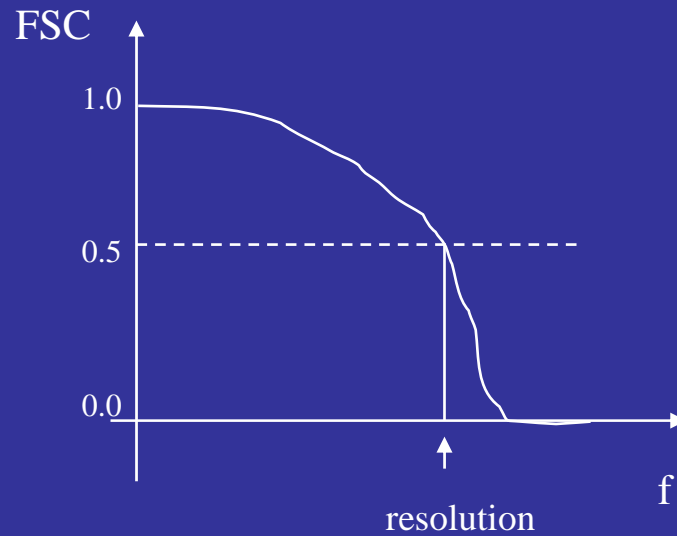
First set of images F



Second set of images G



$$FSC(R) = \frac{\sum_{n \in R} F_n G_n^*}{\left\{ \left(\sum_{n \in R} |F_n|^2 \right) \left(\sum_{n \in R} |G_n|^2 \right) \right\}^{1/2}}$$



Fourier Shell Correlation

WHY DOES IT WORK?

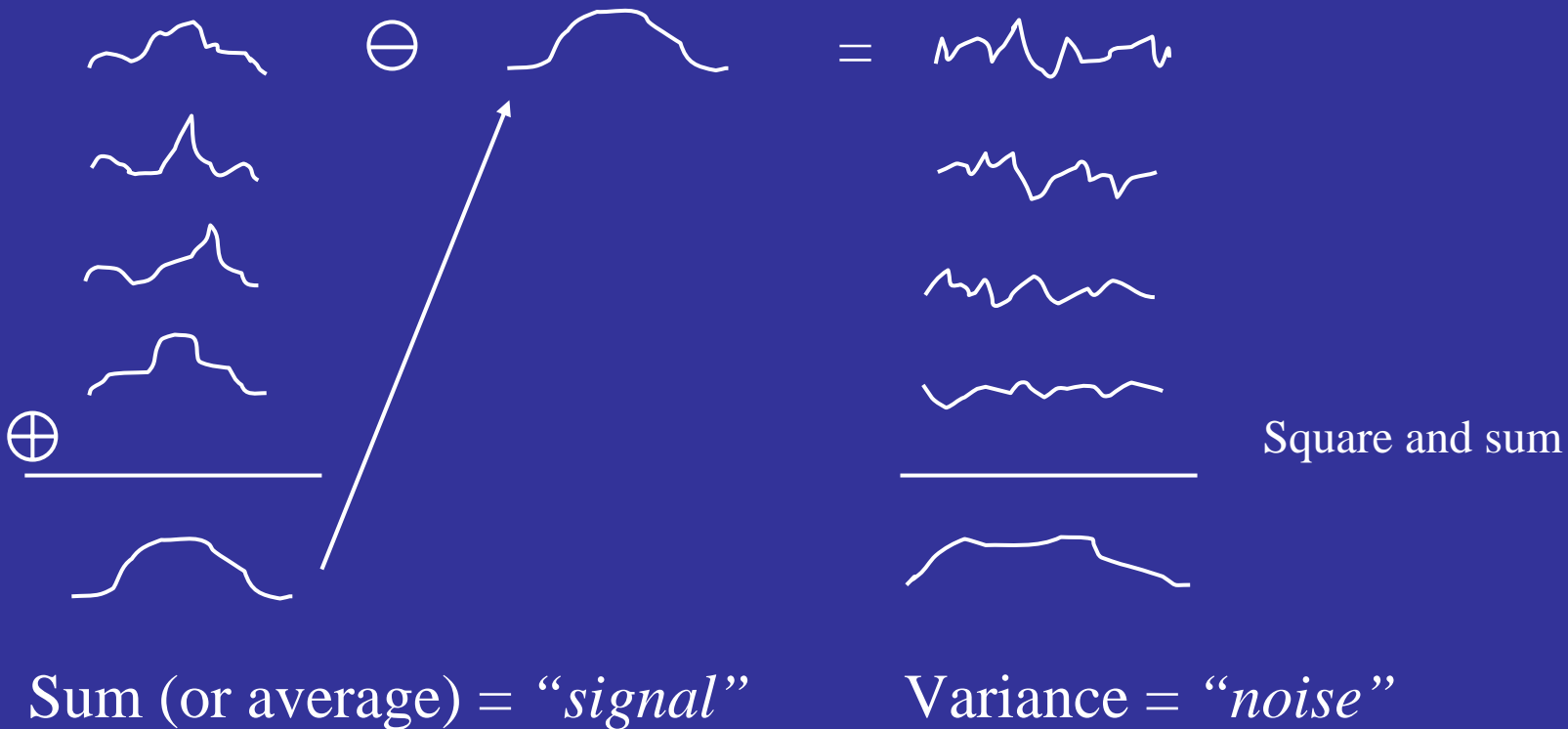
FSC provides a measure of the Spectral Signal-to-Noise Ratio in the reconstruction.

