Challenges in Cryo-electron Tomography

D. N. Mastronarde

Dept. of MCD Biology, University of Colorado, Boulder, CO

NIH
National Institutes of Health

NIGMS

Challenges, Goals, and Needs

- Challenges fall into two categories: acquisition and processing
- Overall goal: Better, faster, and more automatically
- The area of acquisition has two main currents:
  - Methods used to obtain the highest resolution subvolume averages
  - A push for more speed, particularly with novel acquisition methods that are likely to degrade resolution
- Need more clarity on the domain of applicability for rapid acquisition methods
  - Will they just be good for screening or also useful in resolution range that most users hope to achieve?

Challenges in Fast Acquisition

- How well can rapid tilting methods do with current technology?
- What technological improvements are needed to do better?
- Can they be achieved in a robust, commercially available and supportable way?
- To look at specific nearer-term questions, compare popular protocol for getting high-resolution data with current capabilities for rapid tilting

Some Key Ingredients in the Hagen Scheme

- Acquire in dose-symmetric way, with lowest accumulated doses near 0°, by taking images at 0, +2°, -2°, +4°, -4°, etc
- Backlash the tilt changes on one side of zero
  - Essential for returning close to same position
  - Also helps with discontinuity in alignment between two directions: avoids rotation of mechanical tilt axis
- Wait for drift to be under 1 pixel per frame
  - Maintains quality of motion-corrected data
  - Measuring drift is more efficient than waiting a long fixed time
- Keep focus at a consistent value for good CTF correction
  - Autofocus is iterated if necessary until change is below a threshold
  - This may be of lesser importance if defocus can later be determined accurately in individual images

Features in SerialEM to Support Continuous Tilting

- A script command, *BackgroundTilt*, can start a slow continuous tilt and allow the script to go on to acquire and save images
  - Ideal for rapid tilt series from plastic sections with a OneView camera doing drift correction in continuous acquisition mode (Martin Schorb at EMBL)

Features in SerialEM to Support Continuous Tilting

- A script command, *BackgroundTilt*, can start a slow continuous tilt and allow the script to go on to acquire and save images
  - Ideal for rapid tilt series from plastic sections with a OneView camera doing drift correction in continuous acquisition mode (Martin Schorb at EMBL)
- SerialEM has a thread that can do a variety of actions during camera acquisition
  - Tilt or move stage or change image shift after exposure, for extra settling time
  - Continuous changes like dynamic focusing for STEM
- *TiltDuringRecord* command makes that thread start a slow continuous tilt during a long camera acquisition that saves many frames
  - But this is all the procedure can do, and it blocks other script actions

Features in SerialEM to Support Fast Incremental Tilting

- To support rapid incremental tilt series, the thread can now perform a stored sequence of actions including:

  image shift
  focus change
  tilt
  wait for an extra delay
  unblank – wait for exposure time – blank

  - Each action is optional on any given step, including the unblank sequence if tilting twice for backlash, or the tilt in case a continuous tilt was started

Features in SerialEM to Support Fast Incremental Tilting

- To support rapid incremental tilt series, the thread can now perform a stored sequence of actions including:
  image shift
  focus change
  tilt
  wait for an extra delay
  unblank – wait for exposure time – blank
  - Each action is optional on any given step, including the unblank sequence if tilting twice for backlash, or the tilt in case a continuous tilt was started
- *QueueFrameTiltSeries* sets up a sequence with fixed parameters to happen during the next Record
- *QueueFrameSeriesFromVar* sets up a sequence with an array variable specifying separate parameters for each step
  - The variable can be read from a file
  - The sequence could come from a conventional tilt series where X/Y/Z positions are available after the series
  - It could come from analysis of a frame tilt series

How Well Can a Sequence of Changes Work?

- The sequence of image shift and focus changes represents a return to the precalibration approach
  - Can it be used with much smaller fields of view?
- Eucentric height has to be set well for a sequence from one series to apply in another
  - What is the fastest way to find eucentricity with adequate accuracy?
    - CTF fitting in Focus area on opposite sides after setting eucentric focus is worth trying

How Well Can a Sequence of Changes Work?

- The sequence of image shift and focus changes represents a return to the precalibration approach
  - Can it be used with much smaller fields of view?
- Eucentric height has to be set well for a sequence from one series to apply in another
  - What is the fastest way to find eucentricity with adequate accuracy?
    - CTF fitting in Focus area on opposite sides after setting eucentric focus is worth trying
- How well will a sequence in one area work in other areas, particularly different positions lateral to tilt axis?
  - Maybe all that is needed is a lateral tilt axis offset that can be characterized and compensated
  - Or maybe storing calibrations for different lateral positions will work over a longer term
- Would the sequence(s) need to be prepared per session, or will stored calibration(s) work for some time?

Can the Key Aspects of the Hagen Scheme Be Supplied in Fast Tilting?

- Backlash on half the tilts is possible, it just costs some time
- Waiting for drift to settle is possible in the programmed sequence, but is the most problematic aspect
  - Is settling time stereotypical or variable from one area to the next?
  - Does it vary with tilt angle?
  - Is it so large that it becomes the large majority of acquisition time?

