CTF Correction, Resolution and Model Bias

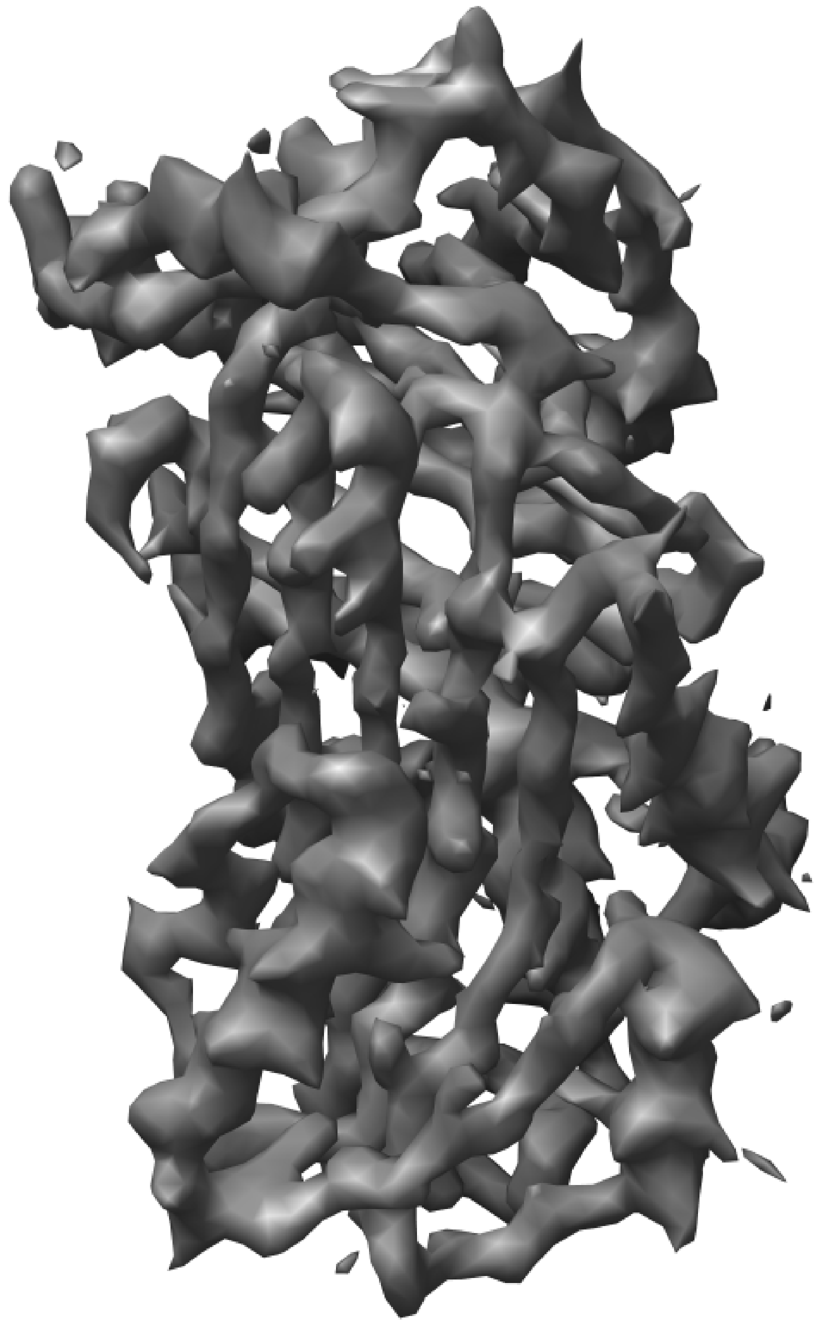
Steve Ludtke
sludtke@bcm.edu

NCRR

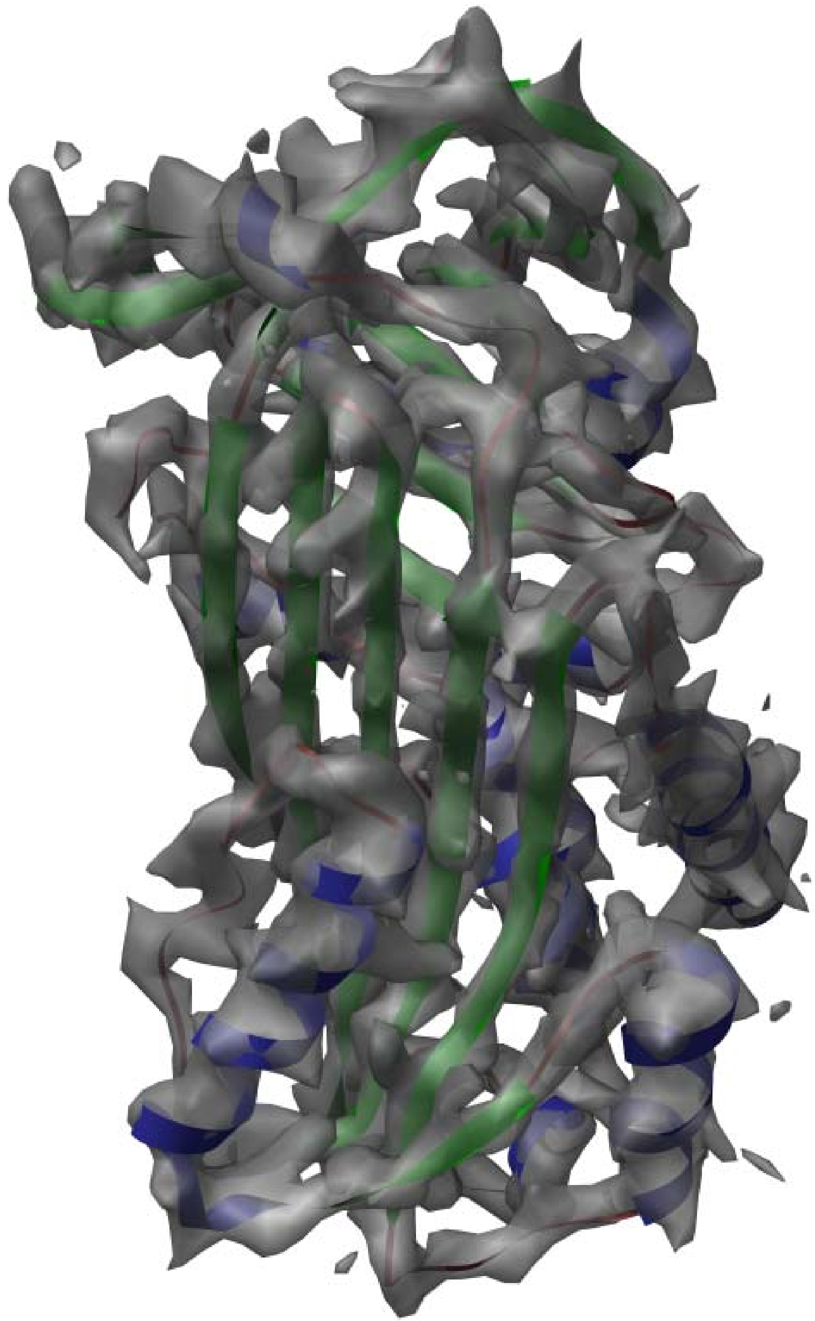
NCMI
National Center for Macromolecular Imaging

4CAA

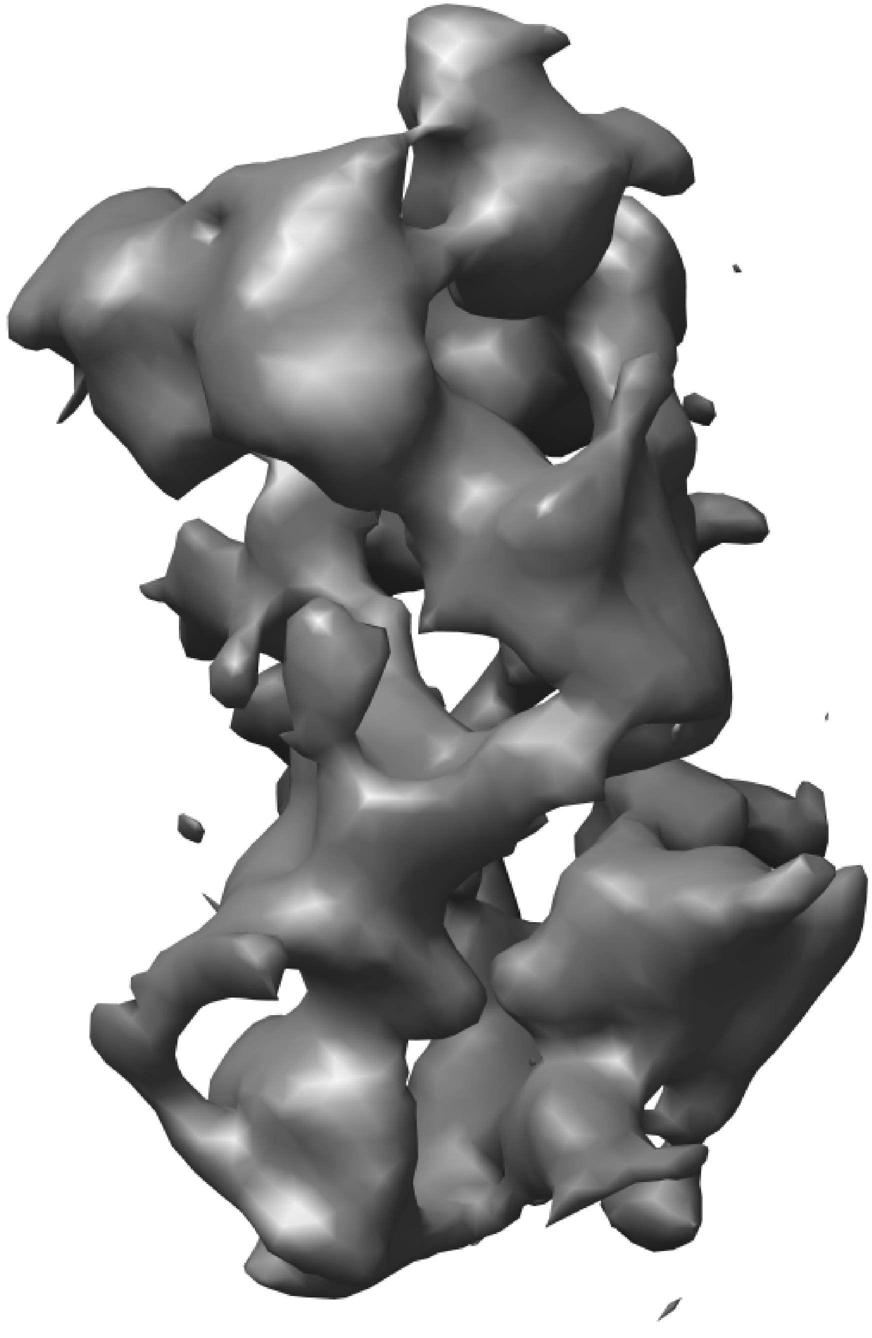
4CAA



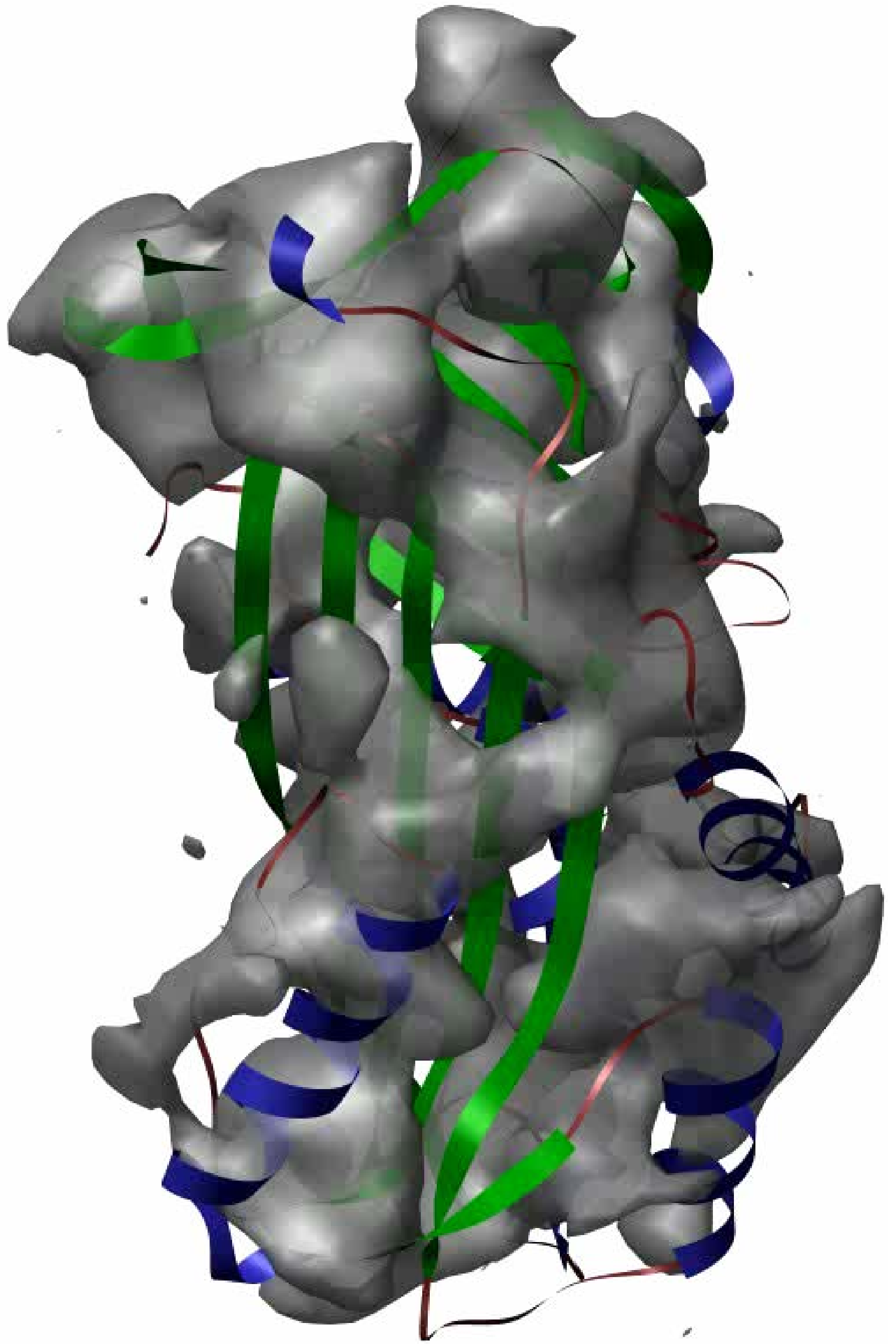
4CAA



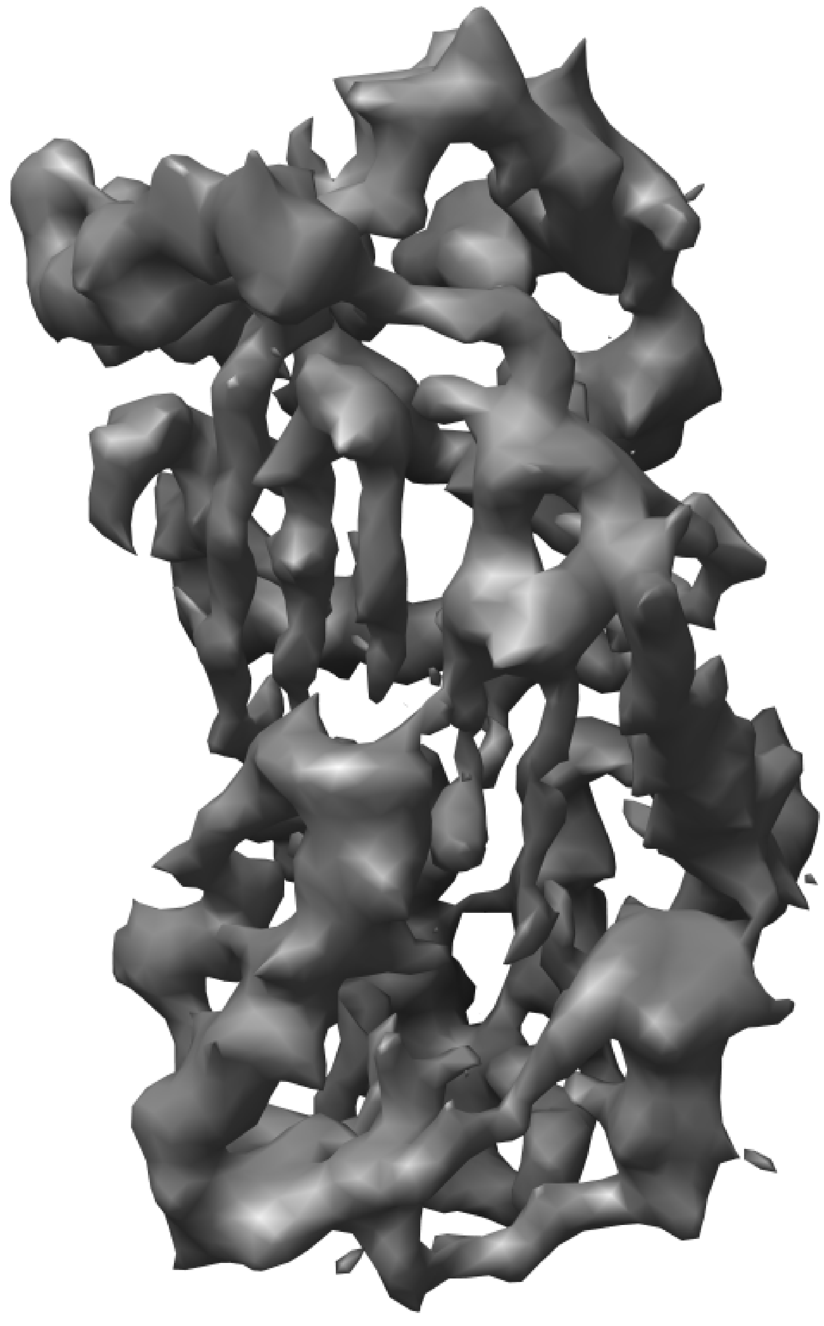
No CTF Corr (1 defocus)



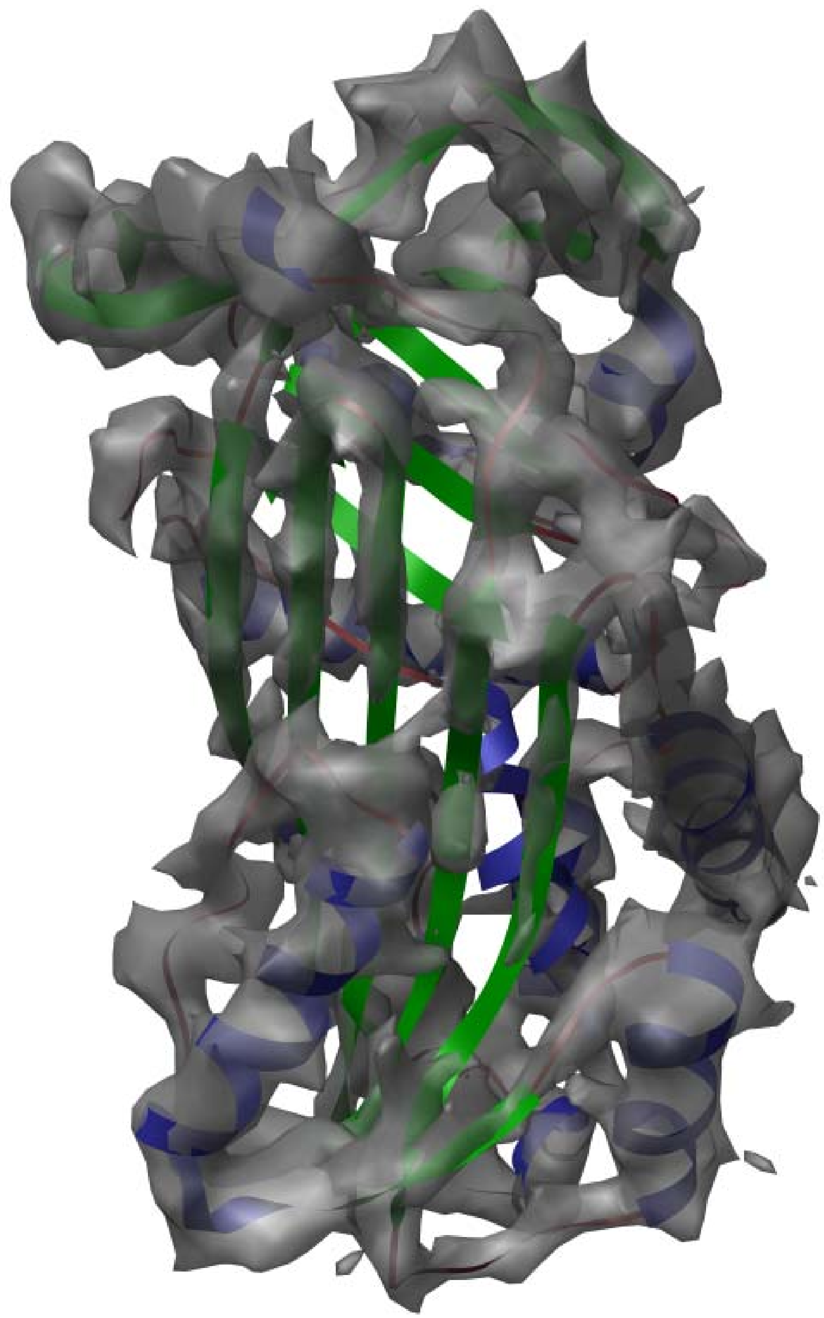
No CTF Corr (1 defocus)