WHAT DOES IT HAVE TO DO WITH RESOLUTION?!?

FSC is directly related to the alignment error.

Signal *versus* noise

When we perform multiple measurements of the same phenomena, we equate the “*signal*” with the part of the measurement that remains the same between measurements, and we assume that the varying part of measurements is the “*noise*”.

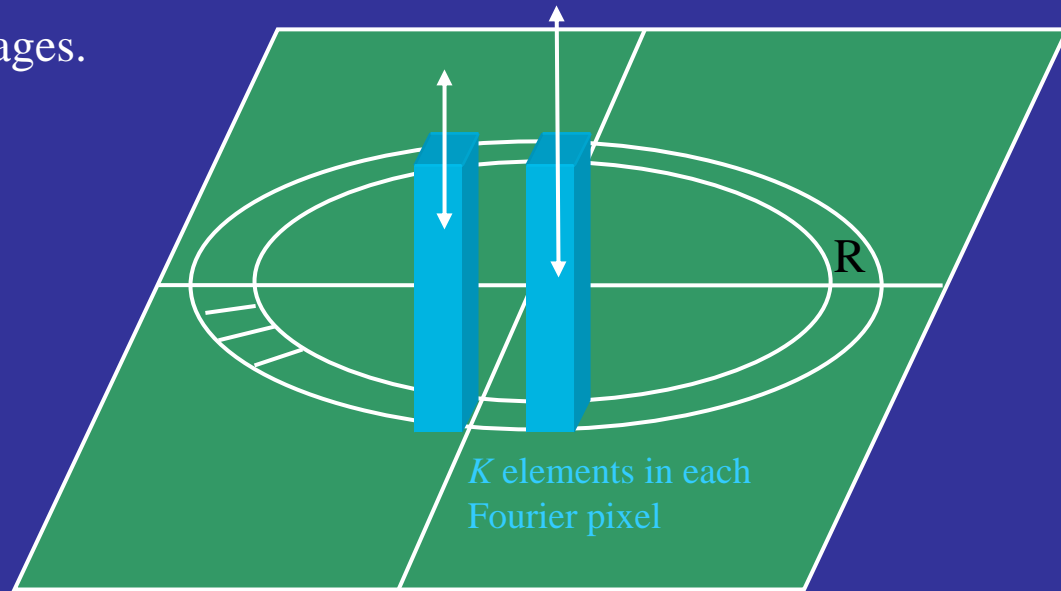
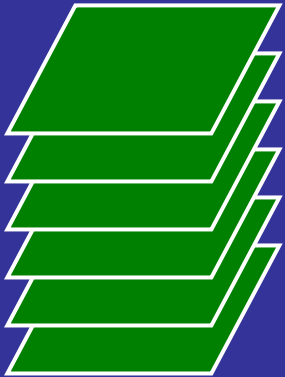


Signal-to-Noise Ratio (SNR)

$$\text{SNR} = \frac{\text{Power of signal}}{\text{Power of noise}}$$

Spectral Signal-to-Noise Ratio (SSNR) in 2D

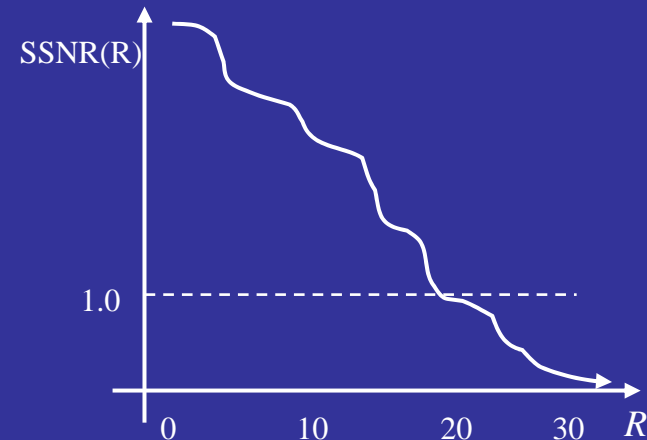
A set of Fourier transforms of 2D images.



Calculate SSNR according to the equation:

$$SSNR(R) = \frac{\sum_{n \in R} \left| \sum_k F_{n,k} \right|^2}{\frac{K}{K-1} \sum_{n \in R} \sum_k \left| F_{n,k} - \langle F \rangle_n \right|^2} - 1$$

where $\langle F \rangle_n = \frac{1}{K} \sum_k F_{n,k}$



Relations between FSC and SSNR

$$SSNR = \frac{FSC}{1 - FSC};$$

$$FSC = \frac{SSNR}{SSNR + 1}$$

For large number of images $Variance(SSNR) \cong Variance(FSC)$

When FSC is calculated for a data set split into halves:

$$SSNR = 2 \frac{FSC}{1 - FSC}$$

FSC is a biased estimate of SSNR.

