SickKids
RESEARCH
INSTITUTE

Optimizing image acquisition
NRAMM - 2017

John Rubinstein

Molecular Medicine Program
The Hospital for Sick Children Research Institute

Departments of Biochemistry and Medical Biophysics
The University of Toronto

www.sickkids.ca/research/rubinstein
@RubinsteinJohn

UNIVERSITY OF TORONTO

Equipment needed for single particle EM

	Voltage	Gun	Stage	Detector
Negative stain	100/120 kV	Thermionic	Side entry RT holder	CCD
Screening Cryo-EM of >300 kDa	100/120 kV	Thermionic	Side entry cryoholder (anticontaminor)	CCD
Screening Cryo-EM of <300 kDa	(100 kV) 200 kV	FEG	Side entry cryoholder (anticontaminor)	DDD
Good resolution cryo-EM	200 kV	FEG	Side entry cryoholder or autoloader-type holder	DDD
Best resolution cryo-EM	300 kV +	FEG	Side entry cryoholder or autoloader-type holder	DDD
Best resolution high-throughput cryo-EM	300 kV +	FEG	autoloader-type holder	DDD

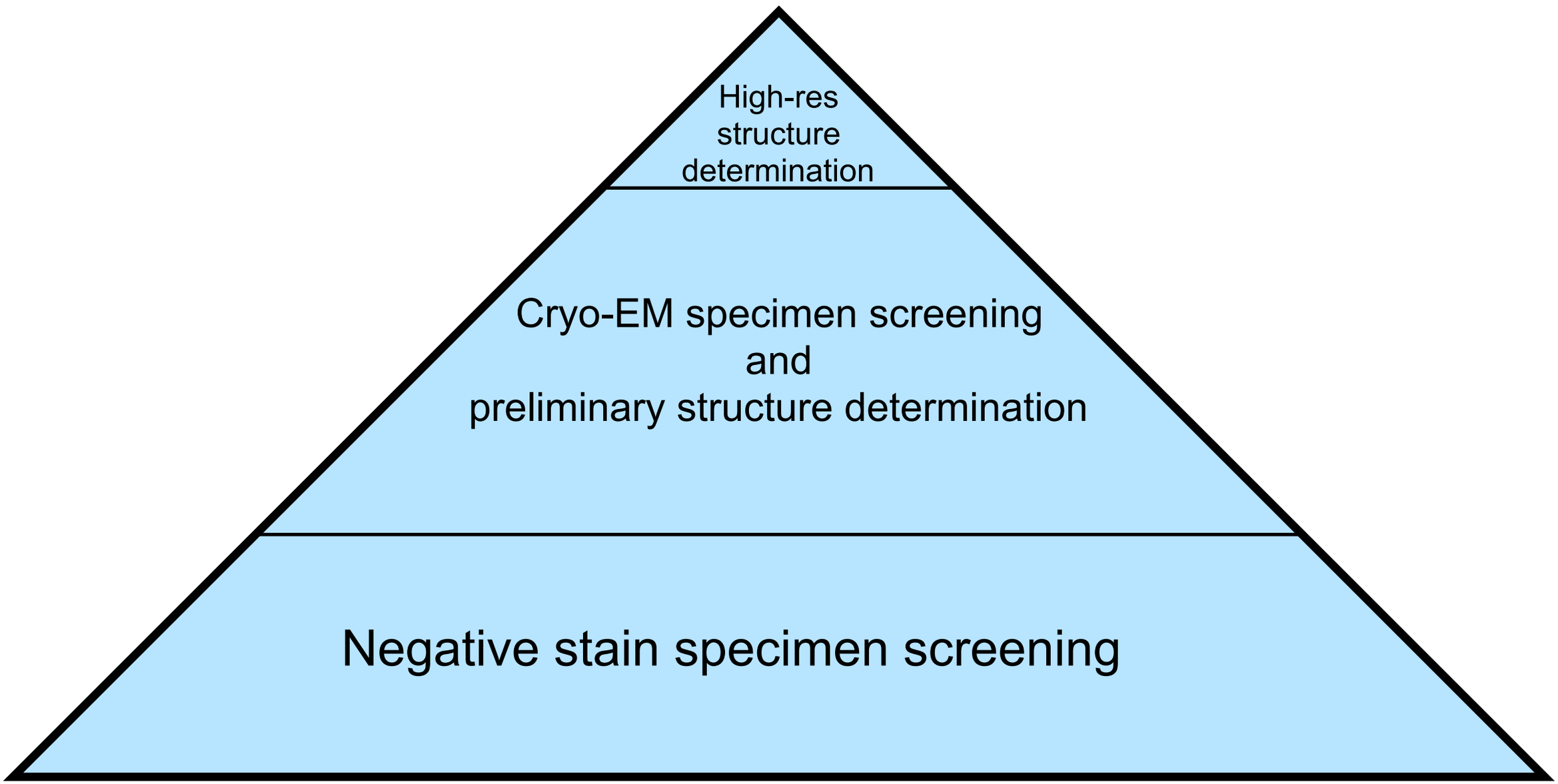
Cryo-EM in Toronto

Everything is easier with a Krios!
(~2014)

2007 to present: Tecnai F20
(K2 summit 2013)

2017: Titan Krios with Falcon3
(Quantum K2/K3??)

Time needed for each experiment



Use your FEG/DDD microscope appropriately…

	Voltage	Stage	Obtaining Parallel beam	Avoiding lens hysterisis
F20	200 kV	Side Entry Cryoholder	Use C2 aperture and lens setting that minimizes beam divergence	Use over-focused diffraction for search mode
Talos/ Talos Arctica/ Glacios	200 kV	Side entry holder/ Autoloader	Use C2 aperture and lens setting that minimizes beam divergence	Constant power lenses
F30/Polara	300 kV	Side entry holder/ Stable stage	Use C2 aperture and lens setting that minimizes beam divergence	Use over-focused diffraction for search mode
Titan Halo	300 kV	Side entry holder	3rd Condenser Lens	Constant power lenses
Titan Krios	300 kV	Autoloader	3rd Condenser Lens	Constant power lenses

Must match exposure/defocus at different voltages

Equivalent exposure calculator

Voltage at which exposure wanted

Voltage at which exposure known

	100 kV	**200 kV**	**300 kV**
100 kV	1	1.5	1.8
200 kV	0.68	1	1.2
300 kV	0.56	0.82	1

Based on linear energy transfers from Glaeser (2007)

Equivalent defocus calculator

Voltage at which defocus wanted

Voltage at which defocus known

	100 kV	**200 kV**	**300 kV**
100 kV	1	1.47	1.88
200 kV	0.678	1	1.27
300 kV	0.532	0.785	1

New defocus must keep product of λ and Δz constant

Decisions you need to make with a DDD

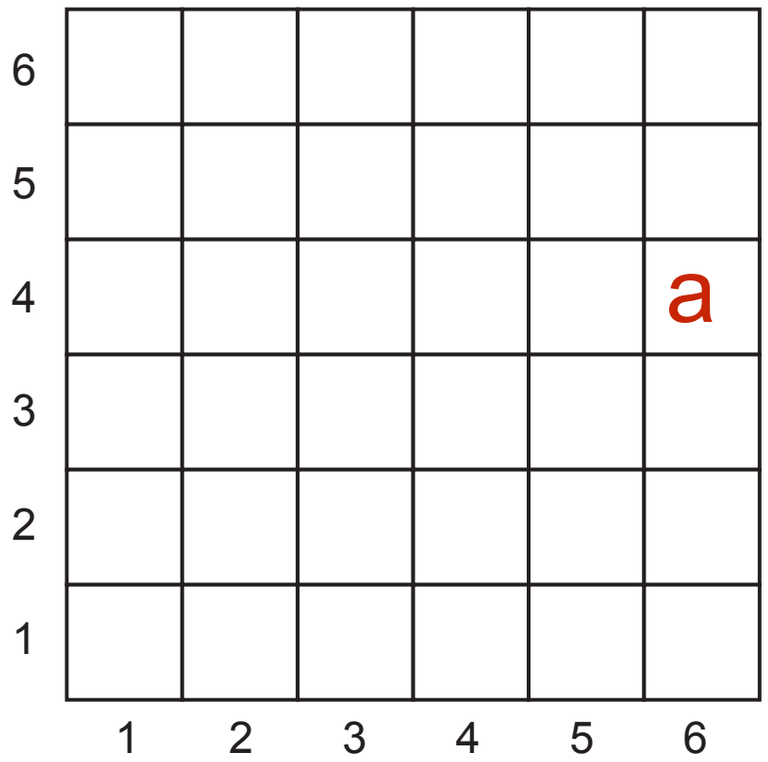
Decision	Relevant concepts
Magnification or pixel size (Å/pixel)	DQE, Nyquist Limit, Aliasing, Anisotropic magnification
Exposure rate (electrons/pixel/second)	Coincidence loss (counting)
Frame rate (frames/sec)	Alignability of frames, movement within frames, radiation damage per frame
Movie length (seconds)	Total signal at different resolutions

frame (Gatan) = **fractions** (FEI)

Fourier basics

Two dimension Fourier transforms

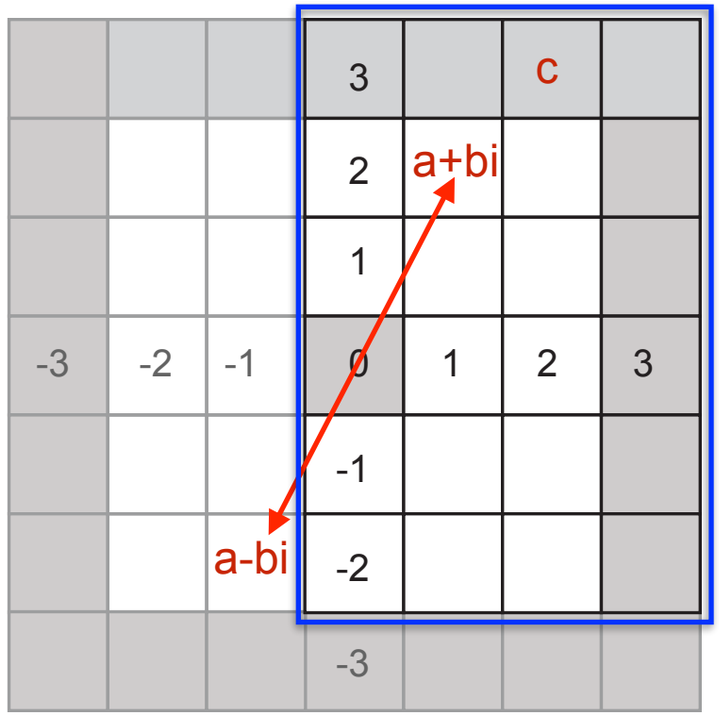
Image array (real value pixels)



FT

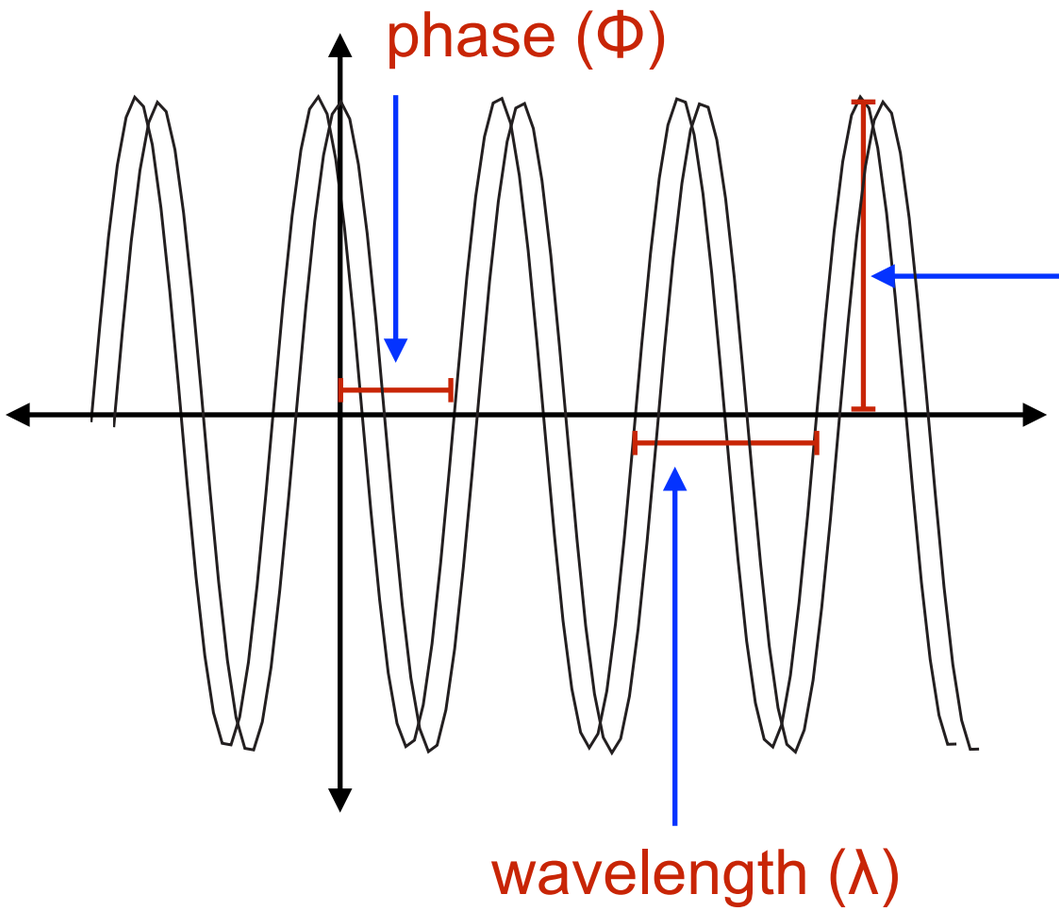
FT^{-1}

Fourier array (complex value pixels)



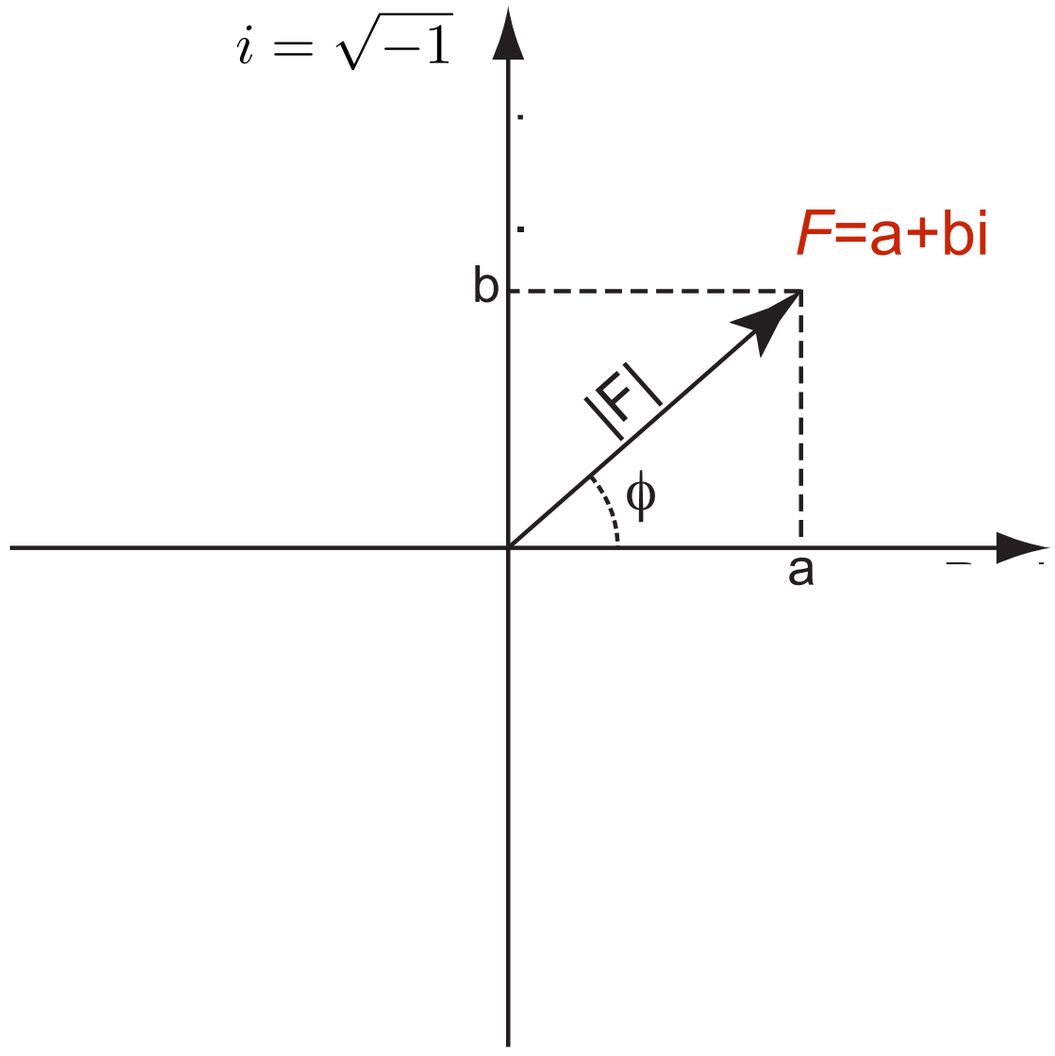
- The FT of real functions (e.g. images) are Hermitian: for every point (a+bi) there is a corresponding point (a-bi)
- For an N $\times$ N pixel image, Fourier transform is N/2+1 $\times$ N