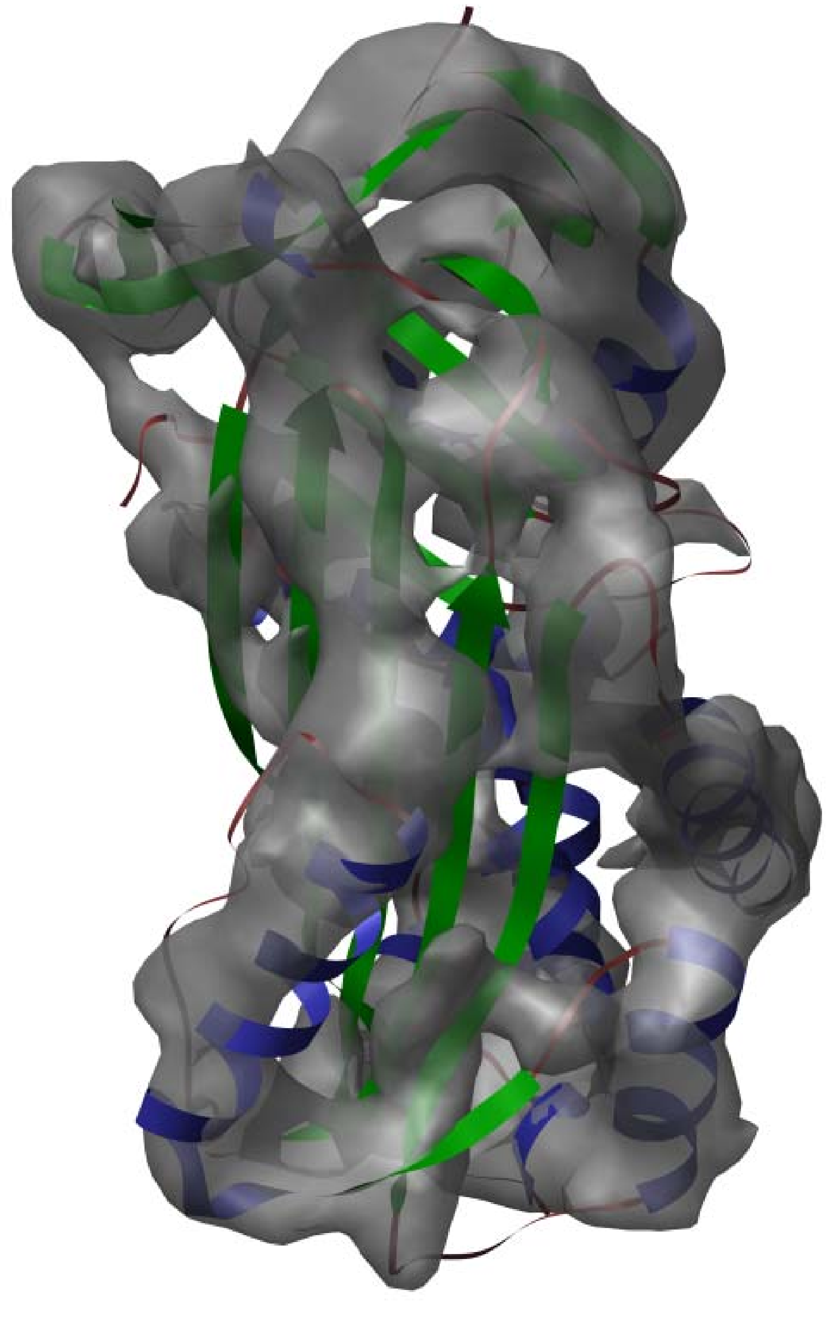
Phase Flipped (1 defocus)



Phase Flipped (1 defocus)

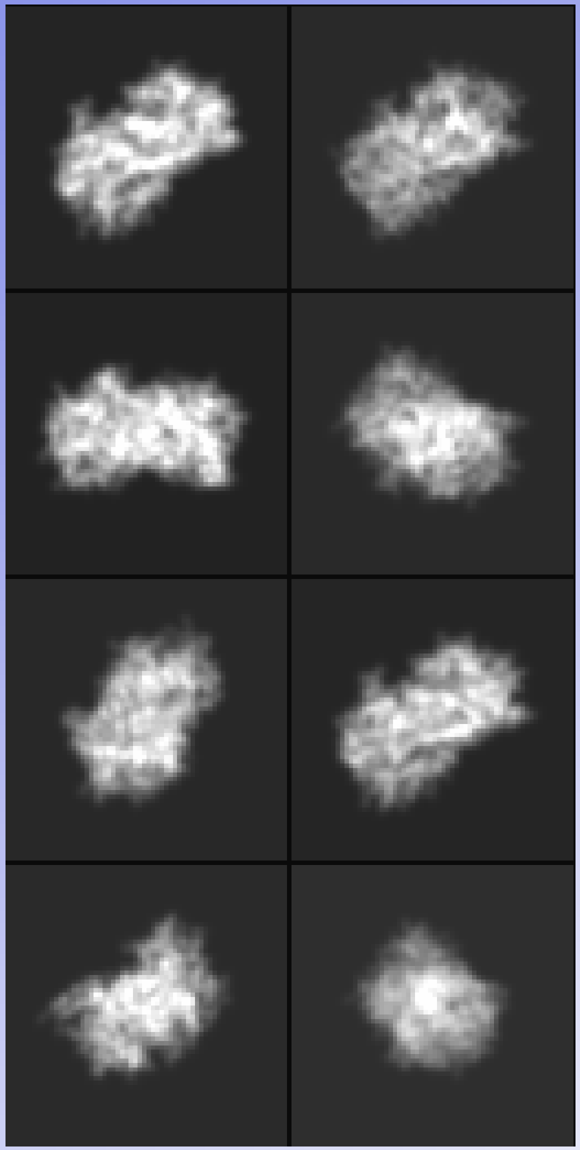


Phase Flipped (mult defocus)

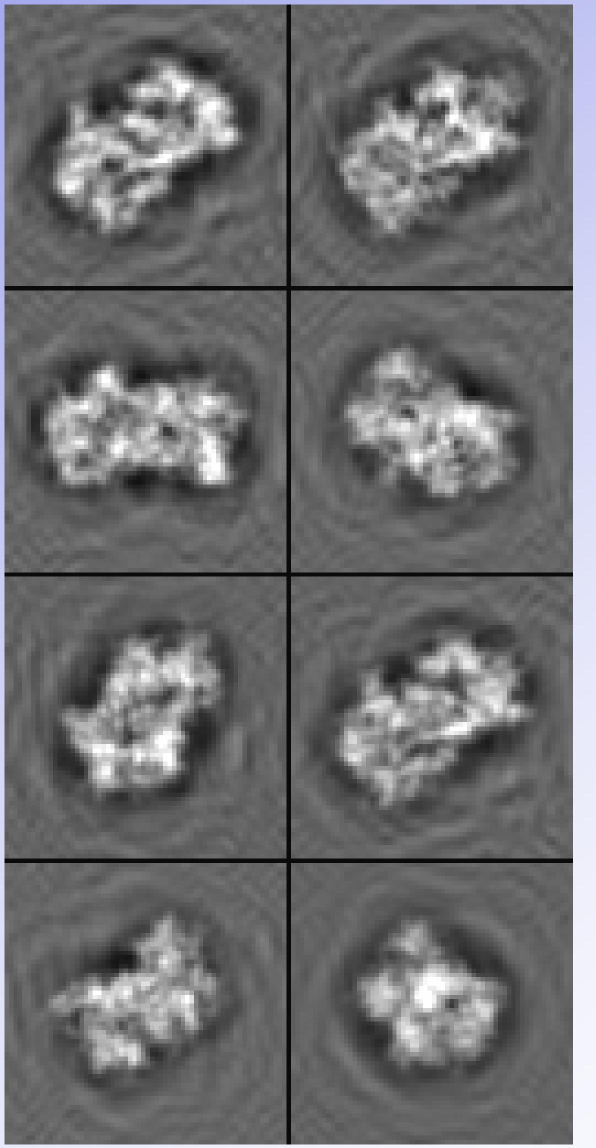


4CAA in 2D

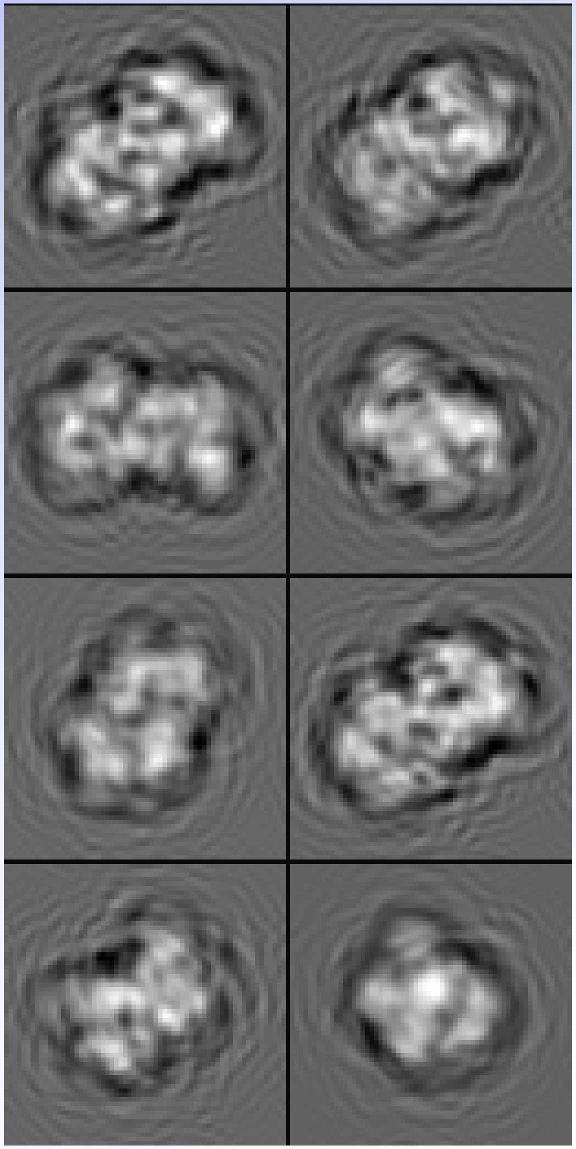
No CTF



CTF Amp



Amp & Pha



CTF Correction

Measured Image → $\overline{M}$; Ideal Particle → $\overline{F}$; Random Noise → $\overline{N}$

$$\overline{M}(s,\theta) = \overline{F}(s,\theta)C(s)E(s) + \overline{N}(s,\theta)$$

$$C(s) = \sqrt{1-Q^2}\sin\gamma + Q\cos\gamma$$

$$\gamma = -\pi\left(\frac{1}{2}C_S\lambda^3 s^4 - \Delta Z\lambda s^2\right)$$

$$E(s) = e^{-Bs^2}$$

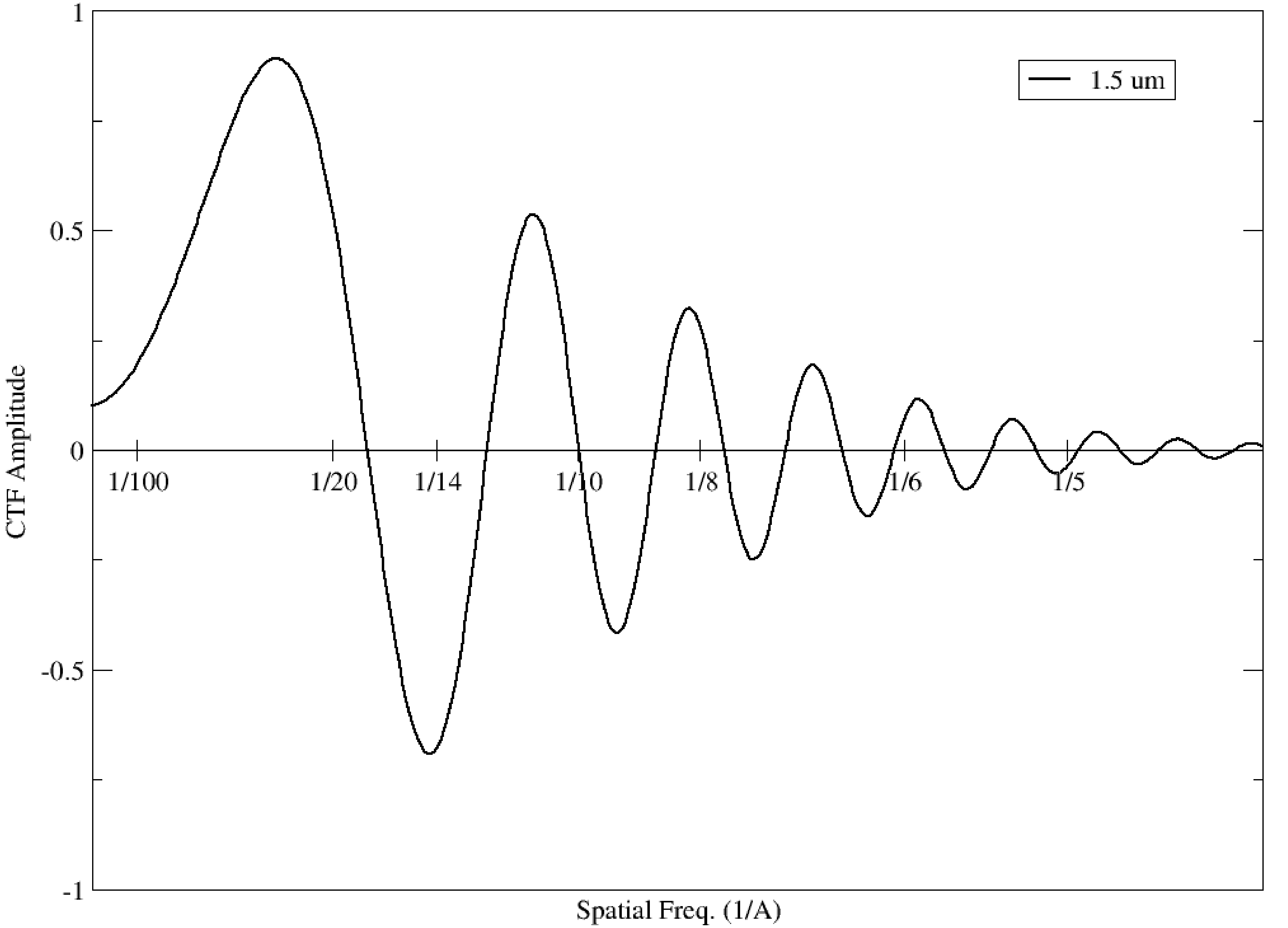
CTF Correction

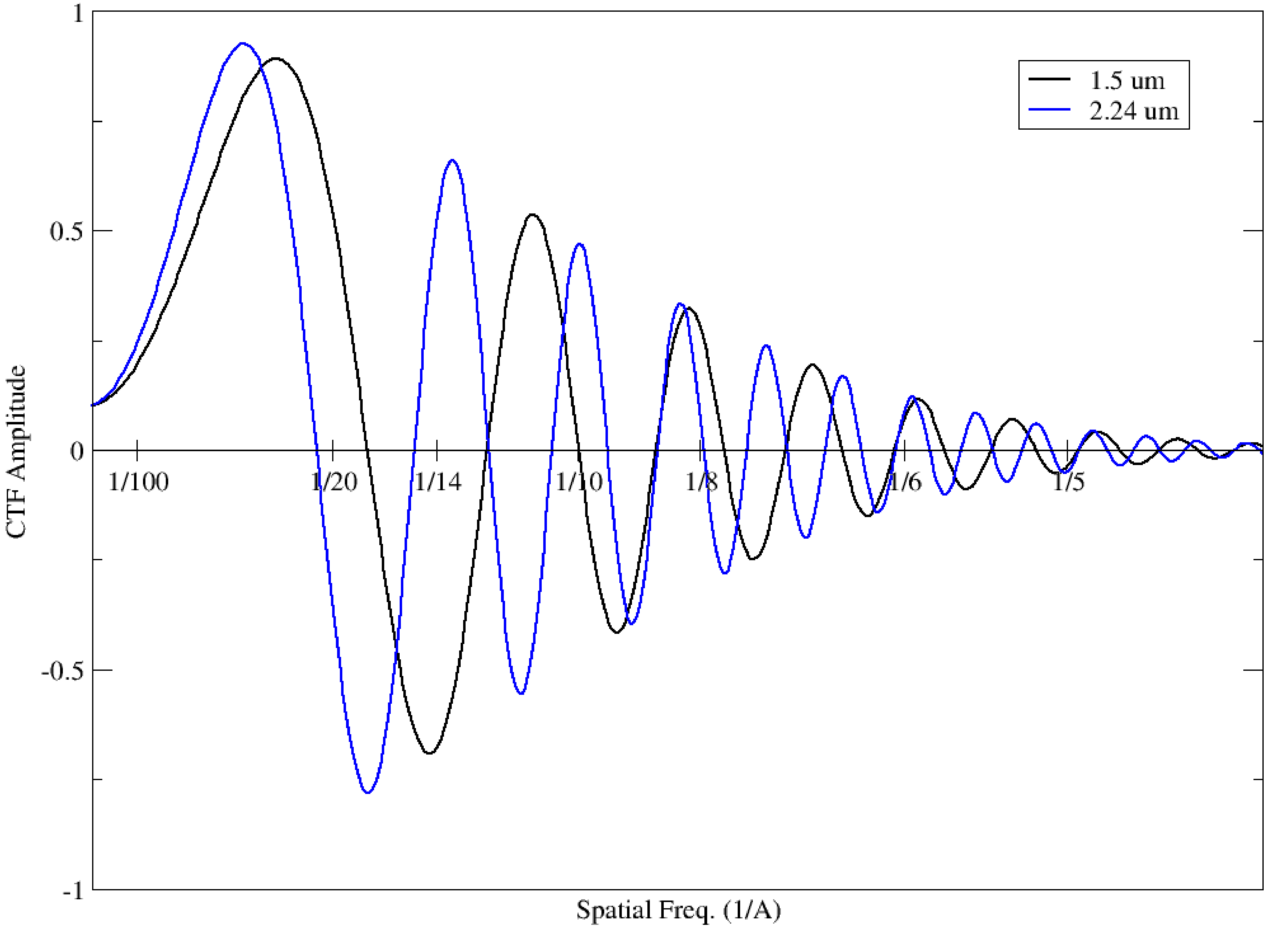
$$\overline{M}(s,\theta) = \overline{F}(s,\theta)C(s)E(s) + \overline{N}(s,\theta)$$