For large number of images, the bias is negligible.

The expectation value of FRC is calculated as

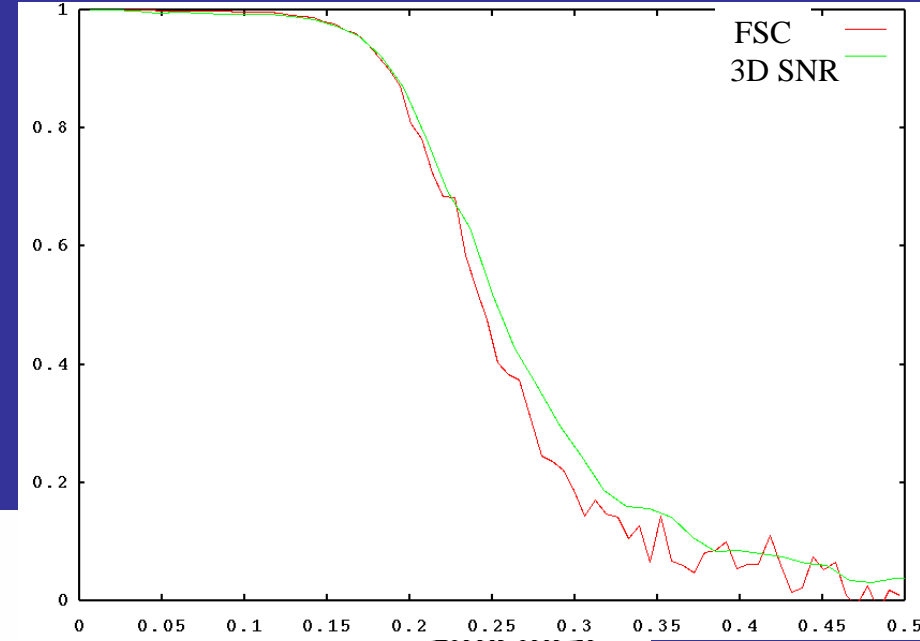
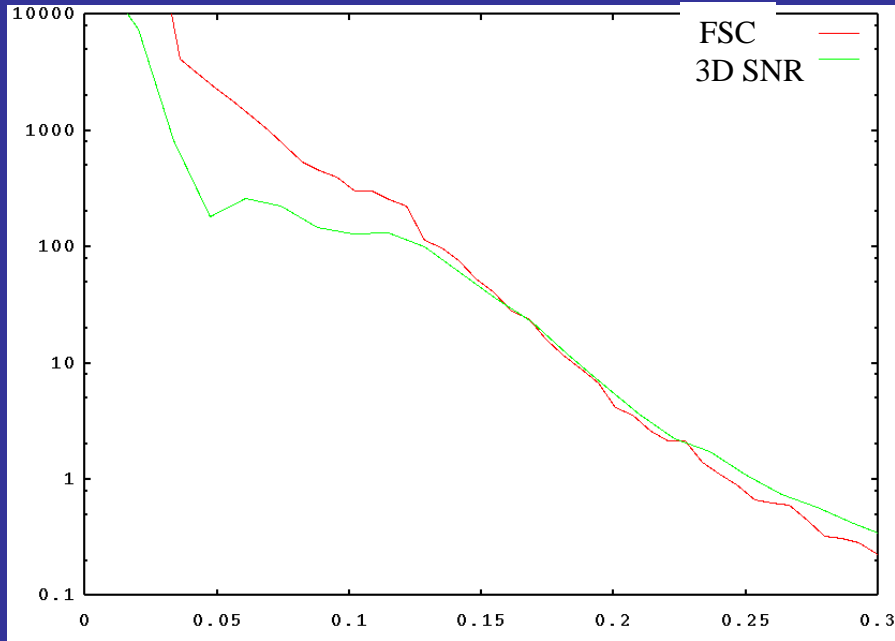
$$E[FRC] \cong \frac{E[\sum_{n=1}^{N_R} F^n G^{n*}]}{E\left[\left\{\left(\sum_{n=1}^{N_R} |F^n|^2\right)\left(\sum_{n=1}^{N_R} |G^n|^2\right)\right\}^{1/2}\right]}$$

$$= \frac{\sum_n F_T^{n^2}}{\sum_n F_T^{n^2} + \sum_n \frac{1}{L} \sigma_N^{n^2}} = \frac{\frac{\sum_n F_T^{n^2}}{\sum_n \frac{1}{L} \sigma_N^{n^2}}}{\frac{\sum_n F_T^{n^2}}{\sum_n \frac{1}{L} \sigma_N^{n^2}} + 1} = \frac{SSNR}{SSNR + 1}$$

Test of 3D SSNR

$$\text{SSNR} = \text{FSC} / (1 - \text{FSC})$$

$$\text{FSC} = \text{SSNR} / (\text{SSNR} + 1)$$



Journal of Structural Biology 138 (2002) 34–46

Structural
Biology

www.academicpress.com

Three-dimensional spectral signal-to-noise ratio for a class of reconstruction algorithms