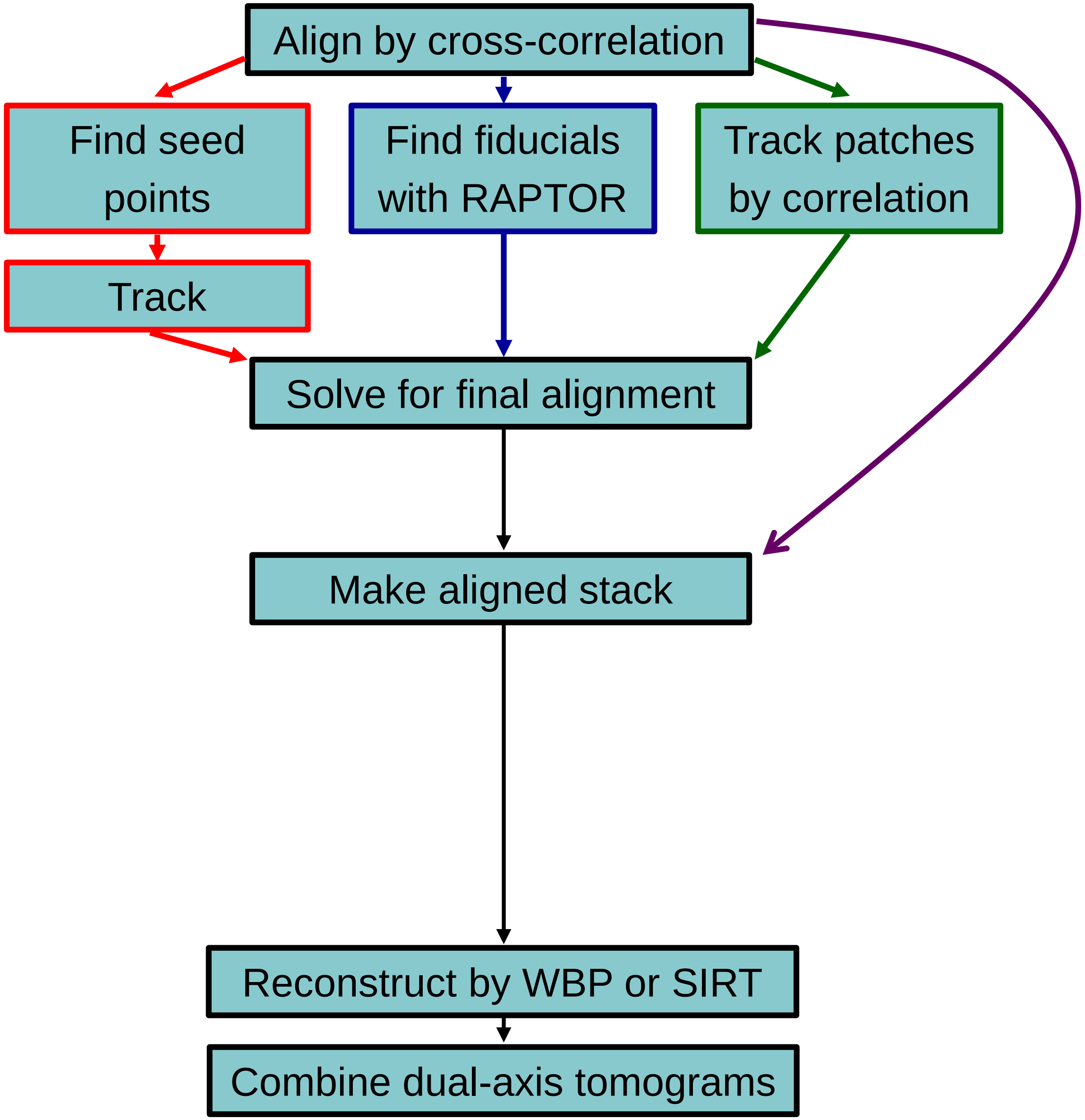
Can the Key Aspects of the Hagen Scheme Be Supplied in Fast Tilting?

- Backlash on half the tilts is possible, it just costs some time
- Waiting for drift to settle is possible in the programmed sequence, but is the most problematic aspect
  - Is settling time stereotypical or variable from one area to the next?
  - Does it vary with tilt angle?
  - Is it so large that it becomes the large majority of acquisition time?
- Focus is unlikely to be as consistently accurate as with iterated autofocusing, even with the programmed sequence
  - Focus changes from incremental tilting at short intervals may be more smoothly varying than in conventional series, allowing adjacent views to be combined in CTF determination when necessary
  - Might need to use bidirectional series from ~20° to get most consistent focus from fast tilting
  - Or it might not matter! Electron counting images generally give good single-image defocus estimates through whole tilt range

Tilt Series Processing

- Automated processing capabilities already exist in various forms
  - Appion/Protomo
  - IMOD
- IMOD batch processing is designed to handle every step of the processing sequence and provide control over numerous options