$$C(s) = \sqrt{1-Q^2}\sin\gamma + Q\cos\gamma$$

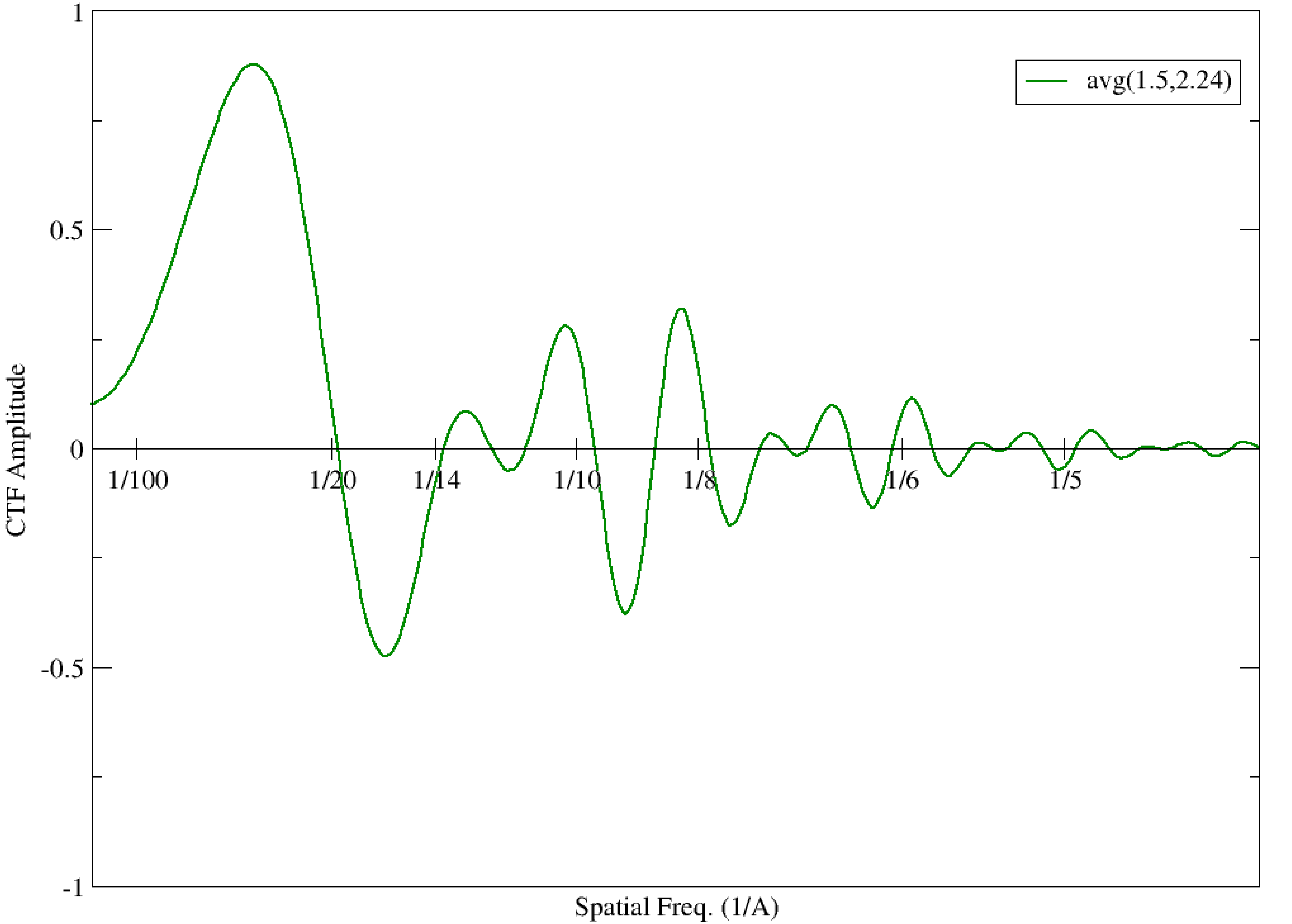
$$\gamma = -\pi(\frac{1}{2}C_S\lambda^3 s^4 - \Delta Z\lambda s^2)$$

$$E(s) = e^{-Bs^2}$$

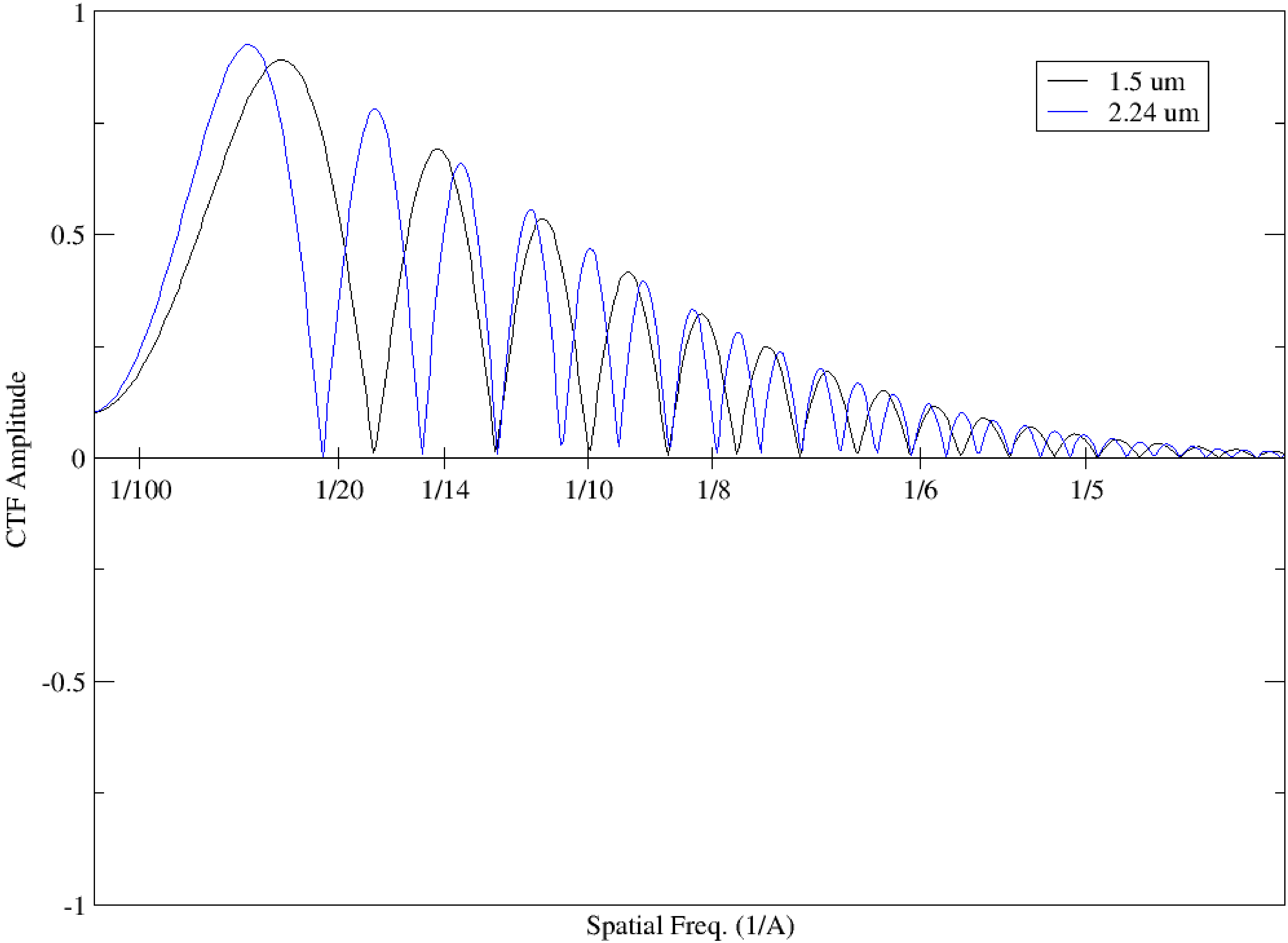

1
0.5
0
-0.5
-1
CTF Amplitude
1.5 um
1/100
1/20
1/14
1/10
1/8
1/6
1/5
Spatial Freq. (1/A)

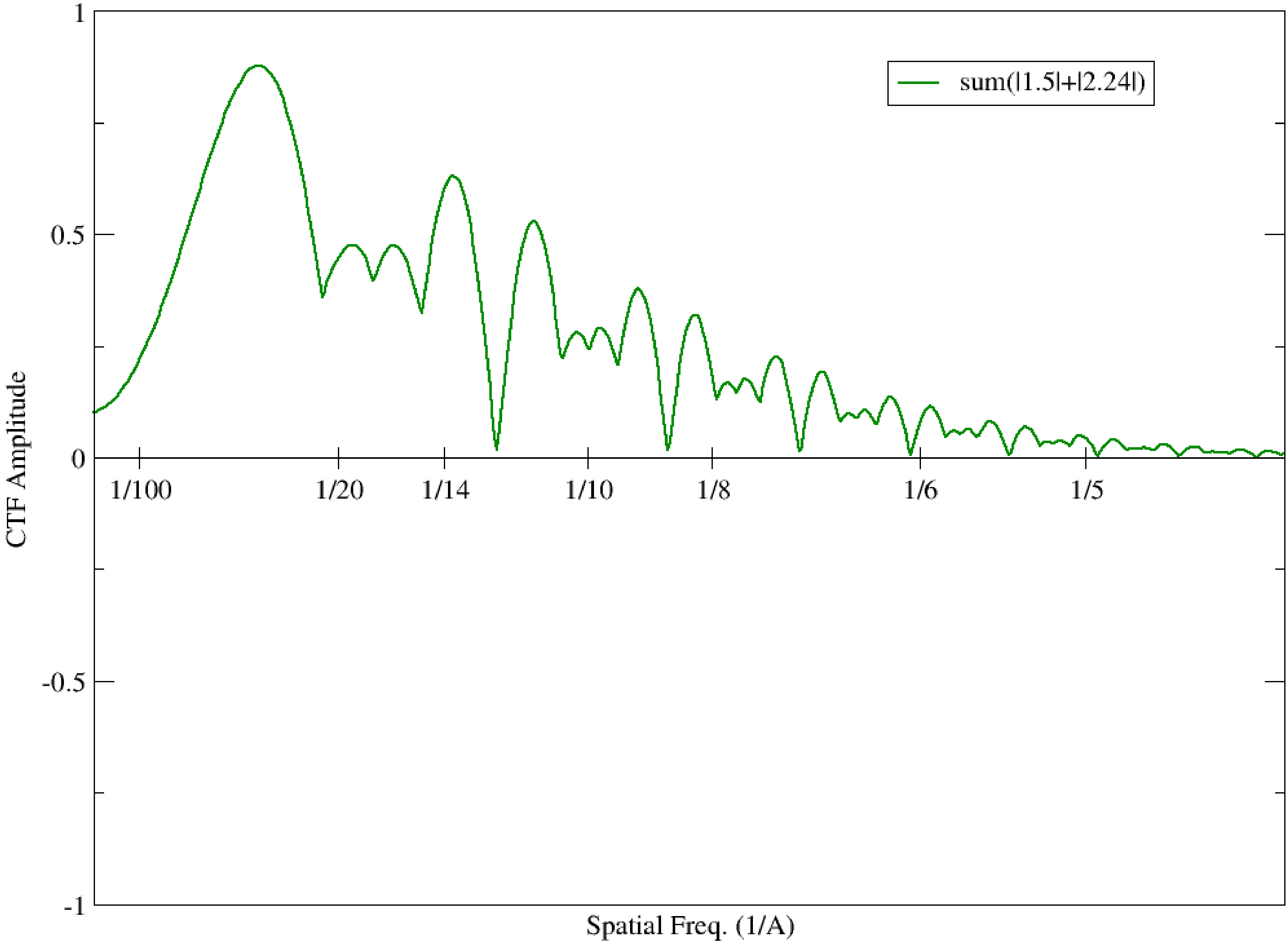


1
0.5
0
-0.5
-1
CTF Amplitude
1/100
1/20
1/14
1/10
1/8
1/6
1/5
Spatial Freq. (1/A)
1.5 um
2.24 um

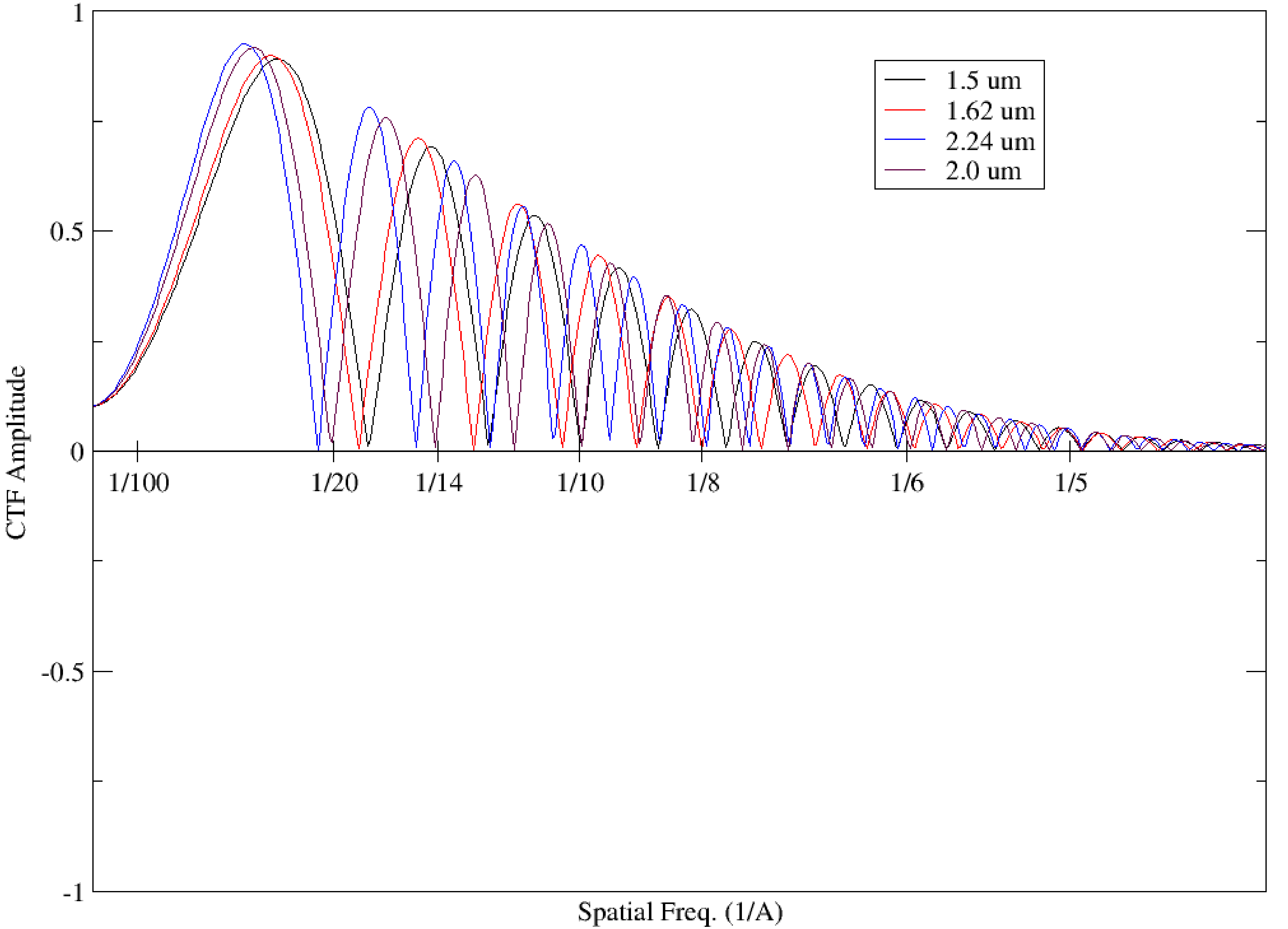


avg(1.5,2.24)
1
0.5
0
-0.5
-1
CTF Amplitude
1/100
1/20
1/14
1/10
1/8
1/6
1/5
Spatial Freq. (1/A)

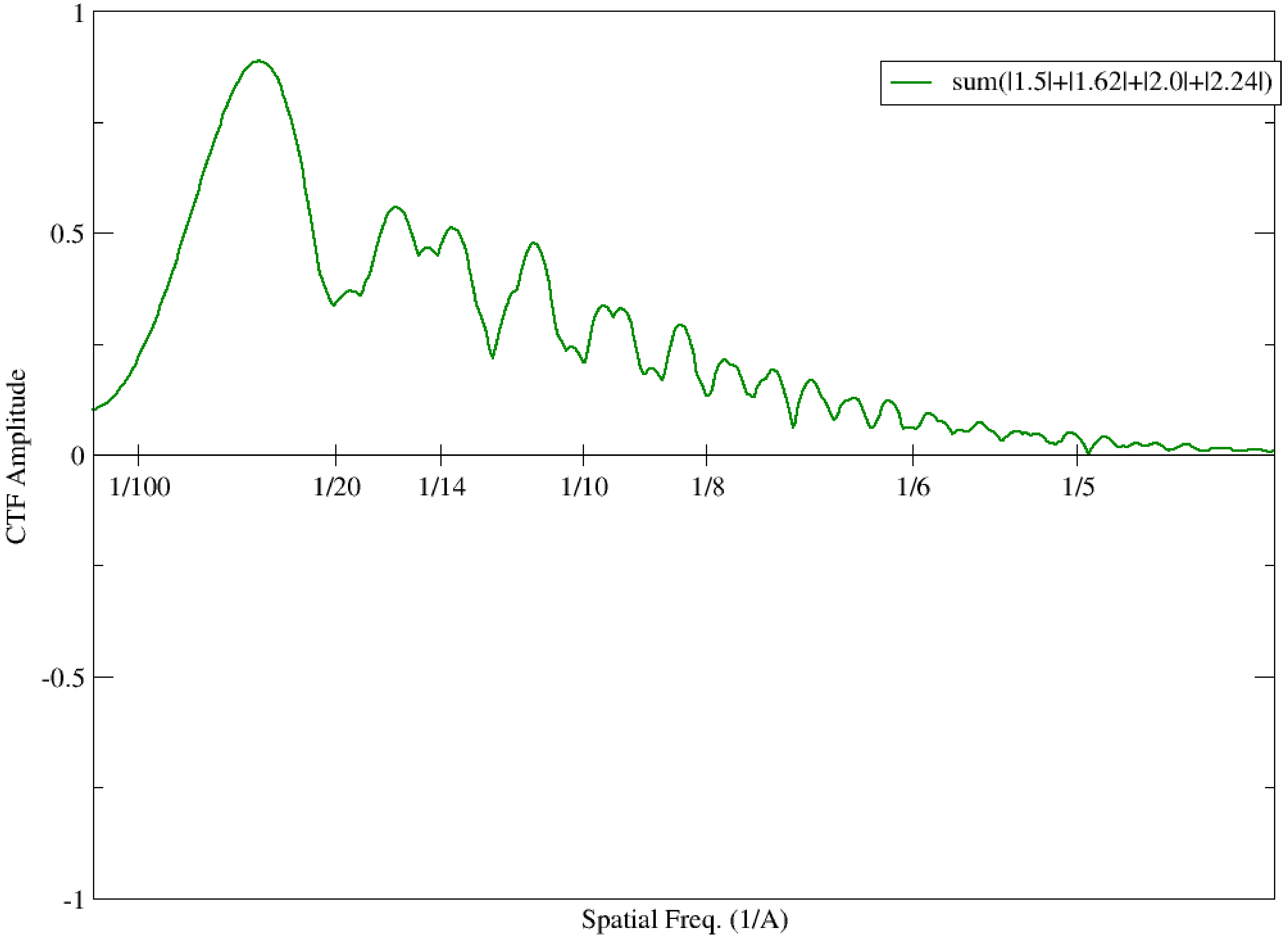

1.5 um
2.24 um
CTF Amplitude
Spatial Freq. (1/A)
1
0.5
0
-0.5
-1
1/100
1/20
1/14
1/10
1/8
1/6
1/5



sum(|1.5|+|2.24|)
1
0.5
0
-0.5
-1
1/100
1/20
1/14
1/10
1/8
1/6
1/5
CTF Amplitude
Spatial Freq. (1/A)