Pawel A. Penczek*

Department of Biochemistry and Molecular Biology, The University of Texas Houston Medical School, 6431 Fannin, MSB6.218, Houston, TX 77030, USA

Received 17 December 2001; and in revised form 13 May 2002

Resolution criteria

should be based on the SNR considerations

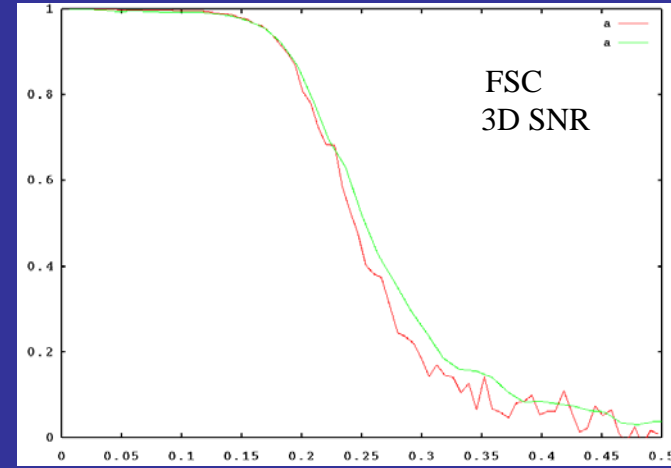
$$SSNR = 2 \frac{FSC}{1 - FSC}$$

Reasonable criterion: include only Fourier information that is above the noise level, i.e., $SSNR > 1$.

$$SSNR=1 \Rightarrow FSC=1/3=0.333$$

Another criterion: (3σ) include Fourier information that is significantly higher than zero, i.e., $SSNR > 0$.

$$SSNR=0 \Rightarrow FSC=0$$



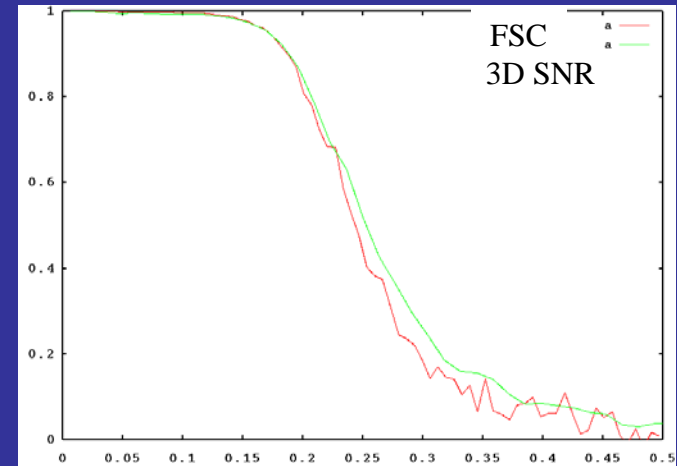
Resolution curve and optimum filtration

$$SSNR = 2 \frac{FSC}{1 - FSC}$$

Wiener filter:

$$G = \frac{SSNR}{SSNR + 1} F$$

$$G = 2 \frac{FSC}{FSC + 1} F$$

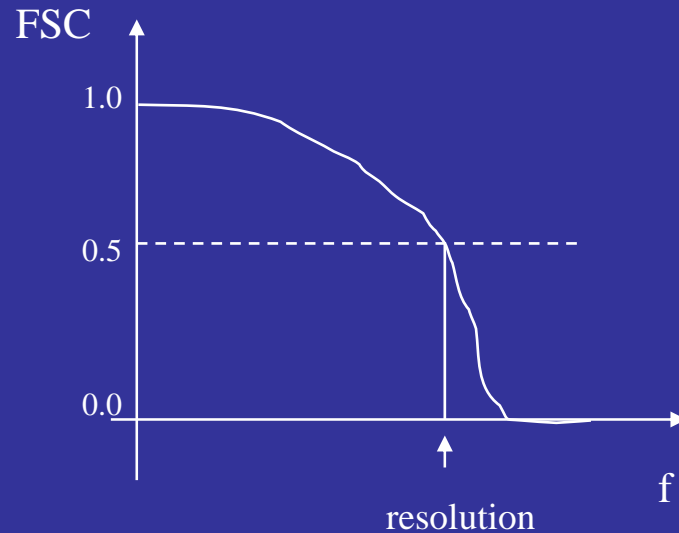


The *FSC* curve should be used for optimum filtration.

Thus, the ‘resolution’ is given by the overall shape of the *FSC*, not by a single number.