Automated Tomogram Generation by Batchruntomo



Align by cross-correlation

Find seed points

Find fiducials with RAPTOR

Track patches by correlation

Track

Solve for final alignment

Make aligned stack

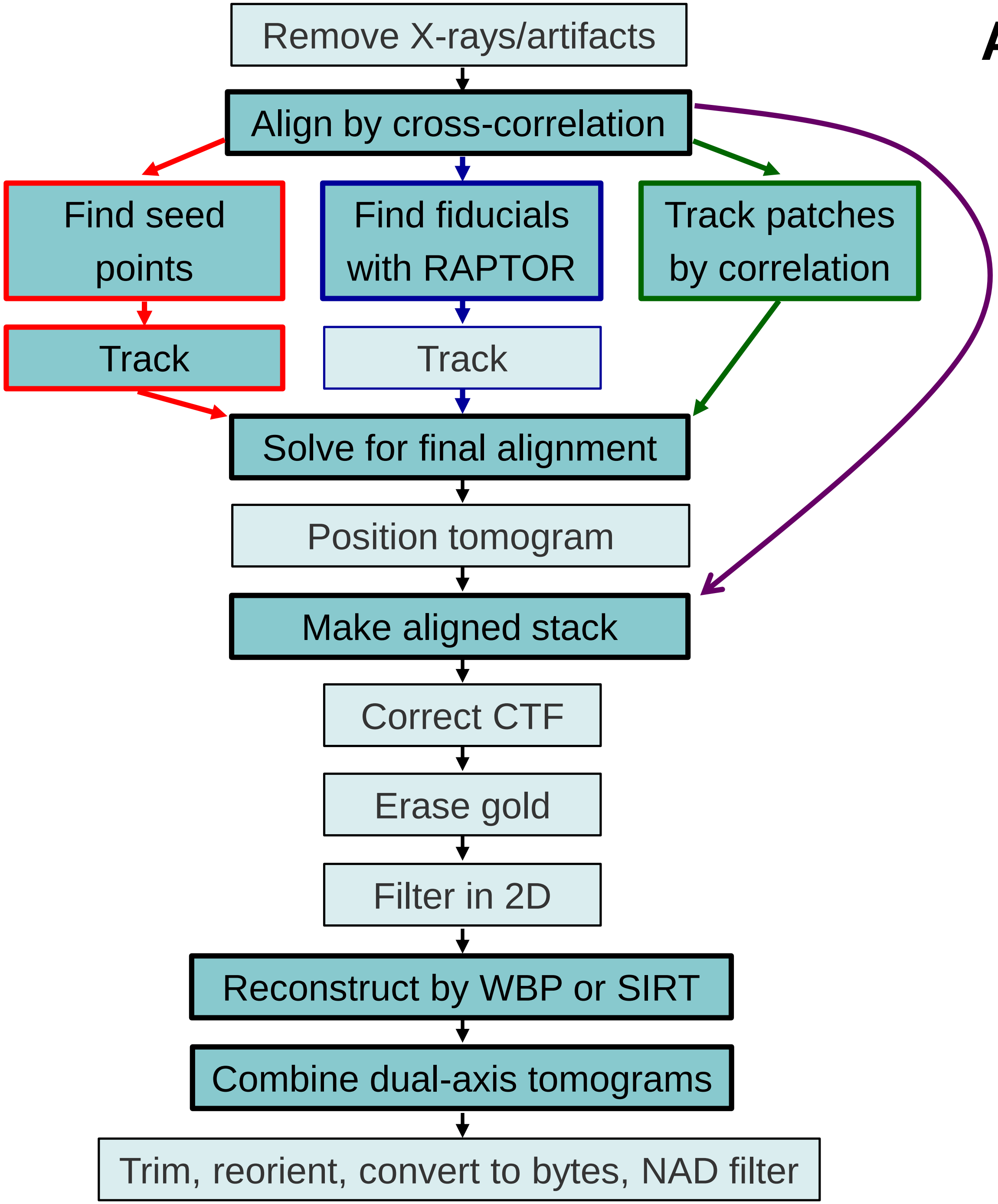
Reconstruct by WBP or SIRT

Combine dual-axis tomograms

Alternative pathways:

- — Autoseed and track
- — RAPTOR
- — Patch tracking
- — Cross-correlation only

Automated Tomogram Generation by Batchruntomo


Remove X-rays/artifacts

Align by cross-correlation

Find seed points

Find fiducials with RAPTOR

Track patches by correlation

Track

Track

Solve for final alignment

Position tomogram

Make aligned stack

Correct CTF

Erase gold

Filter in 2D

Reconstruct by WBP or SIRT

Combine dual-axis tomograms

Trim, reorient, convert to bytes, NAD filter

Alternative pathways:

- — Autoseed and track
- — RAPTOR
- — Patch tracking
- — Cross-correlation only

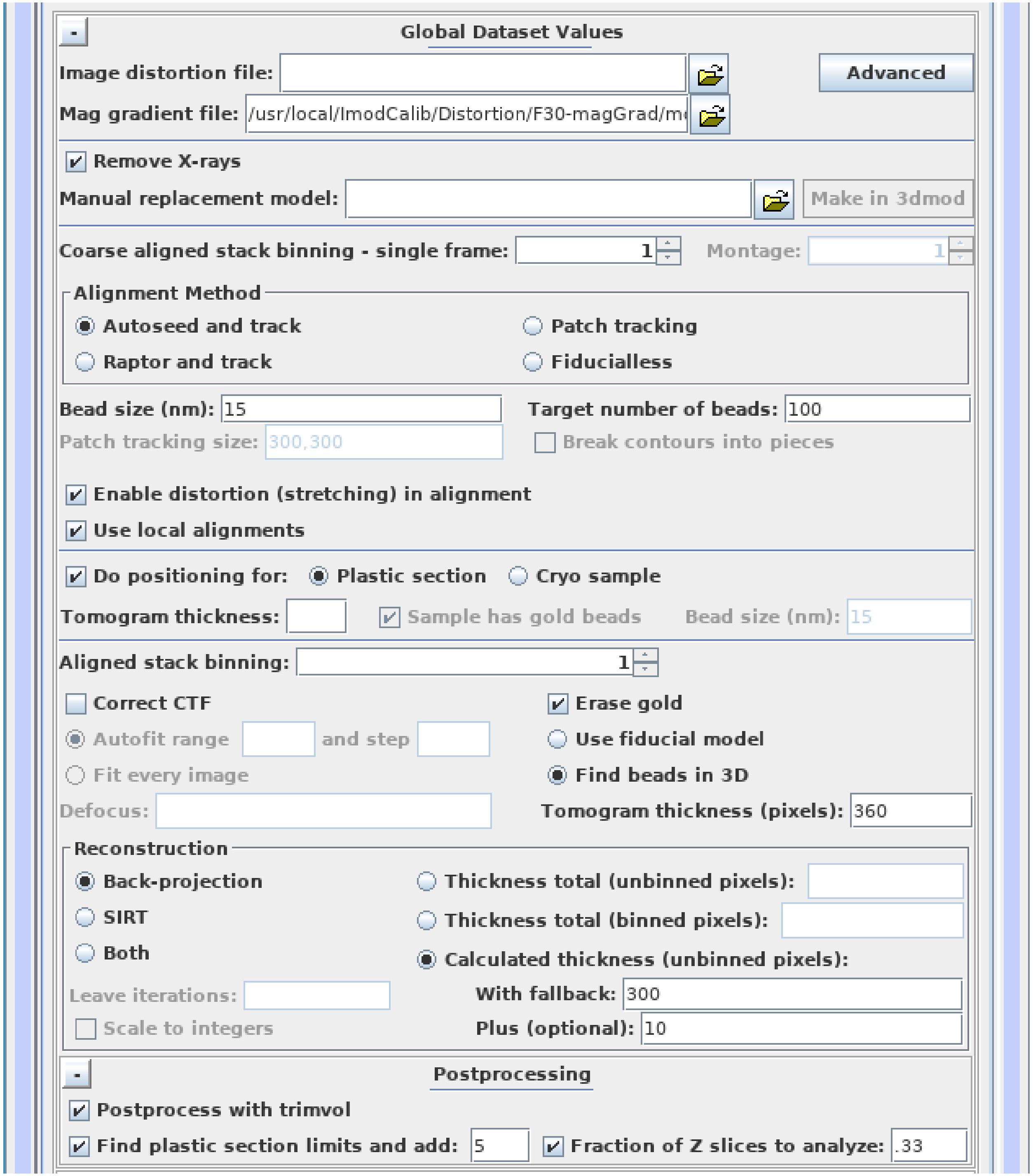
Required steps

Optional steps

After automated processing, the data set can be opened in Etomo for examination and refinement

Setting Parameters for Batch Reconstruction in Basic Interface

- Tried to keep it as simple as possible, so it fits on a screen



Global Dataset Values

Image distortion file:

Advanced

Mag gradient file: /usr/local/ImodCalib/Distortion/F30-magGrad/m

☑ Remove X-rays

Manual replacement model:

Make in 3dmod

Coarse aligned stack binning - single frame: 1

Montage: 1

Alignment Method

● Autoseed and track

○ Patch tracking

○ Raptor and track

○ Fiducialless

Bead size (nm): 15

Target number of beads: 100

Patch tracking size: 300,300

☐ Break contours into pieces

☑ Enable distortion (stretching) in alignment

☑ Use local alignments

☑ Do positioning for: ● Plastic section ○ Cryo sample

Tomogram thickness:

☑ Sample has gold beads

Bead size (nm): 15

Aligned stack binning: 1

☐ Correct CTF

☑ Erase gold

● Autofit range and step

○ Use fiducial model

○ Fit every image

● Find beads in 3D

Defocus:

Tomogram thickness (pixels): 360

Reconstruction

● Back-projection

○ Thickness total (unbinned pixels):

○ SIRT

○ Thickness total (binned pixels):

○ Both

● Calculated thickness (unbinned pixels):

Leave iterations:

With fallback: 300

☐ Scale to integers

Plus (optional): 10

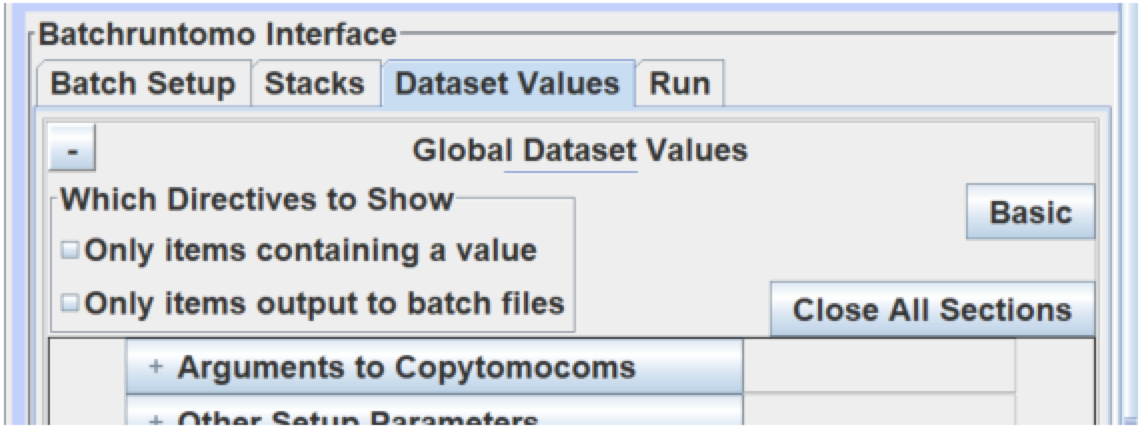
Postprocessing

☑ Postprocess with trimvol

☑ Find plastic section limits and add: 5

☑ Fraction of Z slices to analyze: .33

New Advanced Interface Gives Access to > 150 Other Parameters



Batchruntomo Interface

Batch Setup | Stacks | Dataset Values | Run

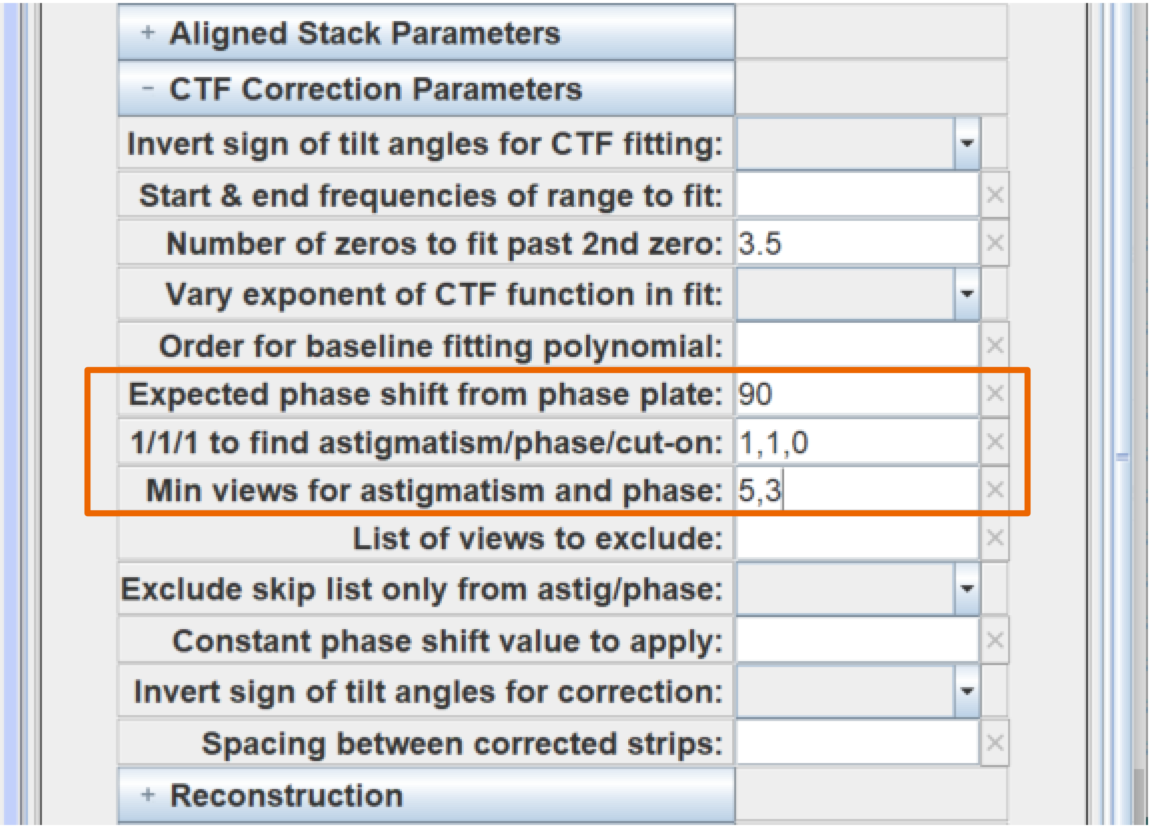
Global Dataset Values

Which Directives to Show

☐ Only items containing a value

☐ Only items output to batch files

. . .