Representing waves as vectors

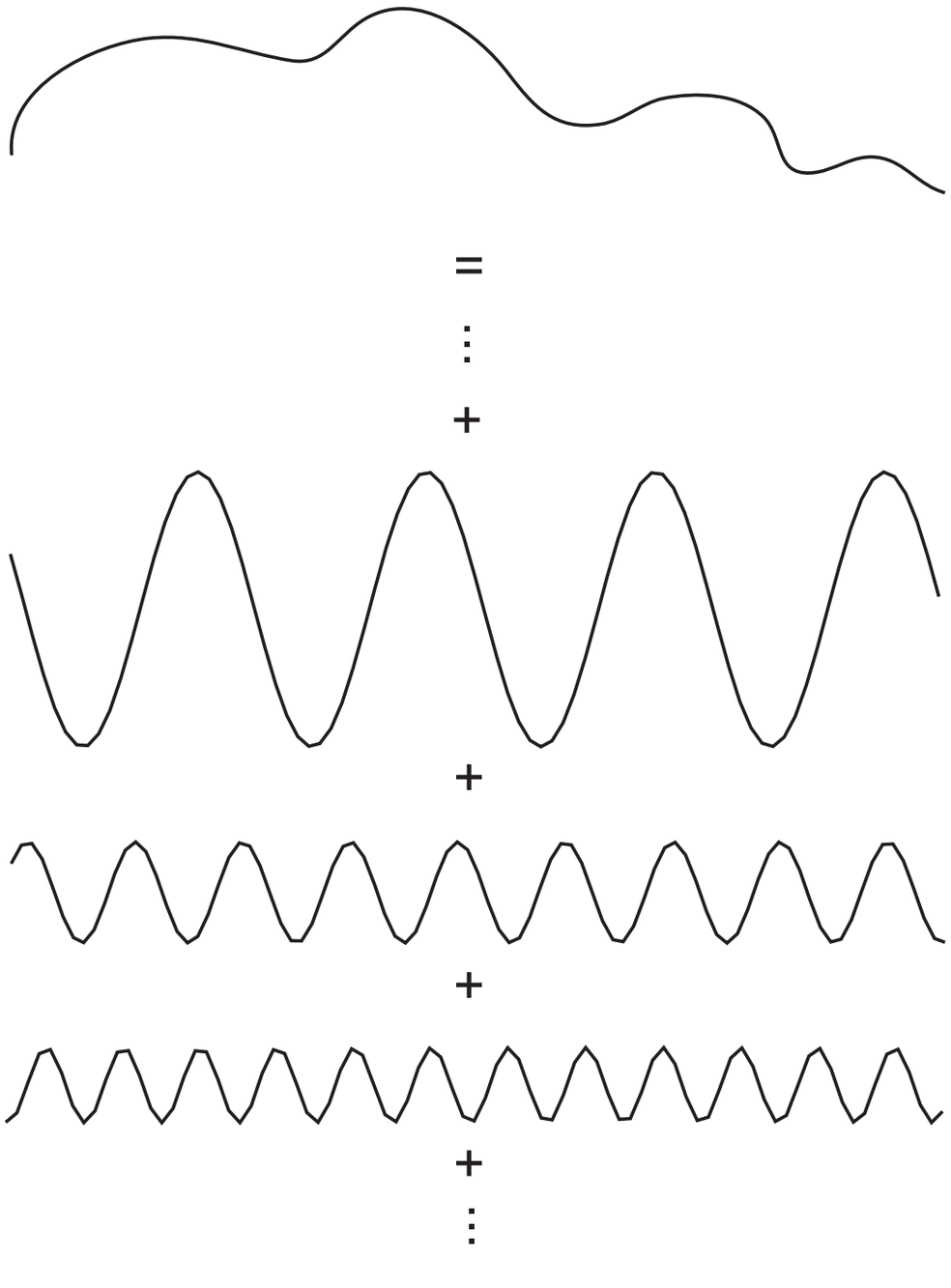


- Wavelength
- Amplitude
- Phase

For waves of a specified wavelength



The FT represents functions in terms of waves



Function

=

wave 27

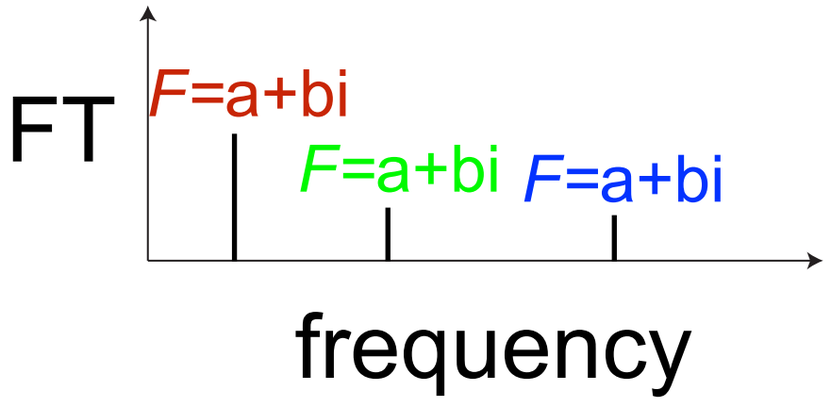
+

wave 28

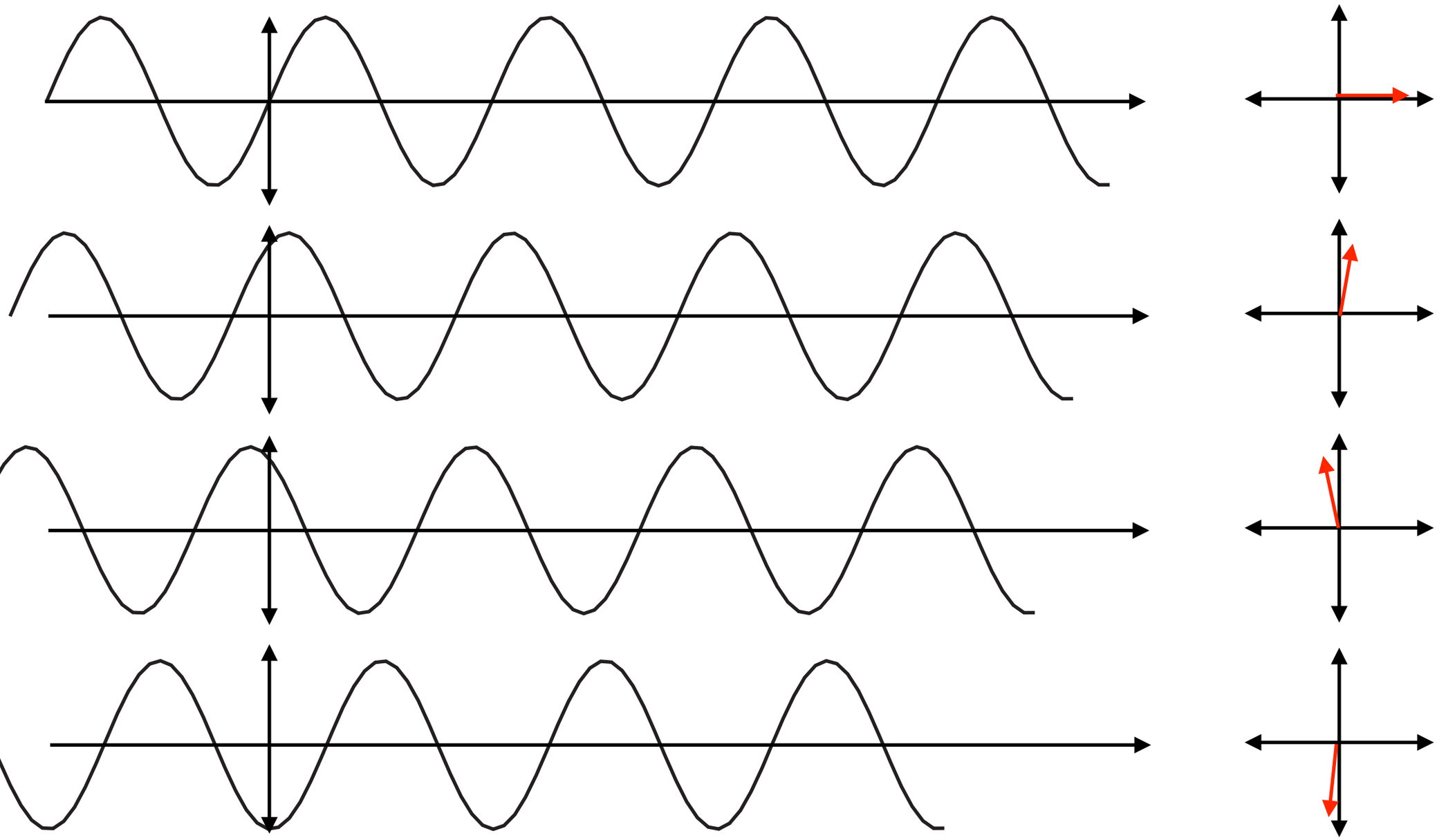
+

wave 29

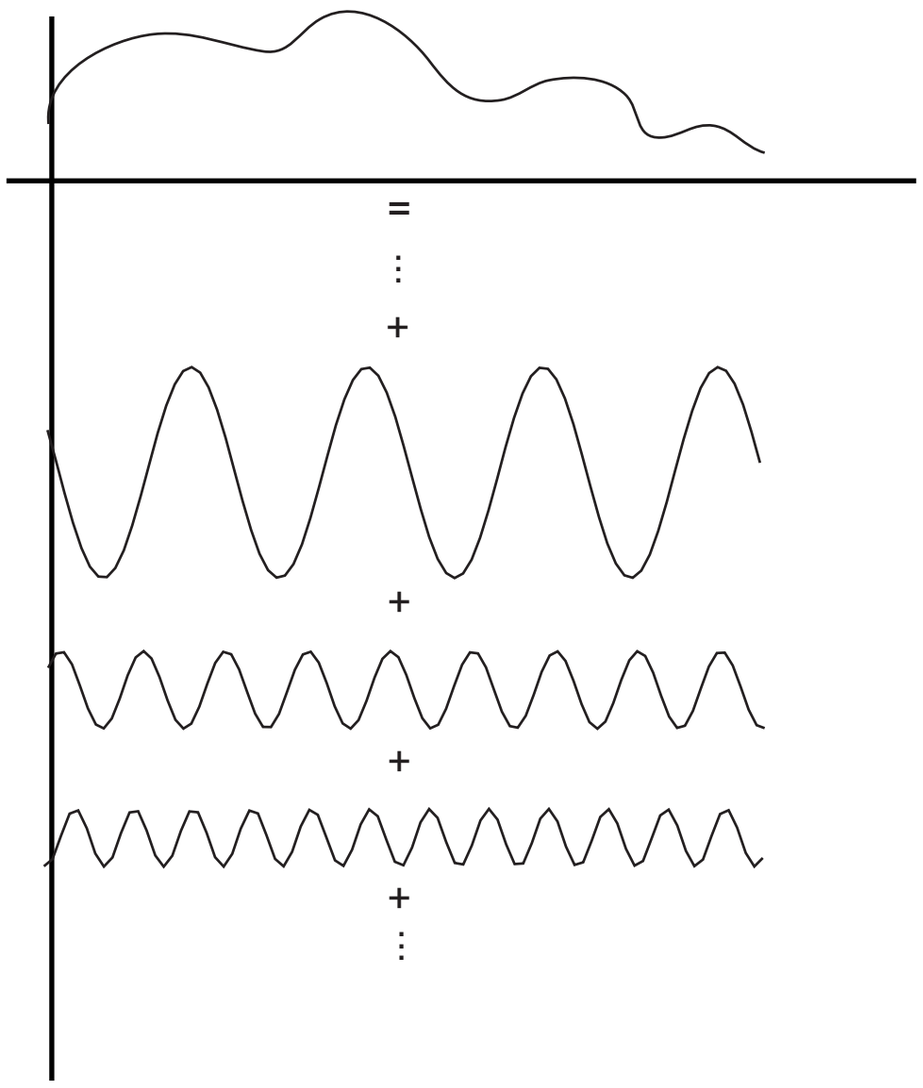
+

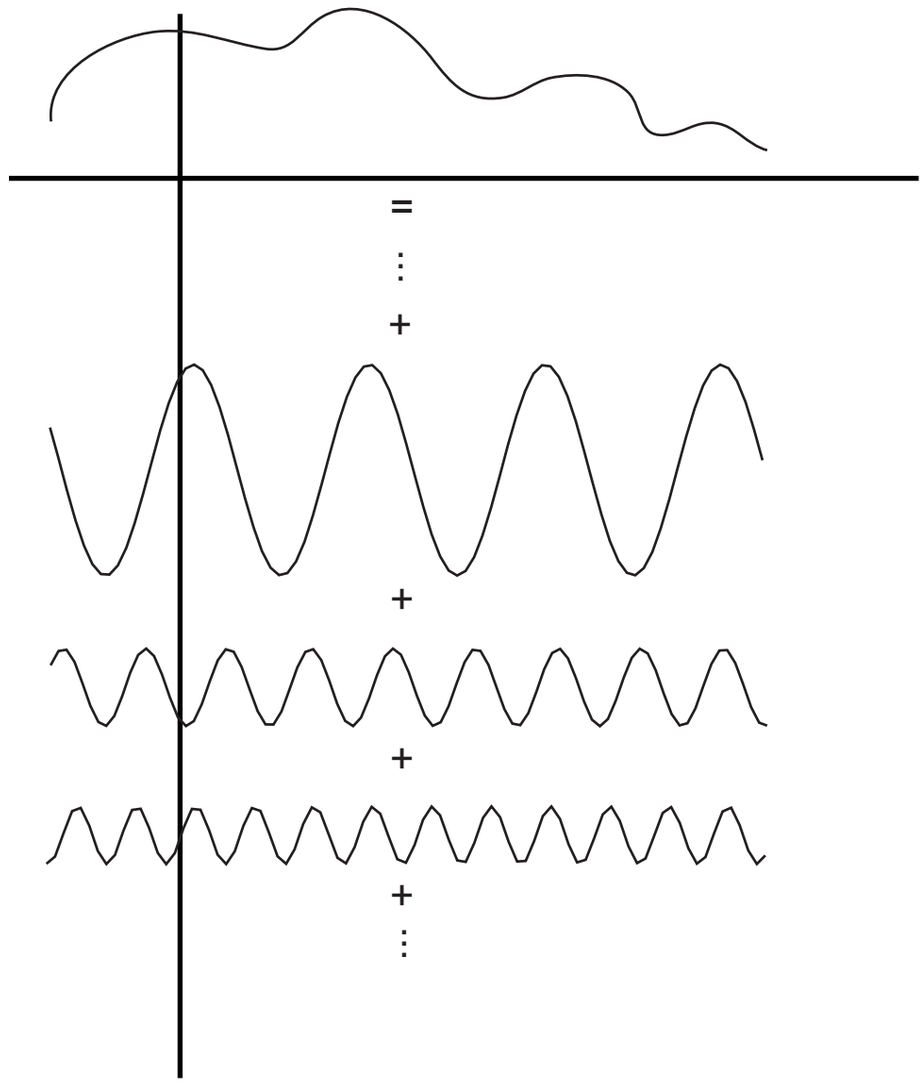


Shifting waves causes a phase change



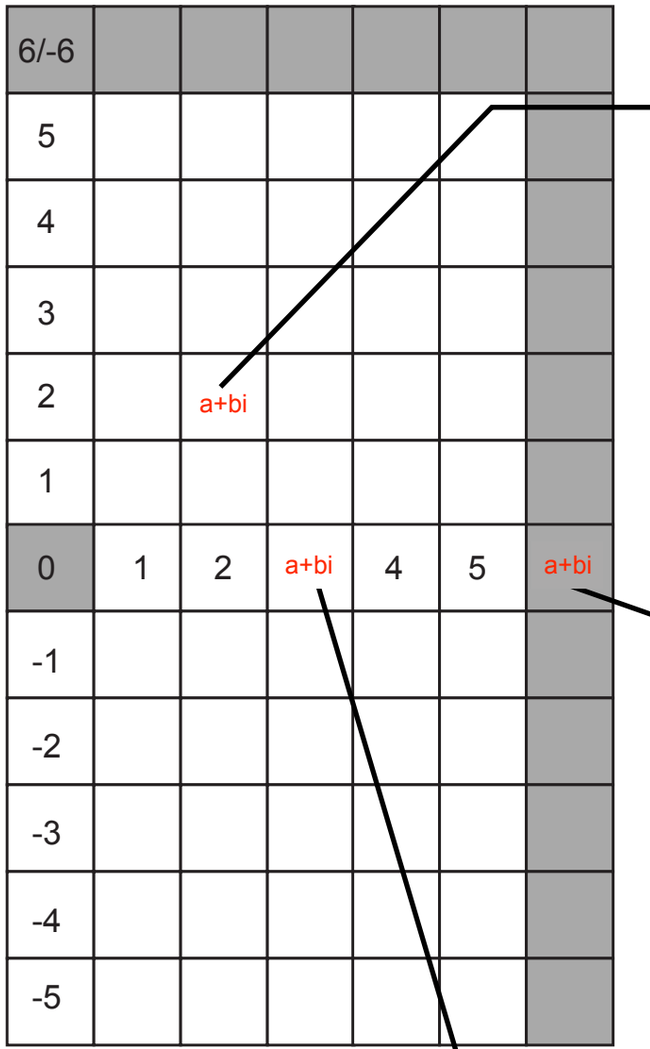
Phase change of Fourier components from shifting

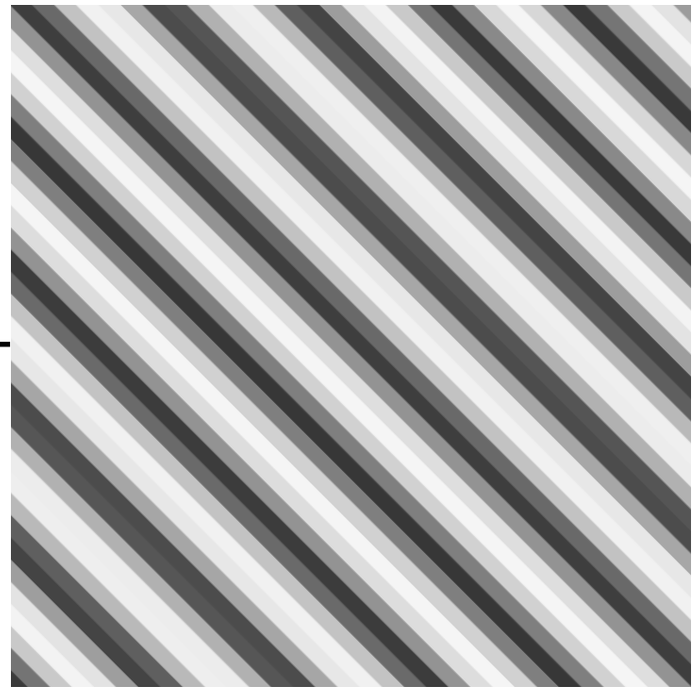


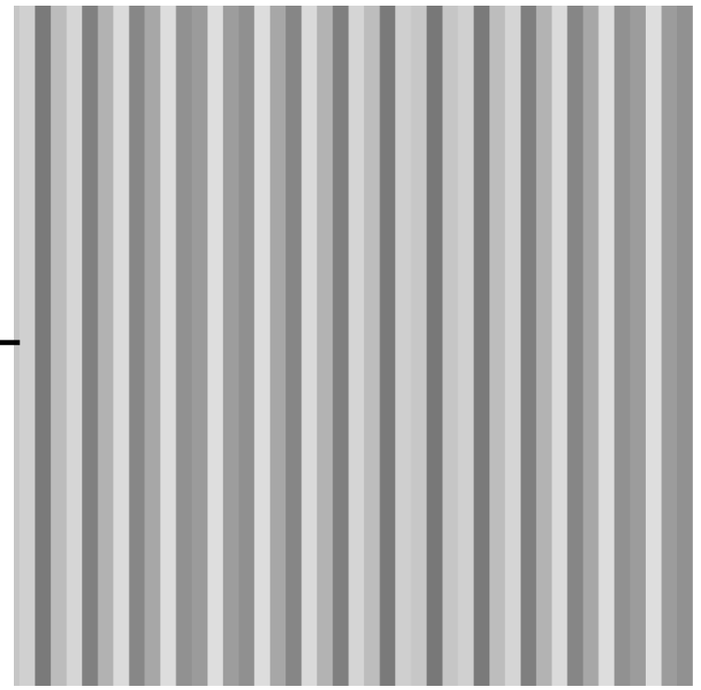


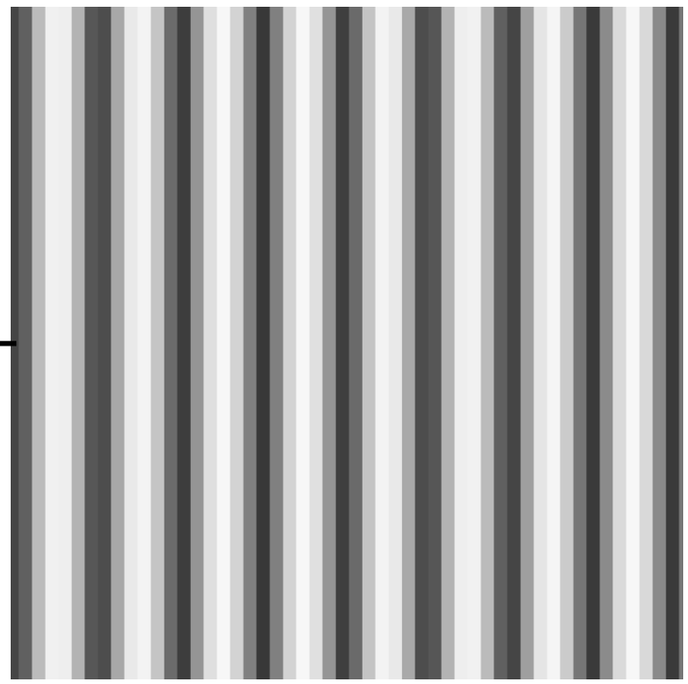
Shifting in real space causes phase changes in Fourier space

Two dimension Fourier transforms







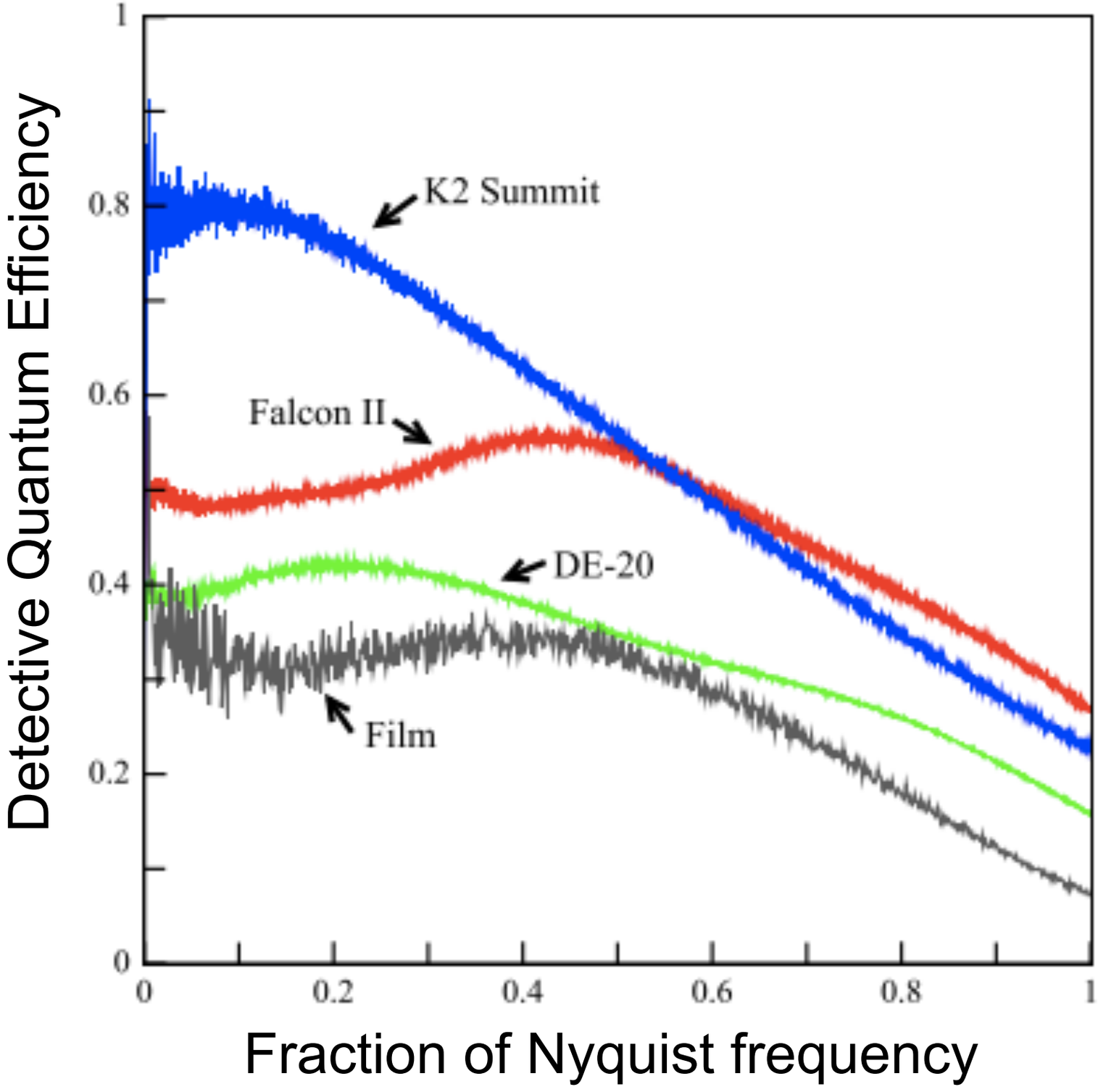


Position of pixel determines
sine wave frequency and direction

Detective quantum efficiency

(What pixel size/magnification should you use?)

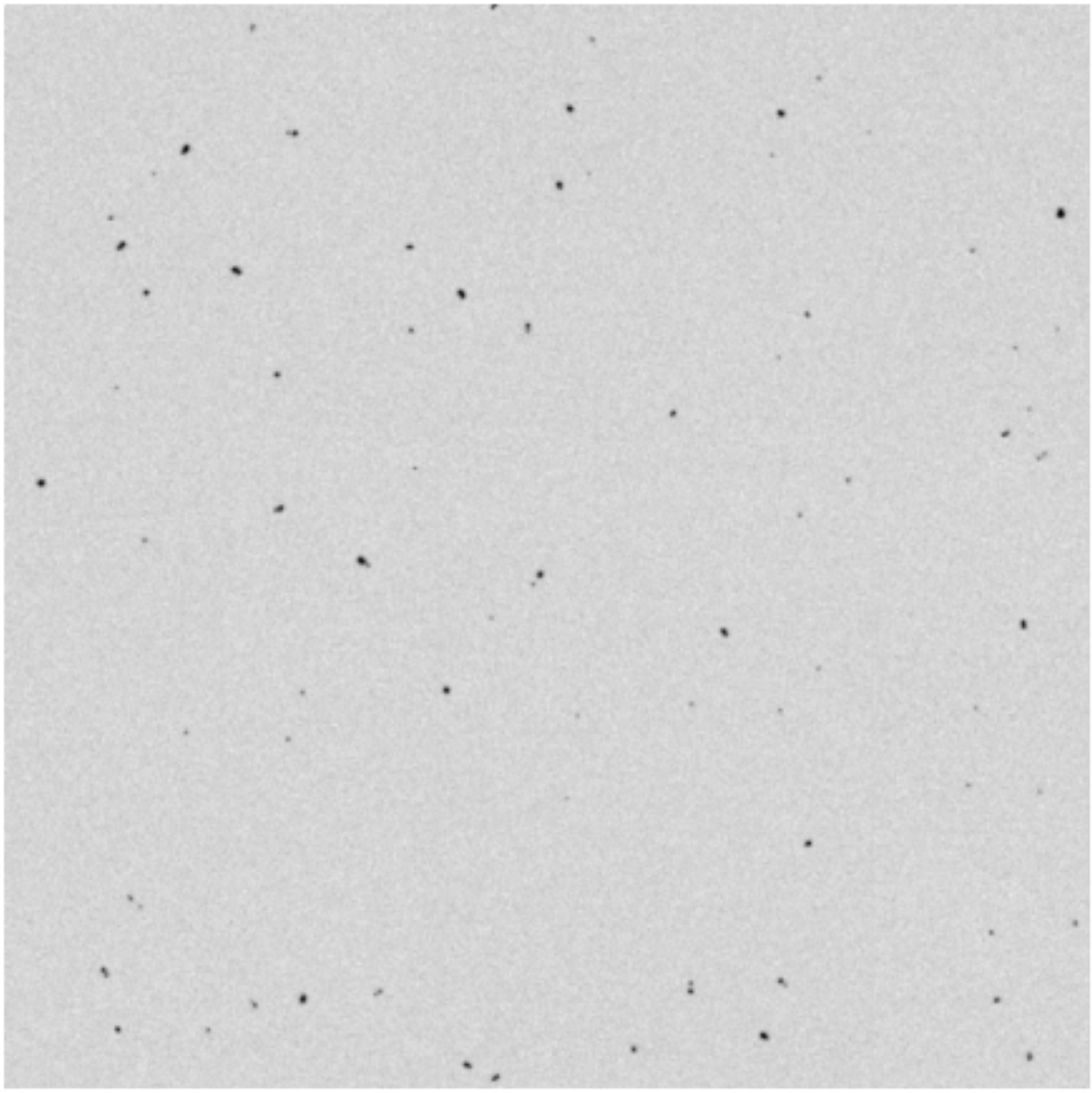
Detective quantum efficiency



$$DQE(res'n) = \frac{[SNR_{out}(res'n)]^2}{[SNR_{in}(res'n)]^2}$$

McMullan et al. (2014), Ultramicroscopy 147, 156-63.

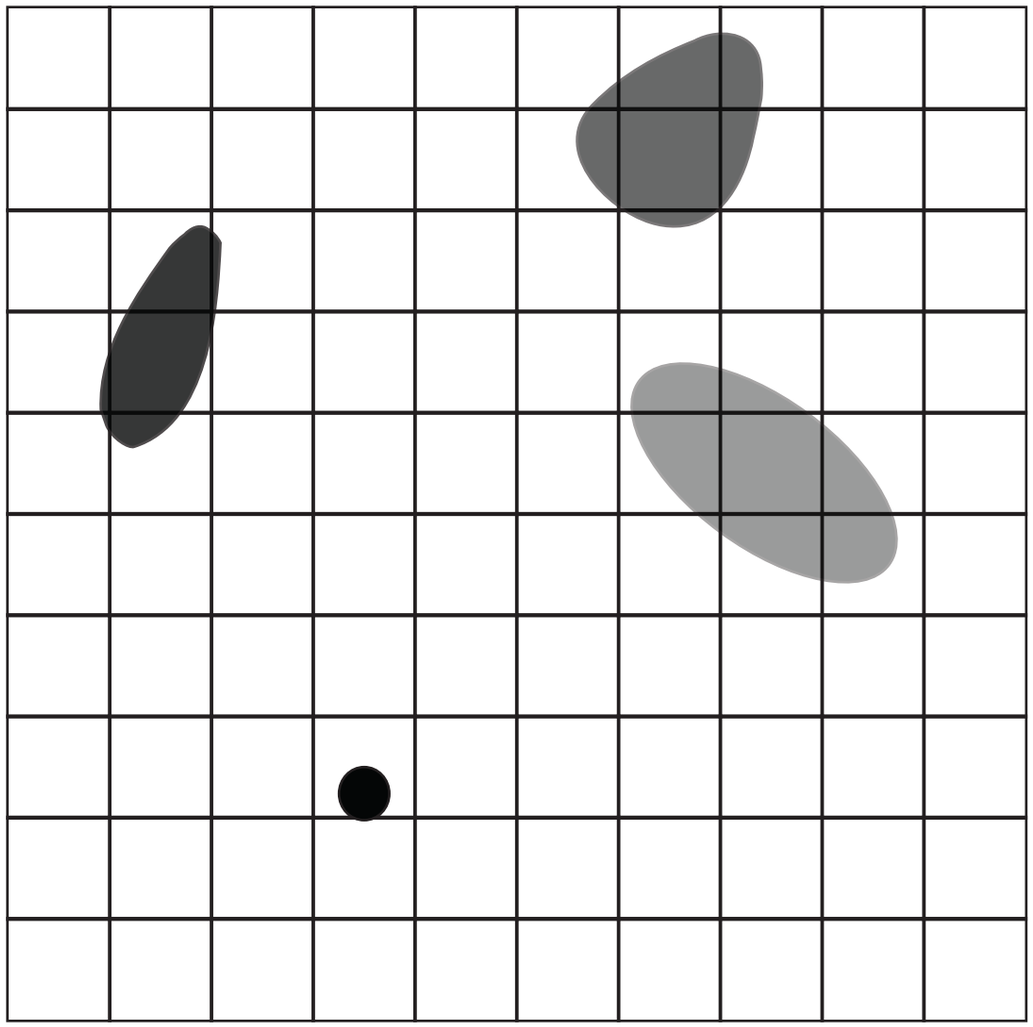
Electron counting can boost DQE



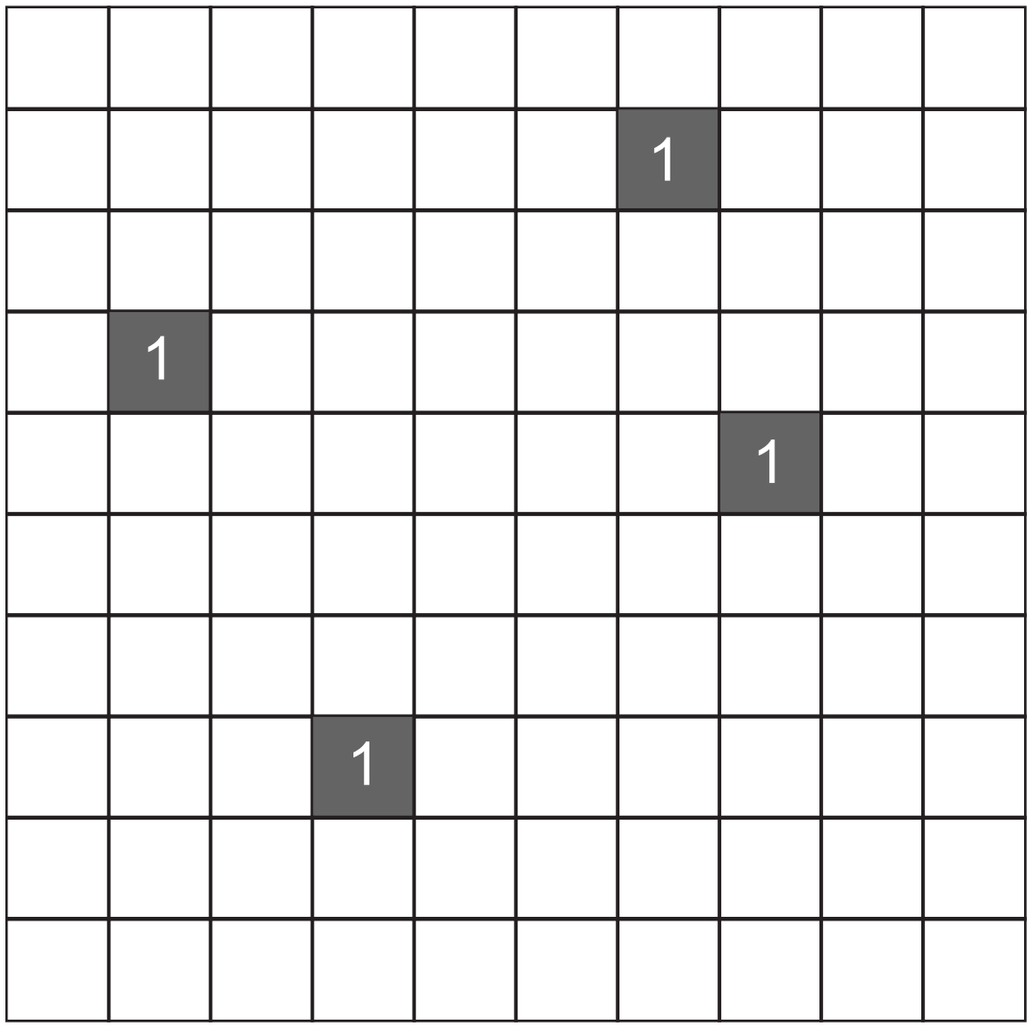
McMullan *et al.* (2009), Ultramicroscopy 109, 1144-7.