Examples of resolution curves

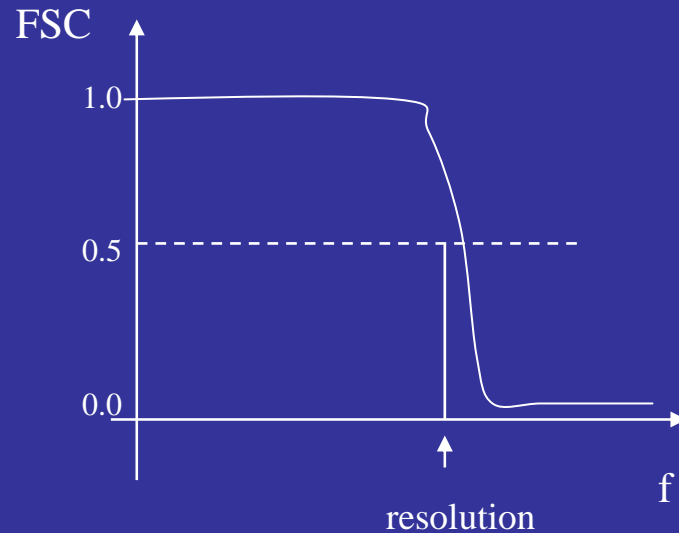
Healthy:



In low frequencies remains one, followed by a semi-Gaussian fall-off, drops to zero at around 2/3 of maximum frequency, in high frequencies oscillates around zero.

Examples of resolution curves

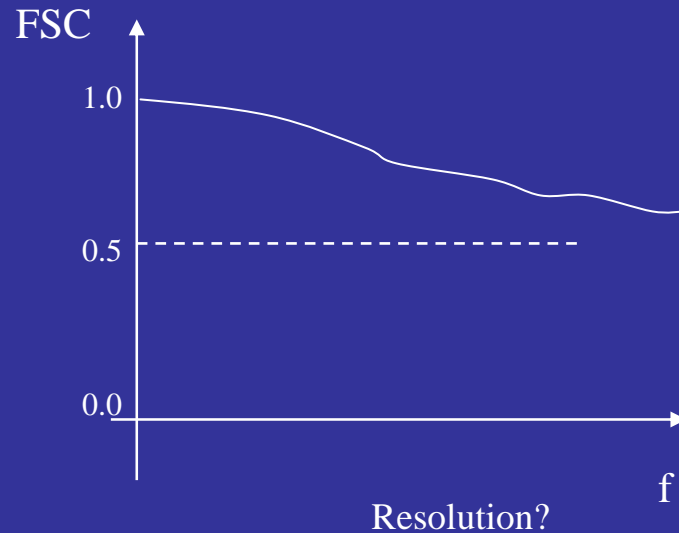
Unhealthy:



“Rectangular”: in low frequencies remains one, followed by a sharp drop, in high frequencies oscillates around zero.
A combination of alignment of the noise and a sharp filtration during the alignment procedure. The result is fake.

Examples of resolution curves

Unhealthy:

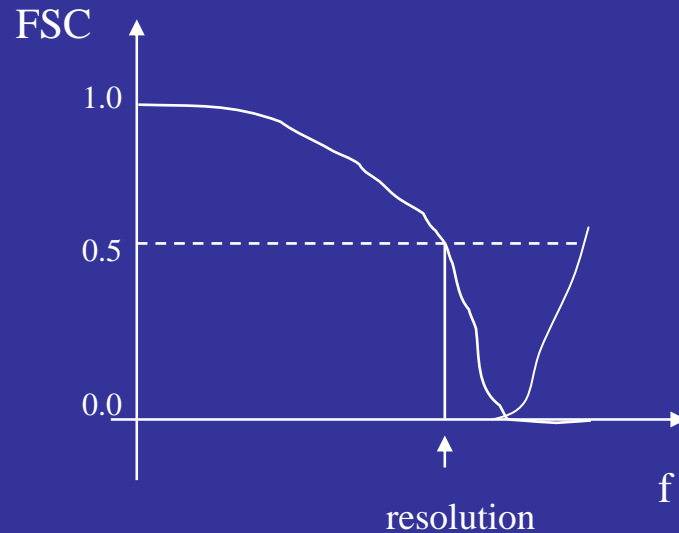


FSC never drops to zero in the whole frequency range.
The noise component in the data was aligned. The result is fake.

In rare cases it could mean that the data was severely undersampled (very large pixel size).

Examples of resolution curves

Unhealthy:

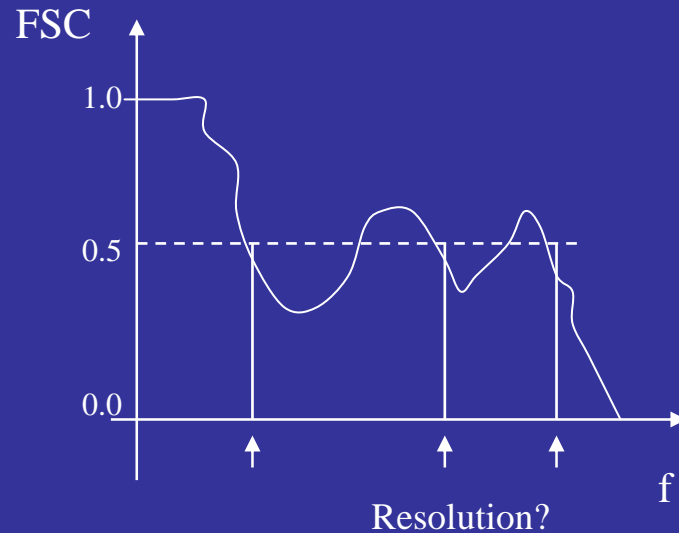


After it drops to zero, increases in high frequencies oscillation.