Invert sign of tilt angles for CTF fitting:

Start & end frequencies of range to fit:

Number of zeros to fit past 2nd zero: 3.5

Vary exponent of CTF function in fit:

Order for baseline fitting polynomial:

Expected phase shift from phase plate: 90

1/1/1 to find astigmatism/phase/cut-on: 1,1,0

Min views for astigmatism and phase: 5,3

List of views to exclude:

Exclude skip list only from astig/phase:

Constant phase shift value to apply:

Invert sign of tilt angles for correction:

Spacing between corrected strips:

Section with CTF parameters opened up

Tilt Series Processing

- Speed of processing is less of an issue than for acquisition
  - If necessary, tilt series can be handled by multiple machines
- IMOD processing is fast, can be ~5-10 minutes for 4K data sets
- Potential challenges:
  - Missing pieces in IMOD fully automated processing: e.g., detection of bad pictures, automatic detection of boundary (carbon hole edge) for fiducial tracking
  - Tilt series alignment
  - CTF correction
  - Hooking it all up into acquisition-reconstruction pipeline
    - Relatively straightforward in a single defined environment, more difficult to provide a distributable tool

Automated Tilt Series Alignment Is Still a Challenge: Fiducial Marker Alignment

- Fiducial markers are still the best way to guarantee a globally correct alignment (consistent across full tilt range)
  - If numerous enough, they allow nonlinear modifications to the alignment that can improve cryotomograms and subvolume averages
    - Local alignments in IMOD: separate fits to overlapping subsets of fiducials
    - Polynomial fits to the full set of fiducials: bundle adjustment (Lawrence et al, 2006) and modeling of beam-induced motion (Fernandez et al. 2018)

Automated Tilt Series Alignment Is Still a Challenge: Fiducial Marker Alignment

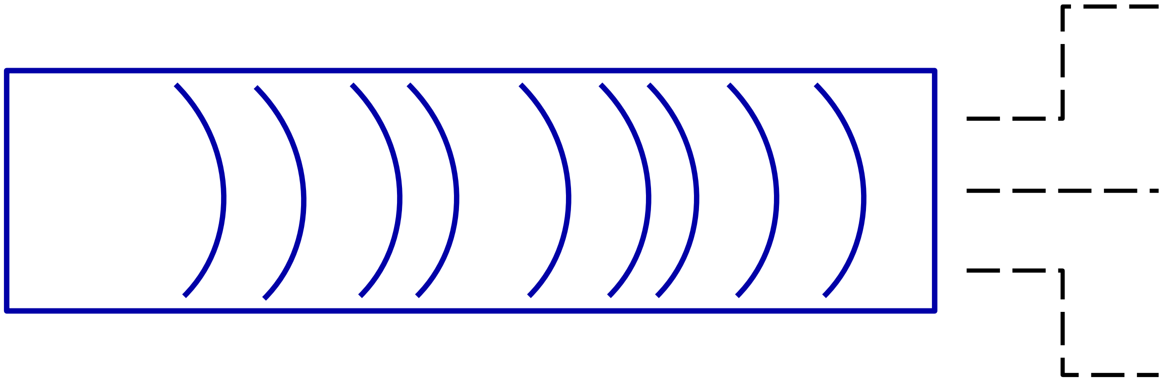
- Fiducial markers are still the best way to guarantee a globally correct alignment (consistent across full tilt range)
  - If numerous enough, they allow nonlinear modifications to the alignment that can improve cryotomograms and subvolume averages
    - Local alignments in IMOD: separate fits to overlapping subsets of fiducials
    - Polynomial fits to the full set of fiducials: bundle adjustment (Lawrence et al, 2006) and modeling of beam-induced motion (Fernandez et al. 2018)
  - But fiducial selection and tracking is still problematic for some fraction of data sets
  - It can be difficult to get gold into the specimen with a good number and distribution, or impossible for cryosections and FIB lamellae
  - The gold may not behave the same as the biological material

Automated Tilt Series Alignment Is Still a Challenge: Correlation-Based Alignment

- The main alternatives to fiducials are two variations of correlation-based approaches
  - Whole image correlations, including iterative alignment with reprojections from tomogram (Protomo from Hans-Peter Winkler)
  - Correlation of image patches through the series, with tracked locations fit to an alignment model
    - Alignator by Castano-Diez, patch tracking in IMOD
    - Allows local alignments if there are enough patches, which is difficult for most cryo-samples with low SNR

Correlation-based Approaches May Not Give Globally Consistent Alignment

- There is no guarantee that correlation between adjacent views is aligning the same features through the whole series
  - Specimens have thickness, and the features that dominate the correlation may shift to different heights during the series



Images near 30º align to put top in tomogram center

Images near 0º align to put middle in tomogram center

Images near -30º align to put bottom in tomogram center

- Iterative alignment to tomogram reprojections has little driving force for fixing such a global misalignment
  - Due to central section theorem, a reprojection will have only information contributed by views at nearby tilt angles

Improving Correlation-Based Approaches

- Despite this fundamental problem, correlation methods yield good results a surprisingly large fraction of the time
- One approach is to include a few real fiducials – points known to represent the same position in 3D through the whole tilt series

Improving Correlation-Based Approaches

- Despite this fundamental problem, correlation methods yield good results a surprisingly large fraction of the time
- One approach is to include a few real fiducials – points known to represent the same position in 3D through the whole tilt series
- A more elaborate but promising approach:
    1. Get initial tomogram
    2. Find particles of some kind and get a subvolume average
    3. Insert average into tomogram at particle locations
    4. Reproject
    5. Align original image locally with reprojection
    6. Make pseudo-fiducial model, use IMOD Tiltalign to get refined alignment that is consistent through the tilt range

Refinement with Reprojections of Particles

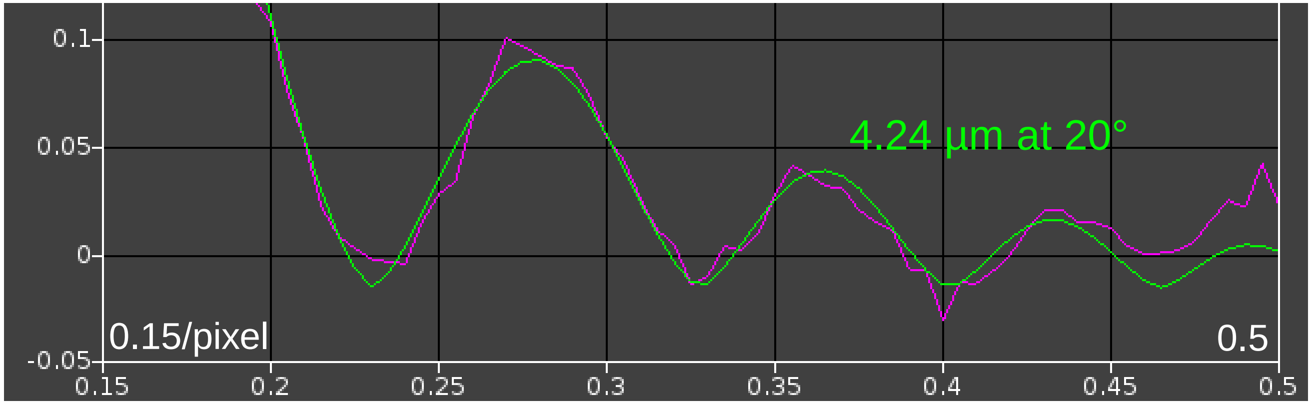
- This method can also be used to refine fiducial-based alignments
- It requires running a much bigger loop of operations
- EmClarity from Ben Himes uses essentially this approach among many other steps
  - But it is oriented toward producing aligned particles, not an aligned tomogram
  - Also takes a long time to run, so not suitable for a pipeline to produce tomograms
- Vojtěch Pražák at Oxford is developing “Flexo” following this approach
  - It may work on whatever is in the specimen (i.e., lots of ribosomes)
  - One of his goals is to incorporate it into IMOD, so it may fit within the IMOD processing stream