Electron counting

Integration



Counting

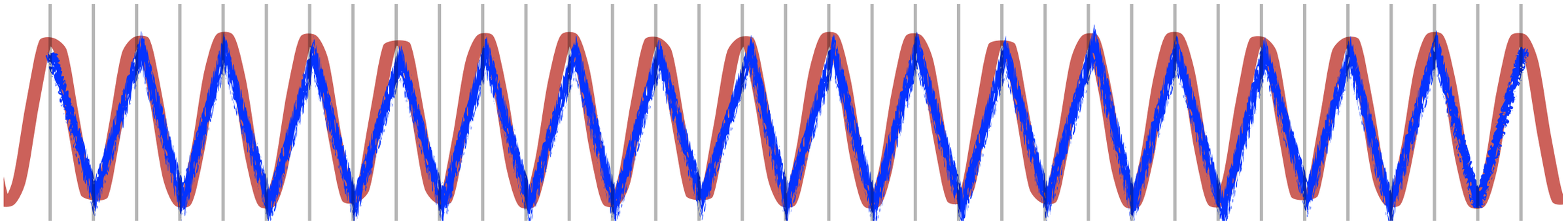


Counting electrons normalizes the signal from each electron on the sensor

Aliasing

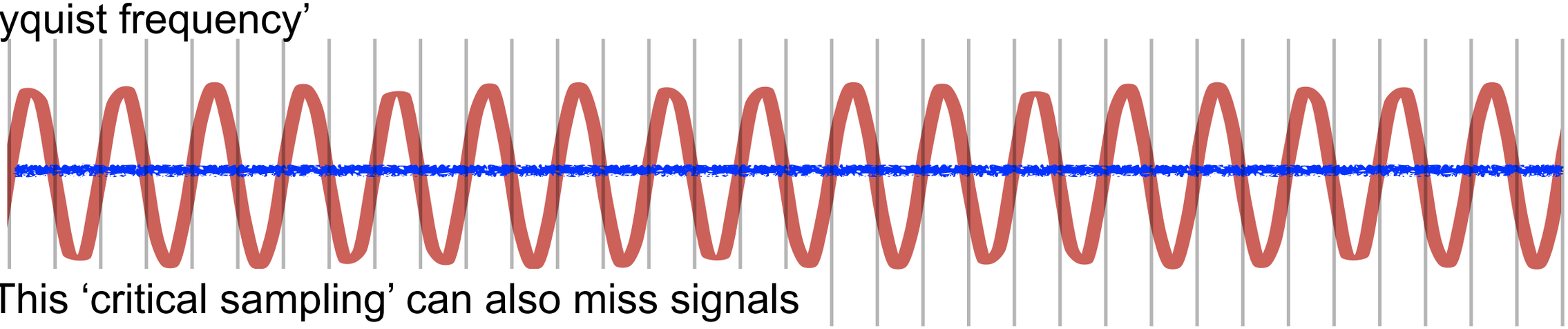
Nyquist frequency and aliasing

Sampling

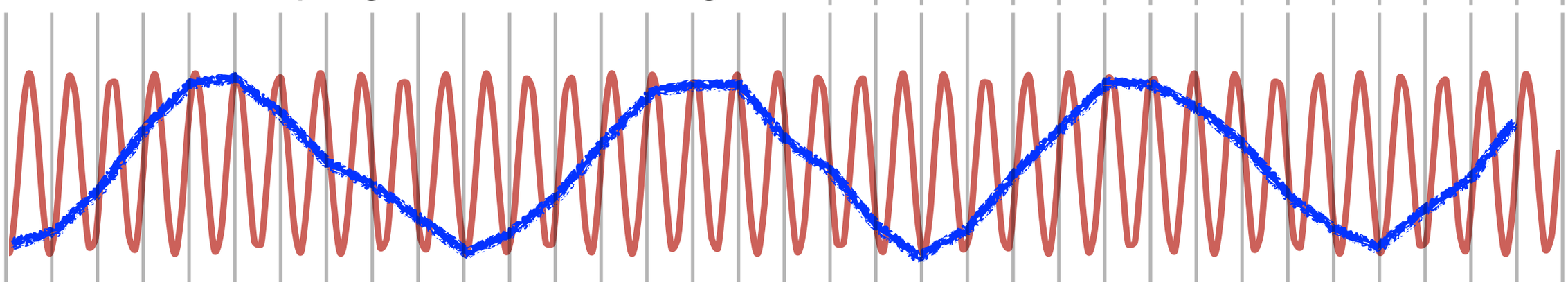


To capture signal of frequency ‘f’, must sample at 2f (e.g. 1 Å pixels allows 2 Å resolution)

‘Nyquist frequency’



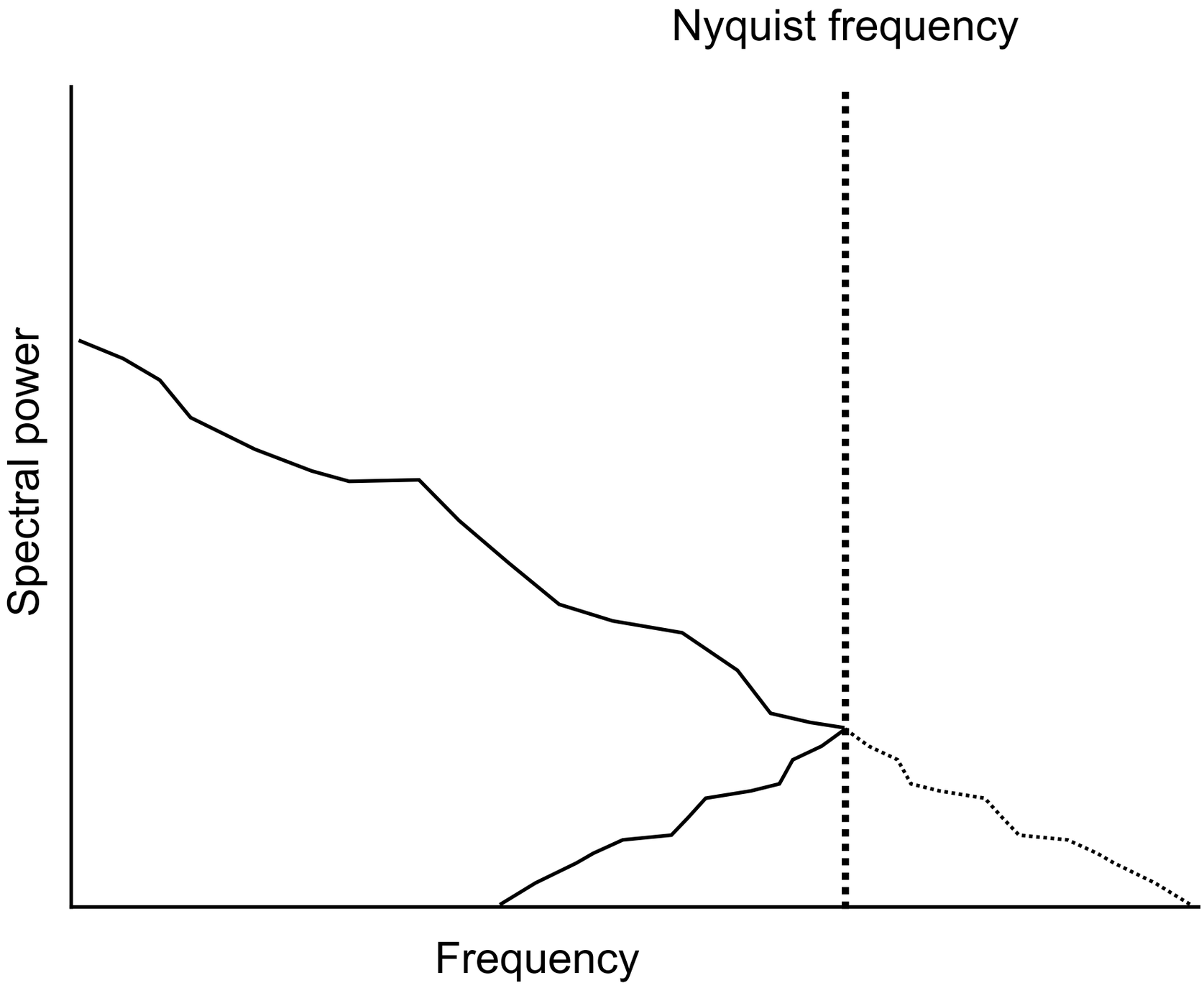
This ‘critical sampling’ can also miss signals



Sampling of signals that are higher-frequency than Nyquist produces lower-frequent power

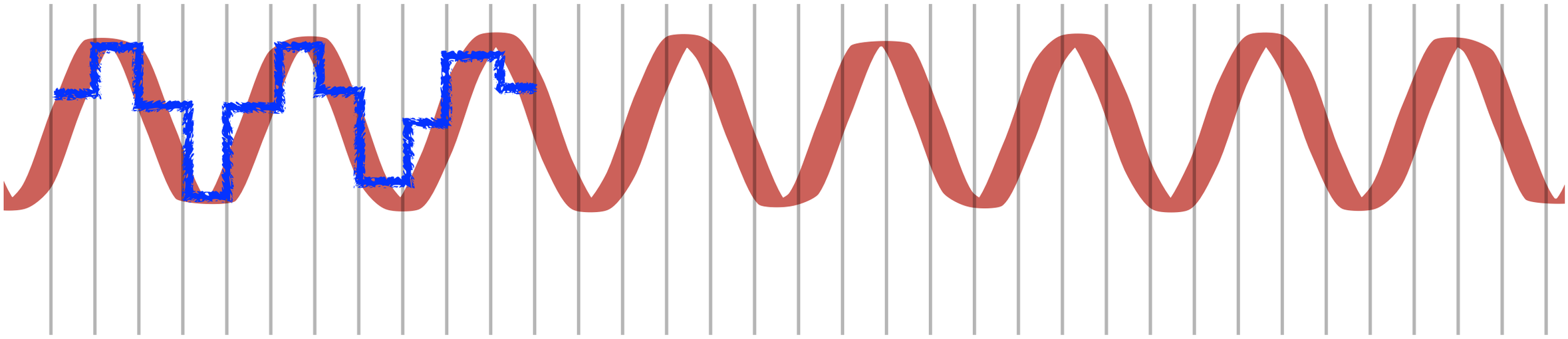
‘Aliasing’

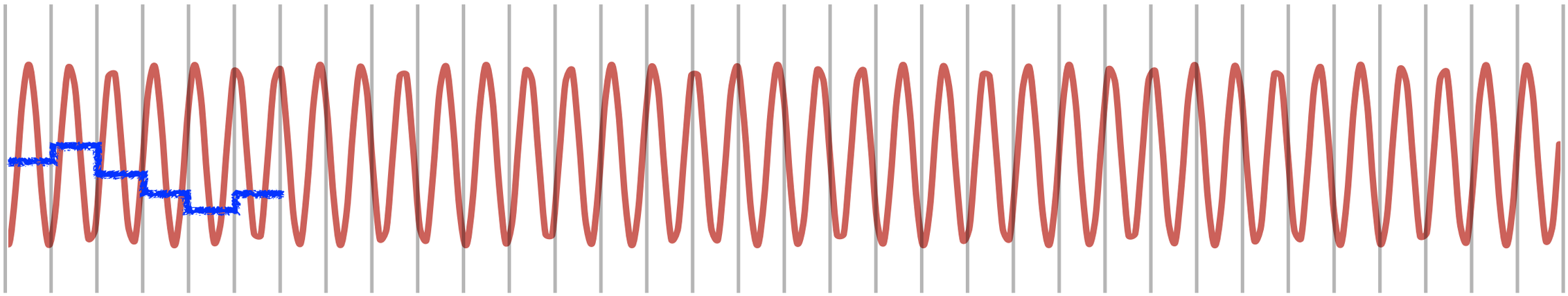
Effect of aliasing on spectral power



Aliasing for real electron sensors

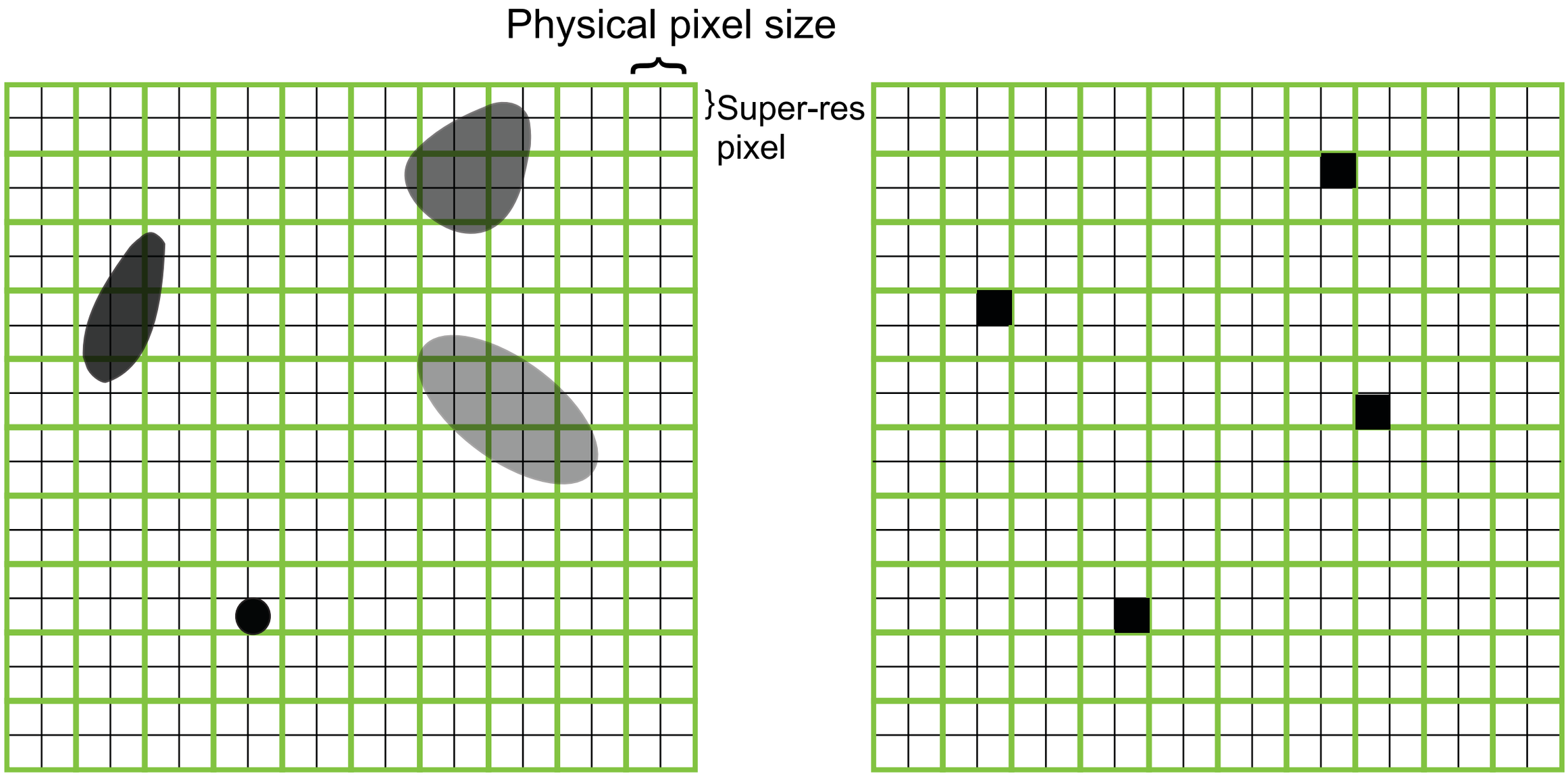
In reality, pixelated detector don’t sample analogue signal but integrate over pixel





Aliasing limits DQE of perfect detector to $(2/\pi)^2$

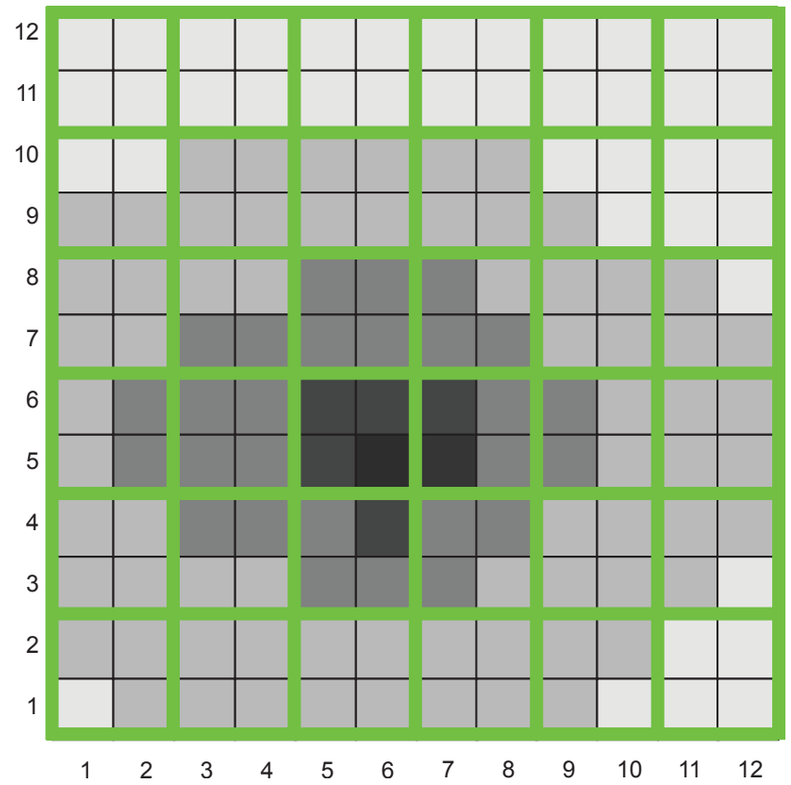
Avoiding the effects of Aliasing - approach 1 (K2/K3)



- Try to localize electron impacts to a corner of a pixel

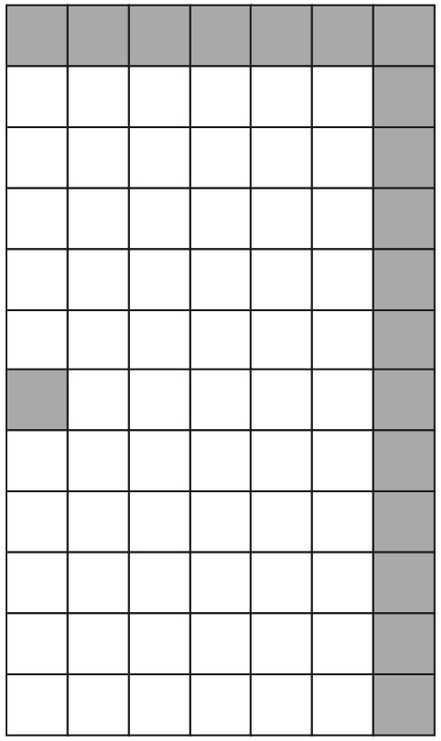
Avoiding the effects of Aliasing - approach 1 (K2/K3)

Collect data in “super-resolution mode”

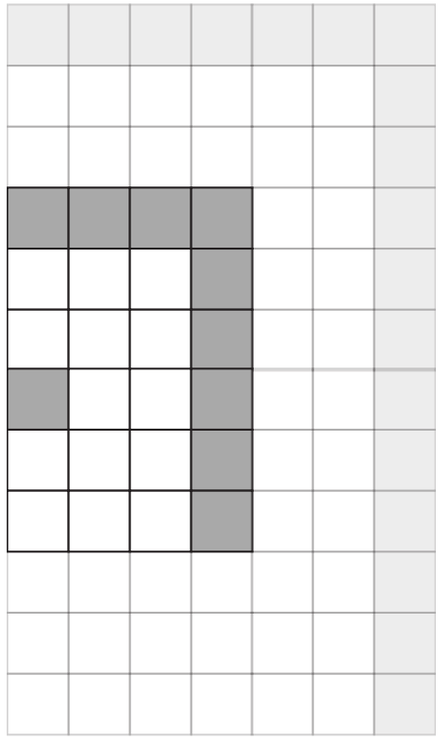


Four ‘super-resolution’ pixels per physical pixel

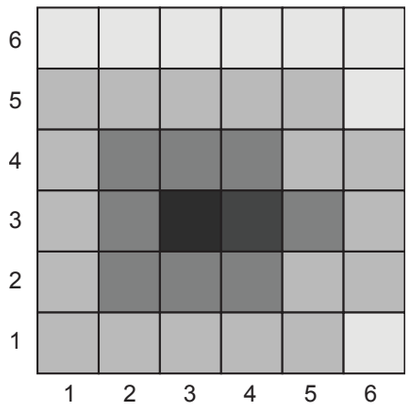
FT



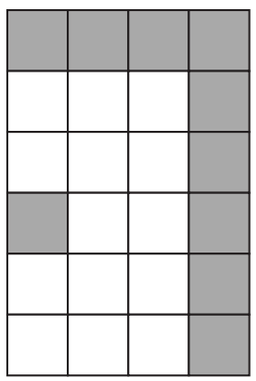
extract central region



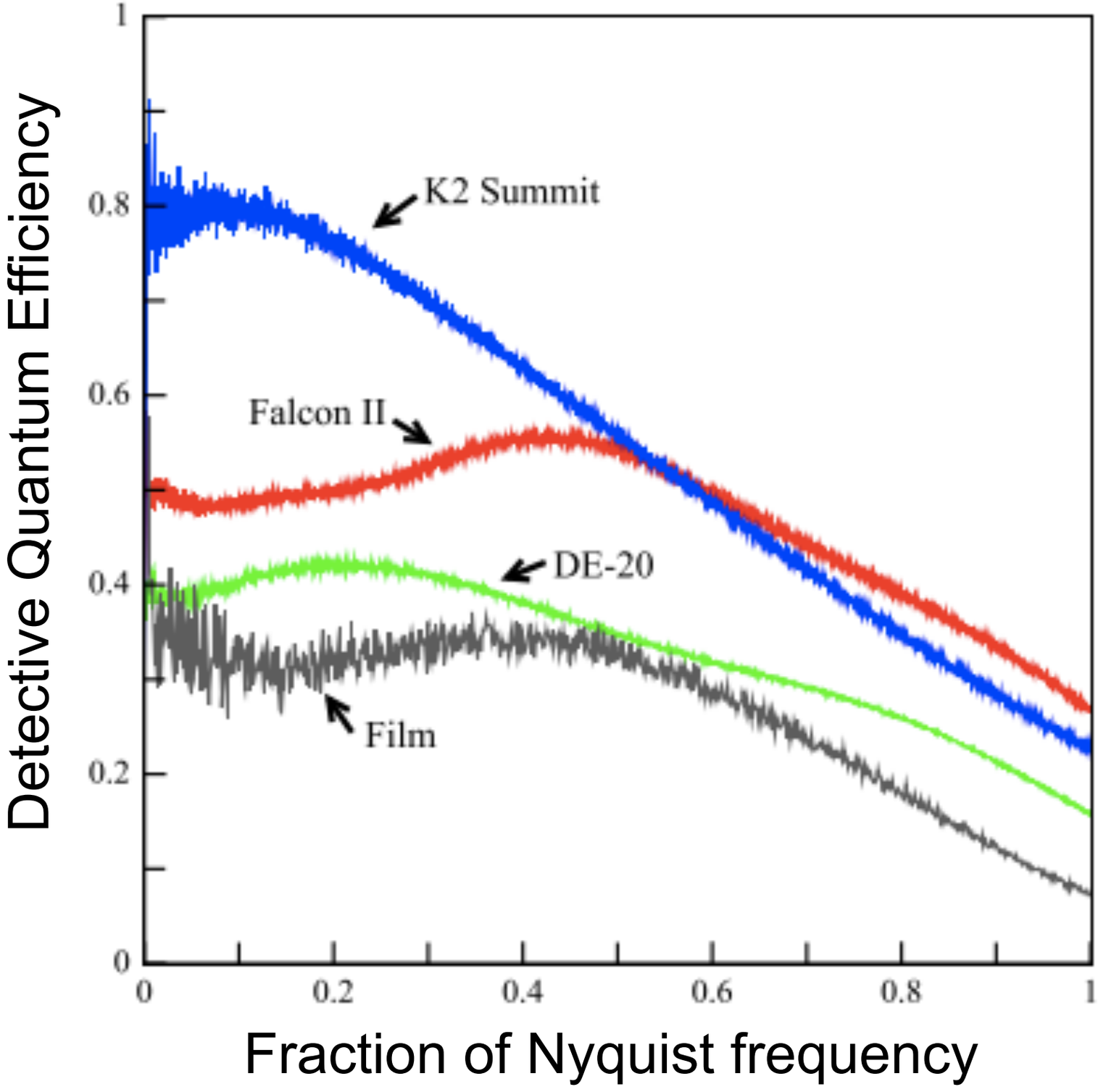
put in new array



FT^{-1}



Super-resolution signal from K2 summit has super-low DQE



McMullan et al. (2014), Ultramicroscopy 147, 156-63.

Don't try to use super-resolution signal for structure determination

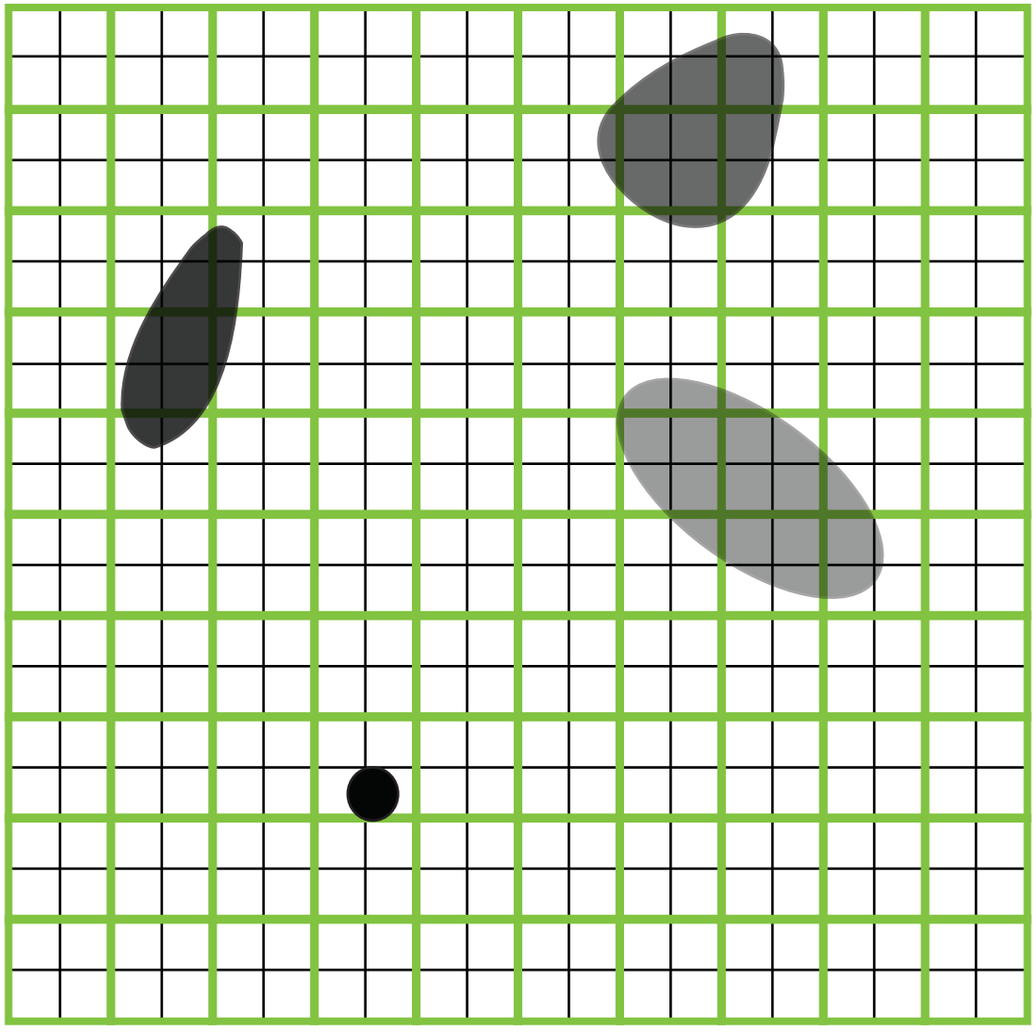
Super-resolution data collection followed by Fourier truncation may help reduce aliasing

Super-resolution data collection takes additional time/disc space

Reduced aliasing probably isn't worth the extra time it takes (get more particle images instead)

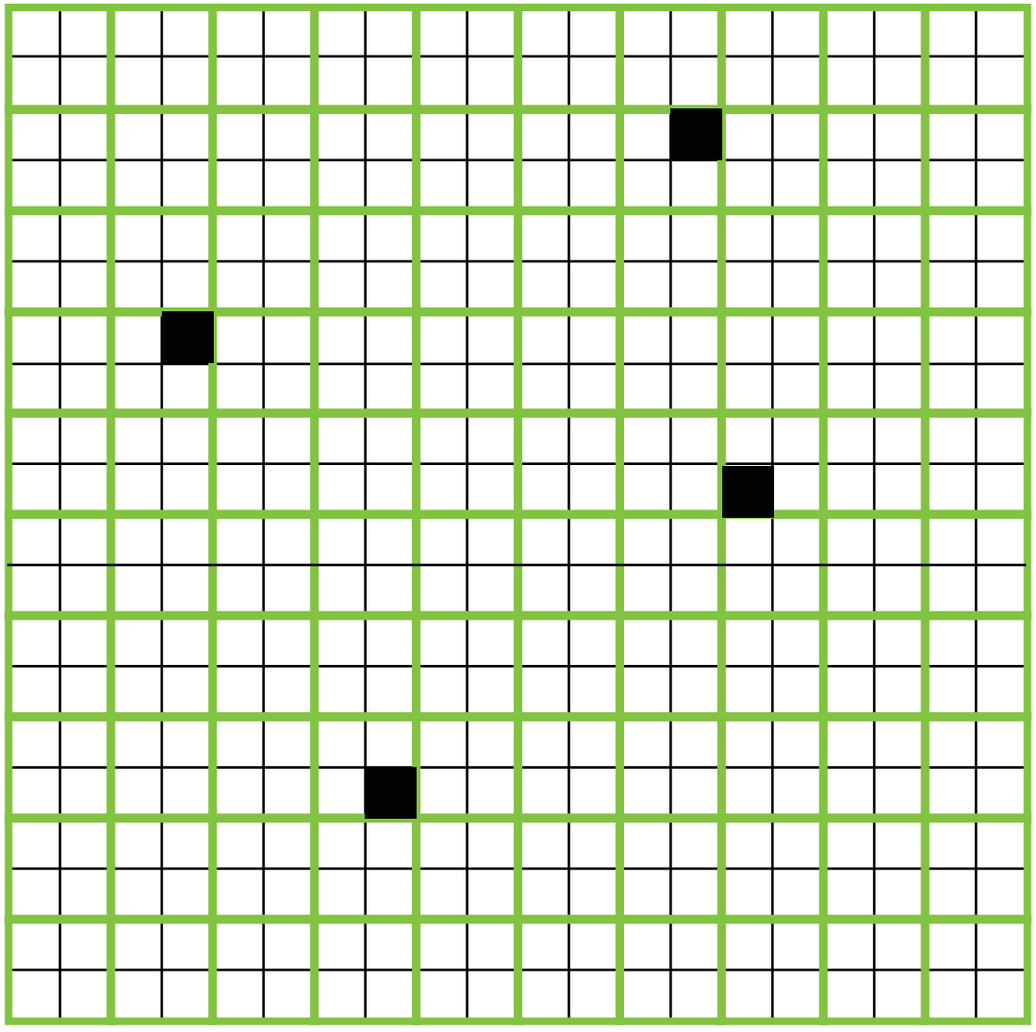
Averaging pixels 2×2 after super-resolution data collection re-aliases image