Data was low-passed filtered; errors in image processing code, mainly in interpolation; all images were rotated by the same angle.

Examples of resolution curves

Unhealthy:



FSC oscillates around 0.5.

The data is dominated by one subset with the same defocus value or there is only one defocus group. It is not incorrect per se, but unclear what is the resolution. Also, will result in artifacts.

FSC can be used to cross-validate EM results

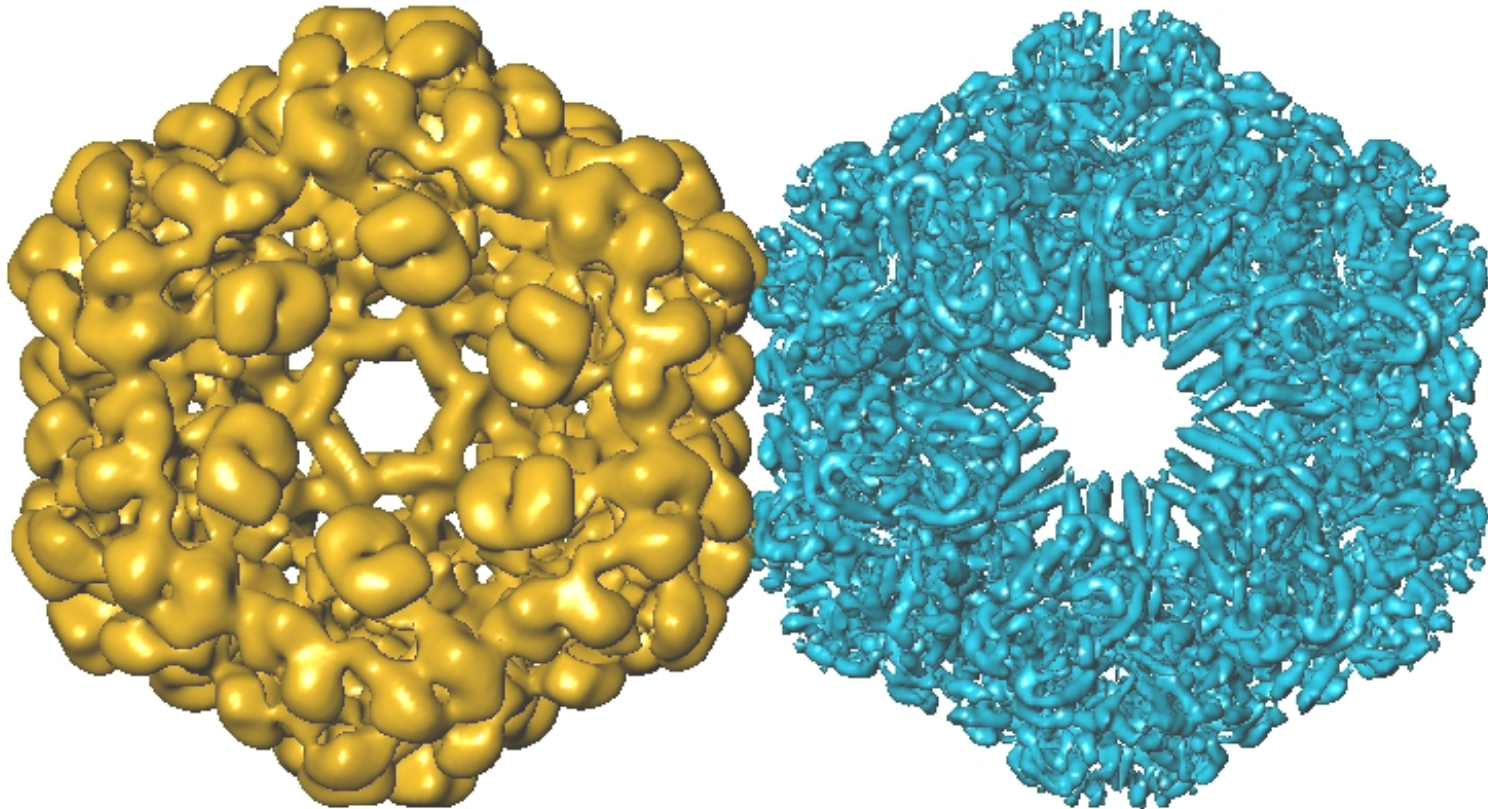
(crossresolution)

EM structure



X-ray crystallographic structure

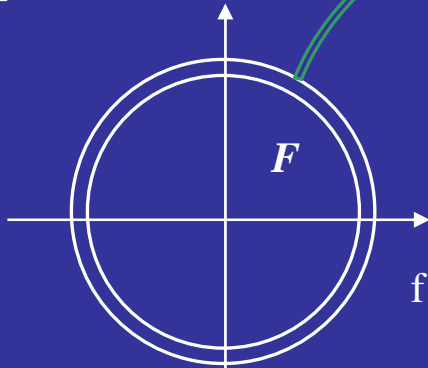
electron density map, the voxel values are
proportional to the Coulomb potentials of atoms