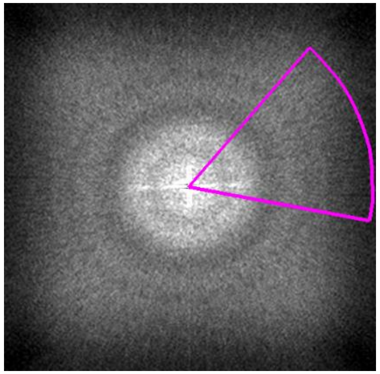
CTF Determination for Tilt Series

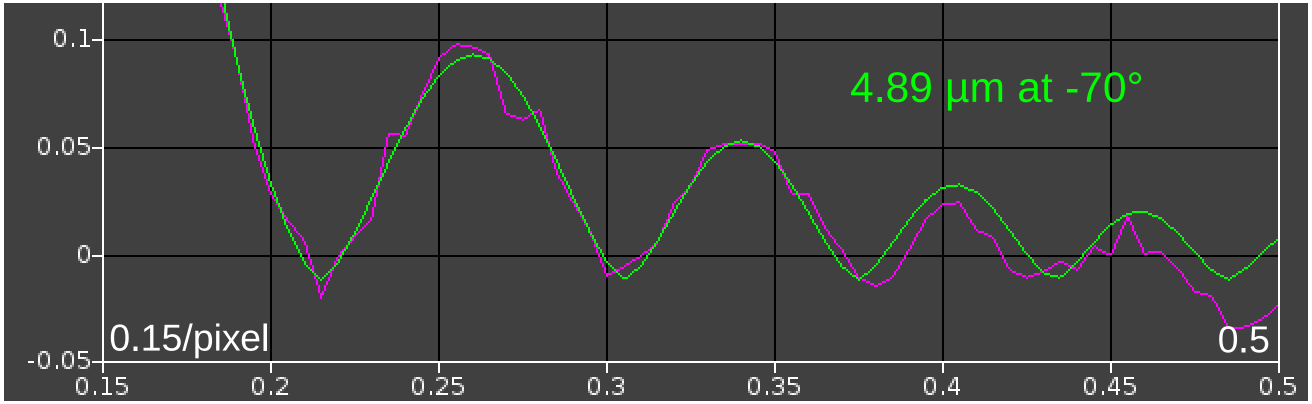
- To get high-resolution subvolume averages, CTF correction needs to include astigmatism
  - IMOD lacked the ability to find astigmatism until recently
  - Users wanting to find astigmatism (and phase for images from phase plates) adopted CTF programs developed for single-particle applications, principally Ctffind4 or gctf
  - These programs find astigmatism by fitting to 2D spectra and do not compensate for tilt
- The CTF determination program in IMOD was recently modernized in various ways, principally to find astigmatism and phase
  - Ctfplotter analyzes 1D spectra from rotational averaging

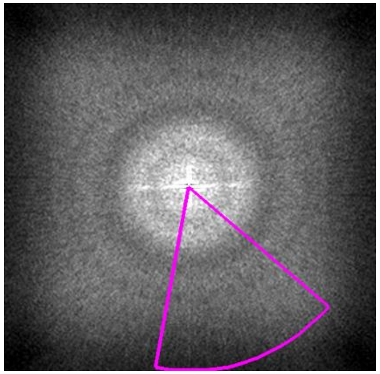
Rotational Averaging from Wedges in Spectrum

- After trying to find astigmatism from 2D spectra with Ctffind4 incorporated into Ctfplotter, I turned to analyzing 1D spectra from wedges of the FFT
- Averaging over restricted angular range (60°) shows clear changes in defocus that can still be measured well
- A series of wedge spectra are analyzed to determine astigmatism