1.5 um
1.62 um
2.24 um
2.0 um
CTF Amplitude
Spatial Freq. (1/A)
1
0.5
0
-0.5
-1
1/100
1/20
1/14
1/10
1/8
1/6
1/5



1
0.5
0
-0.5
-1
CTF Amplitude
sum(|1.5|+|1.62|+|2.0|+|2.24|)
1/100
1/20
1/14
1/10
1/8
1/6
1/5
Spatial Freq. (1/A)

CTF Correction

$$\overline{M}(s,\theta) = \overline{F}(s,\theta)C(s)E(s) + \overline{N}(s,\theta)$$

$$C(s) = \sqrt{1-Q^2}\sin\gamma + Q\cos\gamma$$

$$\gamma = -\pi(\frac{1}{2}C_S\lambda^3 s^4 - \Delta Z\lambda s^2)$$

$$E(s) = e^{-Bs^2}$$

$$N(s)^2 = n_1 e^{n_2 s + n_3 s^2 + n_4\sqrt{s}}$$

$$M(s)^2 = F(s)^2 C(s)^2 E(s)^2 + N(s)^2$$

CTF Correction

Reconstruction — Weight — Measured image

$$\overline{T}(s,\theta)=\sum_{i} k_i \overline{M}_i(s,\theta)$$

$$k_i = ?$$

- Maximize SNR of *T(s,θ)*
- Minimize RMSD between T and F

$$\sqrt{\sum_{x,y}(t(x,y)-f(x,y))^2}$$

CTF Correction

Wiener Filter — CTF Correction — SNR Weight

$$\overline{T}(s,\theta)=\frac{F^2(s)R(s)}{1+F^2(s)R(s)}\sum_i \frac{1}{C_i(s)E_i(s)}\frac{R_i(s)}{R(s)}\overline{M}_i(s,\theta)$$

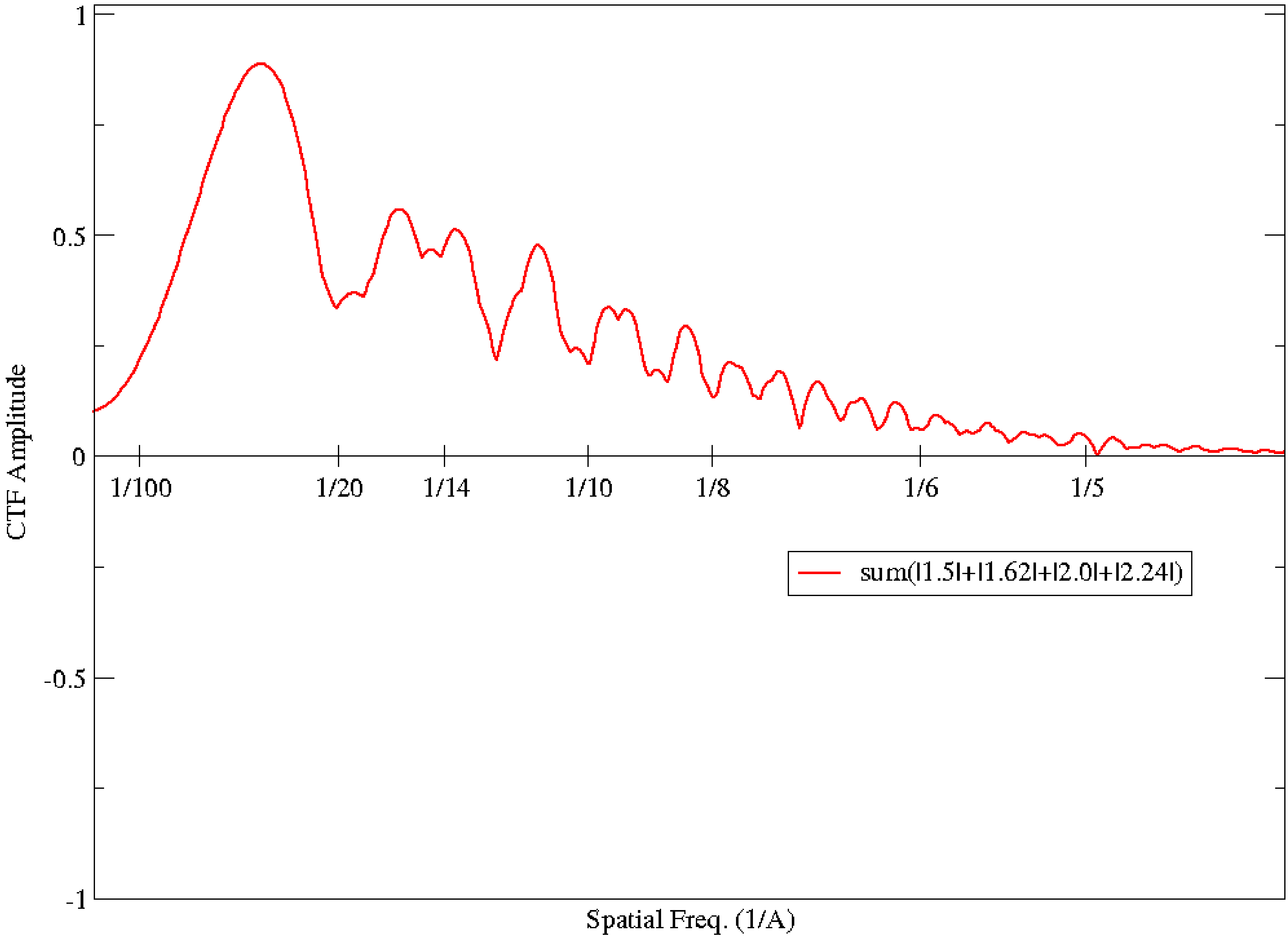
$$R_i(s)=\frac{C_i^2(s)E_i^2(s)}{N_i^2(s)} \qquad R(s)=\sum_i \frac{C_i^2(s)E_i^2(s)}{N_i^2(s)}$$

CTF Correction

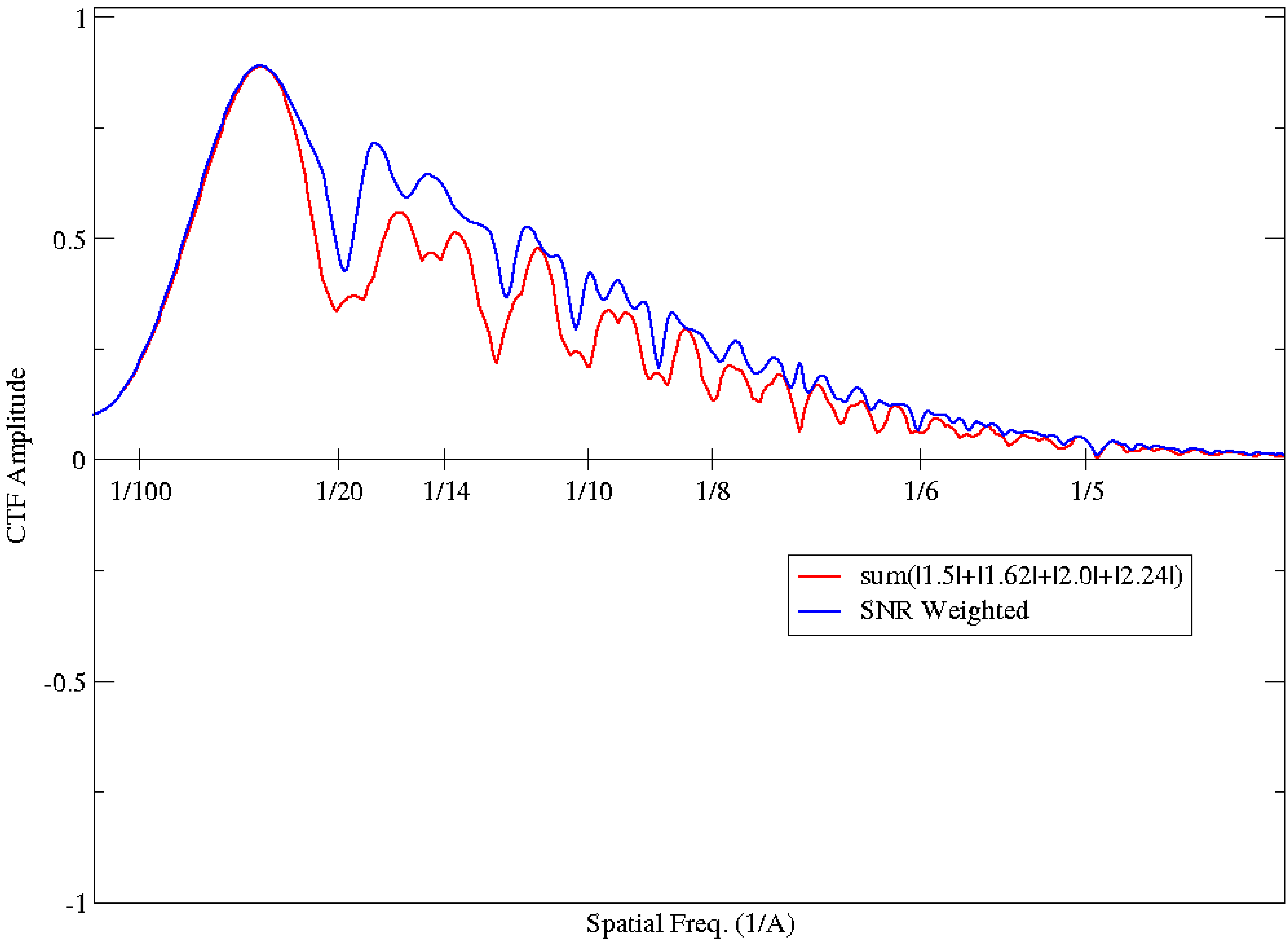
Wiener Filter — CTF Correction — SNR Weight

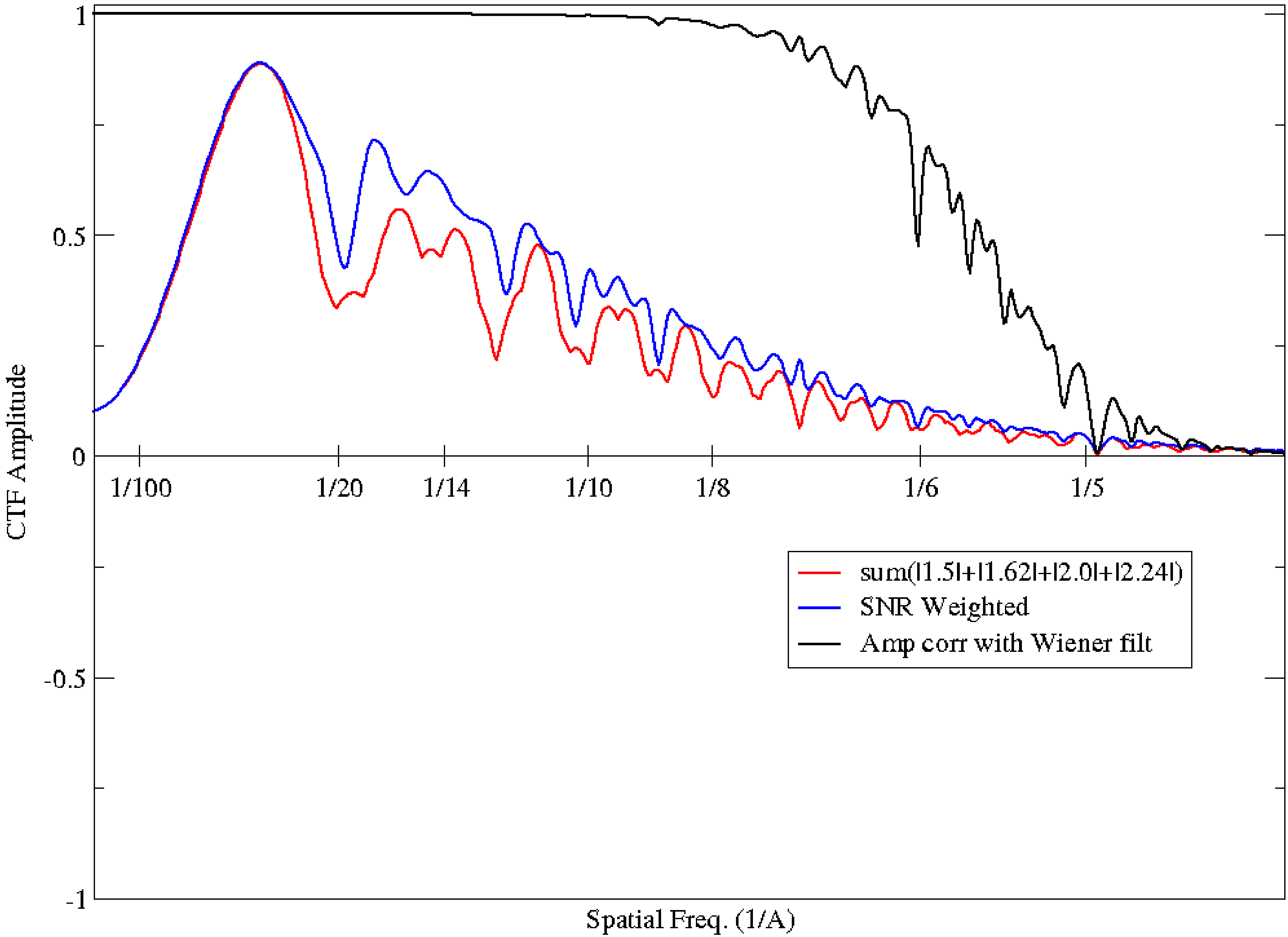
$$\overline{T}(s,\theta)=\frac{F^2(s)R(s)}{1+F^2(s)R(s)}\sum_i \frac{1}{C_i(s)E_i(s)}\frac{R_i(s)}{R(s)}\overline{M}_i(s,\theta)$$

Note that this factor depends on ALL of the data and means you cannot ‘precorrect’ the data then do a reconstruction. You can phase-flip in preprocessing, but Wiener filtration and weighting depend on having all of the data at once.



1
0.5
0
-0.5
-1
CTF Amplitude
1/100
1/20
1/14
1/10
1/8
1/6
1/5
Spatial Freq. (1/A)
sum(|1.5|+|1.62|+|2.0|+|2.24|)


CTF Amplitude
1
0.5
0
-0.5
-1
1/100
1/20
1/14
1/10
1/8
1/6
1/5
sum(|1.5|+|1.62|+|2.0|+|2.24|)
SNR Weighted
Spatial Freq. (1/A)



1
0.5
0
-0.5
-1
CTF Amplitude
1/100
1/20
1/14
1/10
1/8
1/6
1/5
Spatial Freq. (1/A)
sum(|1.5|+|1.62|+|2.0|+|2.24|)
SNR Weighted
Amp corr with Wiener filt

CTF Correction

$$\overline{M}(s,\theta) = \overline{F}(s,\theta)C(s)E(s) + \overline{N}(s,\theta)$$