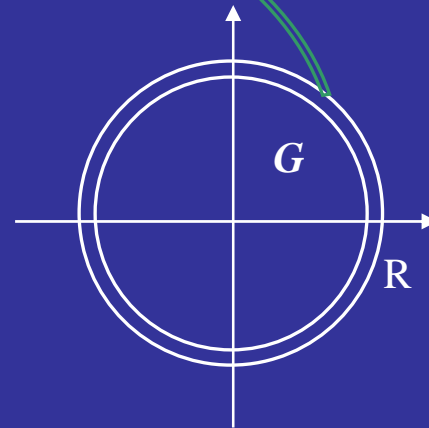
Crossresolution

relation between FRC and SSNR

X-ray map F
(noise-free)



EM map G
(corrupted by noise and
other errors)

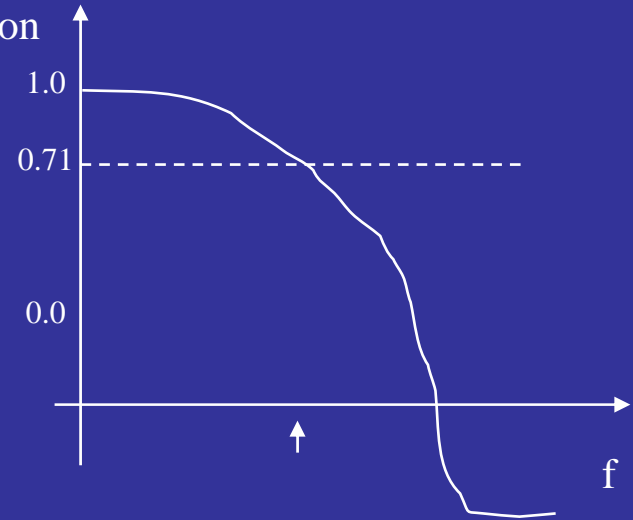


$$FSC(R) = \frac{\sum_{n \in R} F_n G_n^*}{\left\{ \left(\sum_{n \in R} |F_n|^2 \right) \left(\sum_{n \in R} |G_n|^2 \right) \right\}^{1/2}}$$