Avoiding the effects of Aliasing - approach 2 (Falcon3)



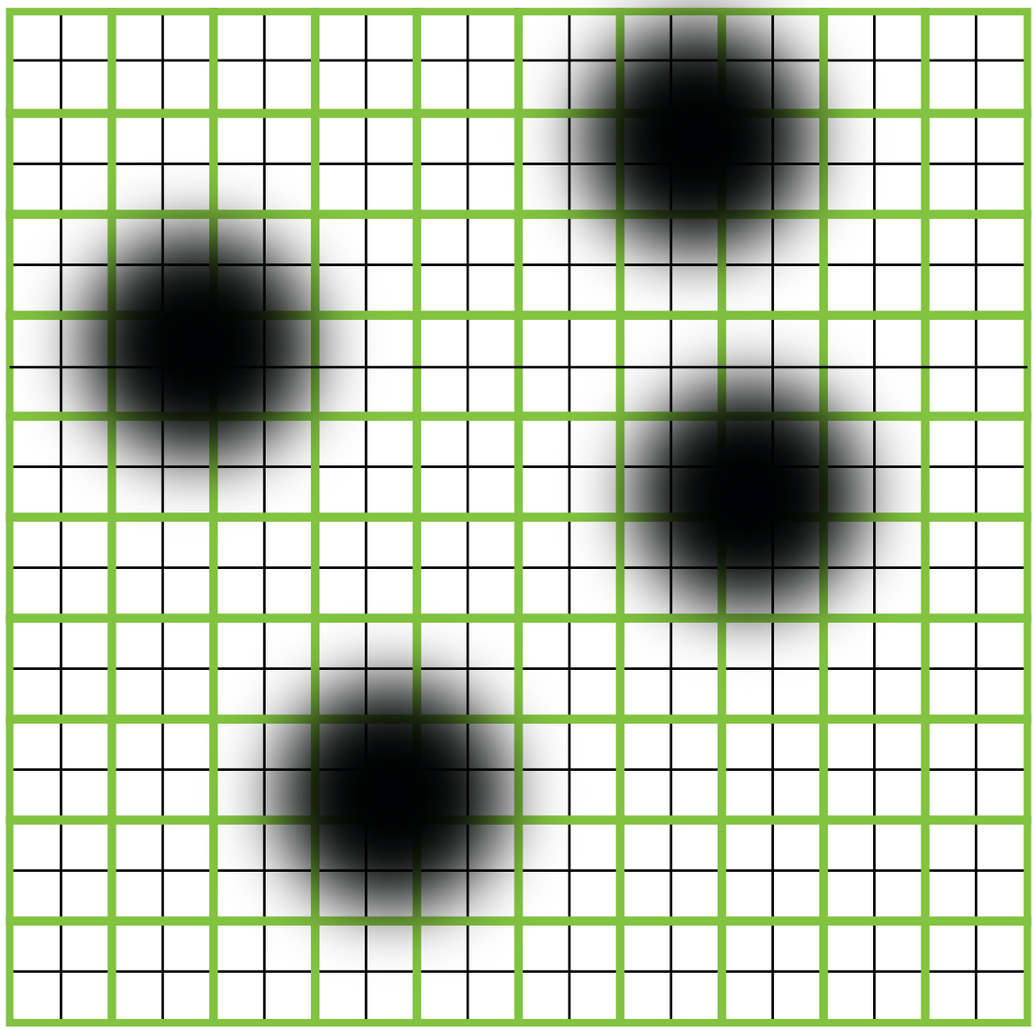
- Acquire a frame with ~1 el/100 pixels

Avoiding the effects of Aliasing - approach 2 (Falcon3)



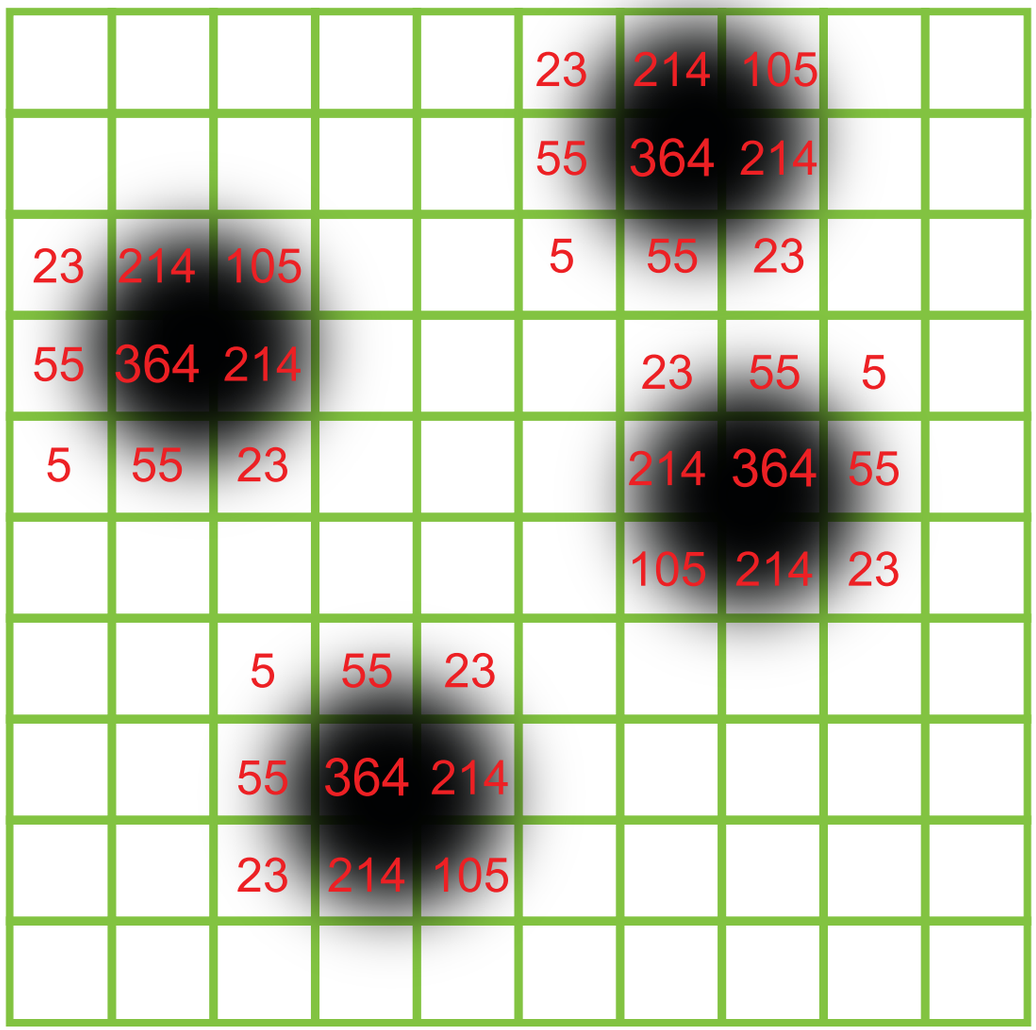
- Acquire a frame with ~1 el/100 pixels
- Localize electron to a sub-pixel (corner of a physical pixel)

Avoiding the effects of Aliasing - approach 2 (Falcon3)



- Acquire a frame with ~1 el/100 pixels
- Localize electron to a sub-pixel (corner of a physical pixel)
- Replace electron with a function (e.g. Gaussian) that covers multiple pixels

Avoiding the effects of Aliasing - approach 2 (Falcon3)



- Acquire a frame with ~1 el/100 pixels
- Localize electron to a sub-pixel (corner of a physical pixel)
- Replace electron with a function (e.g. Gaussian) that covers multiple pixels
- Determine contribute of Gaussian to 9 physical pixels and record

Avoiding the effects of Aliasing - approach 2 (Falcon3)

					23	214	105		
					55	364	214		
23	214	105			5	55	23		
55	364	214				23	55	5	
5	55	23				214	364	55	
						105	214	23	
		5	55	23					
		55	364	214					
		23	214	105					

Advantage: benefits of anti-aliasing without problems of recording super-resolution image

Disadvantage: may complicate data compression

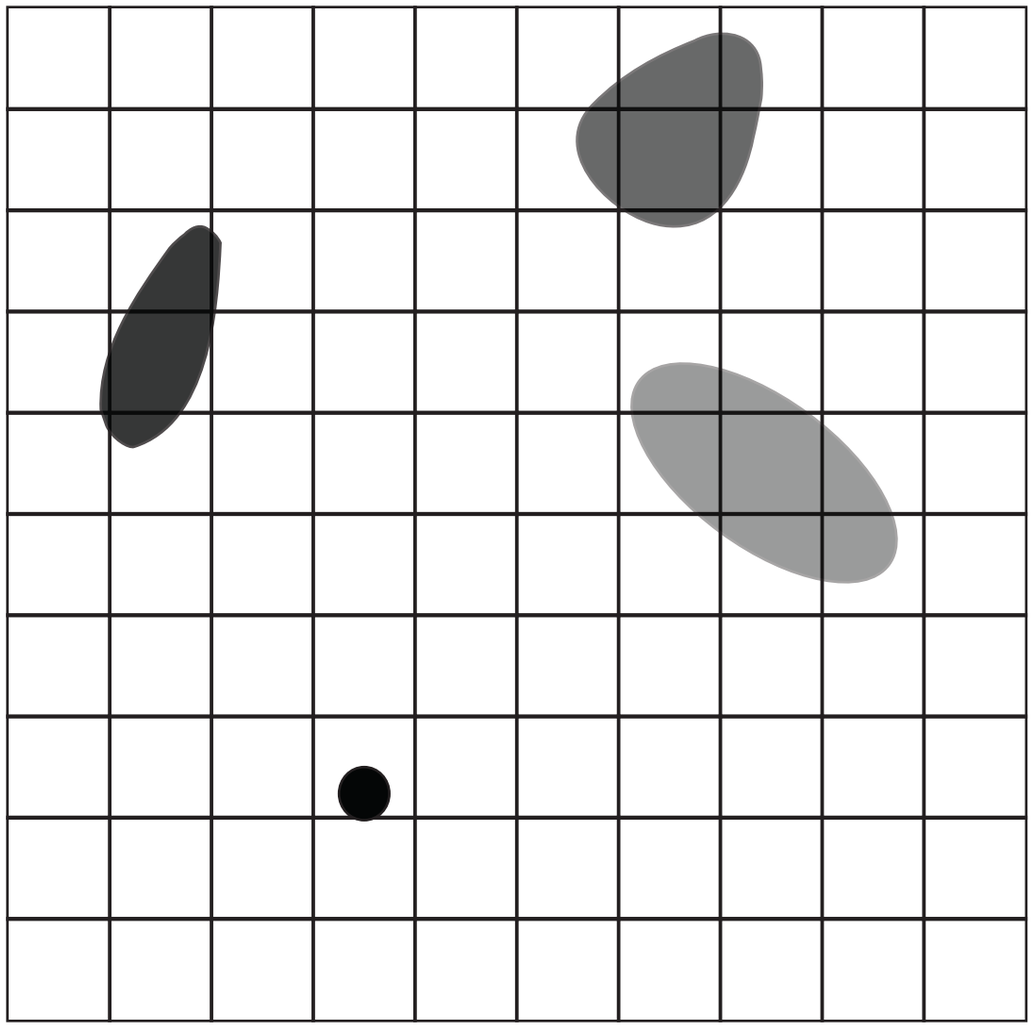
- Acquire a frame with ~1 el/100 pixels
- Localize electron to a sub-pixel (corner of a physical pixel)
- Replace electron with a function (e.g. Gaussian) that covers multiple pixels
- Determine contribute of Gaussian to 9 physical pixels and record

Coincidence loss

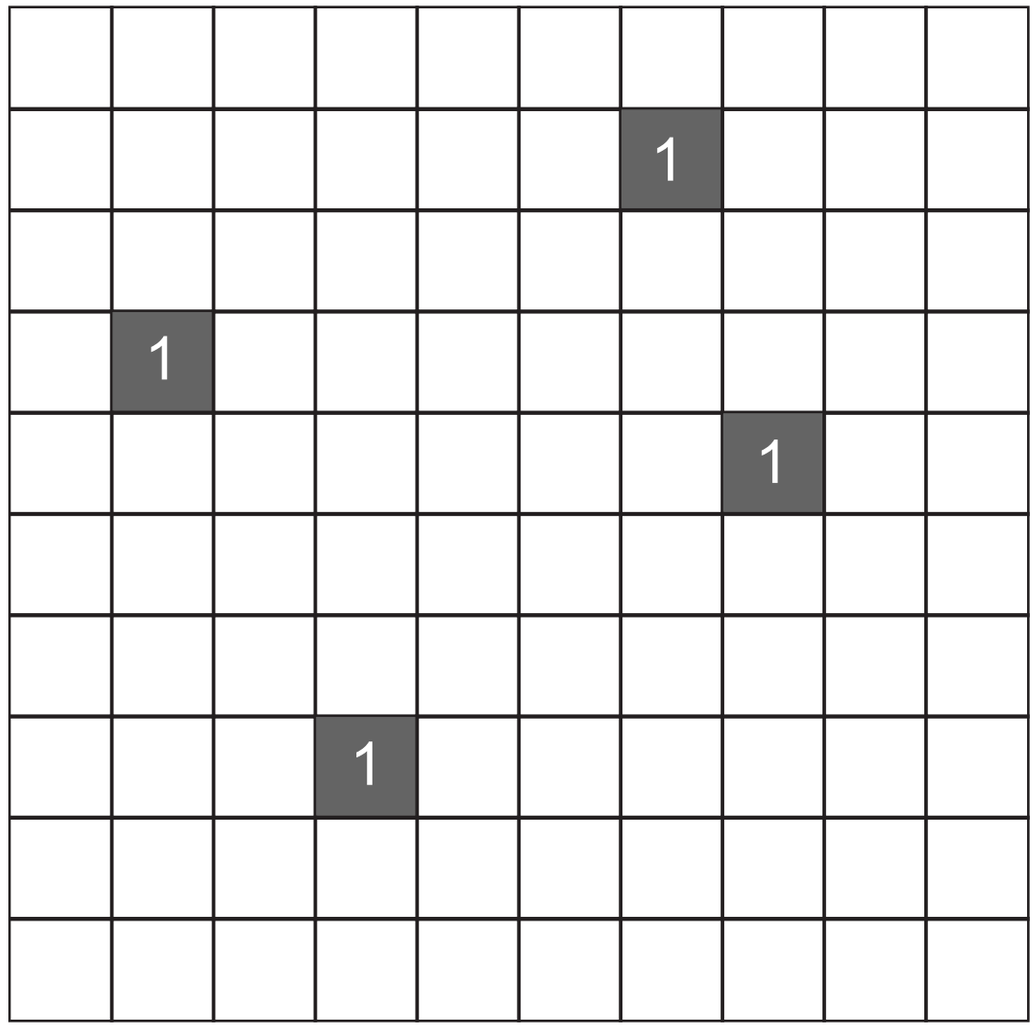
(What exposure rate should you use?)

Electron counting

Integration



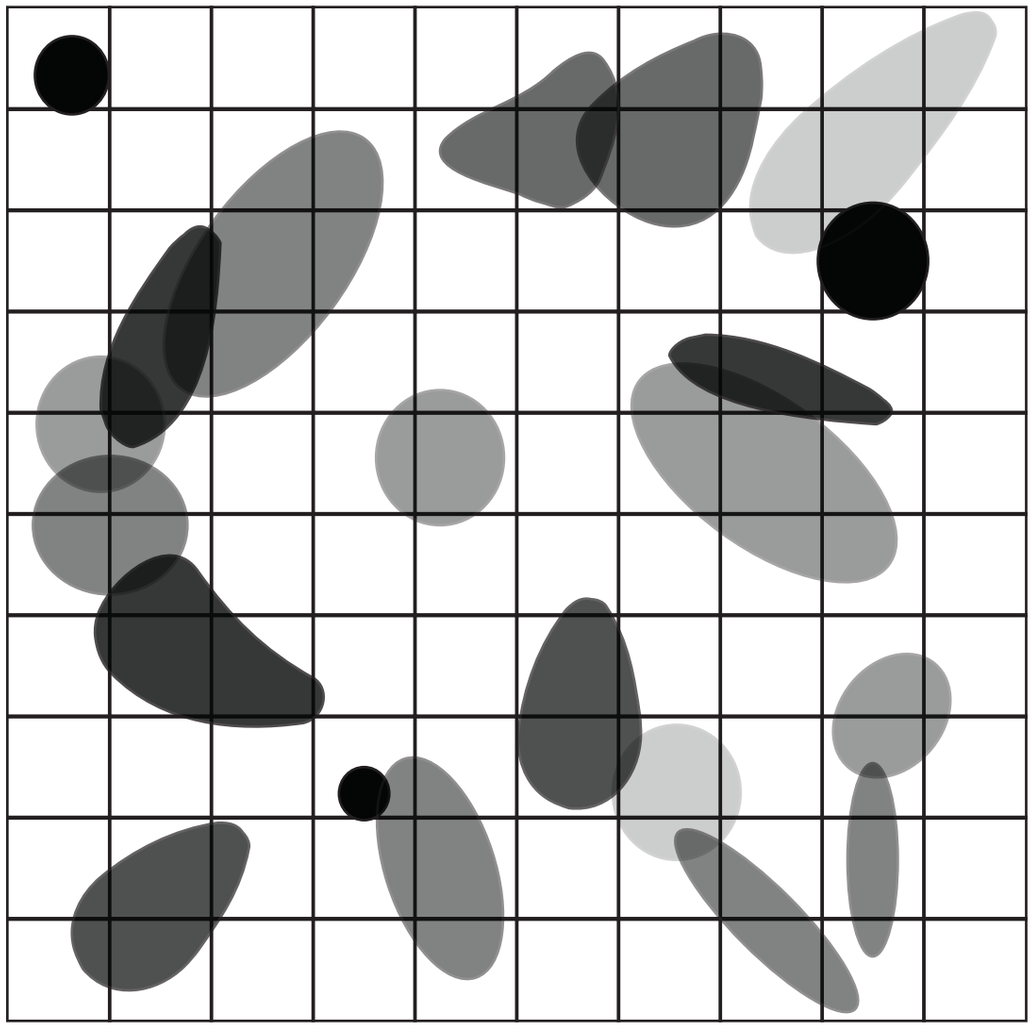
Counting



Counting electrons normalizes the signal from each electron on the sensor

Coincidence loss

Too many electrons in a frame for counting…



Missed electrons = **coincidence loss**

Ideal for counting: 1 electron/100 pixels

Conditions for achieving 1 el/100 pix/frame

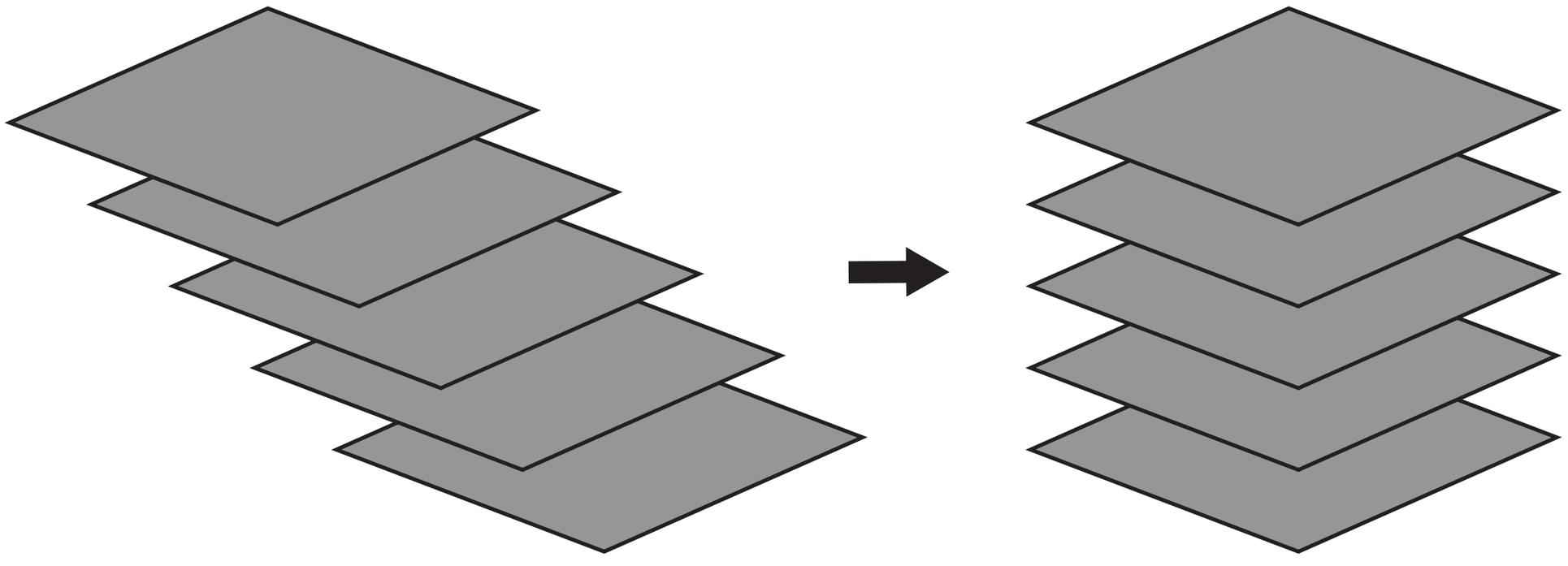
Camera	Frame rate (fps)	Exp rate (e/pix/s)
K2	400	4
K3	1500	15
Falcon 3	40	0.4
DE-20	32	0.32

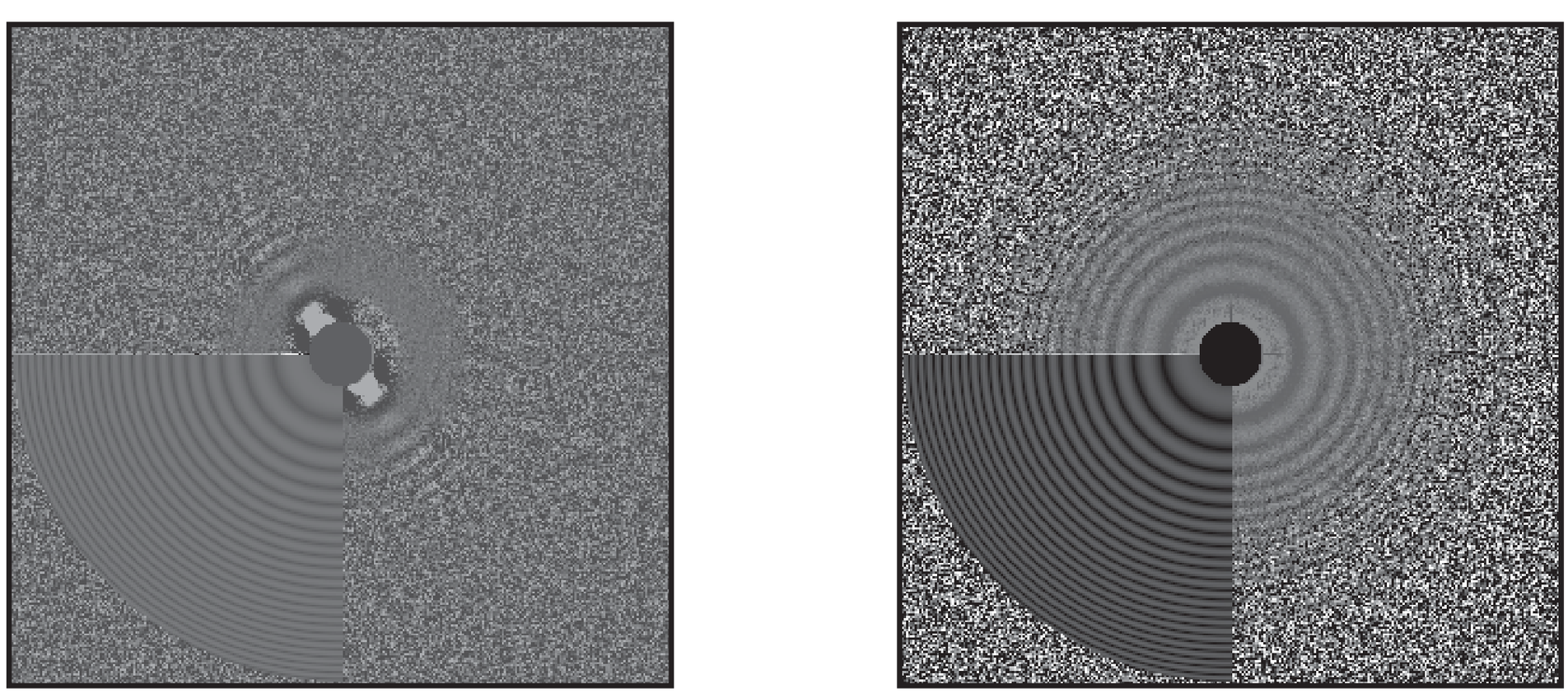
Conditions for achieving acceptable coincidence loss

Camera	Frame rate (fps)	Exp rate (e/pix/s)
K2	400	4-12
K3	1500	15-45
Falcon 3	40	0.4-1.2
DE-20	32	0.32-0.96

Microscope stage must be stable enough to allow counting!

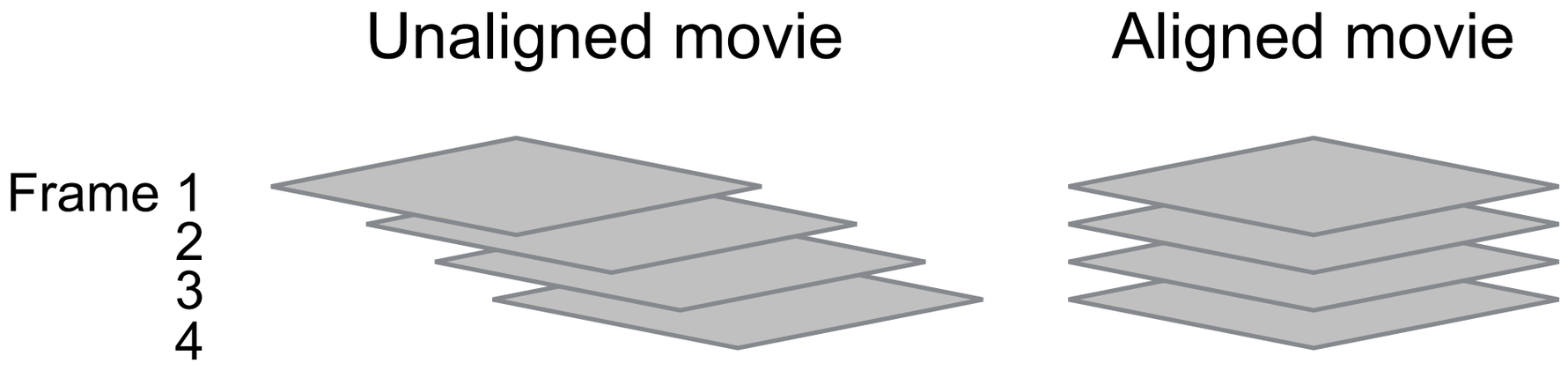
Frame alignment





The MotionCorr algorithm: least squares

Li … Cheng (2013). Nat Methods 10, 584-90.