$$C(s) = \sqrt{1-Q^2}\sin\gamma + Q\cos\gamma$$

$$\gamma = -\pi(\frac{1}{2}C_S\lambda^3 s^4 - \Delta Z\lambda s^2)$$

$$E(s) = e^{-Bs^2}$$

$$N(s)^2 = n_1 e^{n_2 s + n_3 s^2 + n_4\sqrt{s}}$$

$$M(s)^2 = F(s)^2 C(s)^2 E(s)^2 + N(s)^2$$

8 Parameters

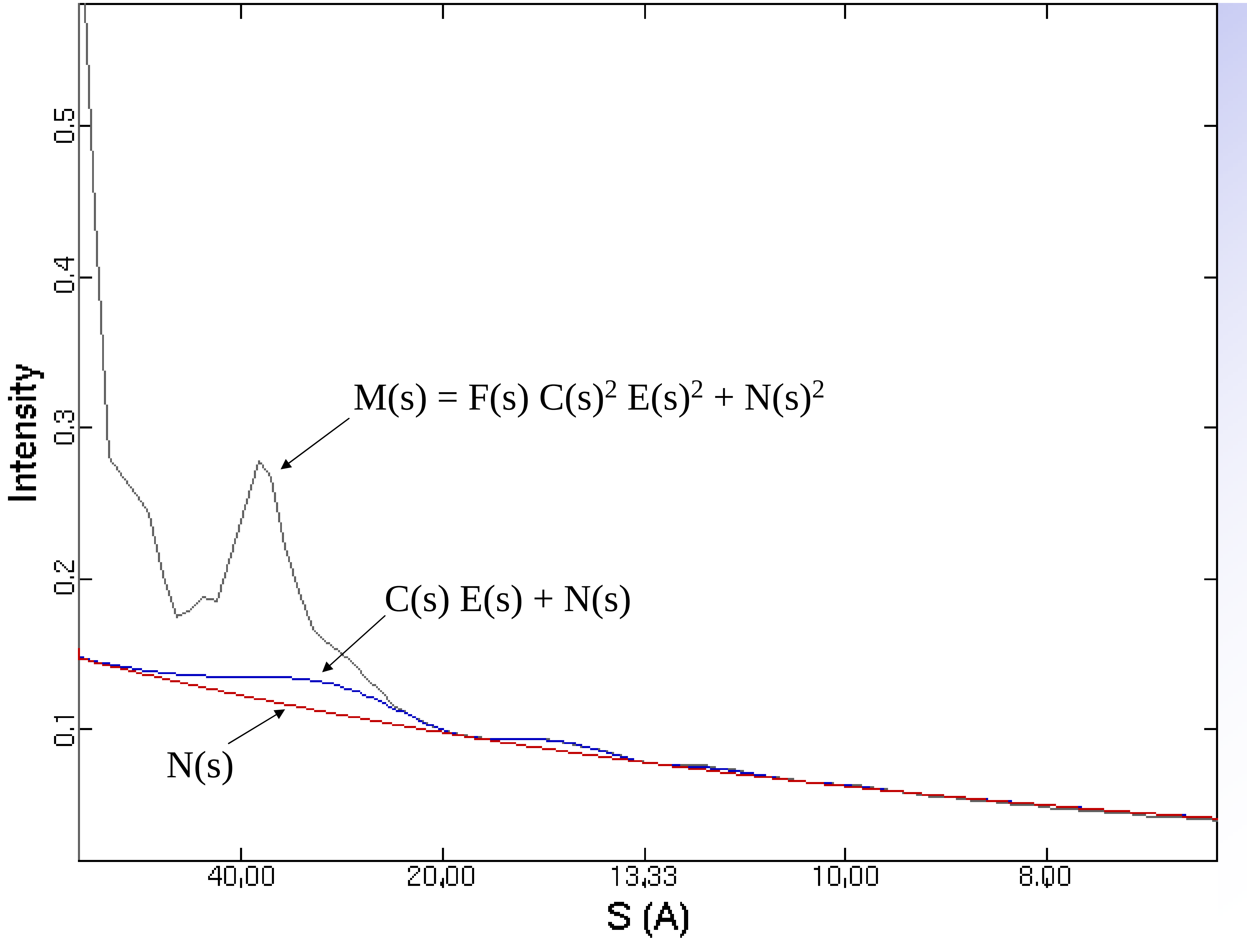
ΔZ - Defocus

Q - Amplitude Contrast

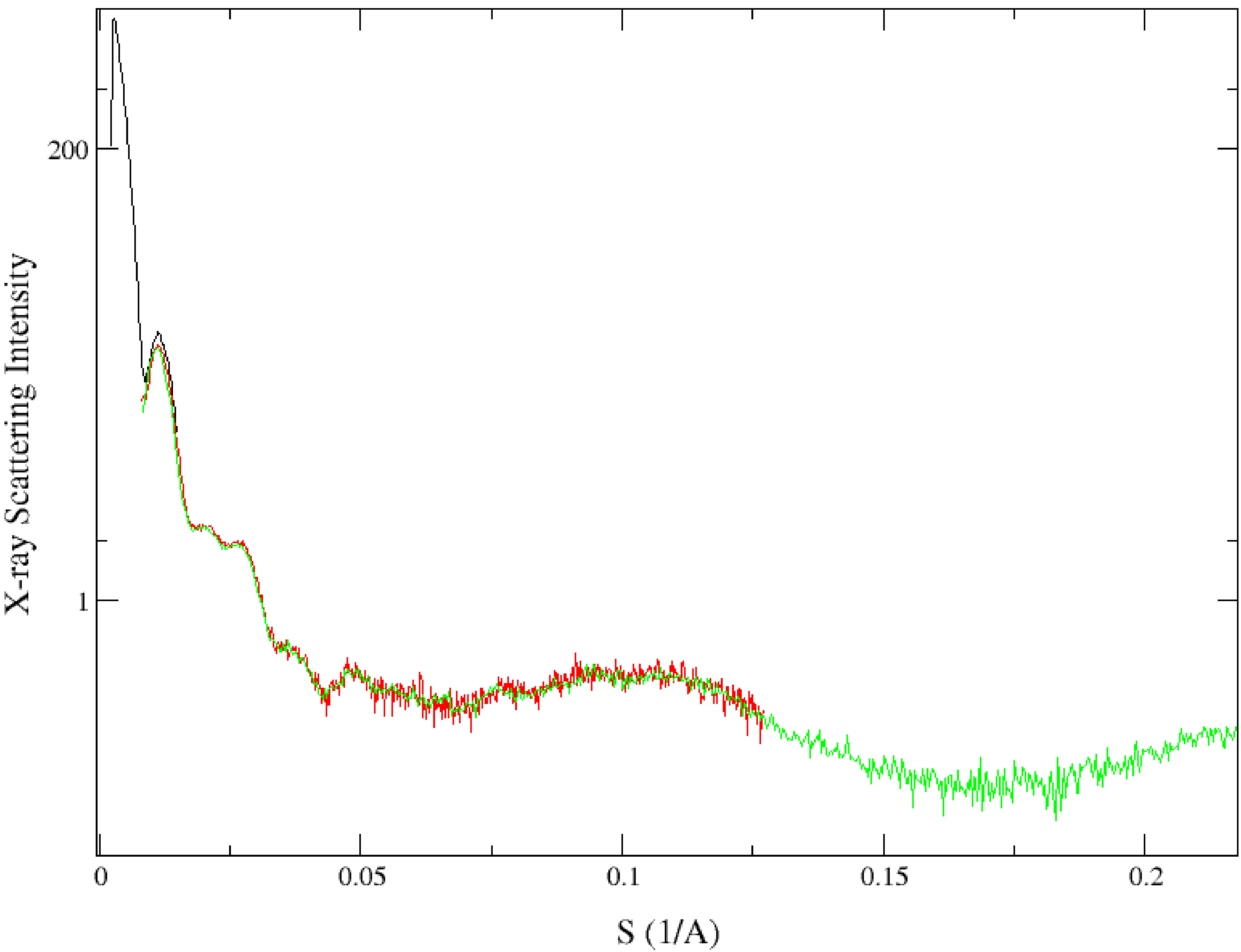
B - Gaussian Envelope Width

k - Signal Amplitude

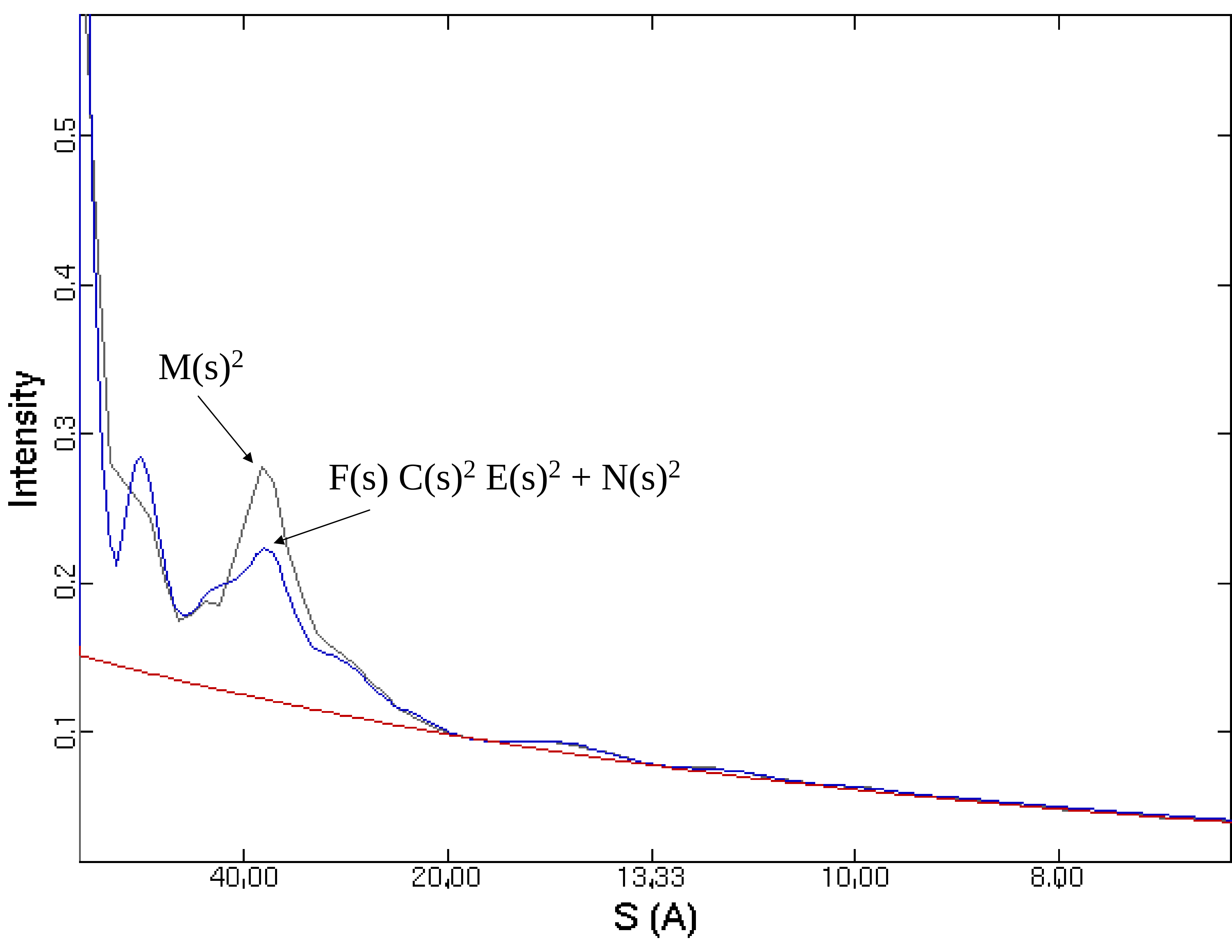
$n_{1\text{-}4}$ - Noise Parameters



M(s) = F(s) C(s)2 E(s)2 + N(s)2
C(s) E(s) + N(s)
N(s)
Intensity
0.5
0.4
0.3
0.2
0.1
40,00
20,00
13,33
10,00
8,00
S (A)



200
1
X-ray Scattering Intensity
0
0.05
0.1
0.15
0.2
S (1/A)


Intensity
0.5
0.4
0.3
0.2
0.1
M(s)^2
F(s) C(s)^2 E(s)^2 + N(s)^2
40,00
20,00
13,33
10,00
8,00
S (A)

Resolution vs. Resolvability

- Resolution – a measure of the ability to distinguish between two close but not identical values of the property being measured; it is expressed as the difference in values of a property necessary to make such a distinction; as, a microscope with a resolution of one micron; a thermometer with a resolution of one-tenth of a degree.

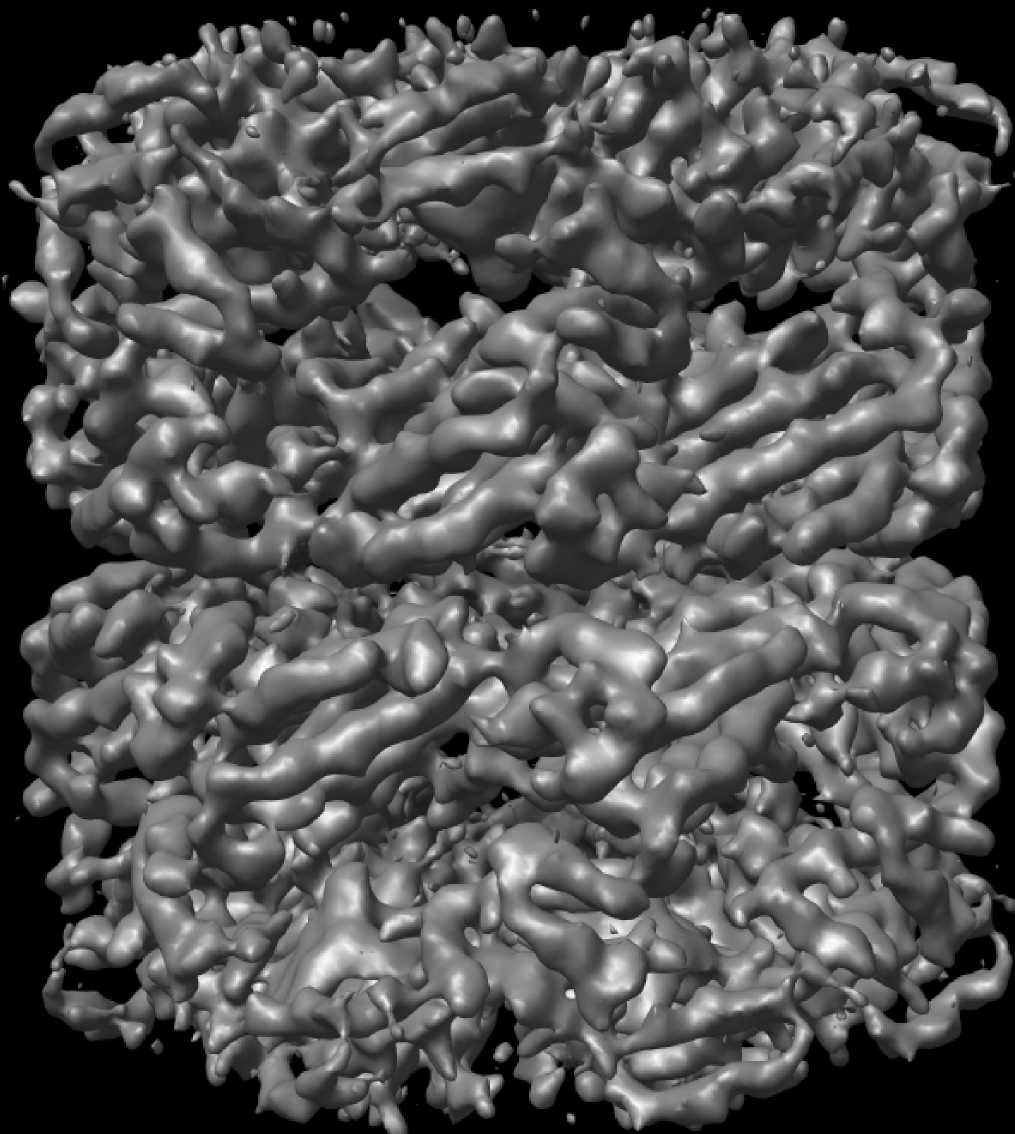
Resolution vs. Resolvability

- In optical microscopes and (most) telescopes Resolution = Resolvability
- In electron microscopes, however, **Resolution ≠ Resolvability**
- In optical microscopes, resolvability is limited by wavelength+optics. This defines the resolution.
- Wavelength of a 100 keV electron is ~0.05 Å
- Electron optics can achieve sub- Å resolvability
- **Noise** is the resolution limiting factor for biological specimens (radiation damage)!
- In SPA, resolution is a measure of Noise, **not** resolvability

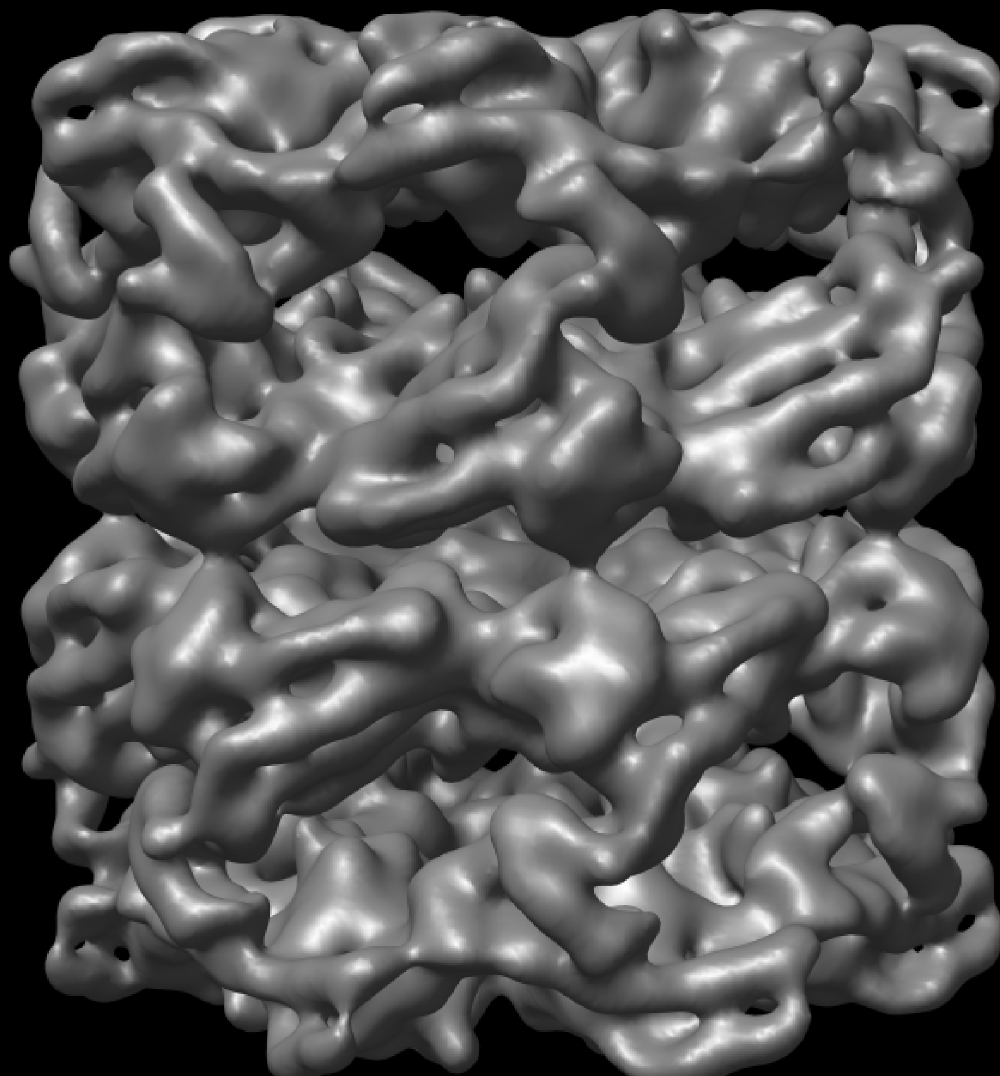
Measure Resolution ?

- ~~Look at the model and see what features you can observe, ie – if you can see α-helices, you must have better than 8-10 Å resolution~~