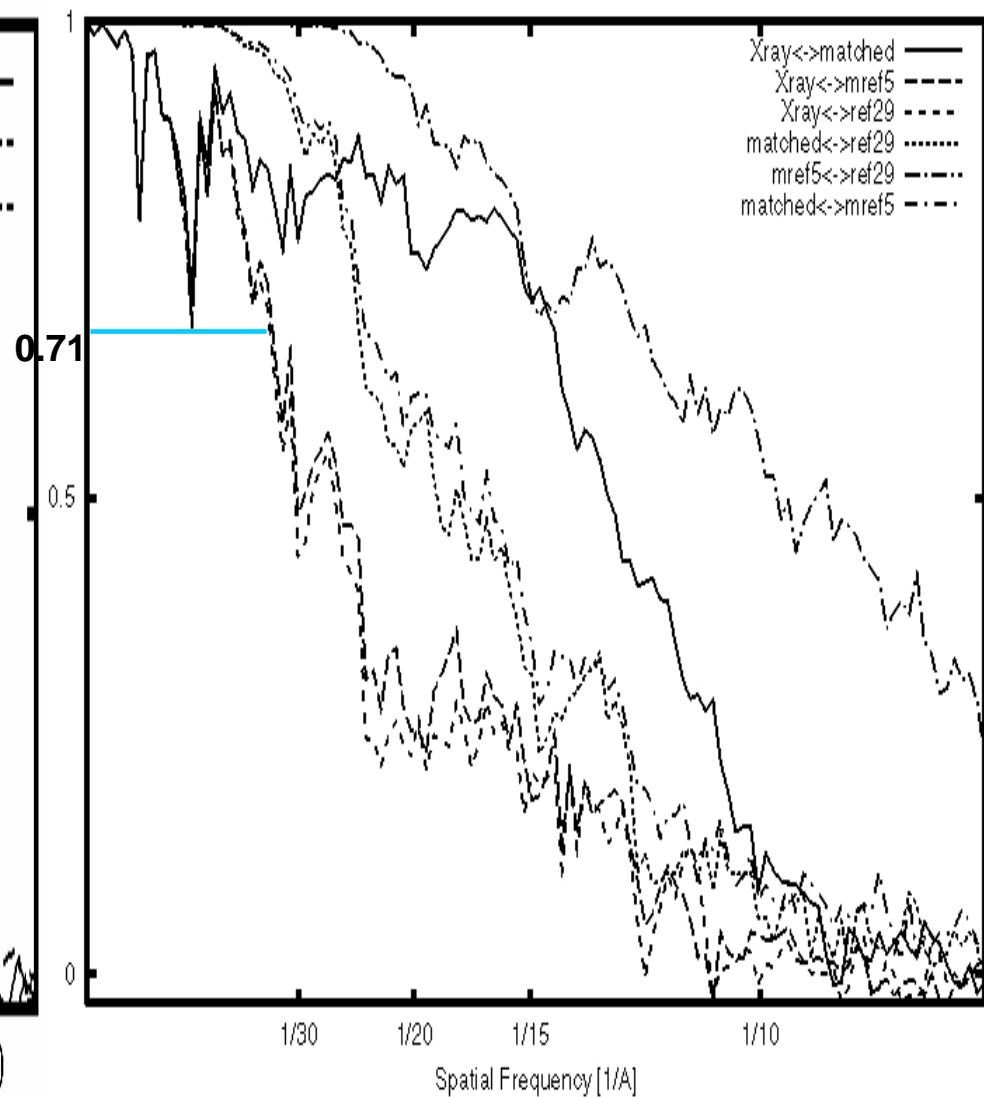
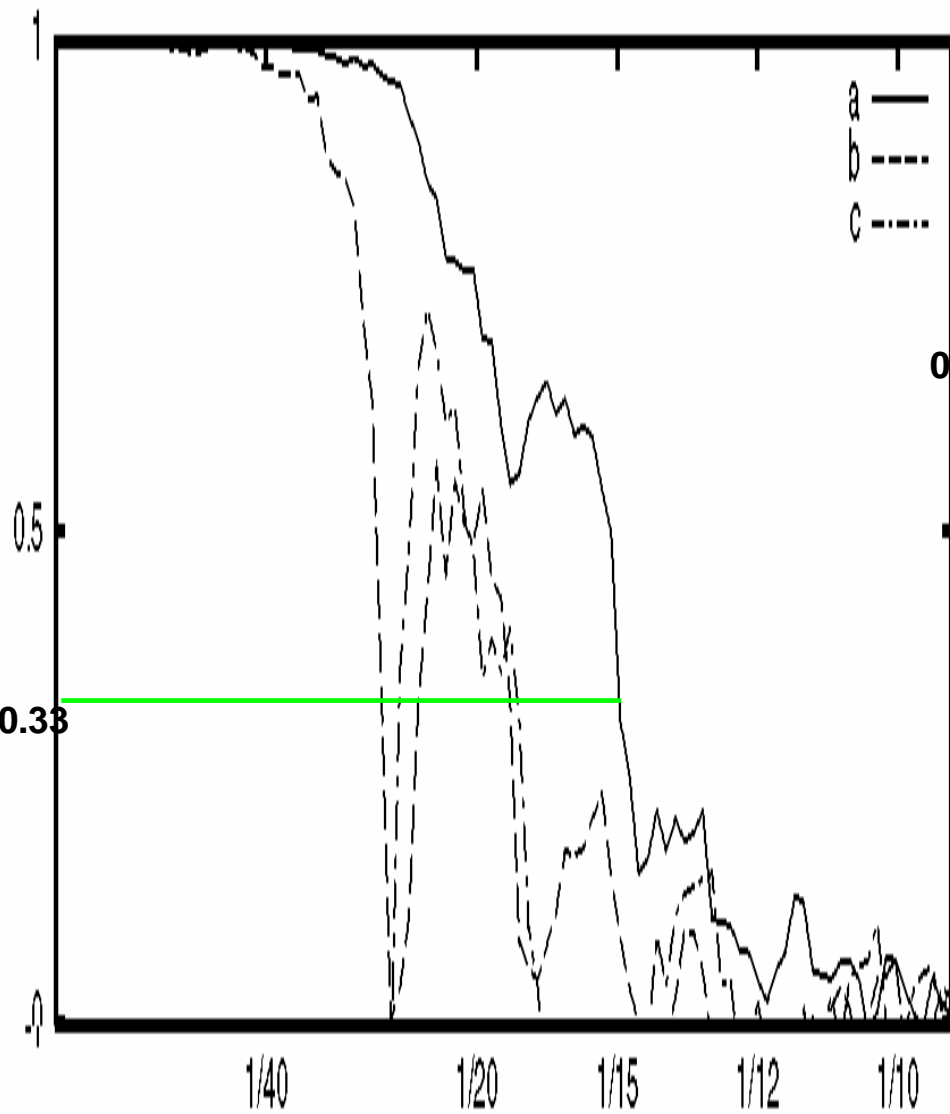
$$SSNR = \frac{FSC^2}{1 - FSC^2}$$

$$SSNR = 1 \quad \Rightarrow \quad FSC = \sqrt{1/2} = 0.71$$

Crossresolution



Resolution *versus* crossresolution



Summary

- ❖ The concept of optical resolution is not applicable to electron microscopy and single particle reconstruction.
- ❖ Resolution measures in EM estimate the “internal consistency” of the results. The outcome is prone to errors. The existing resolution measures cannot distinguish between “true” signal and the aligned (correlated) noise component in the data.
- ❖ FSC and SSNR are mathematically largely equivalent, although the SSNR-based estimate of the spectral signal to noise ratio has lower statistical uncertainty than the FSC-based estimate.
- ❖ The SSNR should be used whenever the number of the input projections is too small to make the division into halves possible (tomography).
- ❖ A reasonable resolution criterion should be based on the SSNR in the data and set such that the Fourier coefficients with a dominant noise component are excluded from the final analysis. For example, $SSNR=1 \Rightarrow FSC=0.333$.
- ❖ The shape of the FSC curve defines an optimum filter for the average/reconstruction.