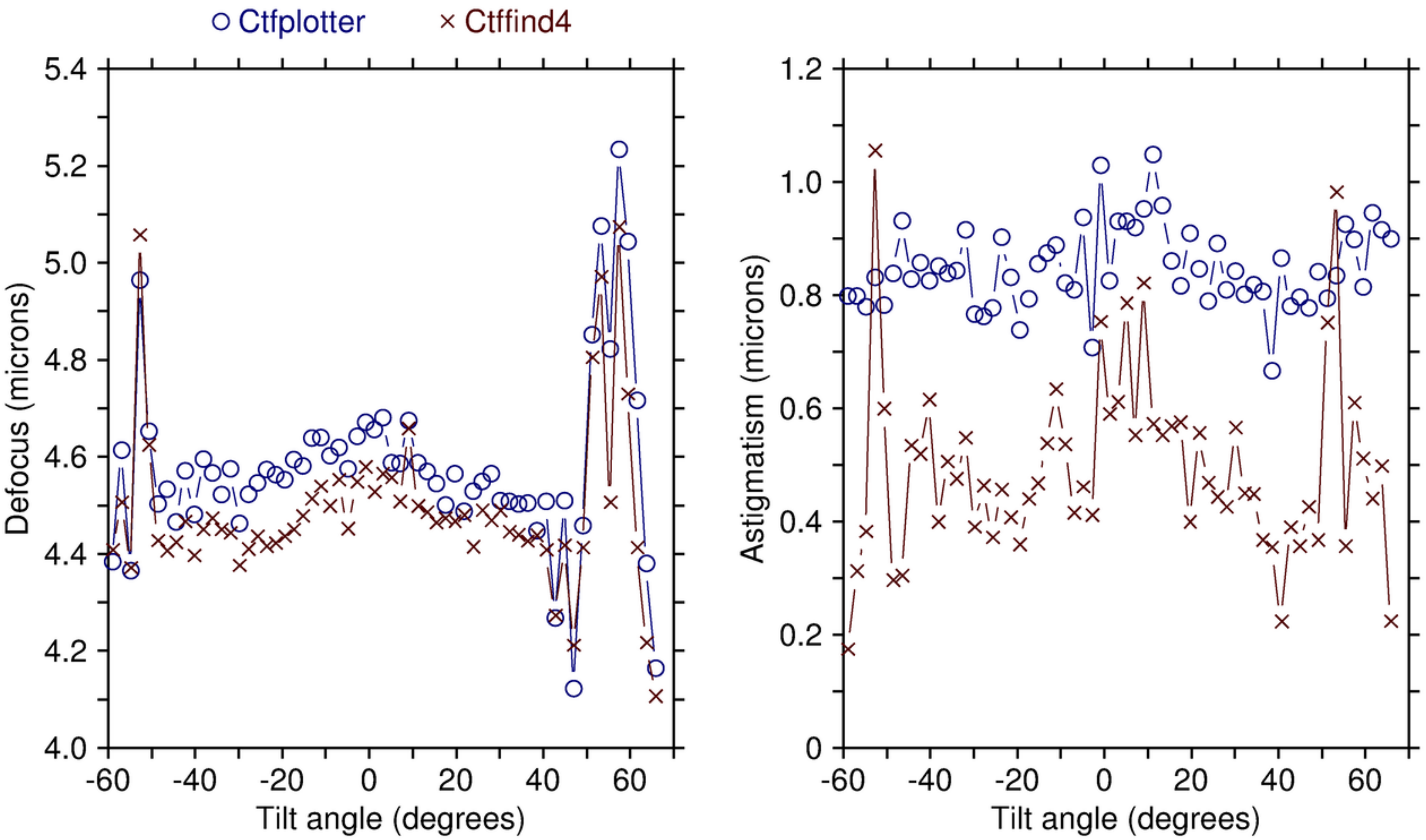






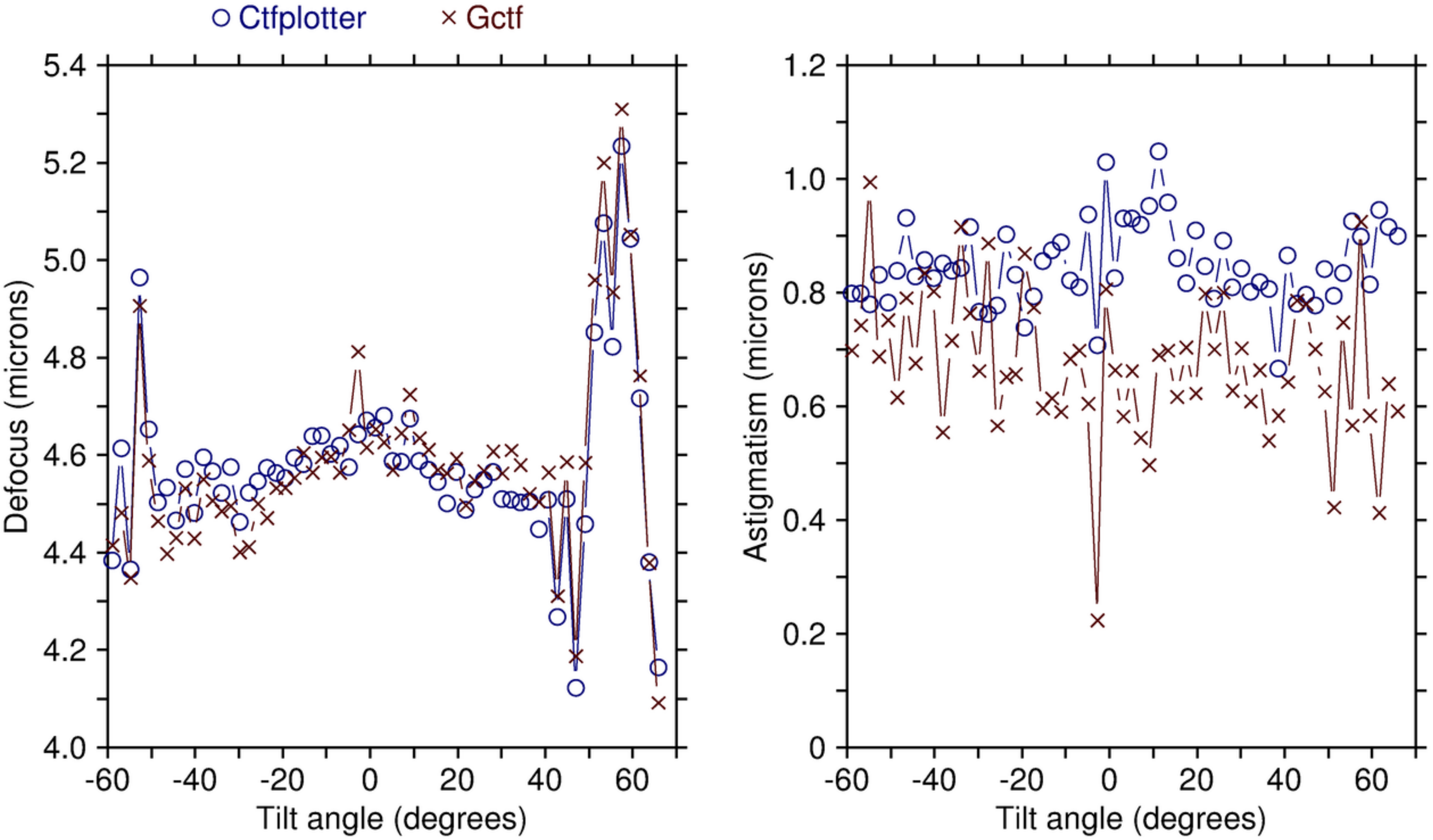
Measuring Astigmatism from Wedge Analysis is More Reliable than Fitting to 2-D Spectra with Ctffind4

- Ctffind4 4.1.8 badly underestimates astigmatism in these images
  - The consistent error in defocus seems to be a side-effect of that