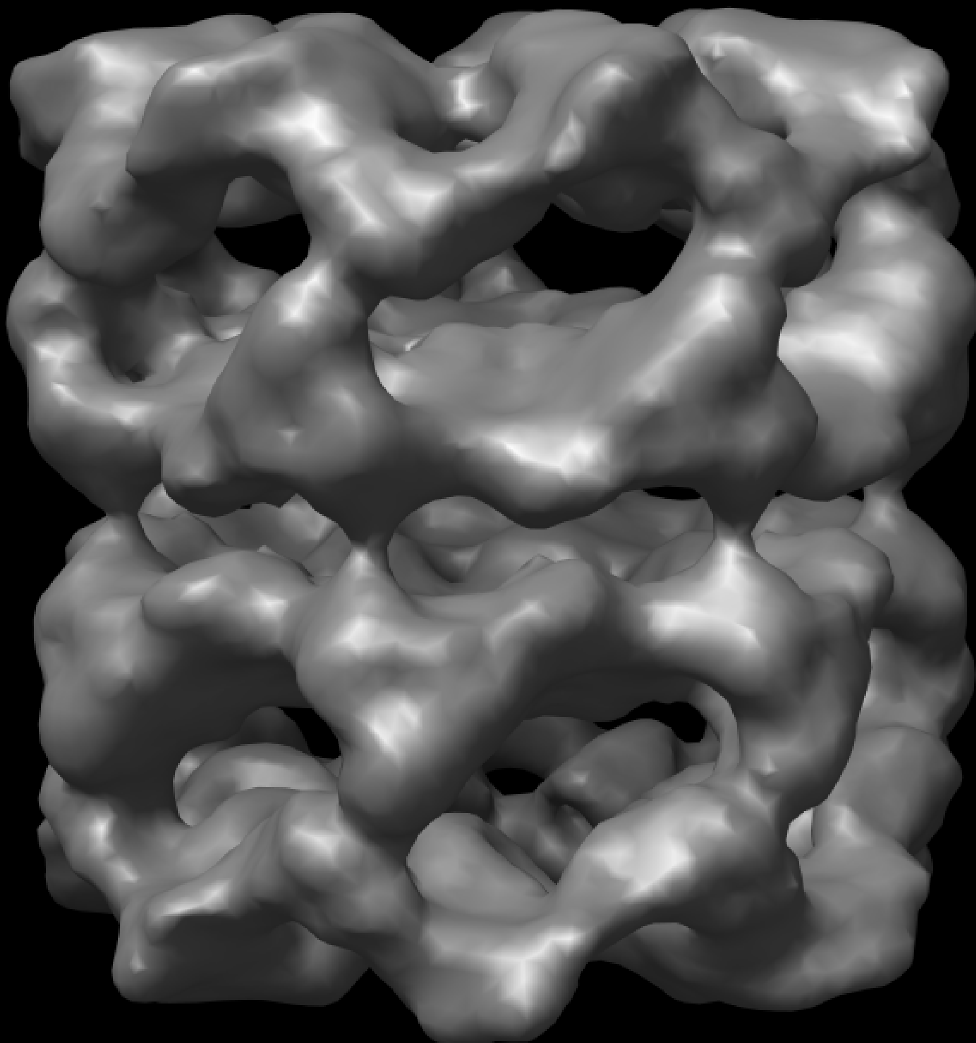
4 Å Resolution GroEL



4 Å → 11.5 Å

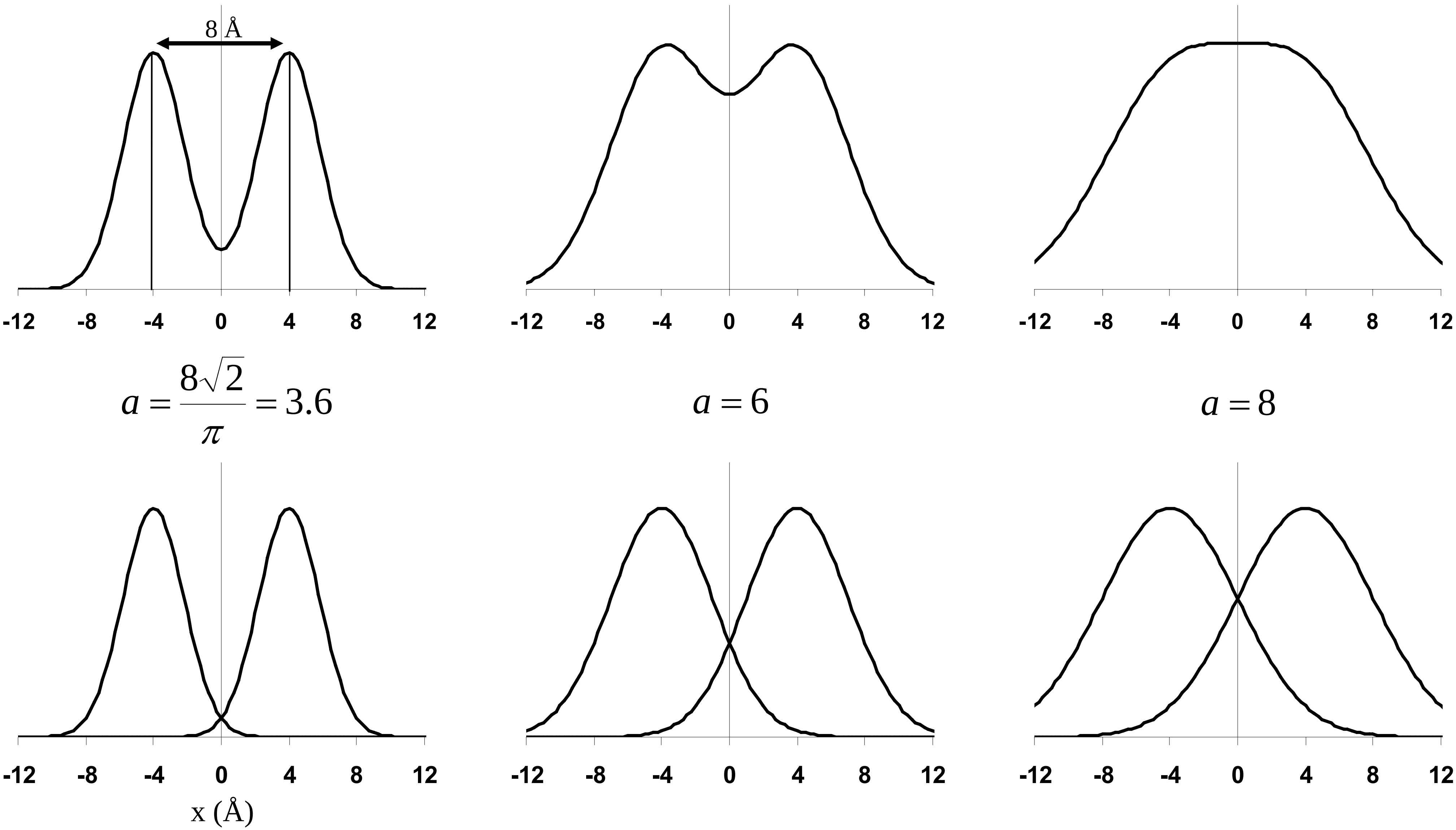


11.5 Å Resolution GroEL



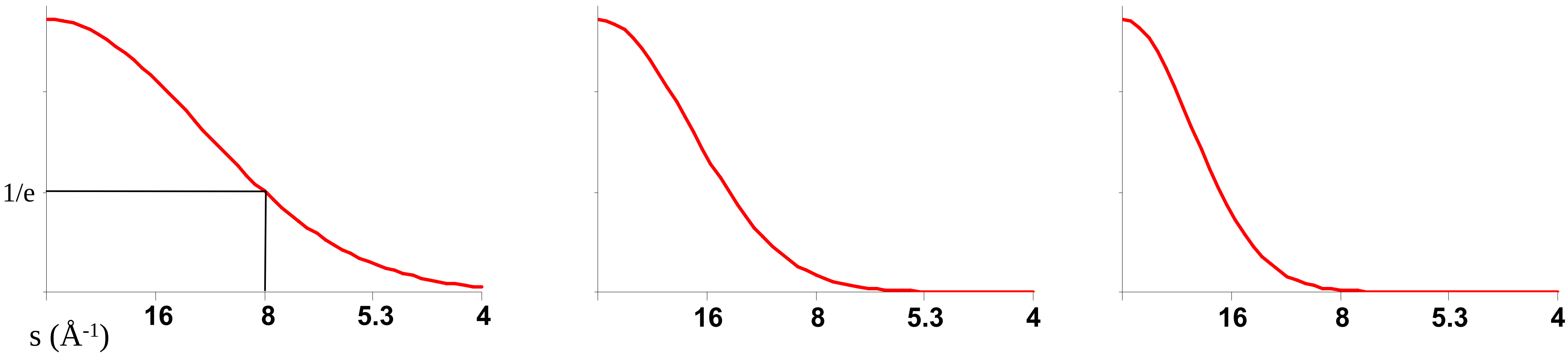
Resolution in Real Space

$$f(x) = e^{-\frac{2x^2}{a^2}}$$

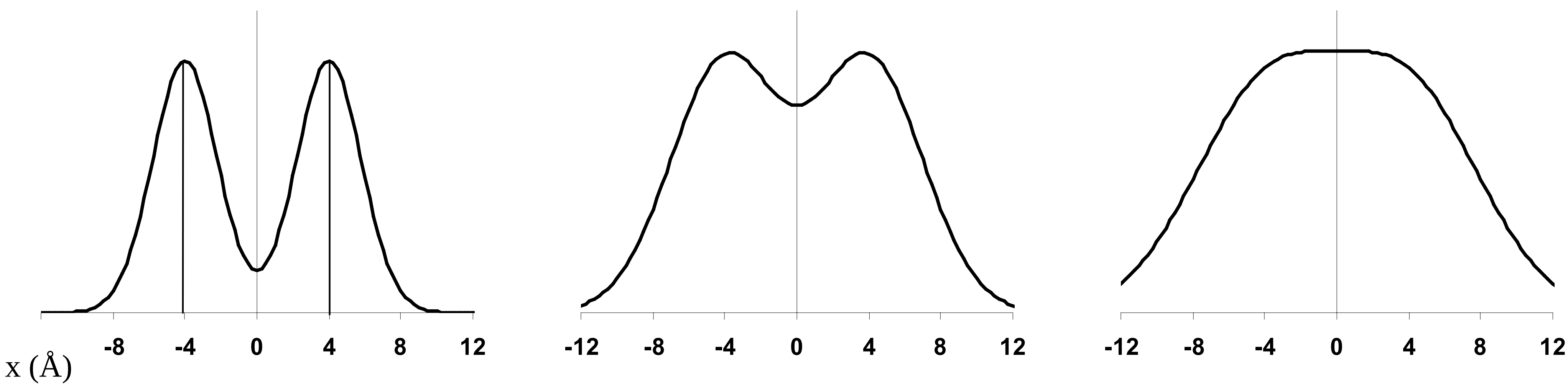


$$a = \frac{8\sqrt{2}}{\pi} = 3.6 \qquad a = 6 \qquad a = 8$$

$$f(s) = e^{-\frac{1}{2}(\pi a s)^2}$$



$$a = \frac{8\sqrt{2}}{\pi} = 3.6 \qquad a = 6 \qquad a = 8$$



$$f(x) = e^{-\frac{2x^2}{a^2}}$$

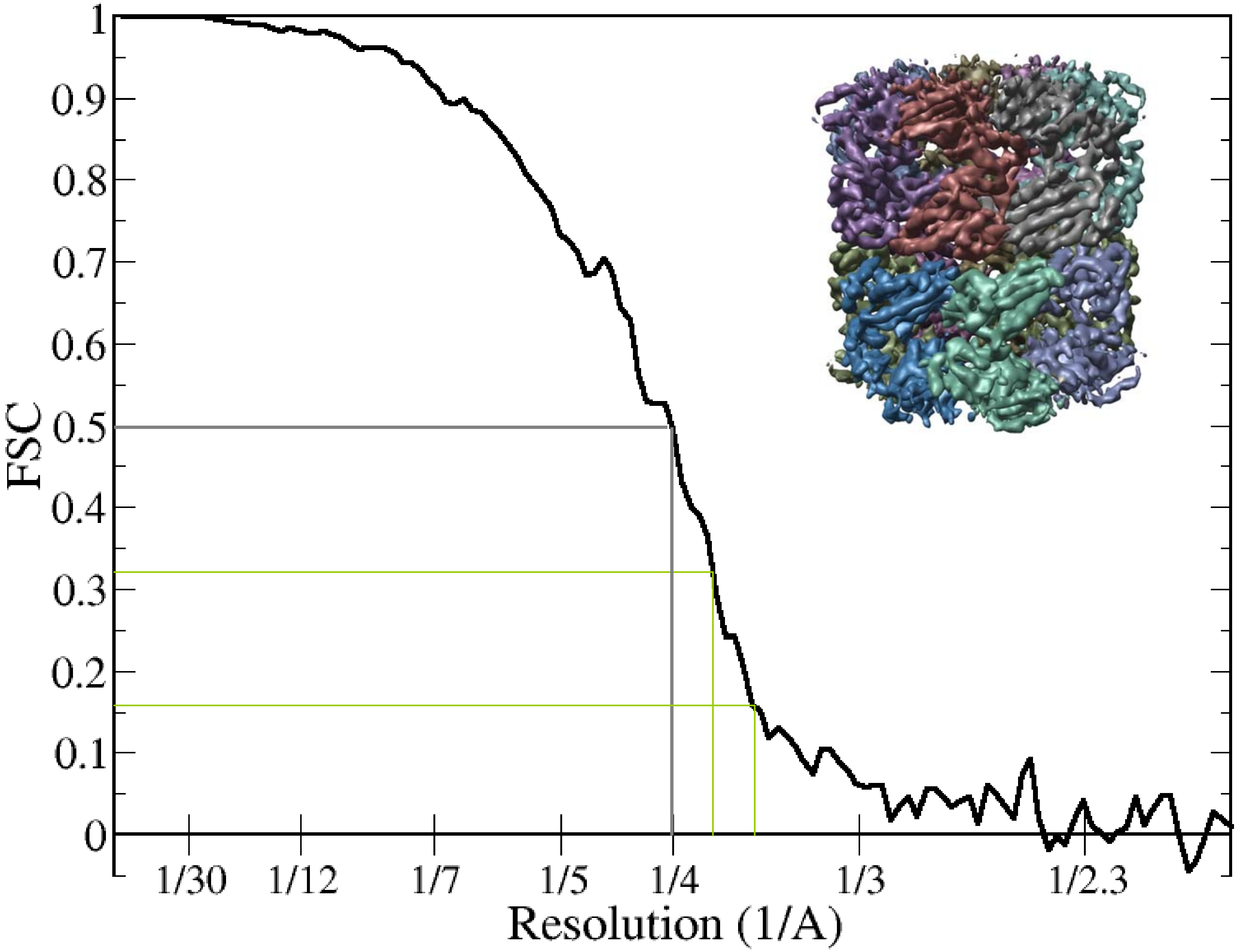
Measure Resolution

- Look at the statistical properties of our reconstruction.
- T-test ? Too expensive, and nonlinear response problems (bootstrapping).
- General practice: split particles into even and odd halves, reconstruct, compare models

Compare Models ?

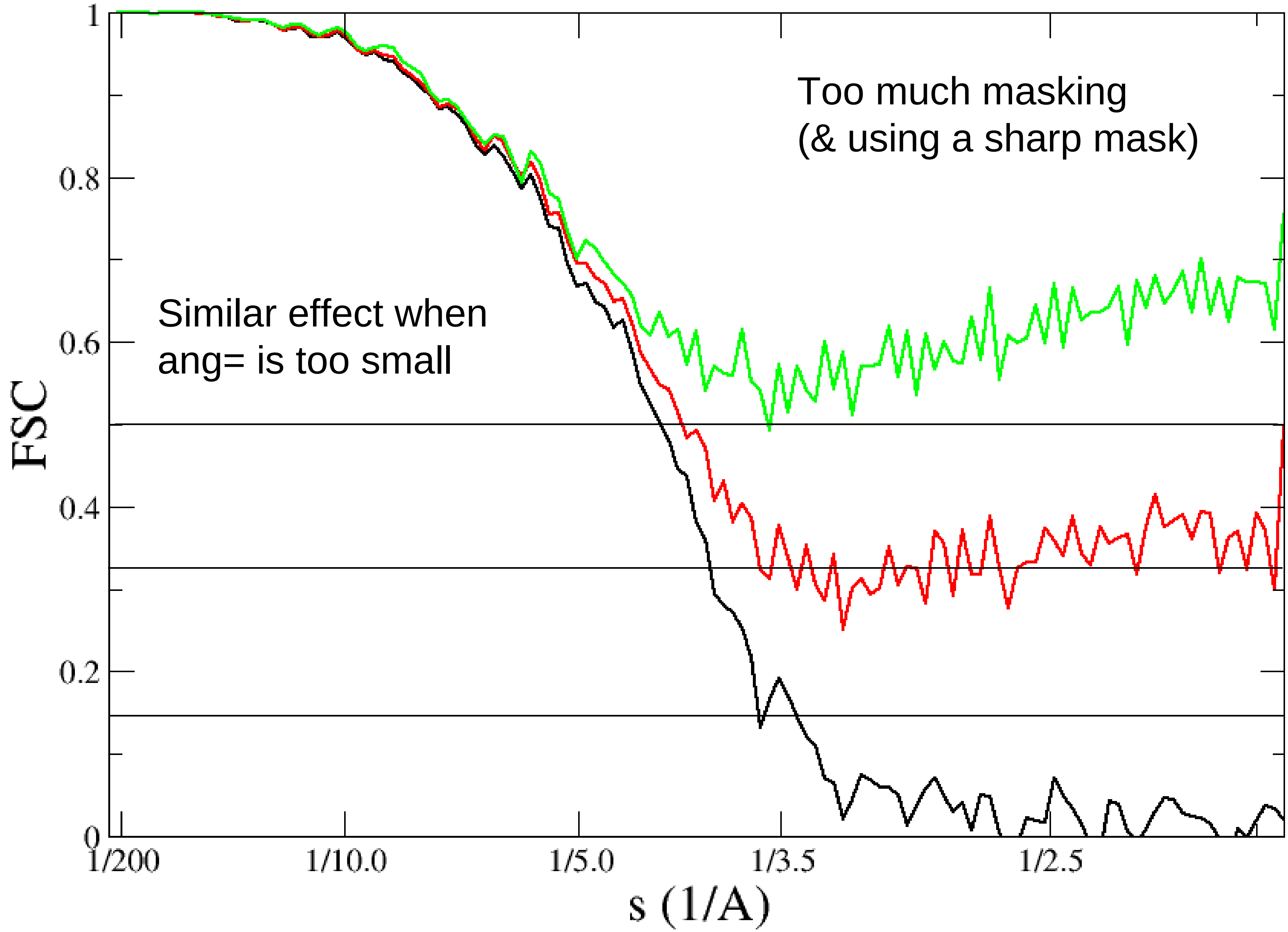
- Phase Residual
  - Definitions vary widely, causes meaning to be ambiguous
- Signal to Noise Ratio
  - Excellent properties. Additive (with perfect alignment). Counterintuitive threshold.
- Fourier Shell Correlation
  - Easy to compute. Sigmoidal curve with clearly defined value at most thresholds
  - What threshold to use ? (0.5, 0.33, 0.13, 3σ)

Good FSC



FSC
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1/30
1/12
1/7
1/5
1/4
1/3
1/2.3
Resolution (1/A)

Bad FSC

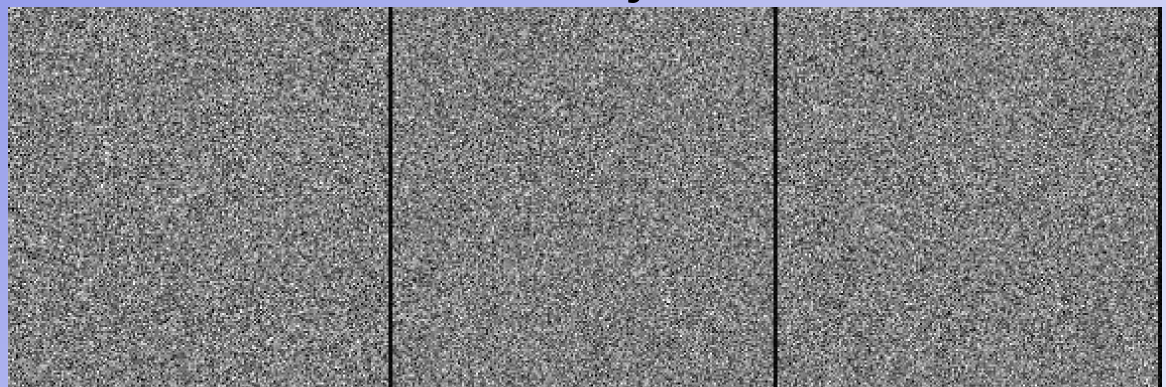


1
0.8
0.6
0.4
0.2
0
FSC
1/200
1/10.0
1/5.0
1/3.5
1/2.5
s (1/A)
Too much masking
(& using a sharp mask)
Similar effect when
ang= is too small

Model Bias

Base

Noisy



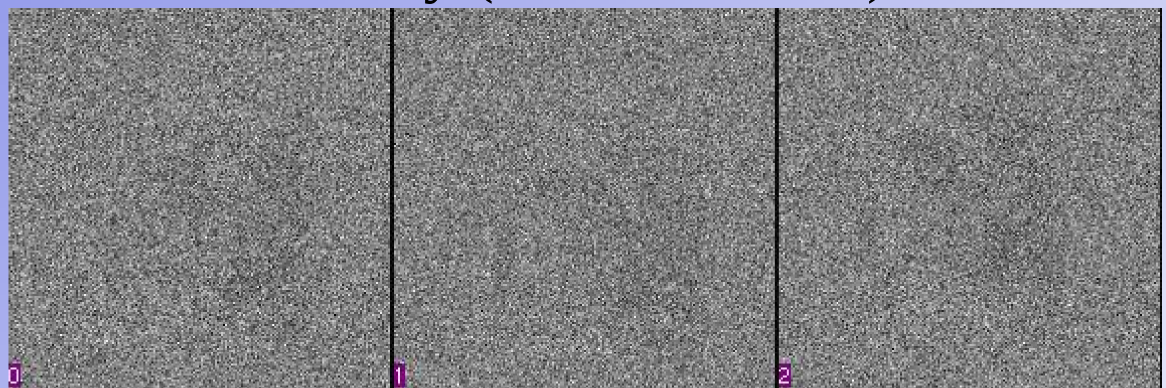
Align to

25 100 250 1000 2000

Model Bias

Base

Noisy (~10% contrast)



Align to

25 100 250 1000 2000

Model Bias

Base

Noisy (~10% contrast)