- Define Frame 1 as “unshifted” (0,0)
- Calculate vectors (*xshift*,*yshift*) that bring two frames into register
- Can use cross correlation to estimate 6 unique vectors for 4 frame movie:

Frame 1 vs Frame 2
Frame 1 vs Frame 3
Frame 1 vs Frame 4
Frame 2 vs Frame 3
Frame 2 vs Frame 4
Frame 3 vs Frame 4

Can calculate (Z/2) × (Z-1) cross-correlation functions for a movie with Z frames
(e.g. 30 frame movie yields 435 CCFs)

MotionCorr: least squares method for aligning frames

t_{NM} means true shift vector between frames N and M

m_{NM} means measured shift vector (by cross correlation) between frames N and M

$$\begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} t_{12} \\ t_{23} \\ t_{34} \end{vmatrix} = \begin{vmatrix} m_{12} \\ m_{13} \\ m_{14} \\ m_{23} \\ m_{24} \\ m_{34} \end{vmatrix}$$

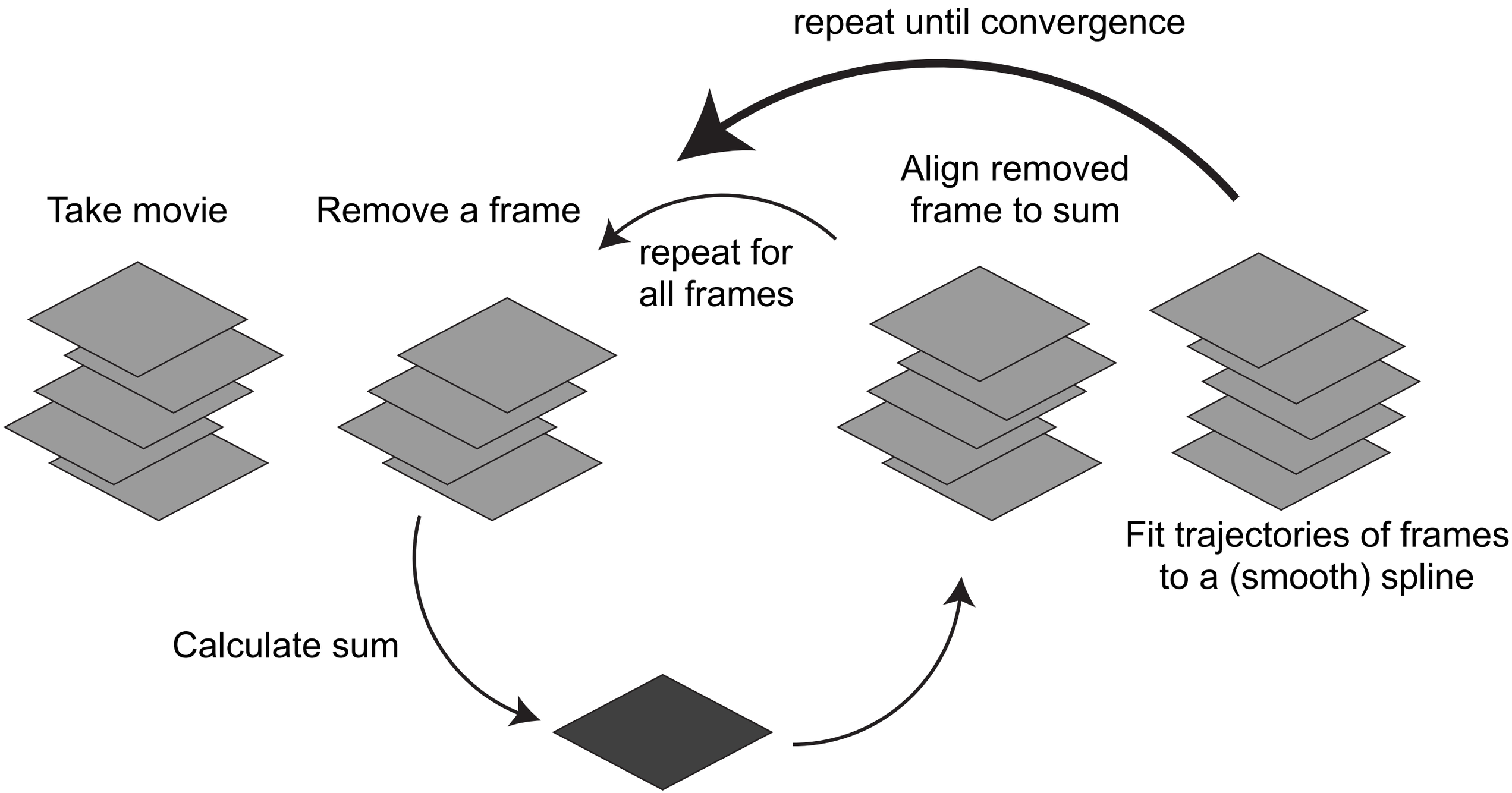
Li…Cheng, Nature Methods

$$m_{12} \simeq 1 \cdot t_{12} + 0 \cdot t_{23} + 0 \cdot t_{34}$$
$$m_{13} \simeq 1 \cdot t_{12} + 1 \cdot t_{23} + 0 \cdot t_{34}$$
$$m_{14} \simeq 1 \cdot t_{12} + 1 \cdot t_{23} + 1 \cdot t_{34}$$
$$m_{23} \simeq 0 \cdot t_{12} + 1 \cdot t_{23} + 0 \cdot t_{34}$$
$$m_{24} \simeq 0 \cdot t_{12} + 1 \cdot t_{23} + 1 \cdot t_{34}$$
$$m_{34} \simeq 0 \cdot t_{12} + 0 \cdot t_{23} + 1 \cdot t_{34}$$

Once matrices are filled in standard linear algebra can be used to find values that best fit the data for t_{12}, t_{23}, t_{34}

The Unblur algorithm: Iterative sum & align (with splines)

Grant & Grigorieff (2015) *eLife* 4:e06980

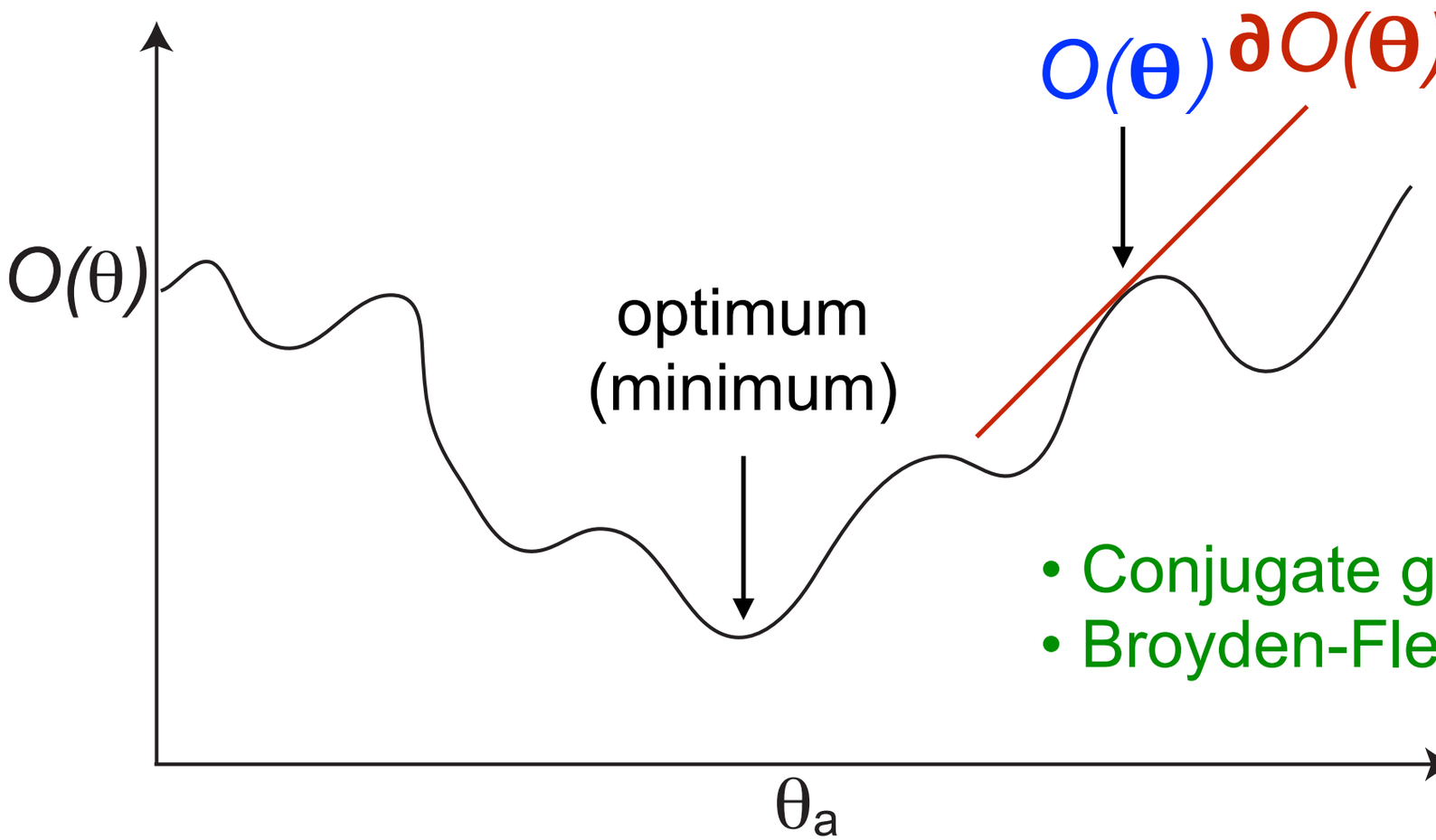


Alignframes_lmbfgs: Gradient-based optimization

Rubinstein and Brubaker (2015) *J Struct Biol* **192**, 188-95. (arXiv 2014)

Use an objective function that, when minimized, maximizes the sum of the correlations of each shifted frame with the sum of the shifted frames.

Calculate the partial derivatives of the objective function to quickly and accurately find the best value of the objective/best shifts.



- Conjugate gradients
- Broyden-Fletcher-Goldfarb-Shanno

Alignframes_lmbfgs: Gradient-based optimization

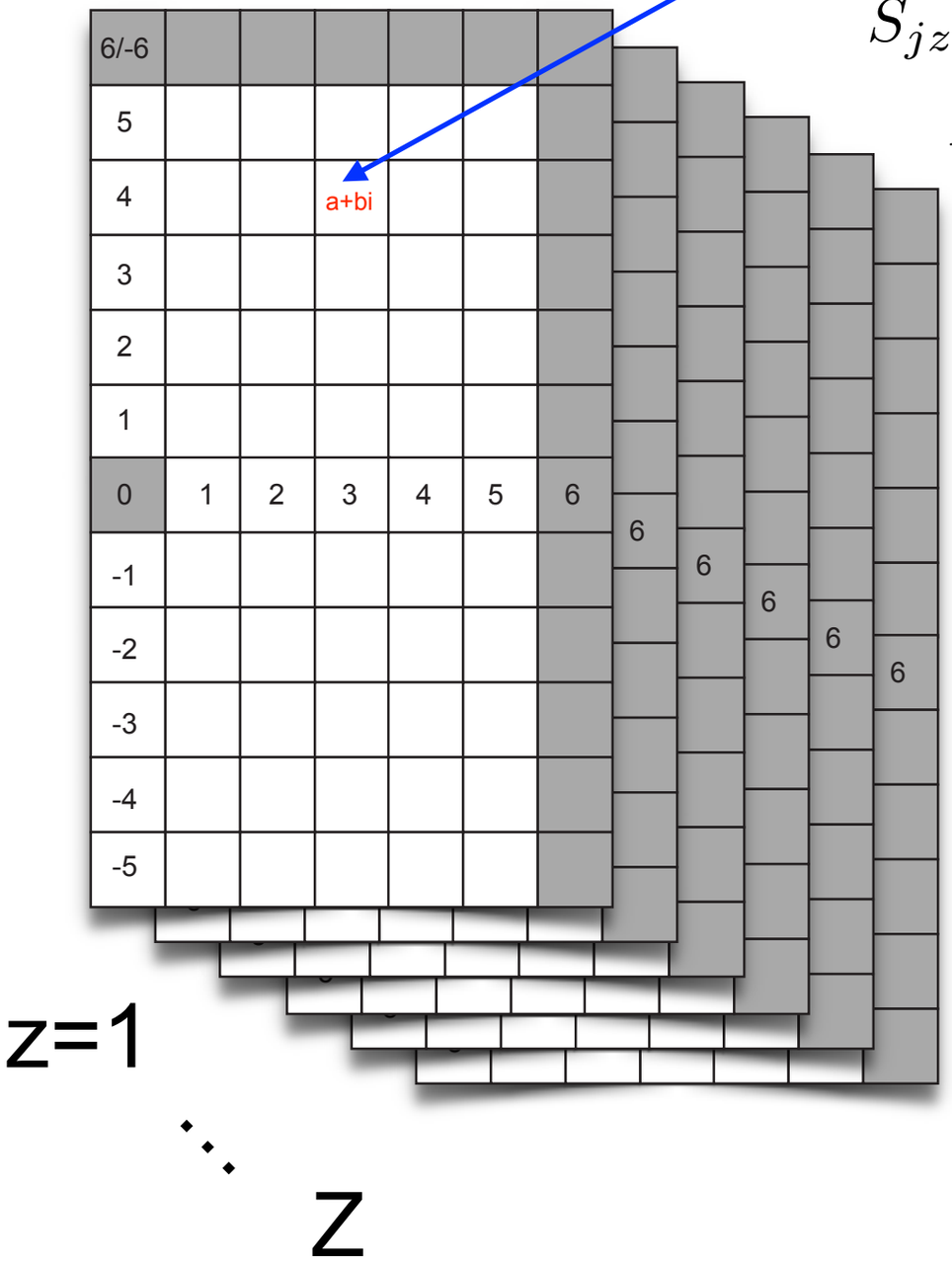
Rubinstein and Brubaker (2015) *J Struct Biol* **192**, 188-95. (arXiv 2014)

FTs of shifted frames

$$F_{jz}(\cos\phi_{jz} + i\sin\phi_{jz})$$

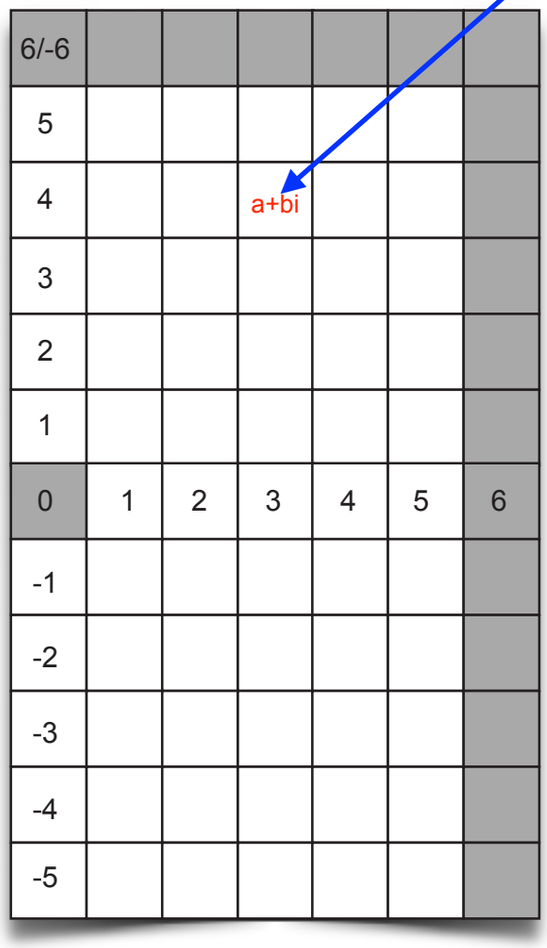
$$S_{jz} = (\cos\phi_{jz} + i\sin\phi_{jz})$$

$$F_{jz}S_{jz}$$



Sum of FTs

$$\sum_{z=1}^{Z} F_{jz}S_{jz}$$



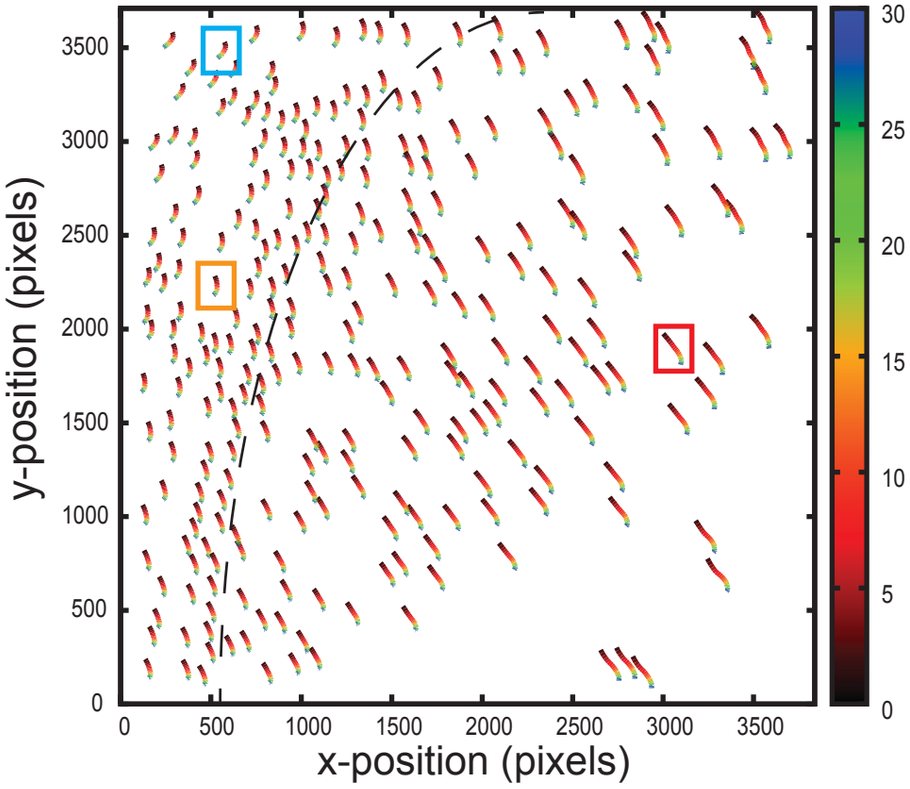
Overall objective function and its derivatives:

$$O(\Theta) = -Re\sum_{z=1}^{Z}\sum_{j=1}^{J}\left[F_{jz}^{*}S_{jz}^{*}\sum_{z'=1}^{Z}F_{jz'}S_{jz'}\right]$$

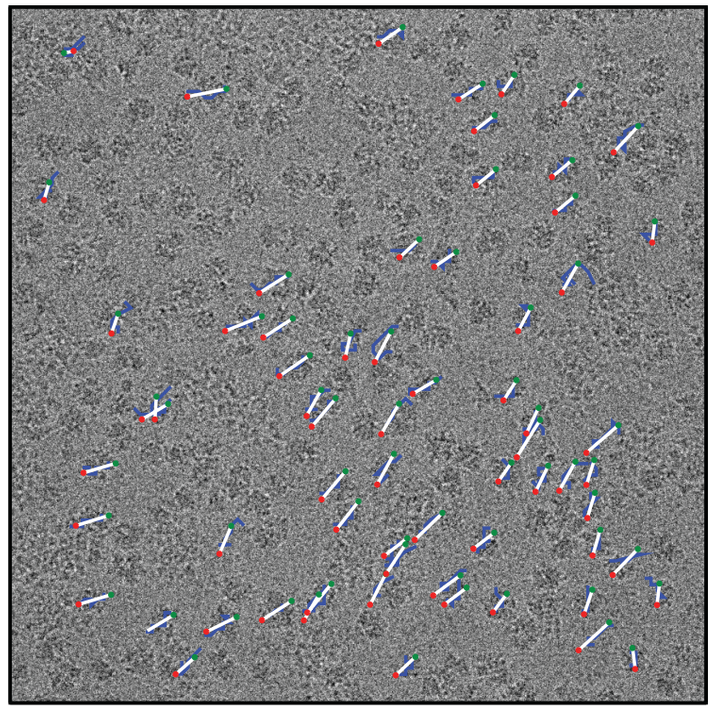
$$\frac{\partial O(\Theta)}{\partial x_a} = -Re\sum_{j=1}^{J}\frac{2\pi i k_x(j)}{N}\left[F_{ja}S_{ja}\sum_{z=1}^{Z}F_{jz}^{*}S_{jz}^{*} - F_{ja}^{*}S_{ja}^{*}\sum_{z=1}^{Z}F_{jz}S_{jz}\right]$$

$$\frac{\partial O(\Theta)}{\partial y_a} = -Re\sum_{j=1}^{J}\frac{2\pi i k_y(j)}{N}\left[F_{ja}S_{ja}\sum_{z=1}^{Z}F_{jz}^{*}S_{jz}^{*} - F_{ja}^{*}S_{ja}^{*}\sum_{z=1}^{Z}F_{jz}S_{jz}\right]$$

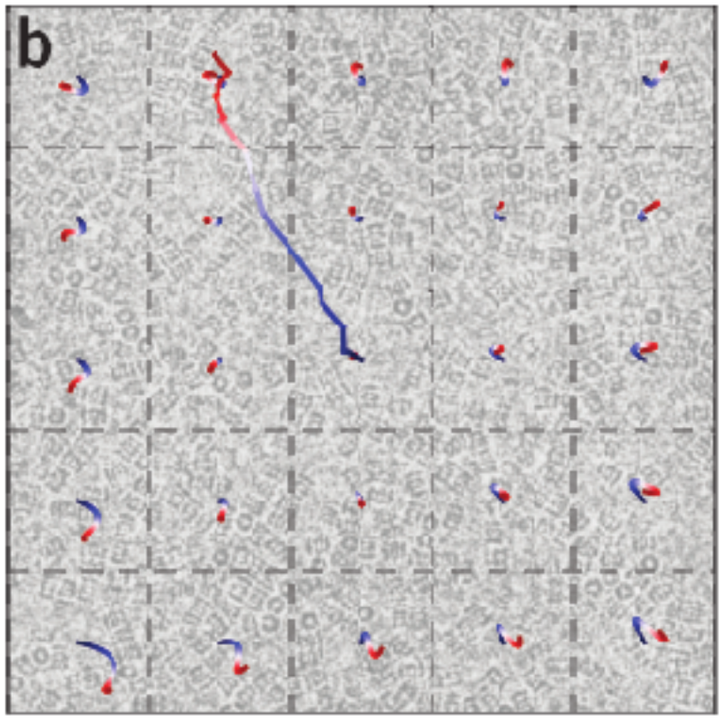
Particle/patch-based motion correction



Alignparts_lmbfgs (Rubinstein & Brubaker, 2015, *JSB* 192, 188-95)



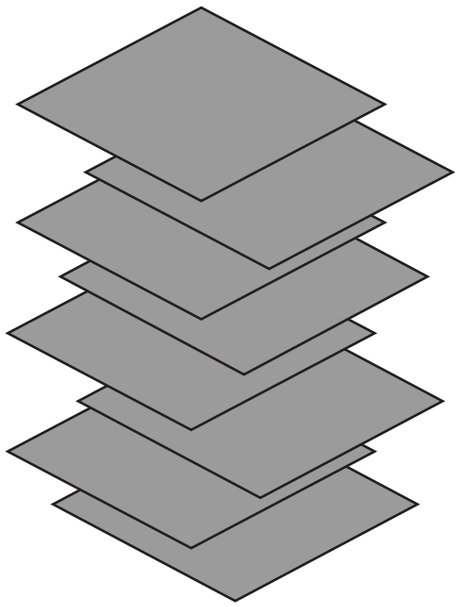
Relion Polishing (Scheres, 2014, *eLife 3:e03665*)



MotionCor2 (Zheng…Agard, 2017, *Nat Meth* 14, 331-2)

Particle polishing in *Relion*

Relion Polishing (Scheres, 2014, *eLife 3:e03665*)



Do whole-frame alignment

↓

Rolling average of frames

Frame2'=ave(Frame1,2,3)

Frame3'=ave(Frame2,3,4)

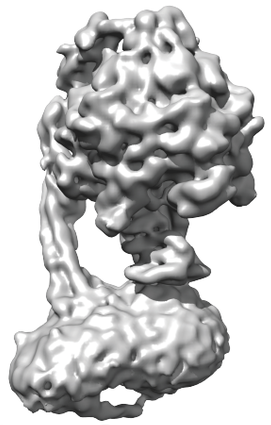
Frame4'=ave(Frame2,3,4)

.

.

.

FrameN'=ave(FrameN-1,N,N+1)

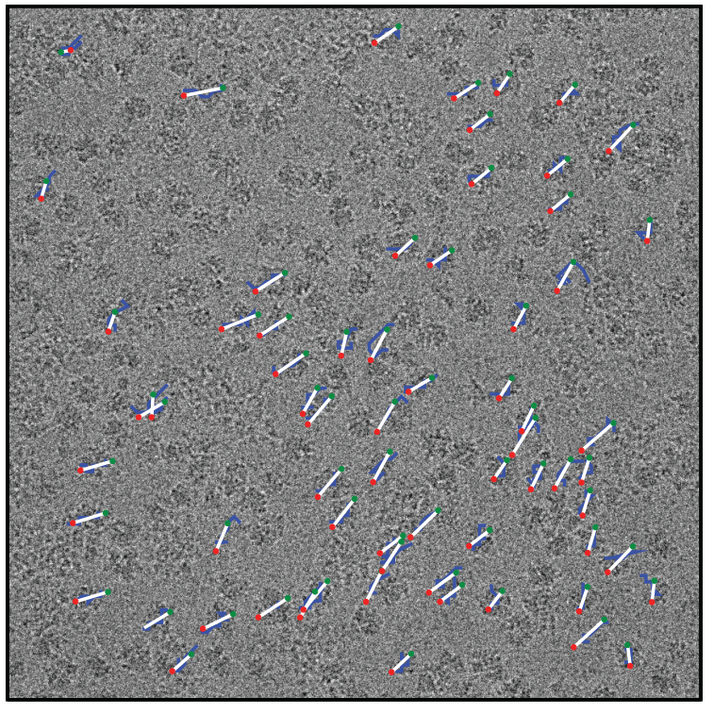


Align rolling frame averages to projections of reference map

Local averaging of speed & direction

Fit trajectories to a straight line

position=$(x_0,y_0)+(\Delta x,\Delta y)$*frame



MotionCor2

Align whole frames with *Unblur*-type approach

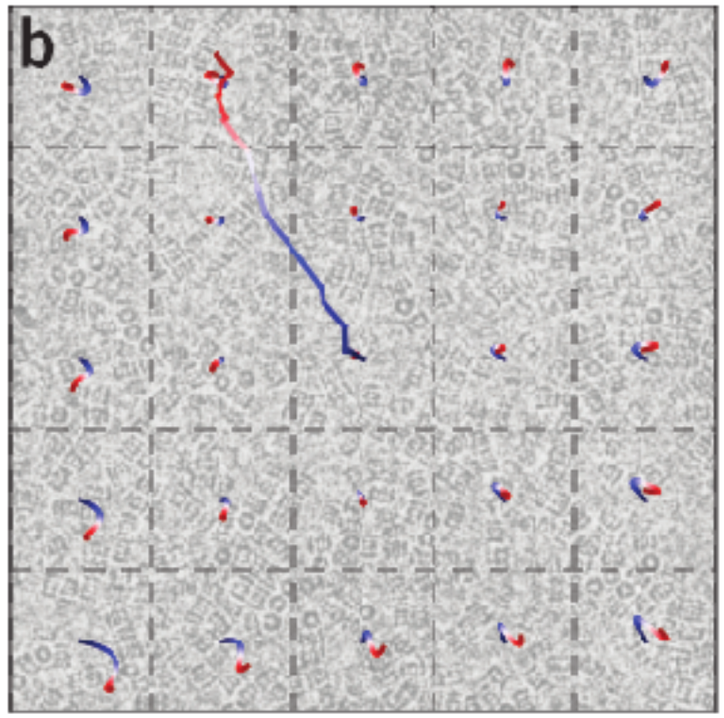
Divide image into grid of patches (e.g. 5×5)