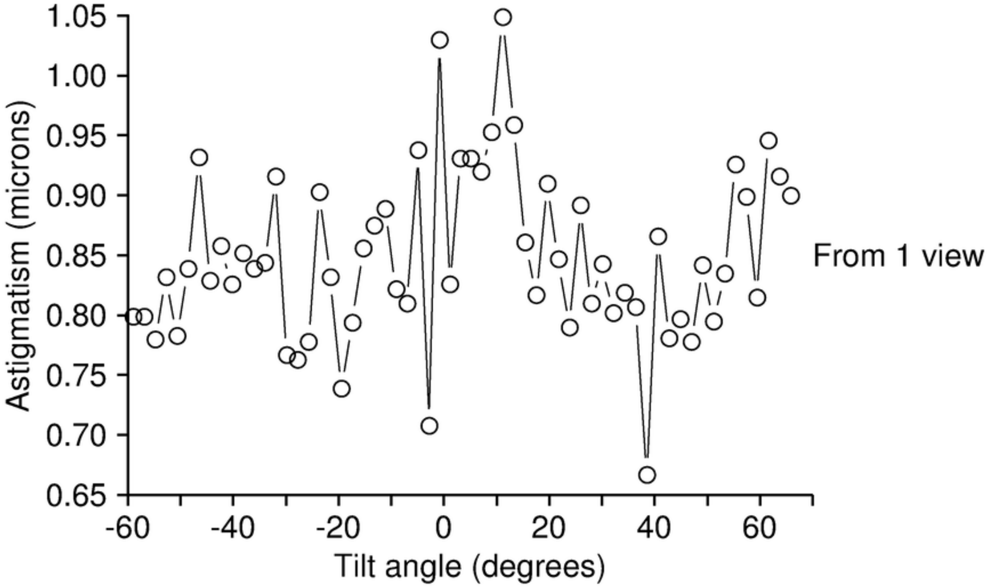


Measuring Astigmatism from Wedge Analysis is More Reliable than Fitting to 2-D Spectra with Gctf

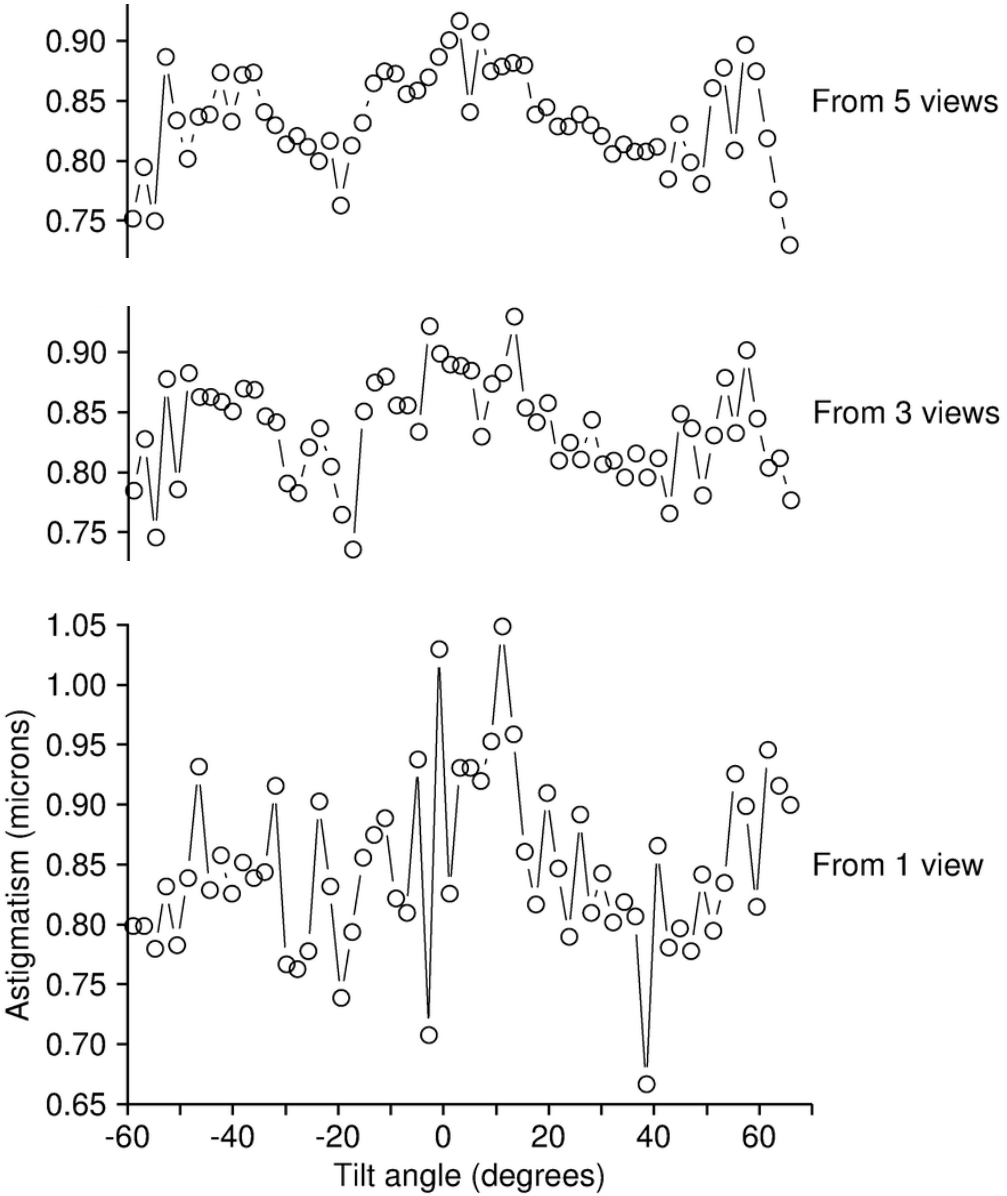
- Gctf does better than Ctffind4 on astigmatism but still underestimates it
- Defocus is also closer, but seems to have a tilt-dependent bias



Summing Spectra over 3 or 5 Views Eliminates Implausible Variations in Astigmatism Estimates



Summing Spectra over 3 or 5 Views Eliminates Implausible Variations in Astigmatism Estimates



Note that defocus is still found on each individual view

Functional Developments Come Before Implementation of Full Automation

- The new methods in Ctfplotter (1-D wedge spectra, independent summing of views for astigmatism, phase, and defocus) are an advance over established programs
- Ctfplotter can be run automatically in batch, but it requires that the proper parameters be set in advance (like everything else in IMOD batch)
  - The other CTF programs have fewer parameters to set and are probably more suitable for running blind
- Another round of development is needed to implement more automatic parameter tuning

Better, Faster, and More Automatic?

- Broader point: “Better” may require “less automatic” and also “slower”, at least in the medium-term
- The path to optimal cryotomography is not a smooth uphill ascent