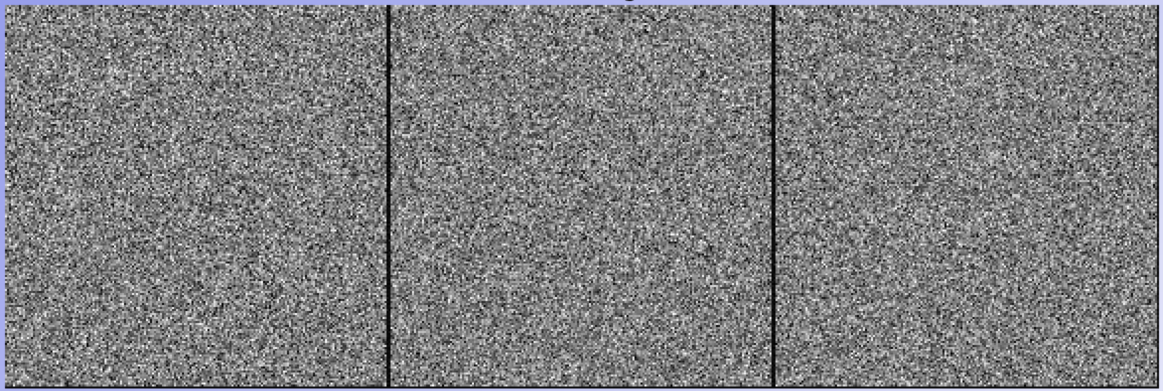
Align to

25 100 250 1000 2000

Model Bias

Base

Noisy



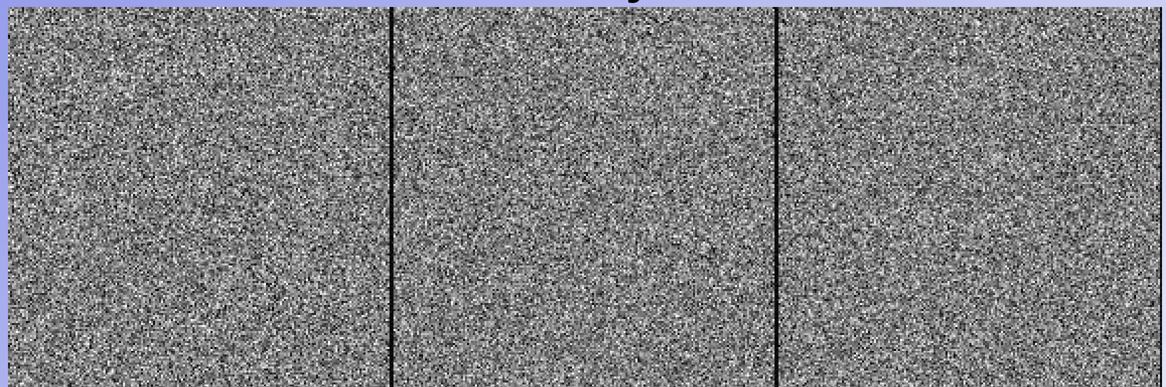
Align to

25 100 250 1000 2000

Model Bias

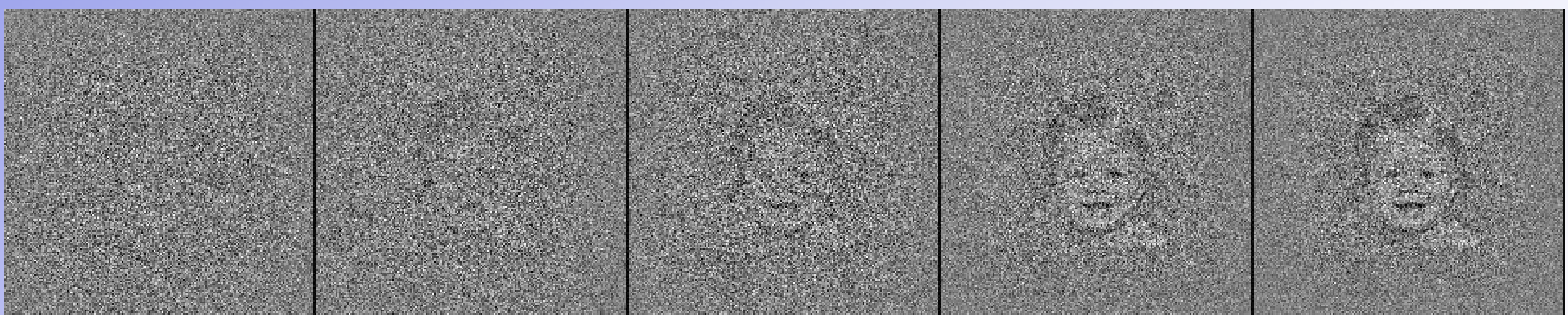
Base

Noisy



Align to

Iter x4

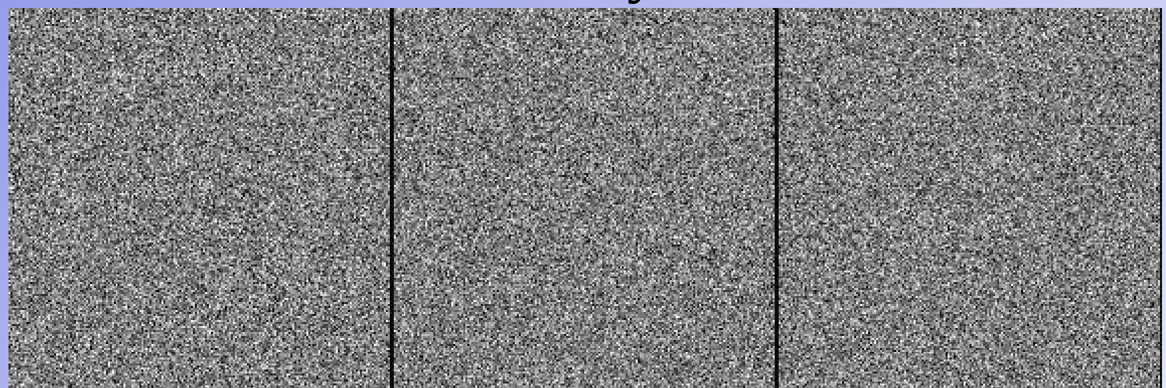


25 100 250 1000 2000

Model Bias

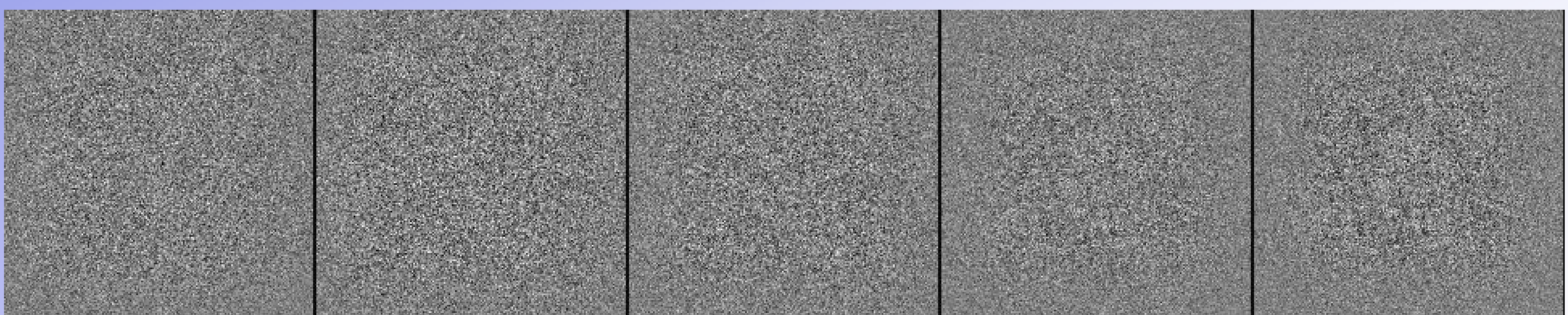
Base

Noisy



Align to

Iter x8

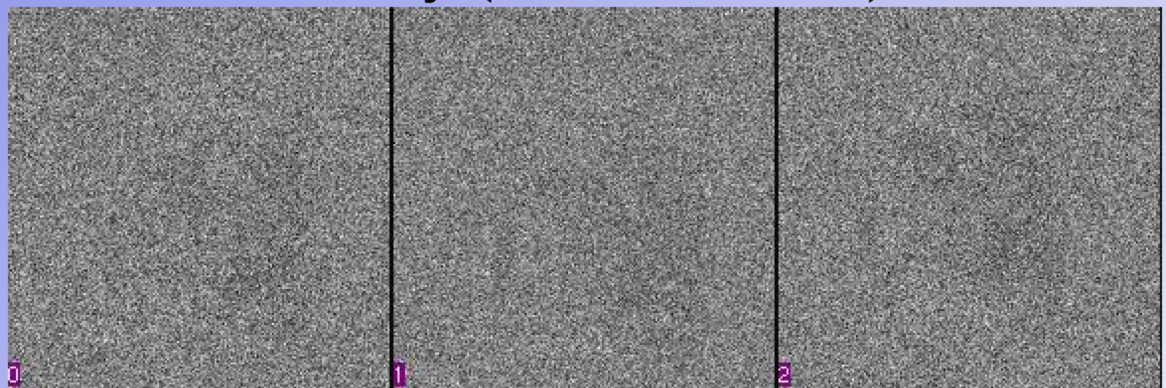


25 100 250 1000 2000

Model Bias

Base

Noisy (~10% contrast)