Align each patch with *Unblur*-type approach

Fit shift at centre of patches to a polynomial model

Use polynomial model to find pixel shifts between patch centres

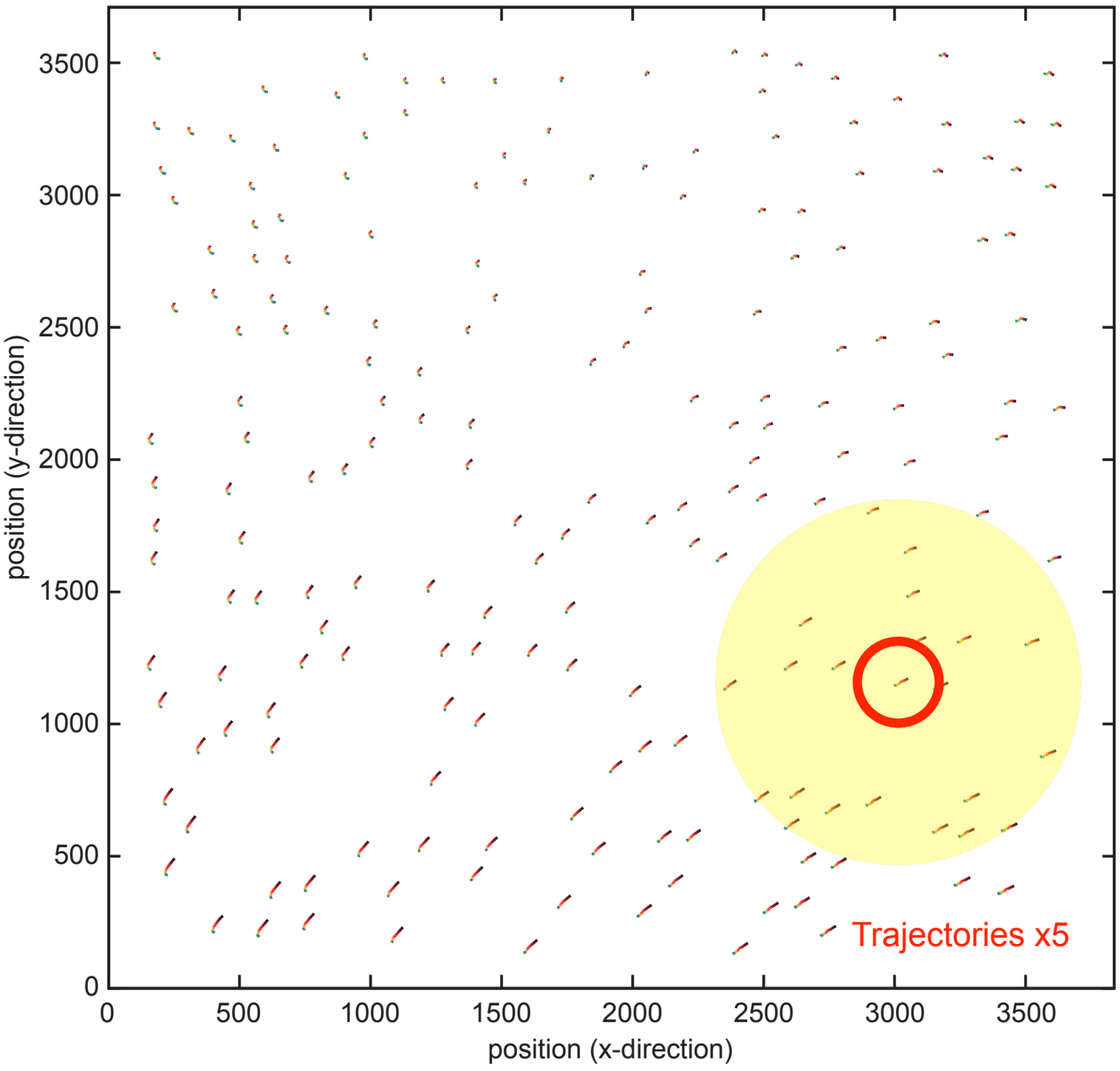
Deform image so that each pixel is appropriately shifted

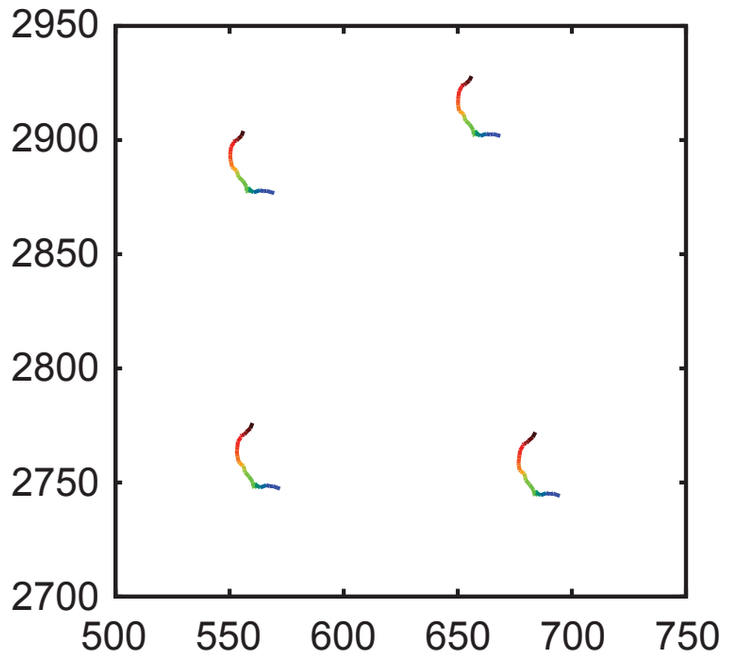


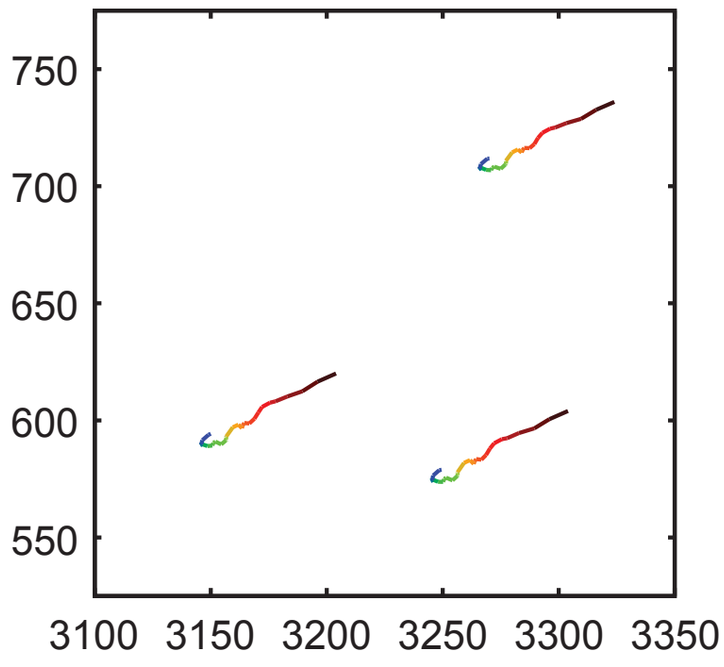
MotionCor2 (Zheng…Agard, 2017, *Nat Meth* 14, 331-2)

alignparts_lmbfgs: gradient-based optimization

Like *alignframes_lmbfgs* but use particle boxes and force smoothness and local averaging







Rubinstein and Brubaker (2015) *J Struct Biol* **192**, 188-95. (arXiv 2014)

High-throughput structure determination pipeline

Sample (from collaborator)

↓

Optimize specimen on F20/K2

↓

Alignframes/Alignparts small dataset

↓

cryoSPARC: ~ 5Å

↓

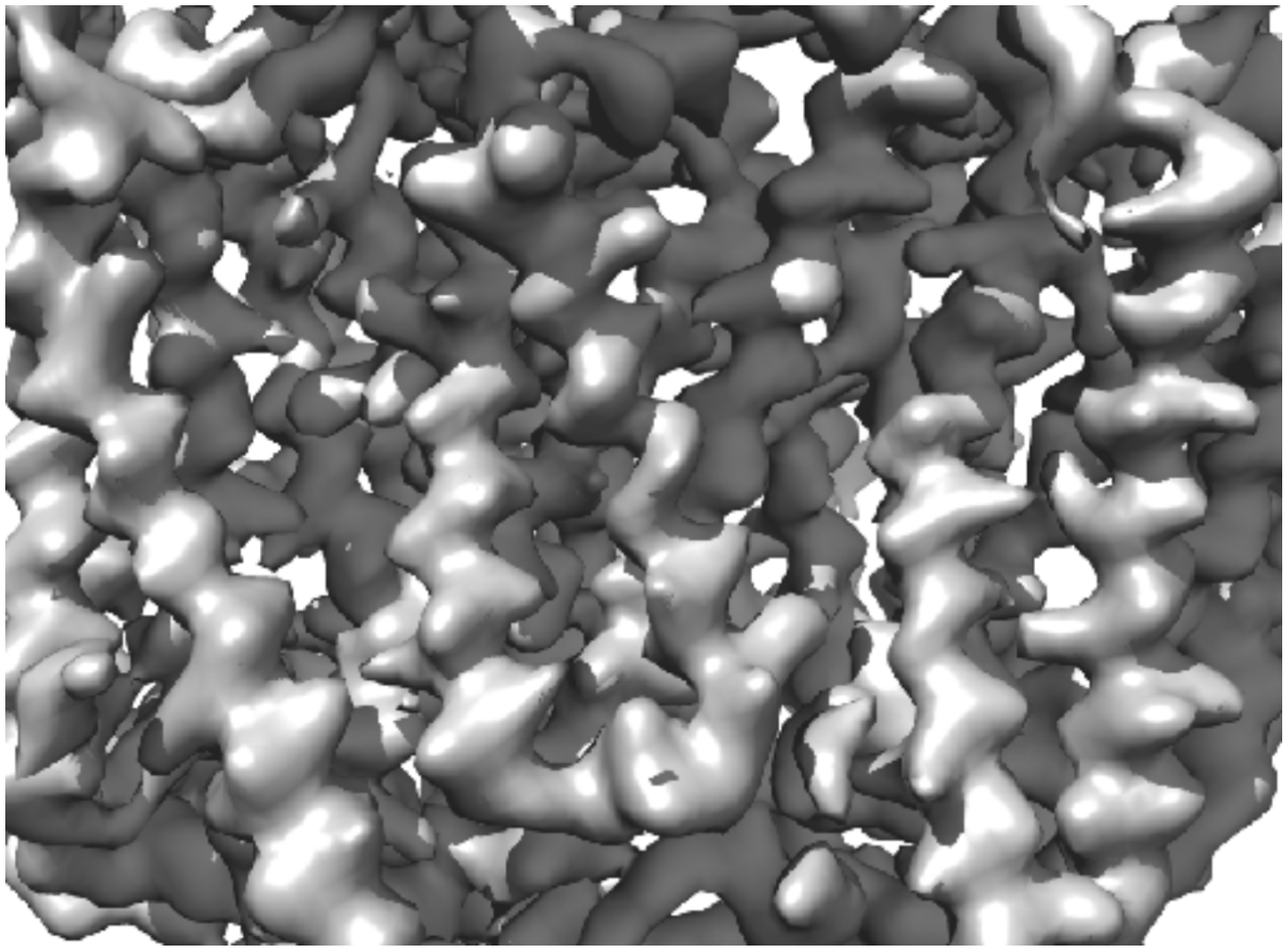
24h Titan Krios/K2 (NRAMM/SEMC)

↓

MotionCor2 large dataset

↓

cryoSPARC: 3.4 Å



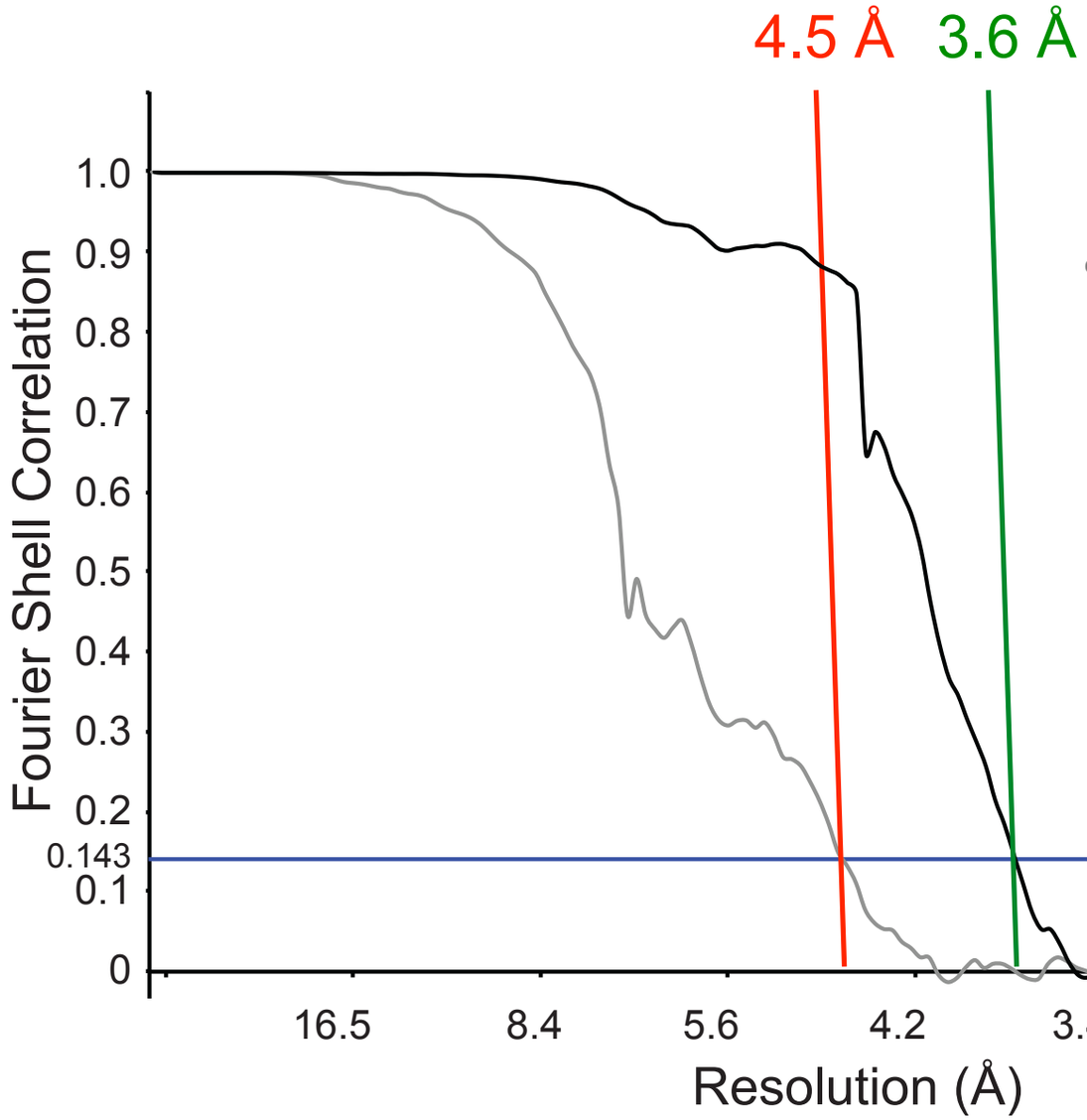
Unpublished

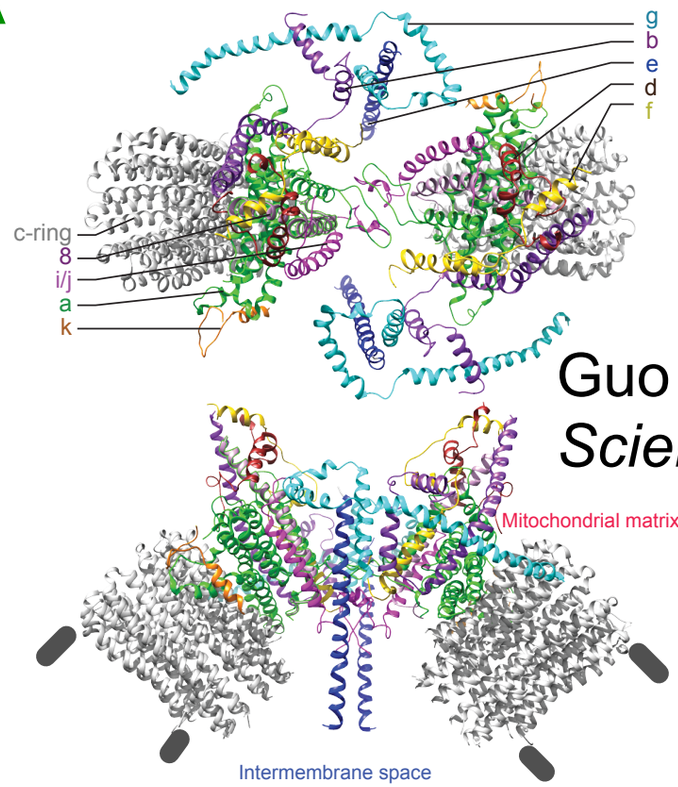
One size does not fit all… (Hui Guo)

Initial dataset: F20/K2, 239159 particle images, *alignparts_lmbfgs* (4.4 Å)

NRAMM dataset: Titan Krios/Quantum K2, 470036 particle images, *MotionCor2* (4.5 Å)

NRAMM dataset: Titan Krios/Quantum K2, 446259 particle images, *alignparts_lmbfgs* (3.6 Å)





Guo *et al.* (2017)
Science, **In Press**

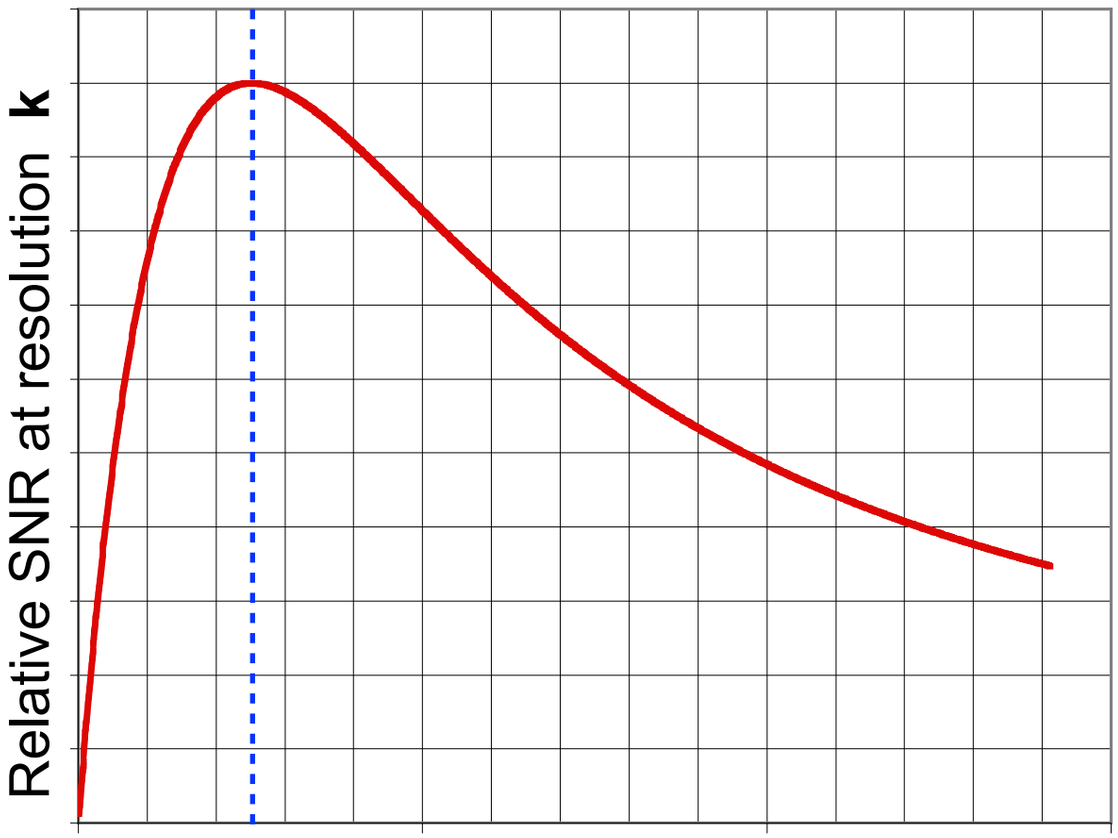
NRAMM: Hui (Alex) Wei, Bridget Carragher, Clint Potter

Comparison of motion correction methods

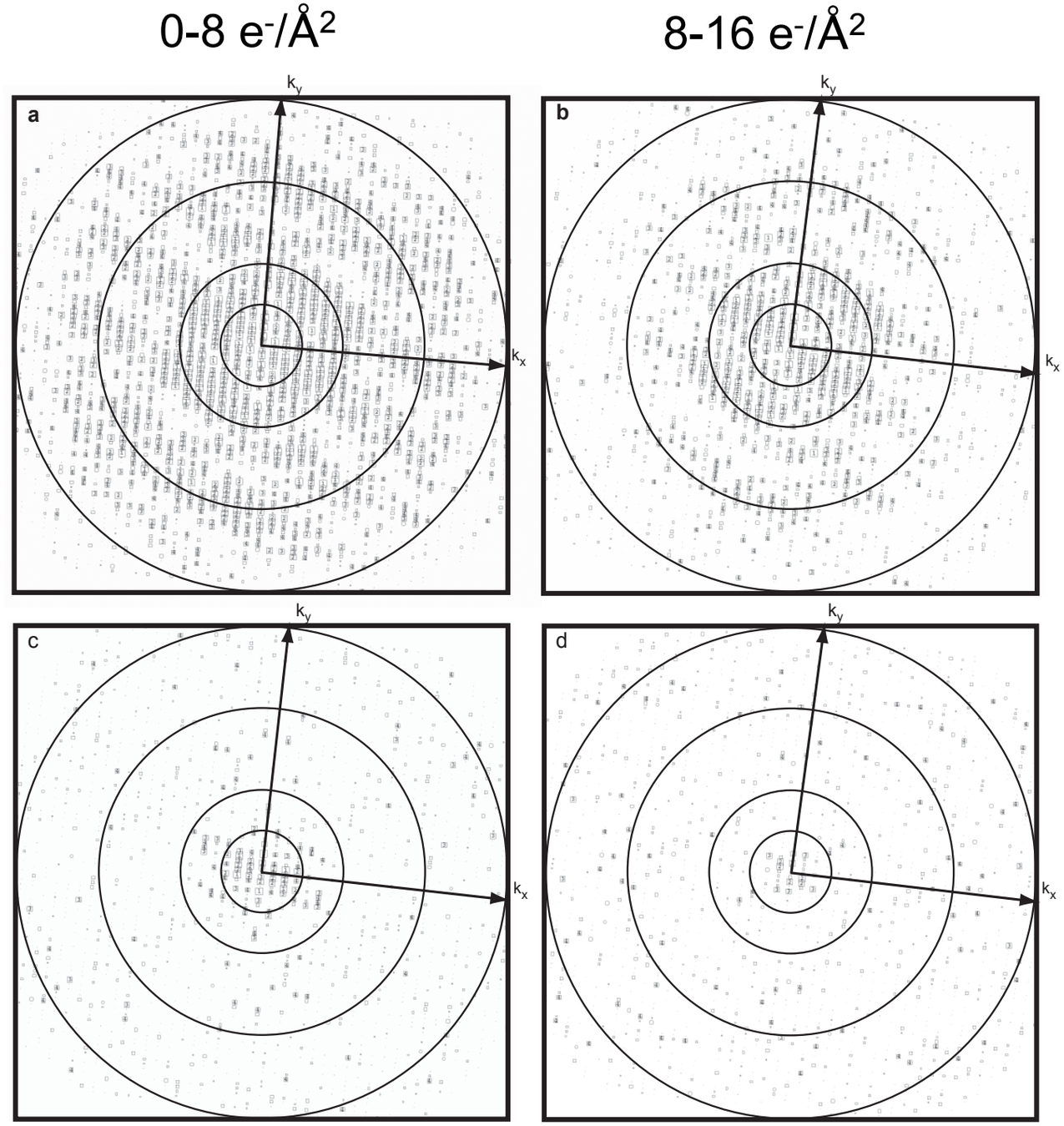
Approach	Frames/ Particles/ Regions	Correlation	Smoothing	Advantages/Disadvantage
Motioncorr (least squares)	Frames	Noisy images to noisy images	Over-determined problem (least squares fit)	Over-determined/low signal-to-noise in comparisons
Unblur (iterative sum & align)	Frames	Noisy images to sums of noisy images	Trajectory fitted to spline	Robust/whole frame only
Alignframes_lmbfgs (gradient-based)	Frames	Noisy images to sums of noisy images	Penalize changes in trajectory (second order smoothing)	Relatively fast and robust/ whole frame only
Polishing in Relion (projection matching)	Particles	Noisy images to high SNR map projections	Linear fit, rolling averages, enforce local correlation	Map projection v. high SNR/ map projection may not match image
Alignparts_lmbfgs (gradient-based)	Particles	Noisy particle images to sums of noisy images	Penalize changes in trajectory, enforce local correlation	Non-linear trajectories/Map projections have higher SNR
MotionCor2 (model-based)	Regions	Noisy patch images to sums of noisy images	Polynomial model	Fast and efficient/May be trying to align empty patches

Exposure weighting

Optimizing signal-to-noise ratios in images

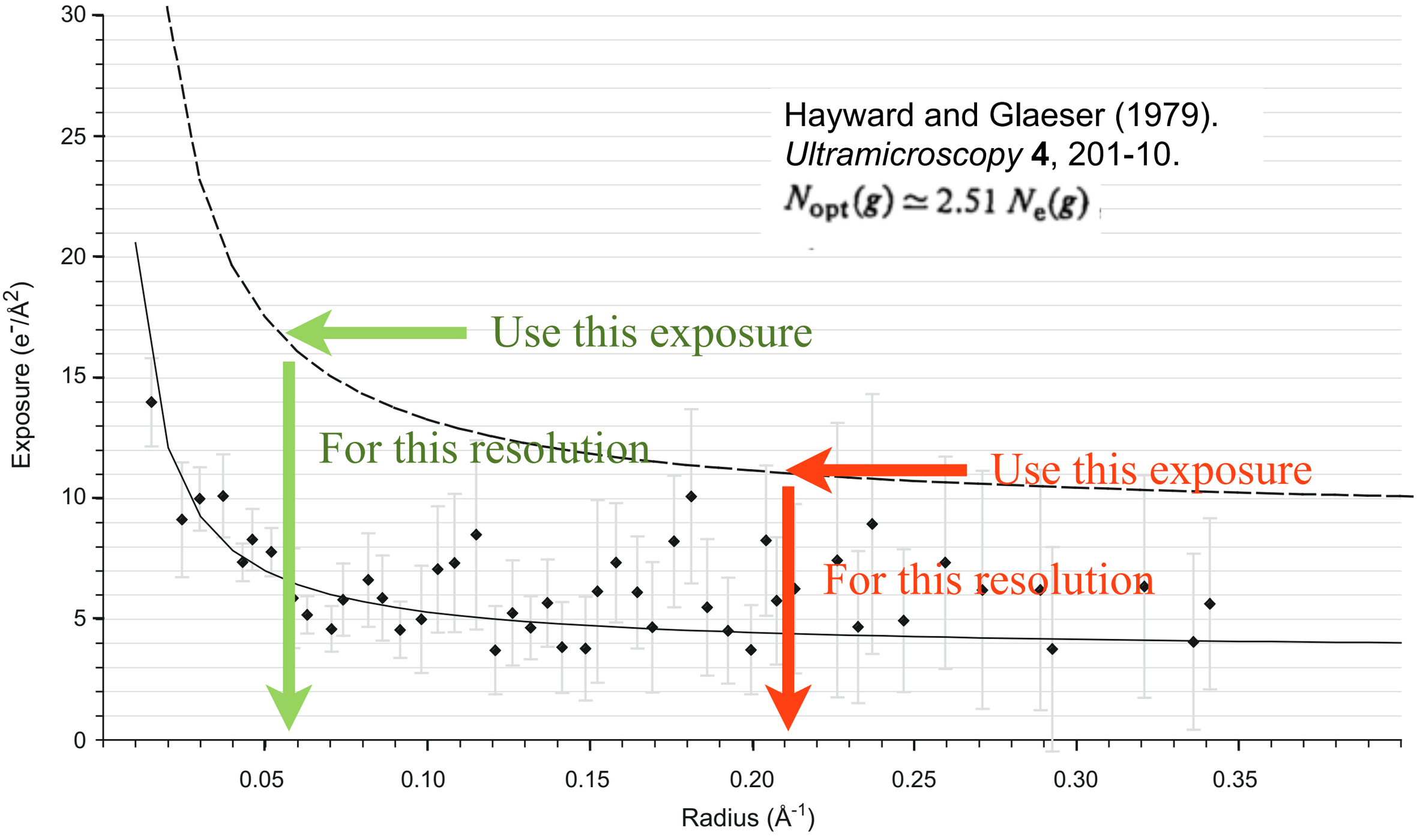


Hayward and Glaeser (1979).
Ultramicroscopy **4**, 201-10.



Baker, Smith, Bueler and Rubinstein (2010).
J Struct Biol **169**, 431-7.