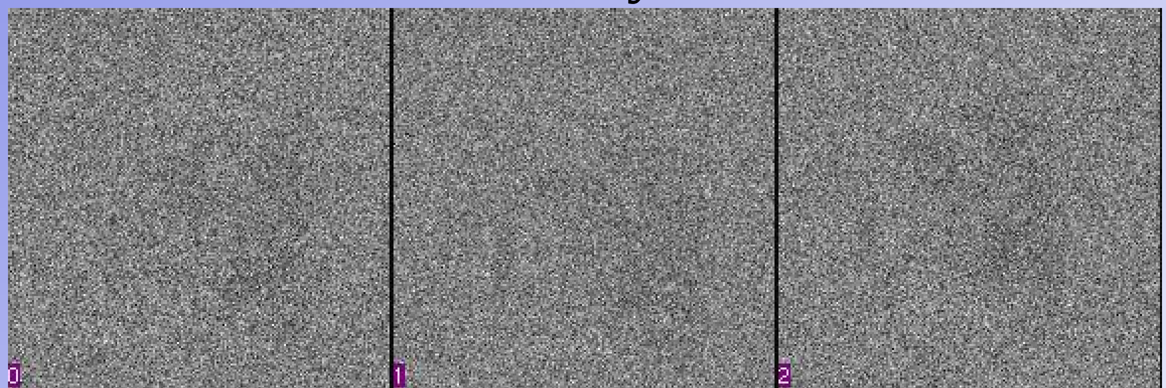
Align to

25 100 250 1000 2000

Model Bias

Base

Noisy

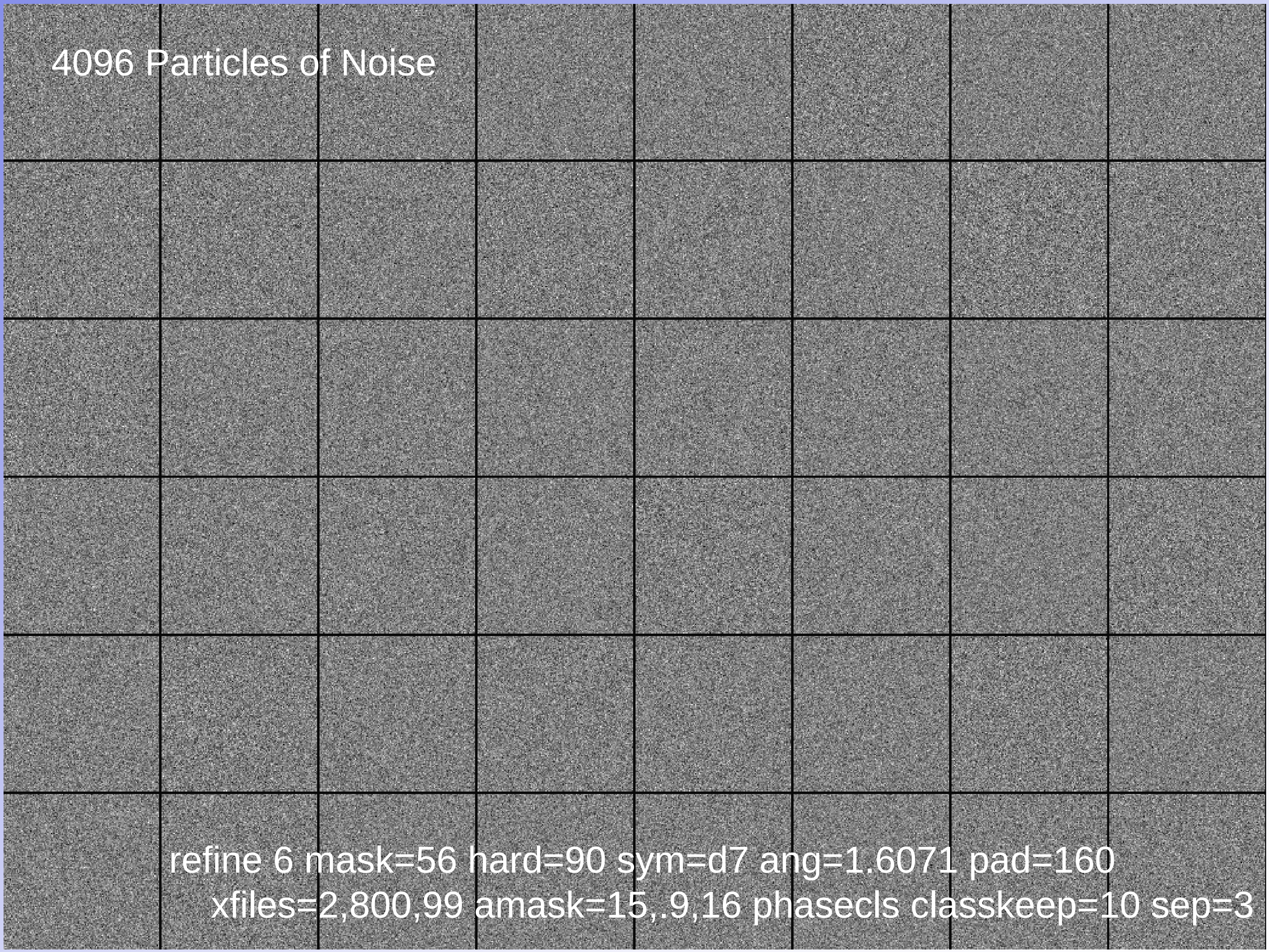


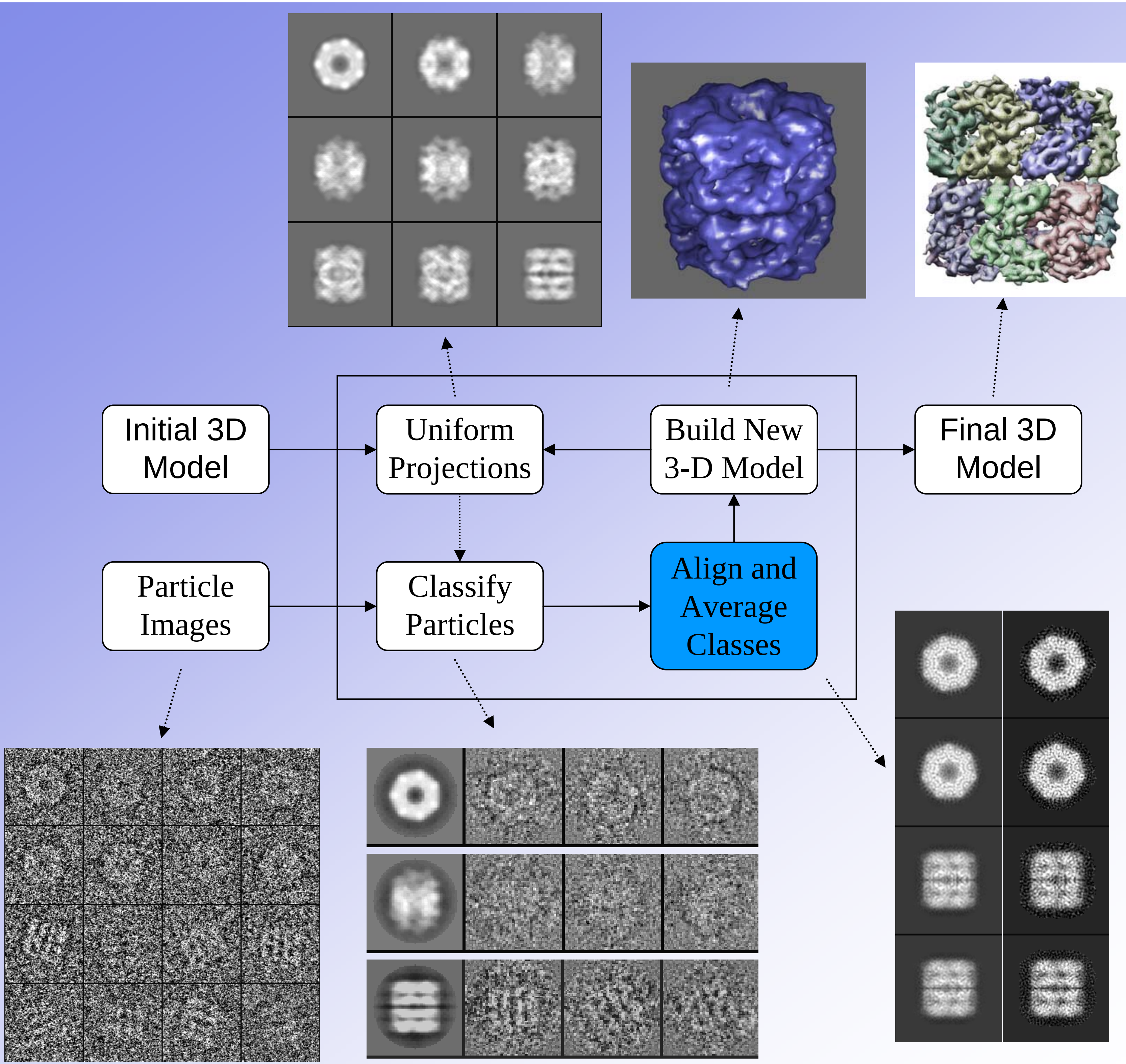
Align to

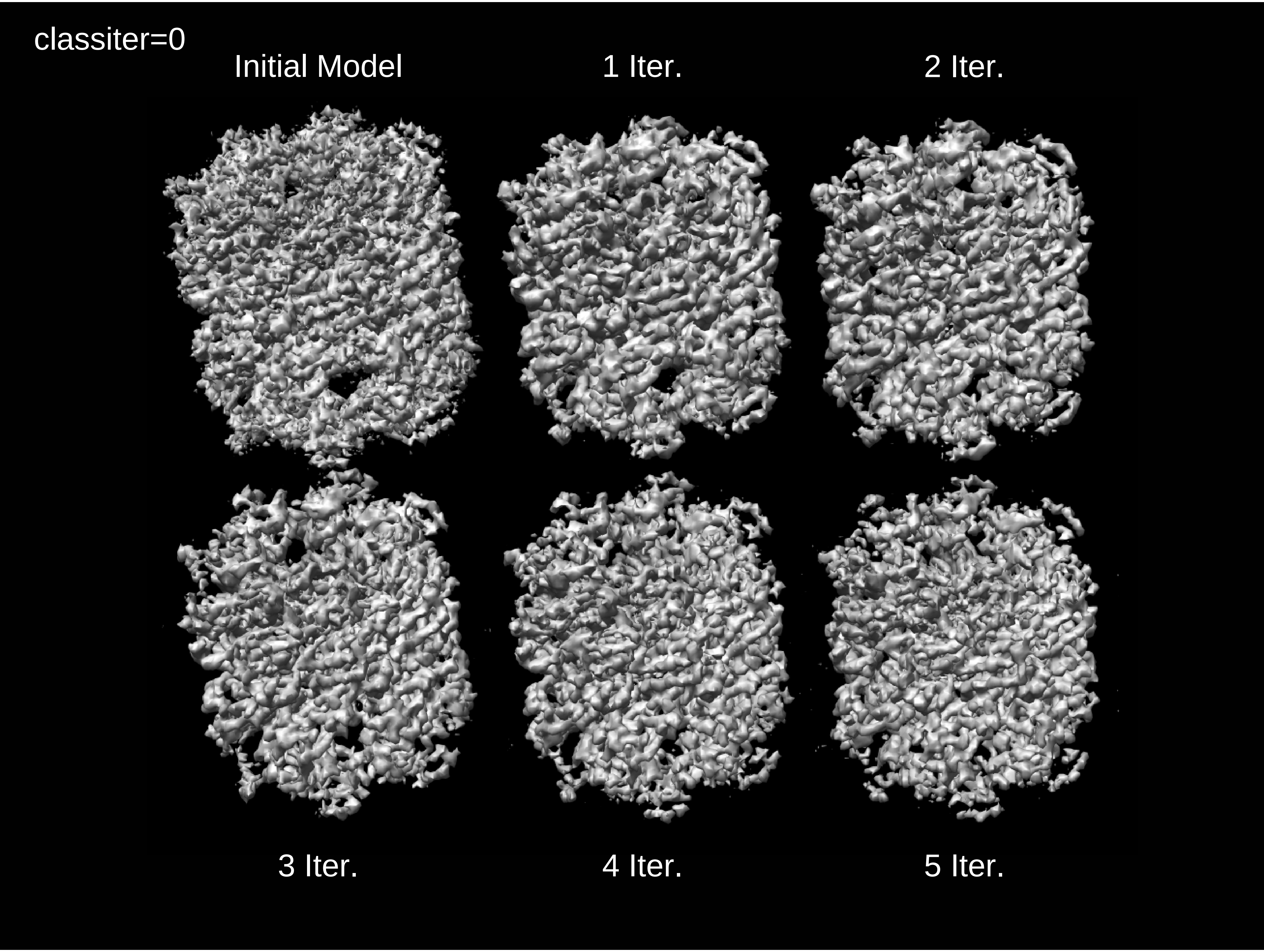
Iter x4

25 100 250 1000 2000

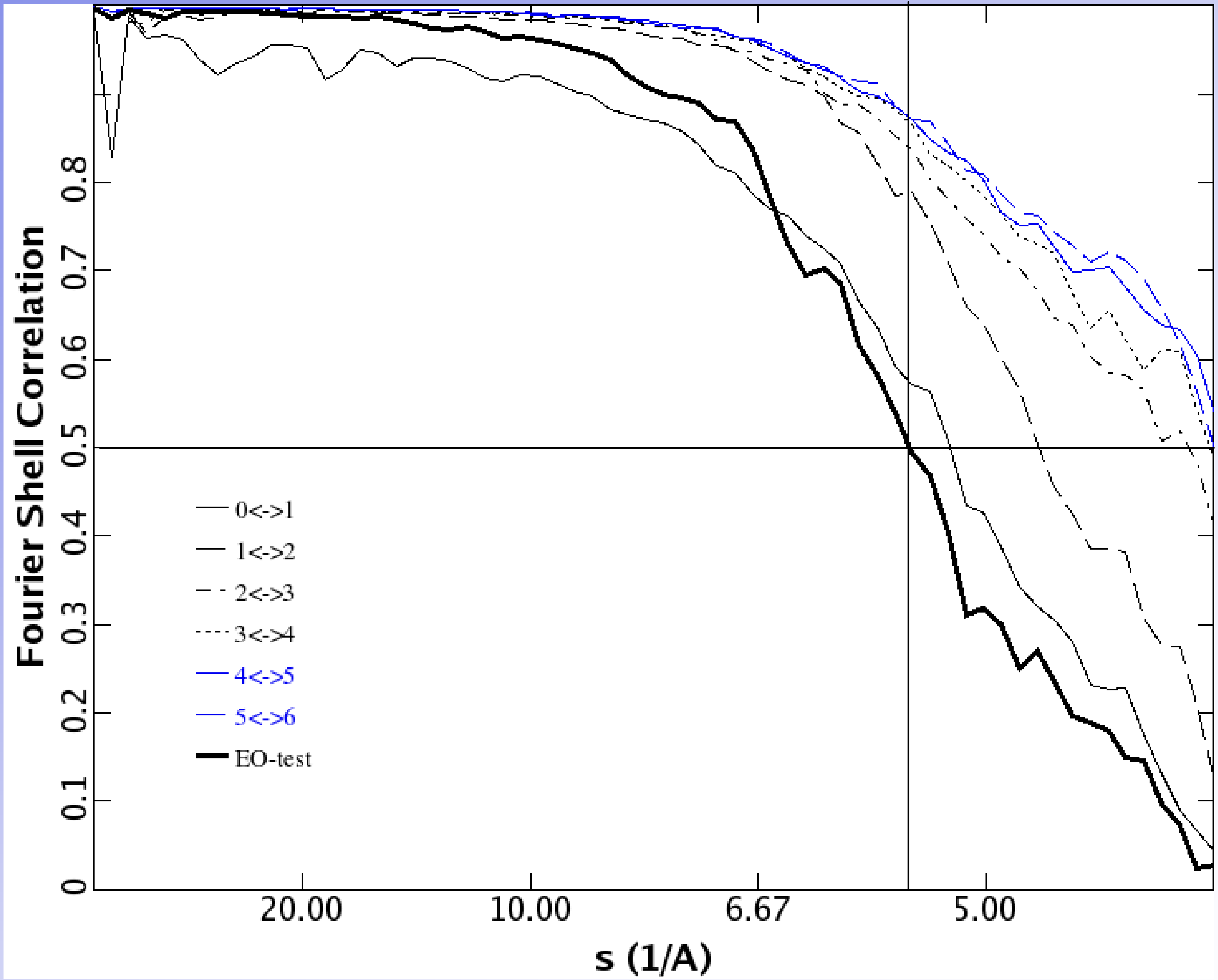
How About 3-D ?


4096 Particles of Noise
refine 6 mask=56 hard=90 sym=d7 ang=1.6071 pad=160
xfiles=2,800,99 amask=15,.9,16 phasecls classkeep=10 sep=3

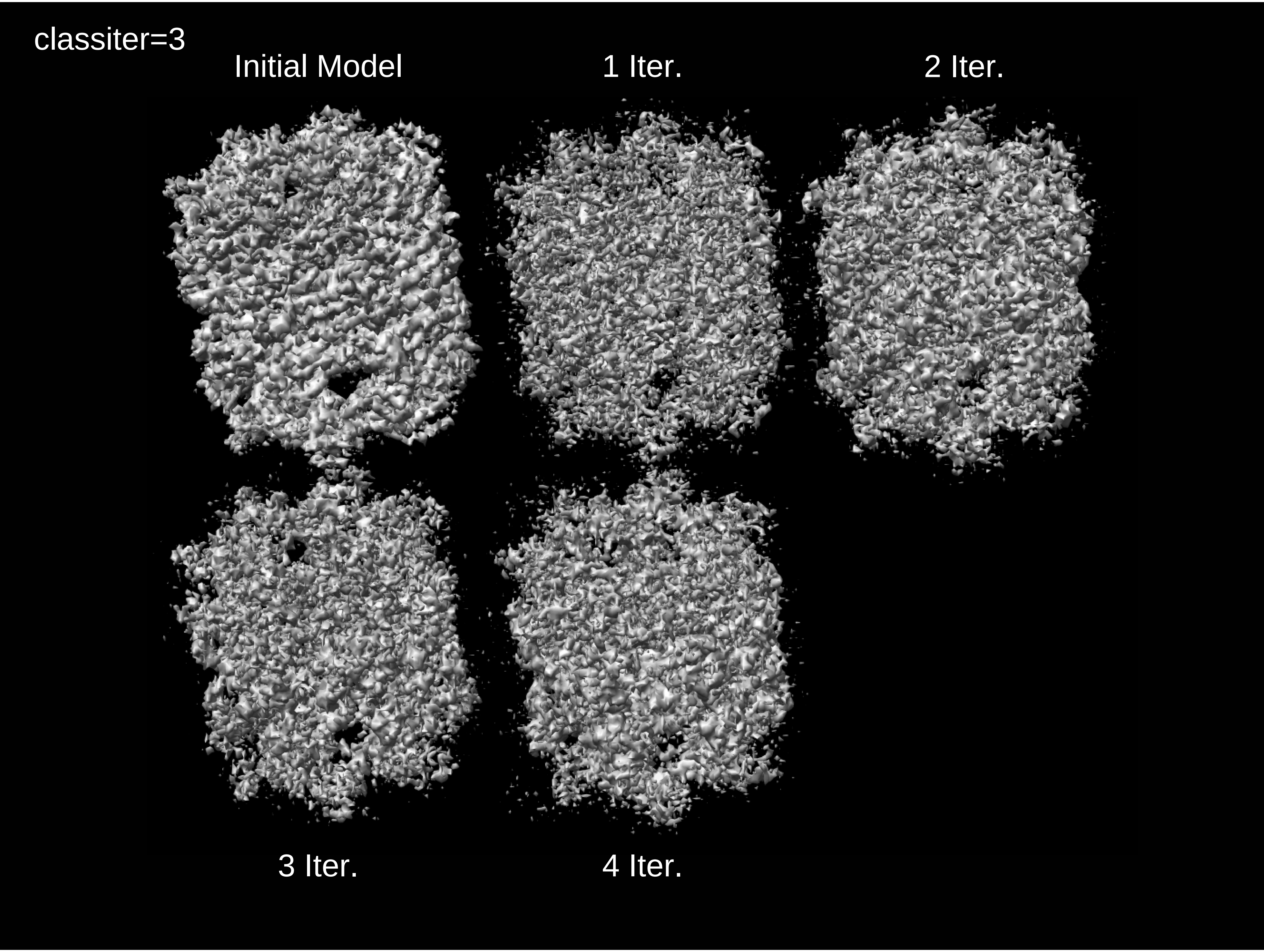

Initial 3D Model
Uniform Projections
Build New 3-D Model
Final 3D Model
Particle Images
Classify Particles
Align and Average Classes

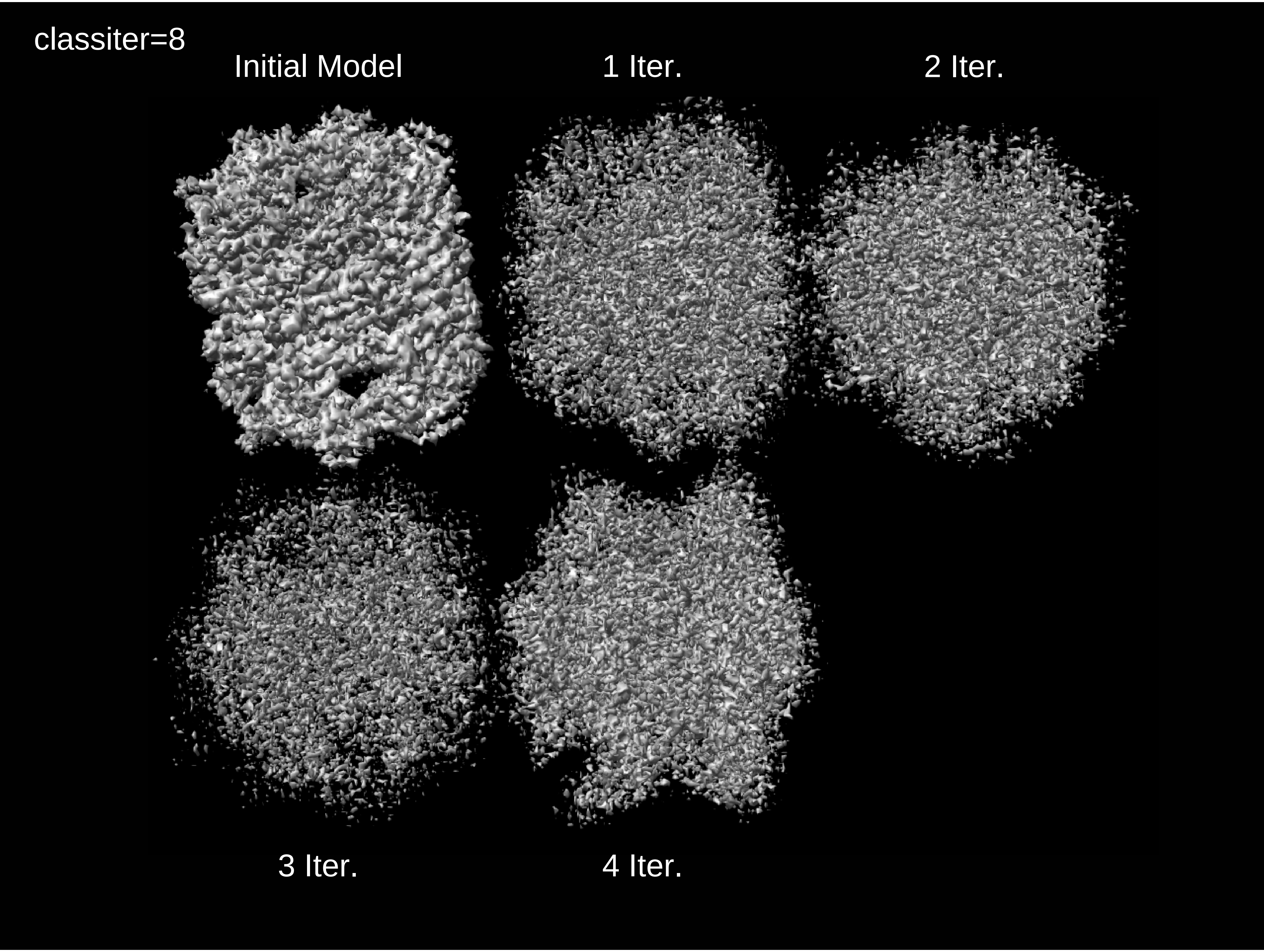

classiter=0
Initial Model
1 Iter.
2 Iter.
3 Iter.
4 Iter.
5 Iter.

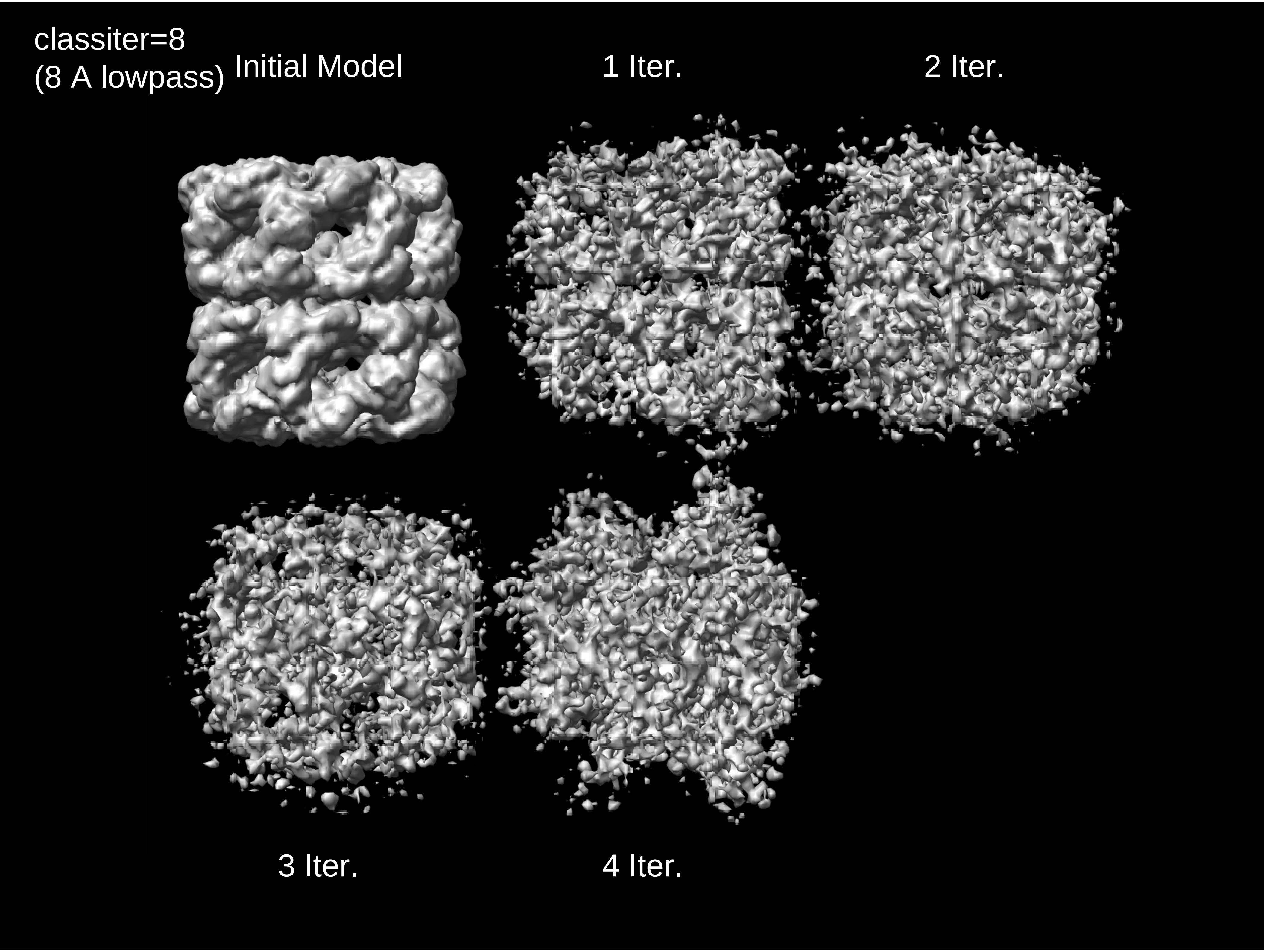
Even Worse



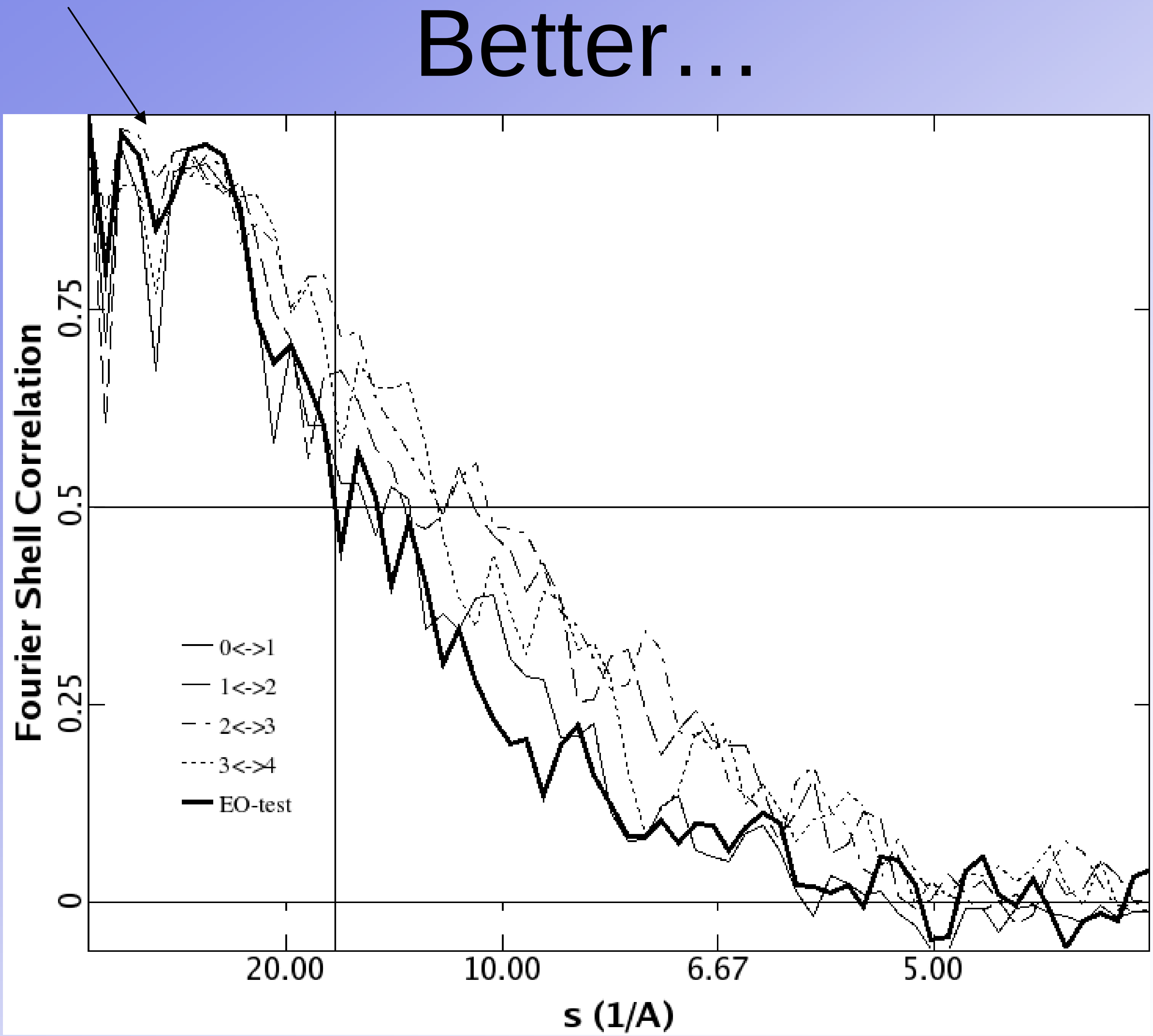
Fourier Shell Correlation
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
20.00
10.00
6.67
5.00
s (1/A)
0<->1
1<->2
2<->3
3<->4
4<->5
5<->6
EO-test


classiter=3
Initial Model
1 Iter.
2 Iter.
3 Iter.
4 Iter.


classiter=8
Initial Model
1 Iter.
2 Iter.
3 Iter.
4 Iter.


classiter=8
(8 A lowpass)
Initial Model
1 Iter.
2 Iter.
3 Iter.
4 Iter.

Better…



Fourier Shell Correlation
0.75
0.5
0.25
0
20.00
10.00
6.67
5.00
s (1/A)
0<->1
1<->2
2<->3
3<->4
EO-test

How Do we Stop This ?

- Always start out with classiter>3 for a few rounds
- Always refine from multiple starting models
- If the results are not effectively the same, try to establish which one is correct by looking at self consistency of projections/class-averages
- Make sure the features you are interpreting come out of all good refinements

NCMI
National Center for Macromolecular Imaging

Research supported by the NCRR and NIGMS

Graphics produced using Chimera (CGL at UCSF) (supported by NIH P41 RR-01081) and EMANimator.