Optimal weighting for radiation damage

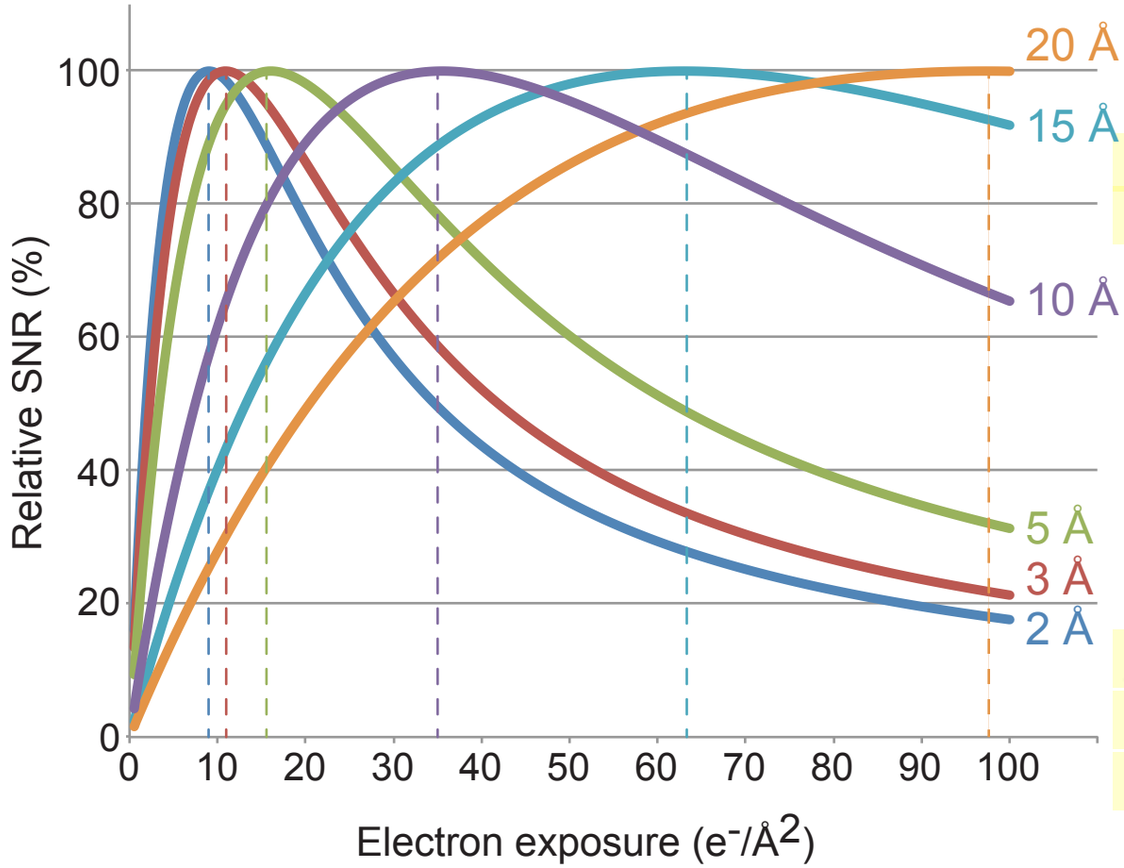


Hayward and Glaeser (1979).
Ultramicroscopy **4**, 201-10.

$$N_{opt}(g) \simeq 2.51\, N_e(g)$$

Baker, Smith, Bueler, and Rubinstein (2010), *J. Struct. Biol.*, **169**, 431-7.
Baker and Rubinstein (2010), *Method Enzymol* **481**, 373-90.

Physics-based exposure weighting



Baker, Smith, Bueler and Rubinstein (2010). *J Struct Biol* **169**, 431-7.

Physics based exposure weighting

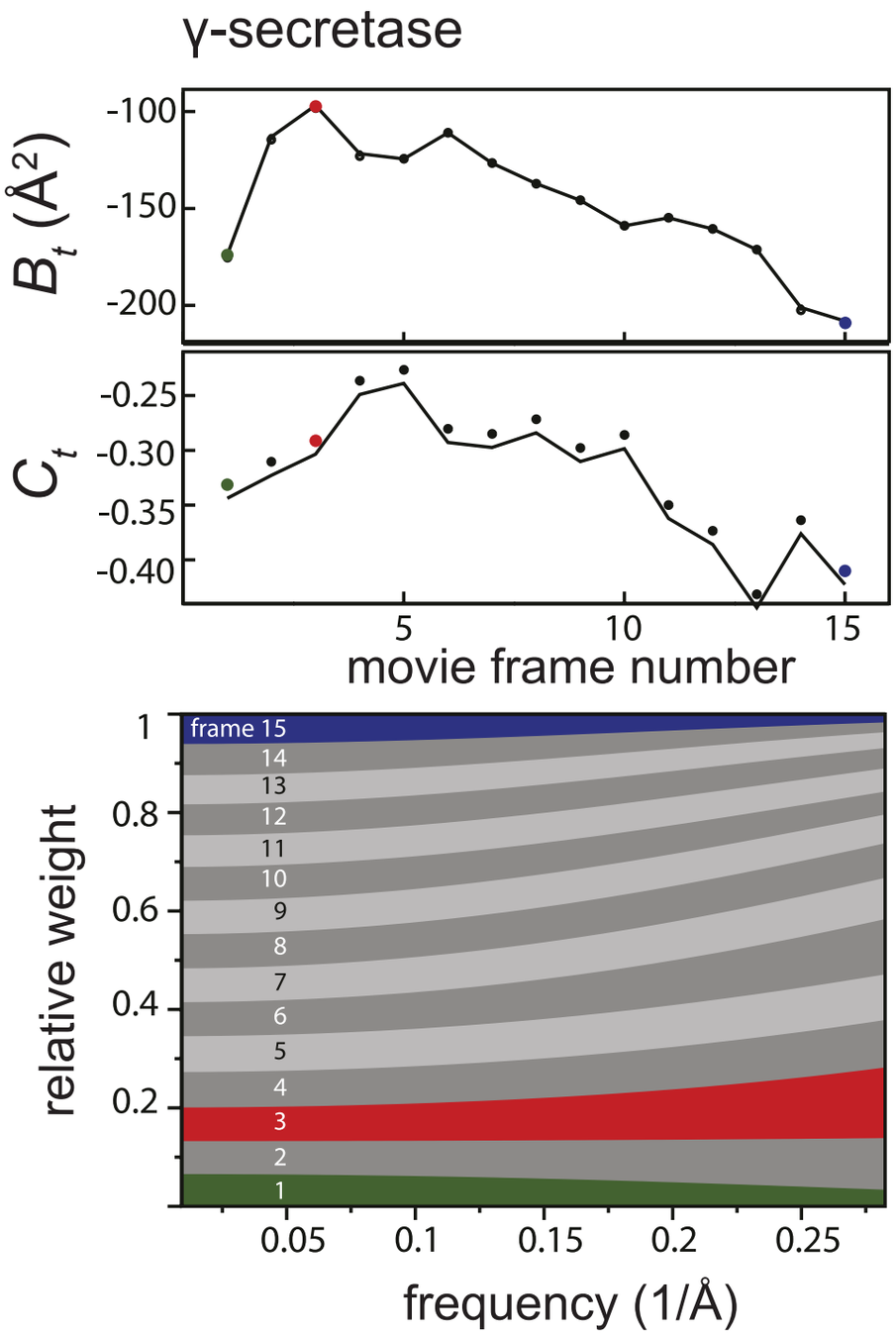
These results also have an additional application. New image detectors offer the possibility of capturing many frames during the imaging process (Faruqi and Henderson, 2007), making each image an exposure series. With these detectors, the first few frames would record high spatial frequencies from the specimen with maximal SNRs. In comparison, an average of many frames would have low SNRs for the high-resolution information but would record low spatial frequencies with high SNRs. This average of many frames would allow for accurate alignment of particle images while the first few frames could be used to build high-resolution models. An even more sophisticated approach would be to use the optimal exposures measured here to calculate weighted averages of frames in order to maximize the SNR at each spatial frequency.

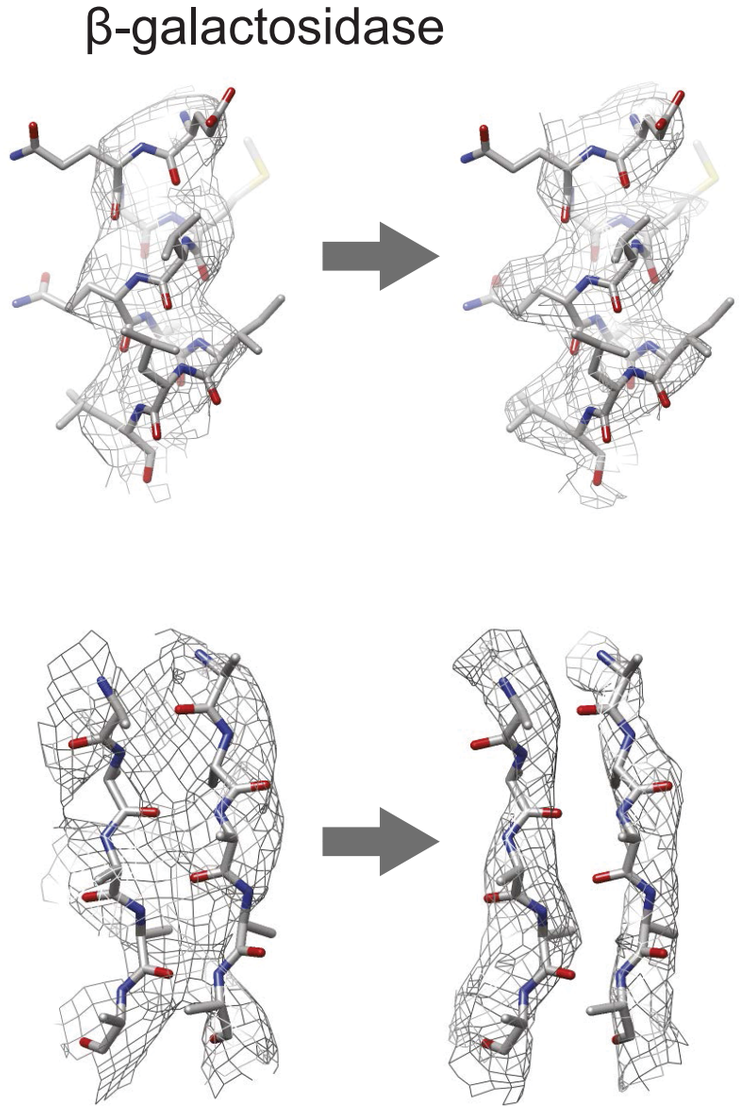
Better exposure curves measured by: Grant & Grigorieff (2015) *eLife* 4:e06980

Used in: *Alignparts_lmbfgs* (Rubinstein & Brubaker, 2015, *JSB* 192, 188-95)
Unblur (Grant & Grigorieff, 2015, *eLife* 4:e06980)
MotionCor2 (Zheng…Agard, 2017, *Nat Meth* 14, 331-2)

20S Proteasome

Data-driven exposure weighting





Scheres (2014). *eLife 3:e03665*.

Anisotropic magnification

magnification(x) ≠ magnification(y)

Anisotropic magnification detected numerous microscopes at low magnification

Affects small pixel detectors (Gatan K2/K3, DirectElectron)

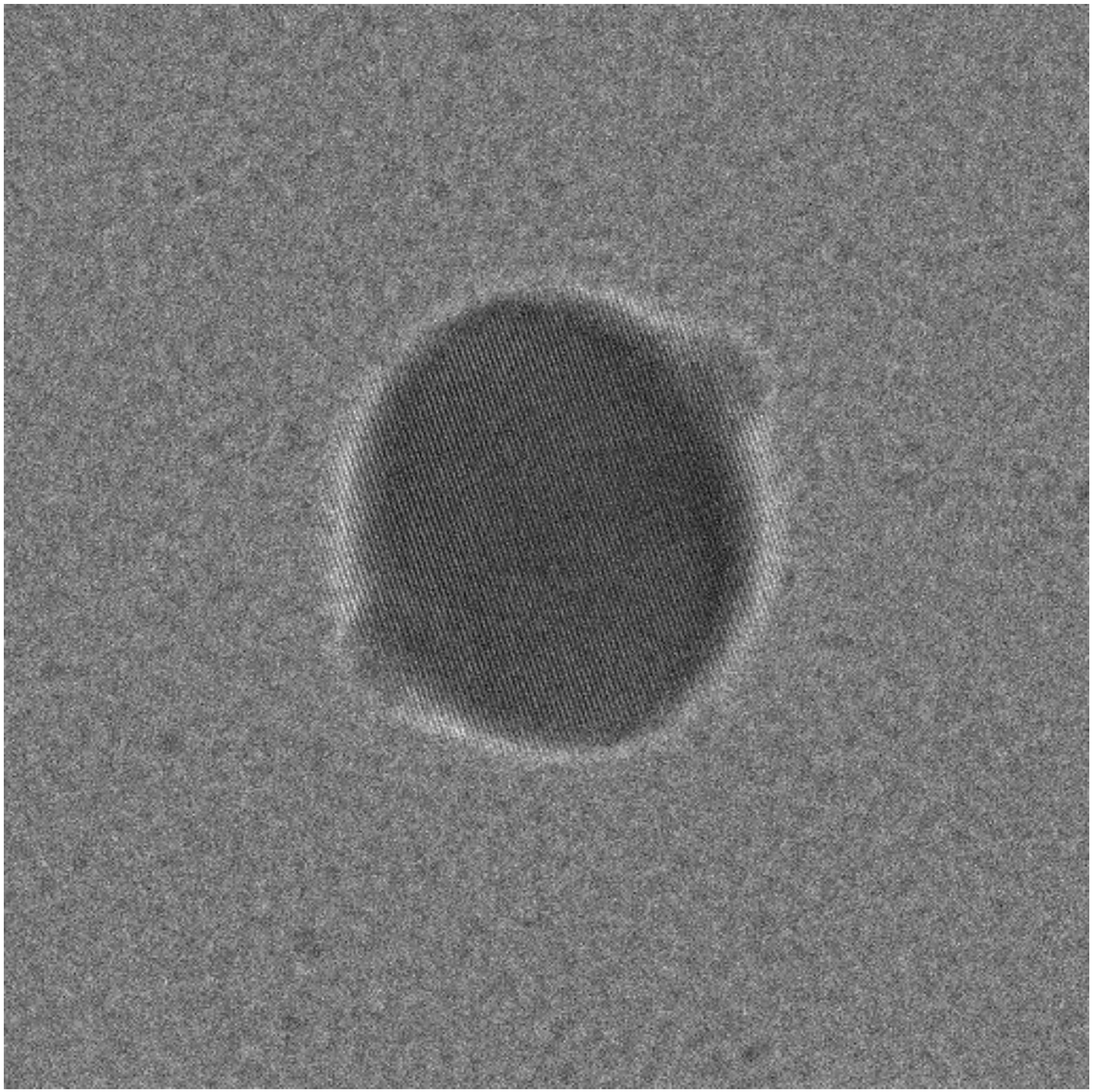
Several papers describe (e.g.)

Baldwin, J., Henderson, R., 1984. Ultramicroscopy 14 (4), 319–335.

Grant, T., Grigorieff, N (2015) J Struct Biol (2015) 192(2):204-8.

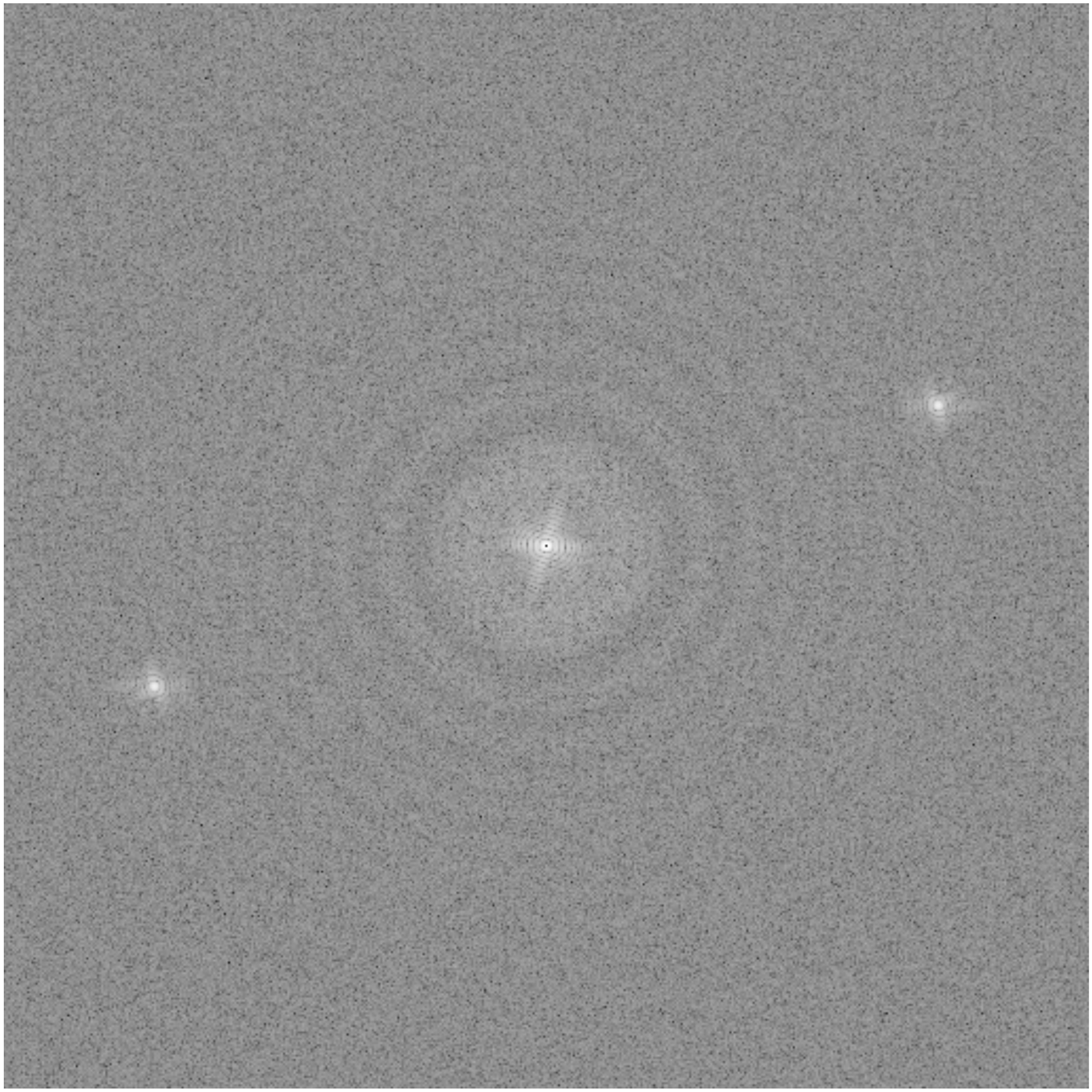
Zhao, J., Brubaker, M. A., Benlekbir, S., Rubinstein, J.L. (2015). J Struc Biol 192(2):209-15

Thallous chloride crystal - 25 kx magnification setting

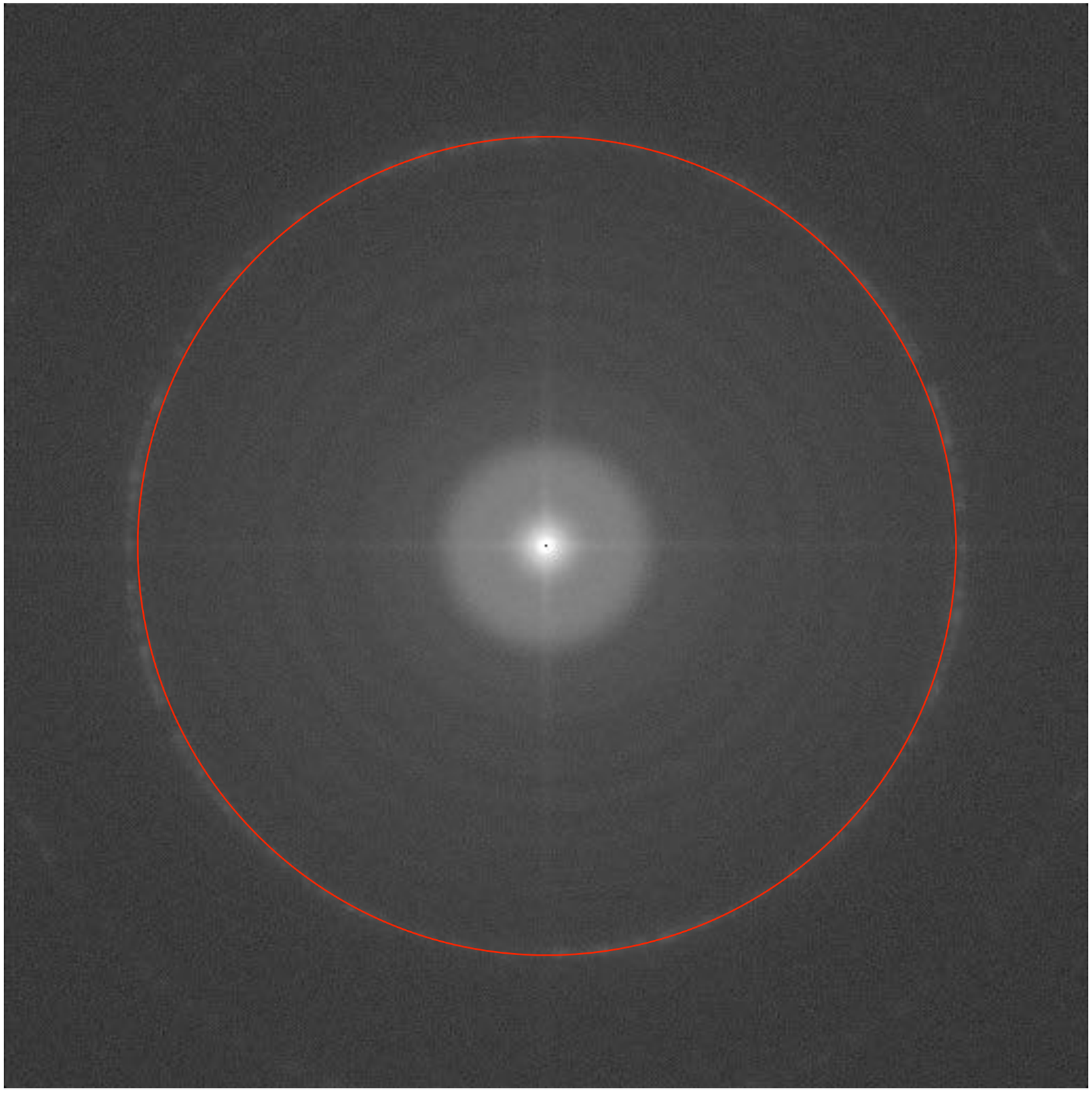


d=3.842 Å

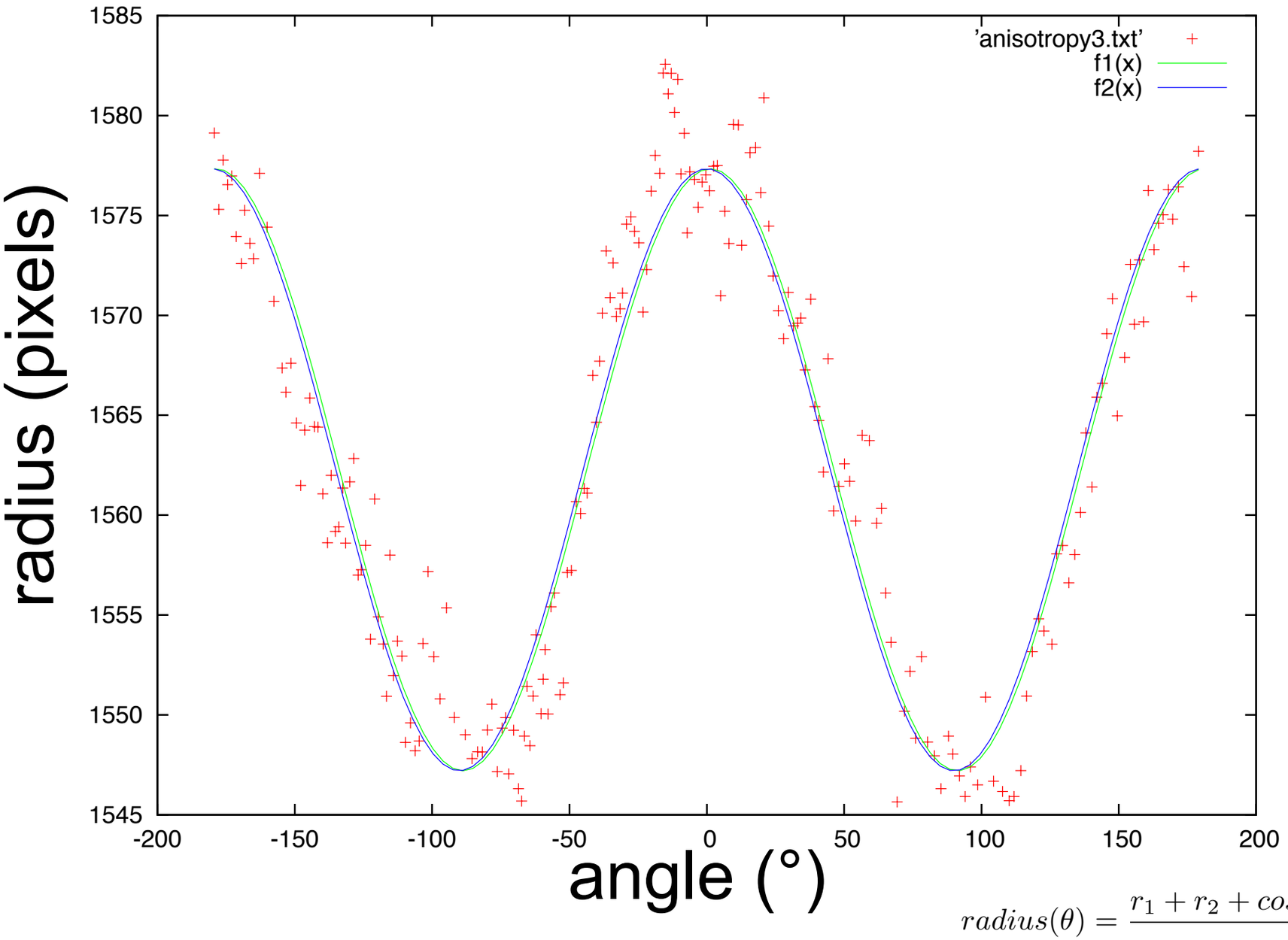
FT of thallous chloride crystal image



Average of many thallous chloride FTs



Fitting of measured radius to an ellipse



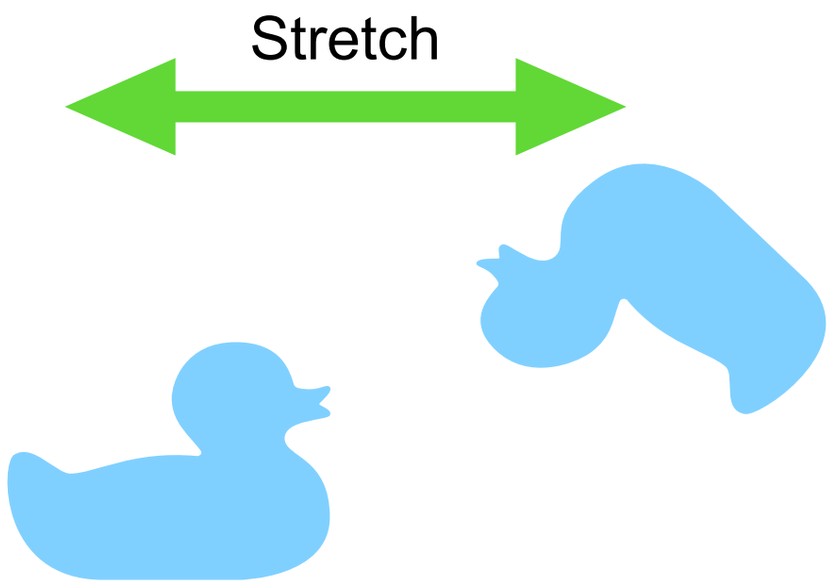
r_1=1577 pixels

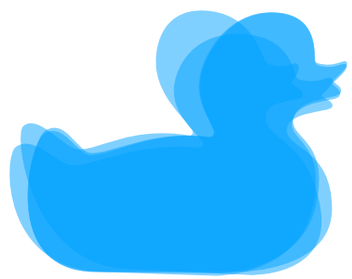
r_2=1547 pixels

θ_{off}=1.3°

$$radius(\theta) = \frac{r_1 + r_2 + cos\left(2 \cdot (\theta - \theta_{off})\right)(r_1 - r_2)}{2}$$

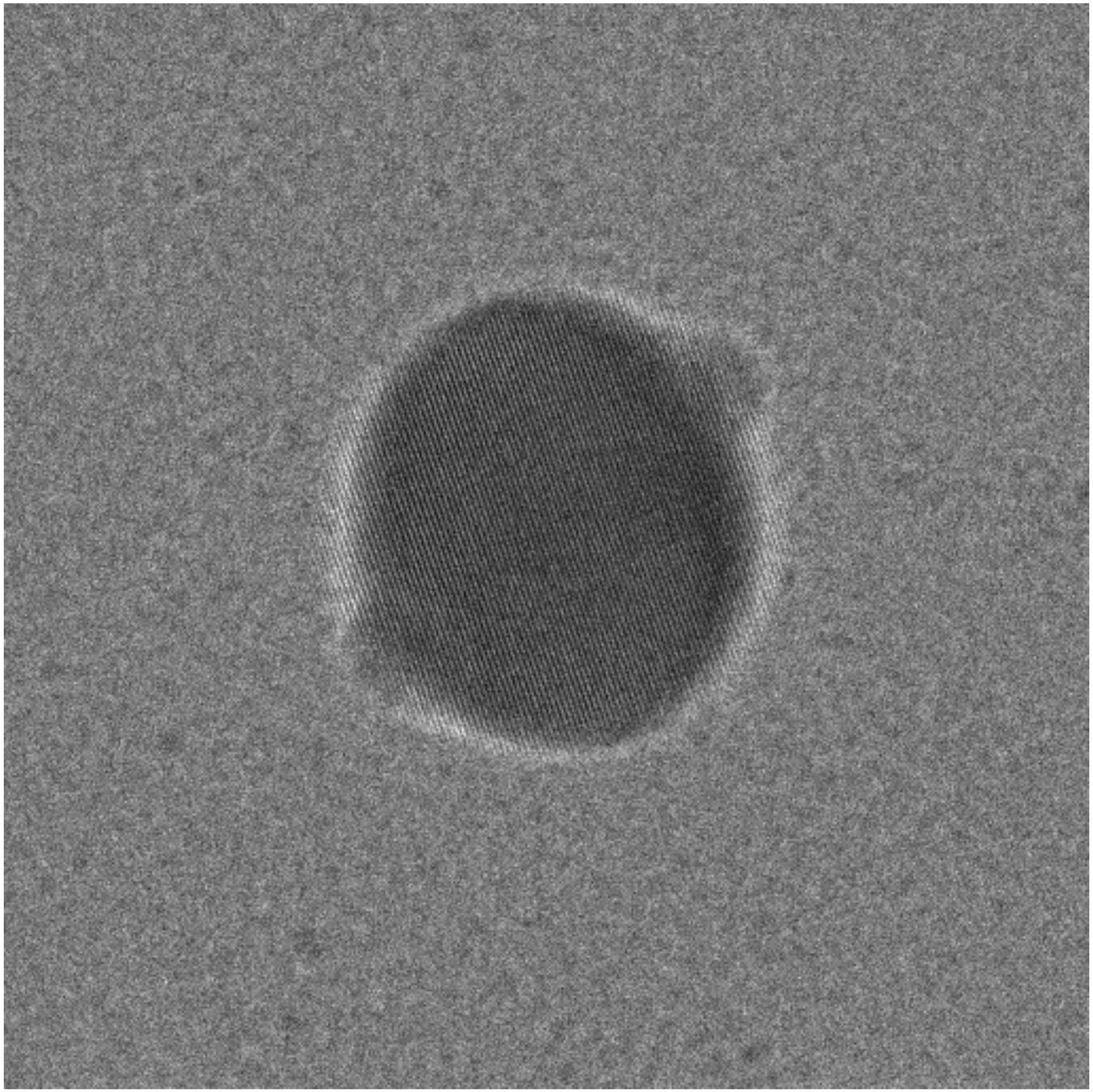
Anisotropic magnification degrades resolution



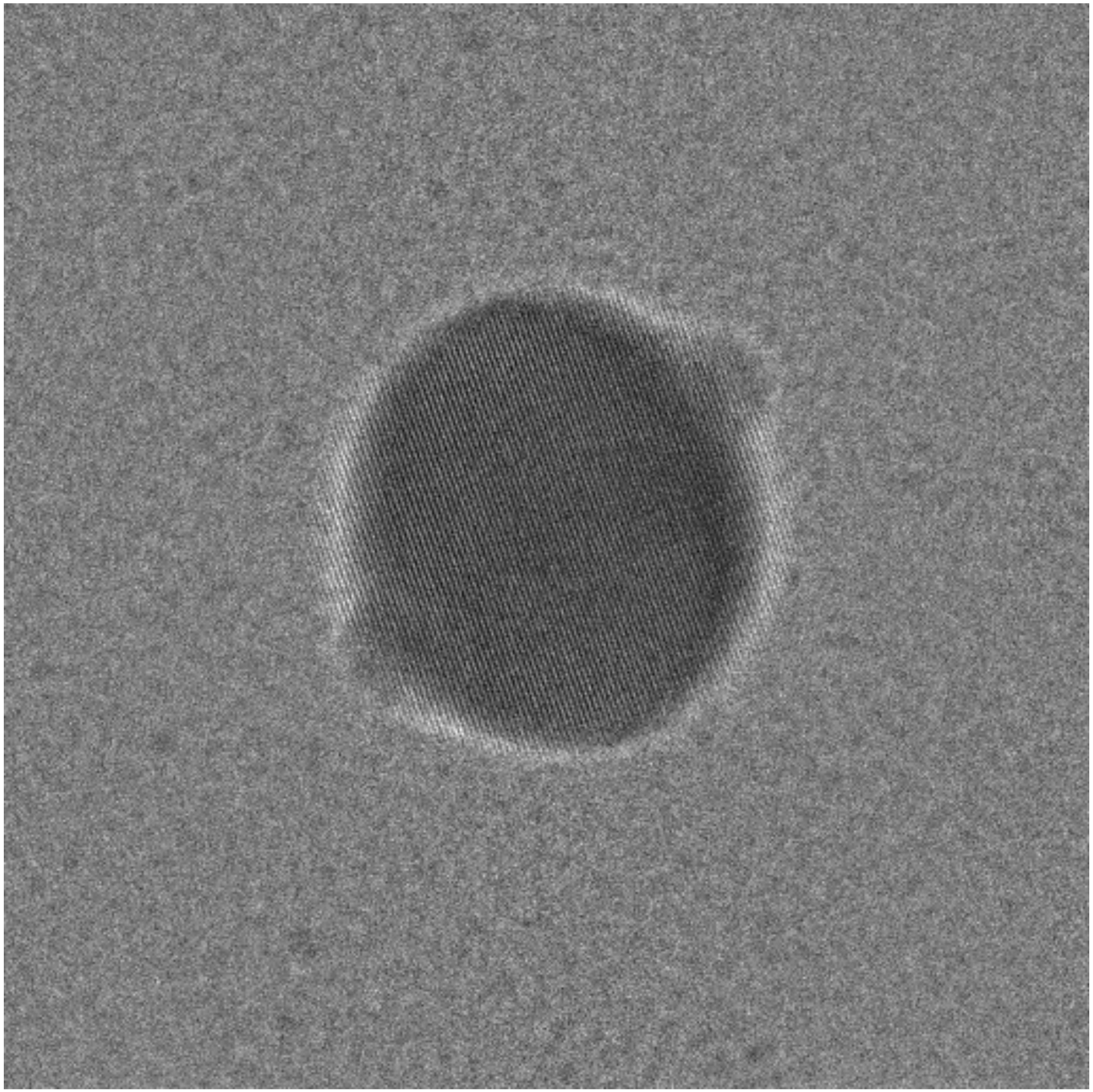


Distorted particles are
no longer identical

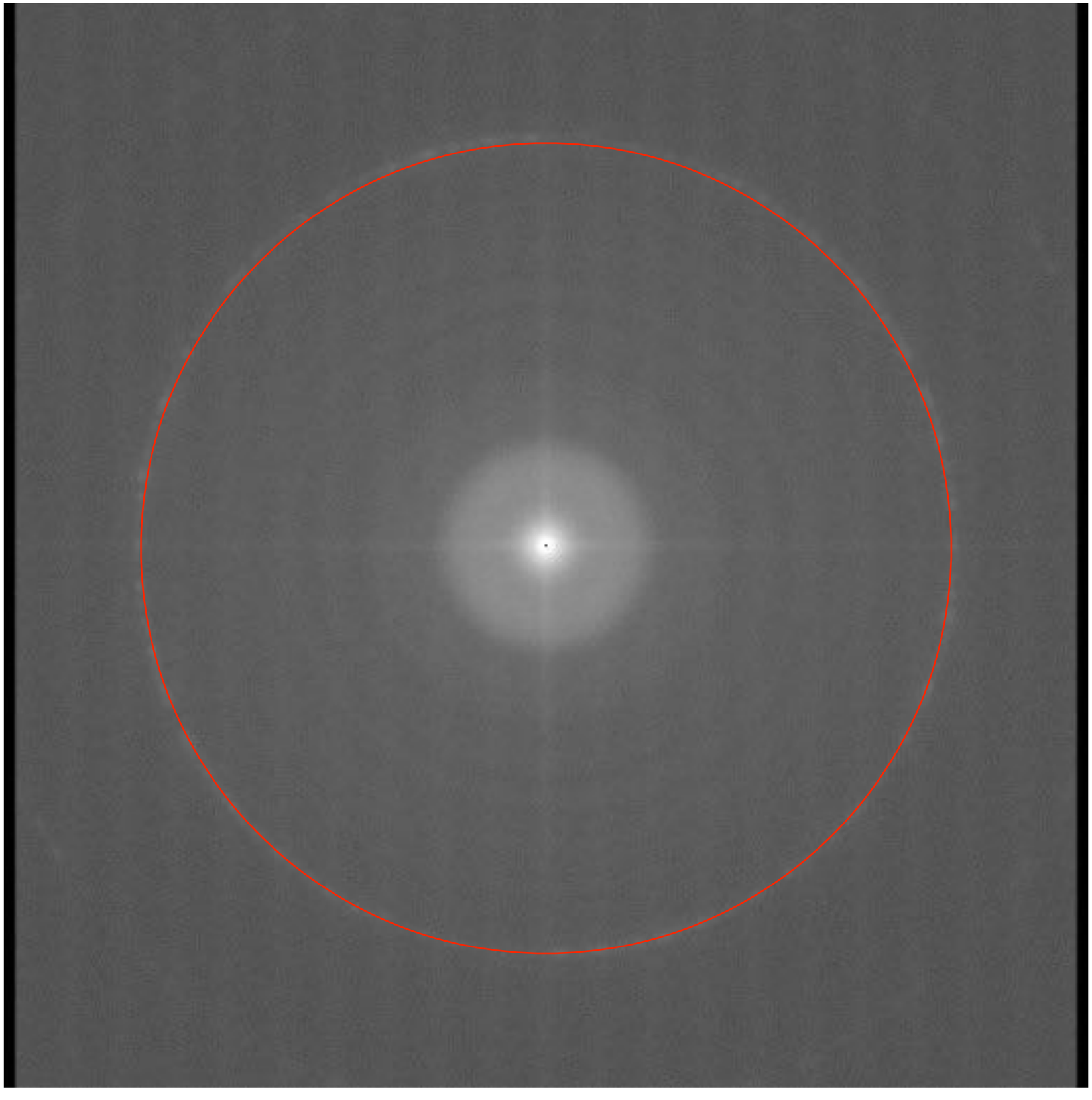
Thallous chloride crystal



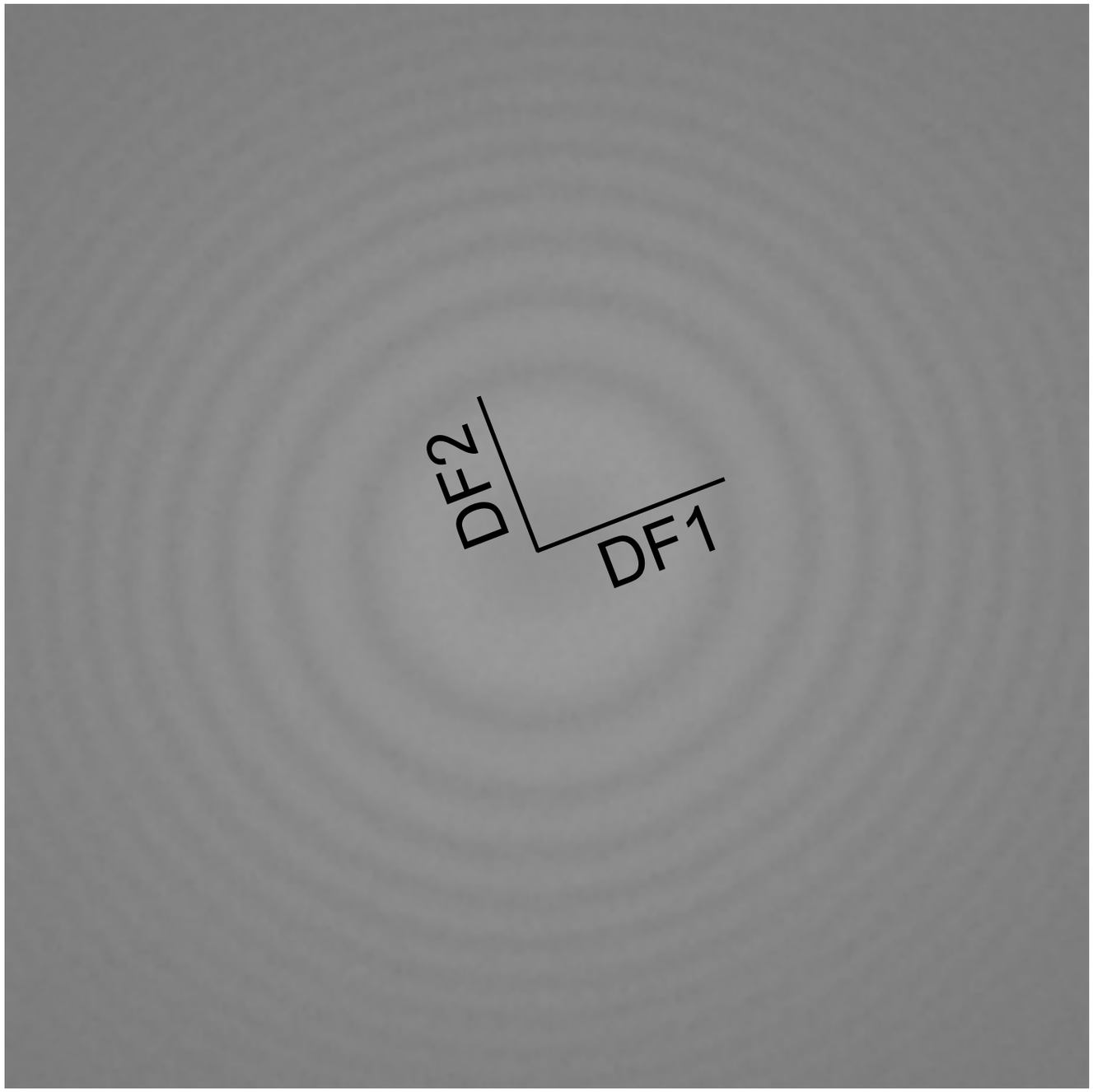
Corrected thallous chloride crystal



Average of many corrected thallous chloride FTs

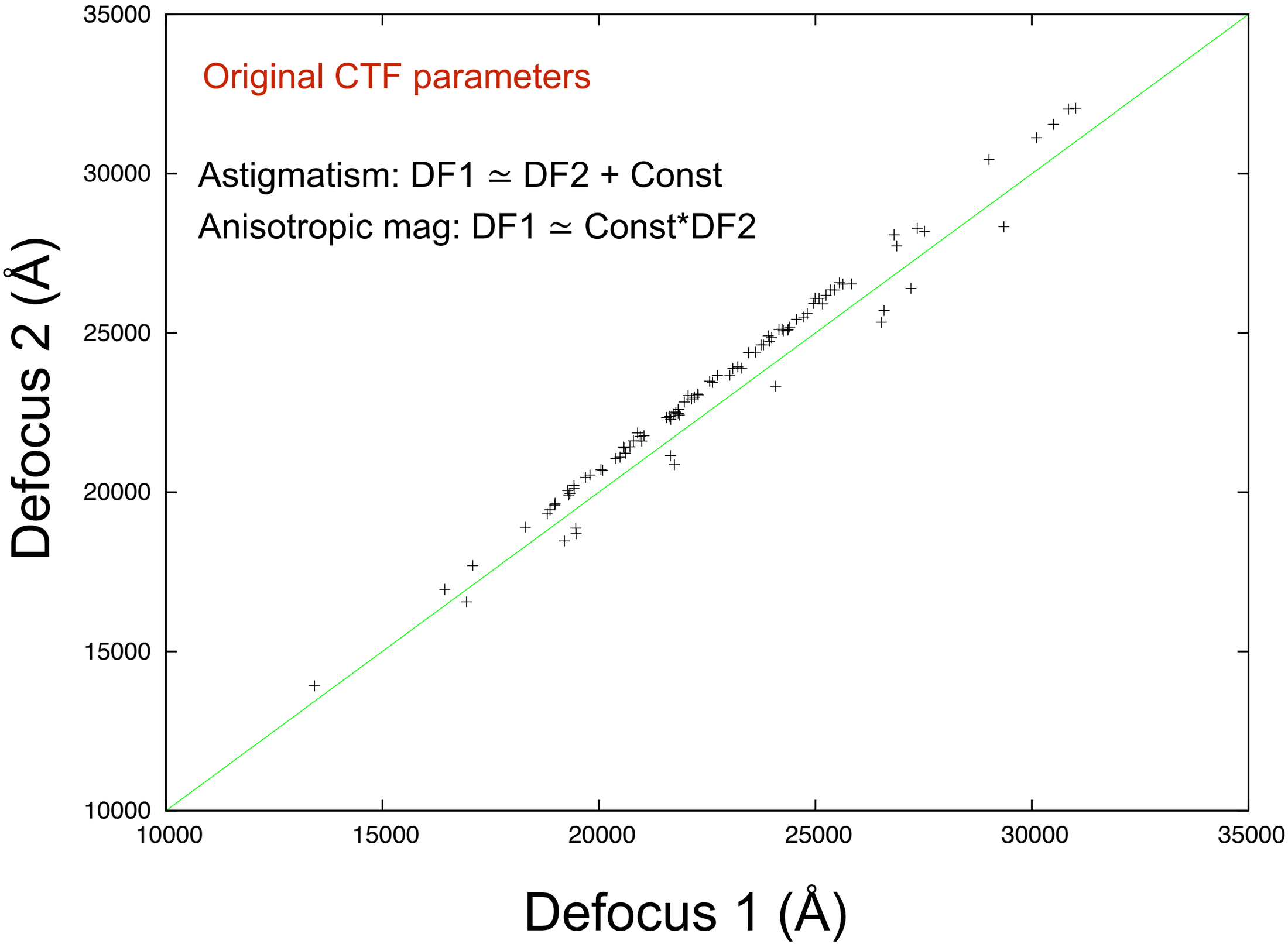


Anisotropic magnification affects CTF estimation

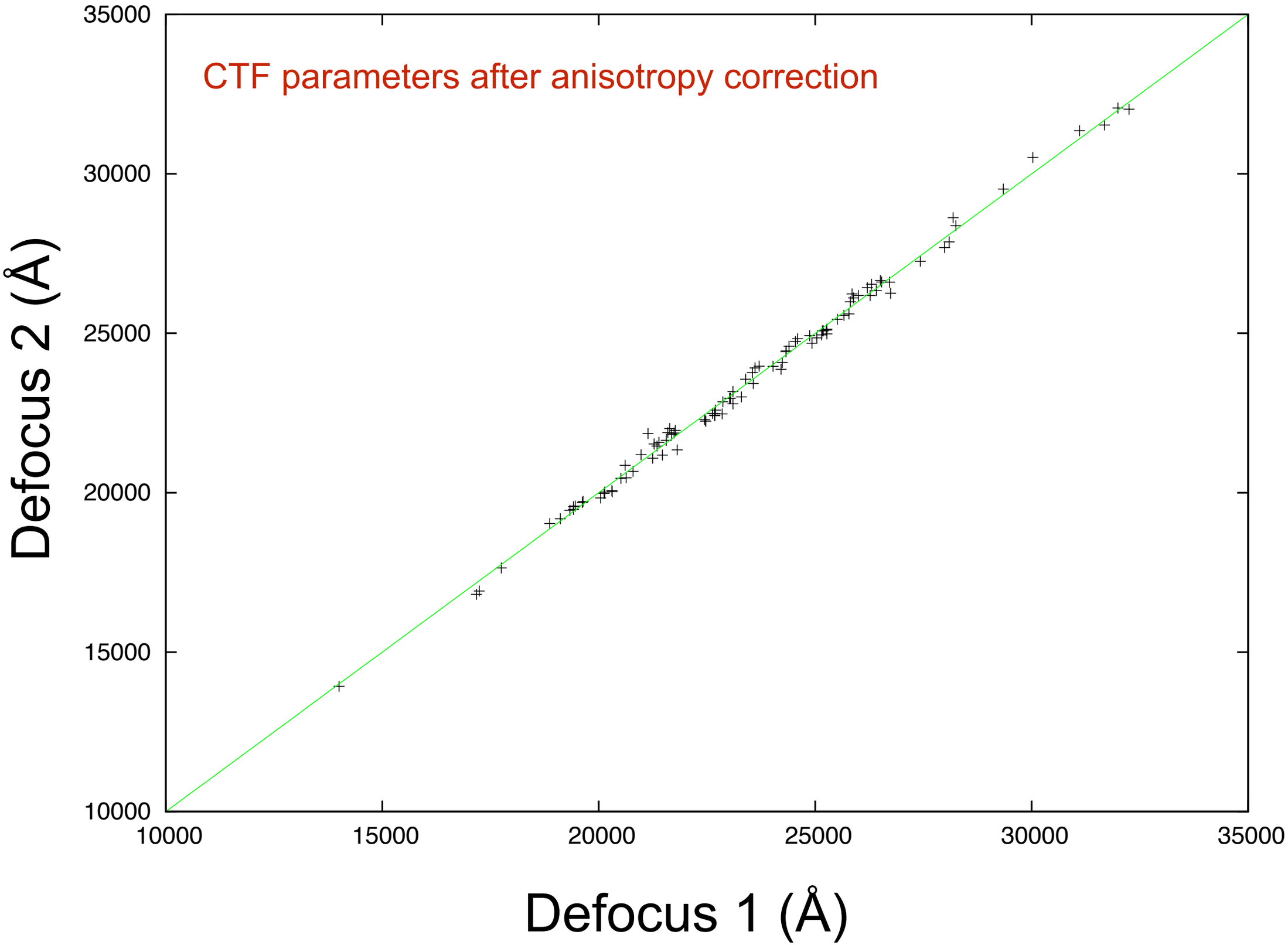


- Anisotropic magnification appear different (worse) at low magnification
- Will look like objective lens astigmatism in power spectra

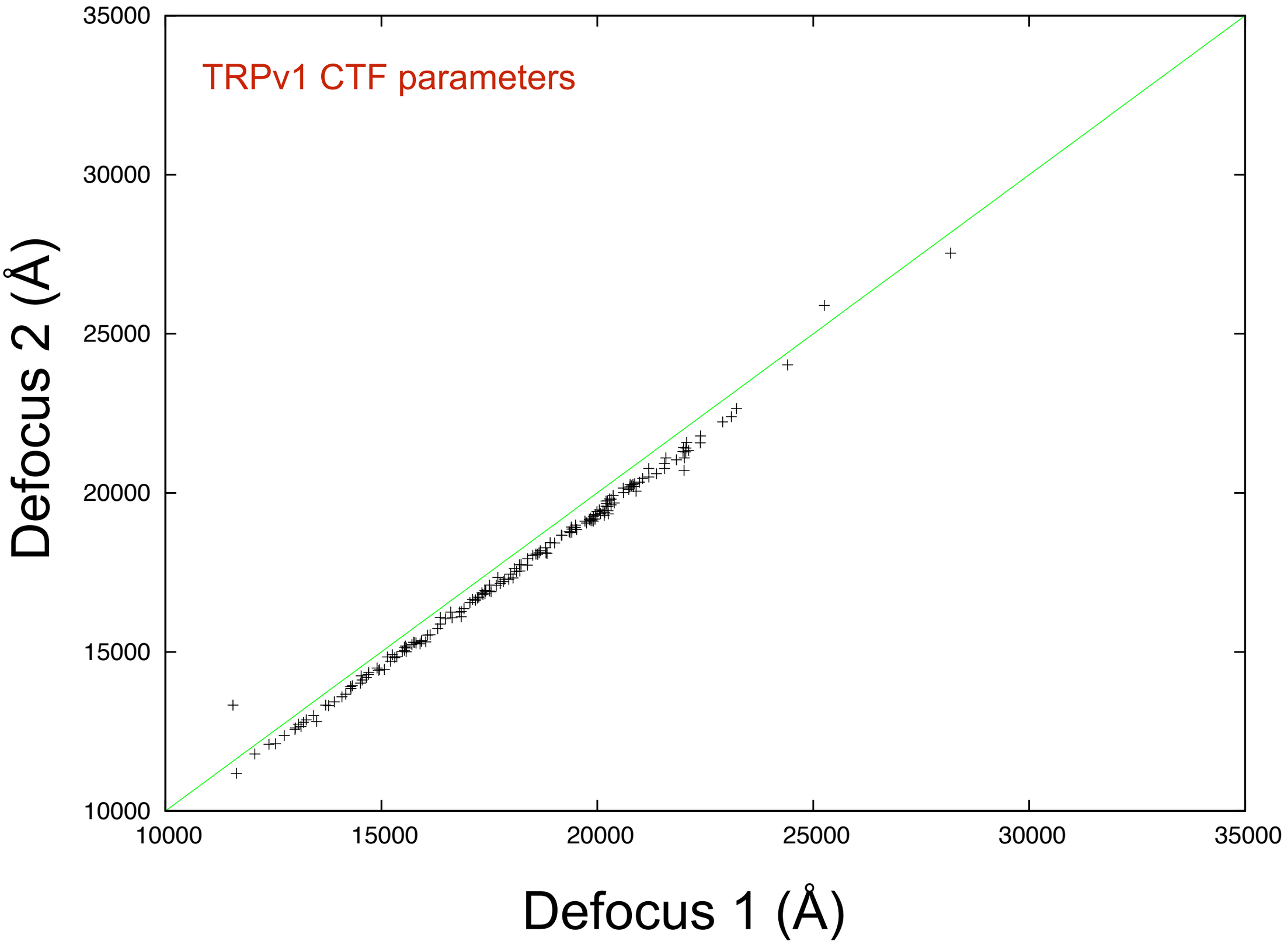
Easy way to check for anisotropic magnification (Jianhua Zhao)



Easy way to check for anisotropic magnification (Jianhua Zhao)



Is the problem widespread? (Yifan Cheng/Jianhua Zhao)



Decisions you need to make with a DDD

Decision	Relevant concepts	Conclusion
Pixel size/magnification (Å/pixel)	DQE, Nyquist Limit, Aliasing, Anisotropic magnification	Resolution of interest must be within Nyquist limit and have good DQE
Exposure rate (electrons/ pixel/second)	Coincidence loss	Minimize coincidence loss while keeping exposure time reasonable
Frame rate (frames/sec)	Alignability of frames, movement within frames, radiation damage per frame	Ensure enough signal to align frames but exposure weight properly
Movie length (seconds)	Total signal at different resolutions	Ensure enough signal to align and classify particles

Acknowledgements:

Current members:

Yazan Abbas
Samir Benlekbir
Stephanie Bueler
Hui Guo
Zev Ripstein
Thamiya Vasanthakumar

Past members:

Jianhua Zhao
Lindsay Baker

Collaborators

Hui (Alex) Wei (NRAMM)
Bridget Carragher (NRAMM)
Clint Potter (NRAMM)
Marcus Brubaker (York U)