Tomogram reconstruction with SCIPION: from raw movies to a tomogram

November 2022

Contents

Introduction

Overview

This tutorial aims to familiarize you with initial steps of cryo-electron tomography (cryo-ET) analysis using SCIPION (de la Rosa-Trevín et al., 2016) framework. We assume you are familiar with SCIPION GUI and have some experience with cryo-ET data processing. Here we will demonstrate the workflow starting from raw data up to the tomogram reconstruction, highlighting the features of our software framework. After that, you should be able to run SCIPION with your own data.

Software requirements

To follow this tutorial you need to have SCIPION version 3.x already installed in your system. Also, make sure you have the following plugins and their corresponding binaries installed and configured:

- scipion-em-tomo
- scipion-em-motioncorr / motioncor2
- scipion-em-cistem / ctffind4
- scipion-em-imod / imod
- scipion-em-aretomo / aretomo
- scipion-em-novactf / novactf

Test data

We will use a small subset of HIV-1 virus like particles collected on a 300 kV TFS Titan Krios microscope at EMBL, Heidelberg, Germany. The files can be downloaded from EMPIAR-10164 with the following commands :

```
mkdir empiar-10164
cd empiar-10164
url="ftp://ftp.ebi.ac.uk/empiar/world_availability/10164/data"
wget -m -q -nd ${url}/mdoc-files/TS_01.mrc.mdoc
wget -m -q -nd ${url}/mdoc-files/TS_03.mrc.mdoc
wget --show-progress -m -q -nd ${url}/frames/TS_01_*.mrc
wget --show-progress -m -q -nd ${url}/frames/TS_03_*.mrc
```

The empiar-10164 folder now contains movies for 2 complete tilt-series in MRC format, as well as 2 MDOC files with metadata. Throughout the tutorial we will refer to this folder as `$DATA/empiar-10164`.

1 Import of tilt-series movies

We will start by launching SCIPION GUI:

```
> scipion3
```

In the project window click on Create Project button and enter a project name, keeping the default location. Press Create, so that a new project window appears.

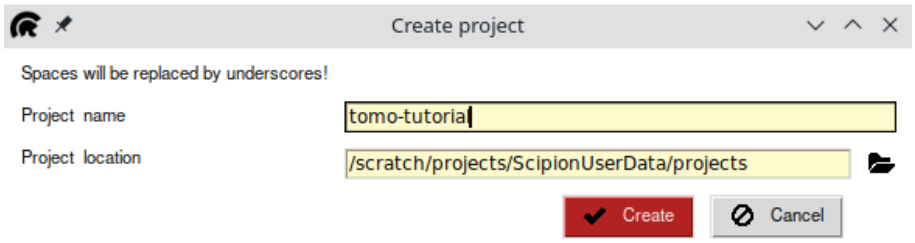

Figure 1: Create project dialog.

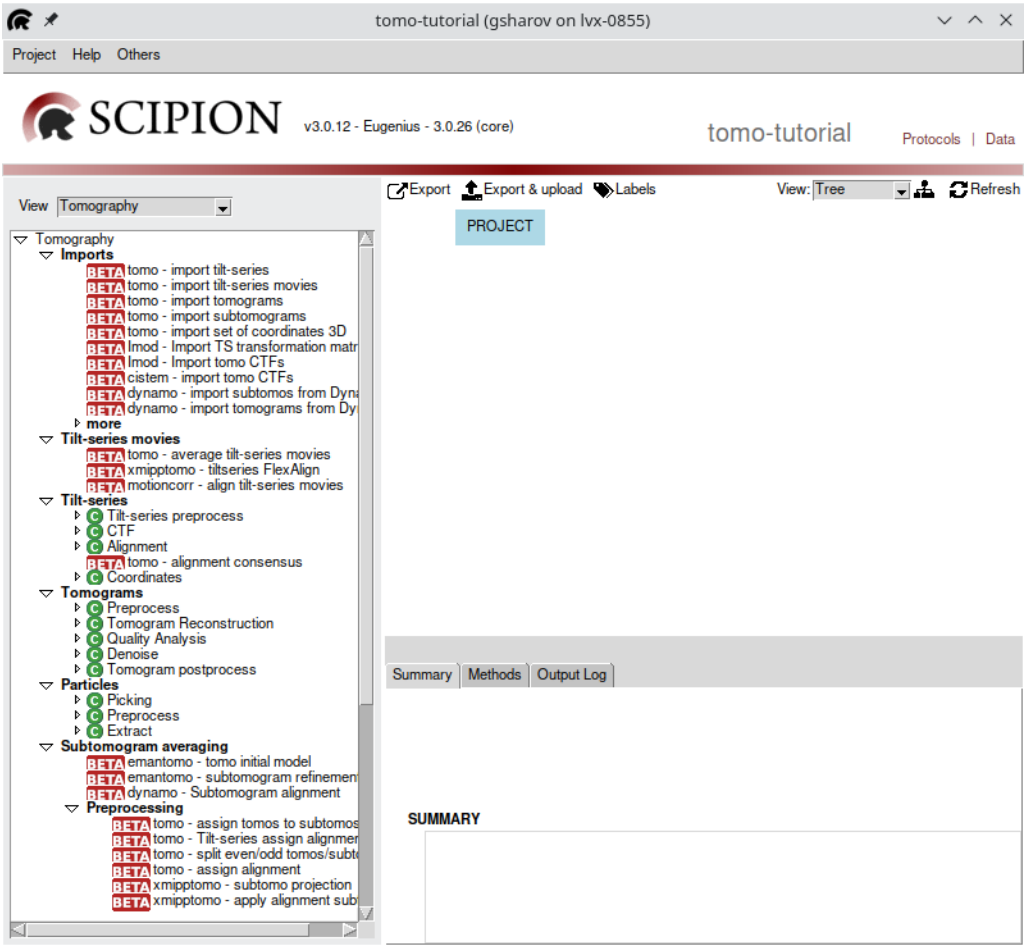

Figure 2: Main project window.

On the left panel select View > Tomography and double-click on the import tilt-series movies protocol. On the *Import* tab fill in the parameters as listed below:

- **Files directory**: `$DATA/empiar-10164`
- **Pattern**: *.mdoc
- **Microscope voltage (kV)**: 300
- **Pixel size (Å/px)**: 0.675. These movies are in the super-resolution format. The pixel size listed on the EMPIAR website is not correct.
- **Dose per tilt image (e/Å^2)**: 3.0

Here we will not import data in streaming (on-the-fly) mode, so just ignore the second protocol tab.

TIP

If you have a defects file from your detector, you will need to input it later, during motion correction. Same applies to the gain image orientation.

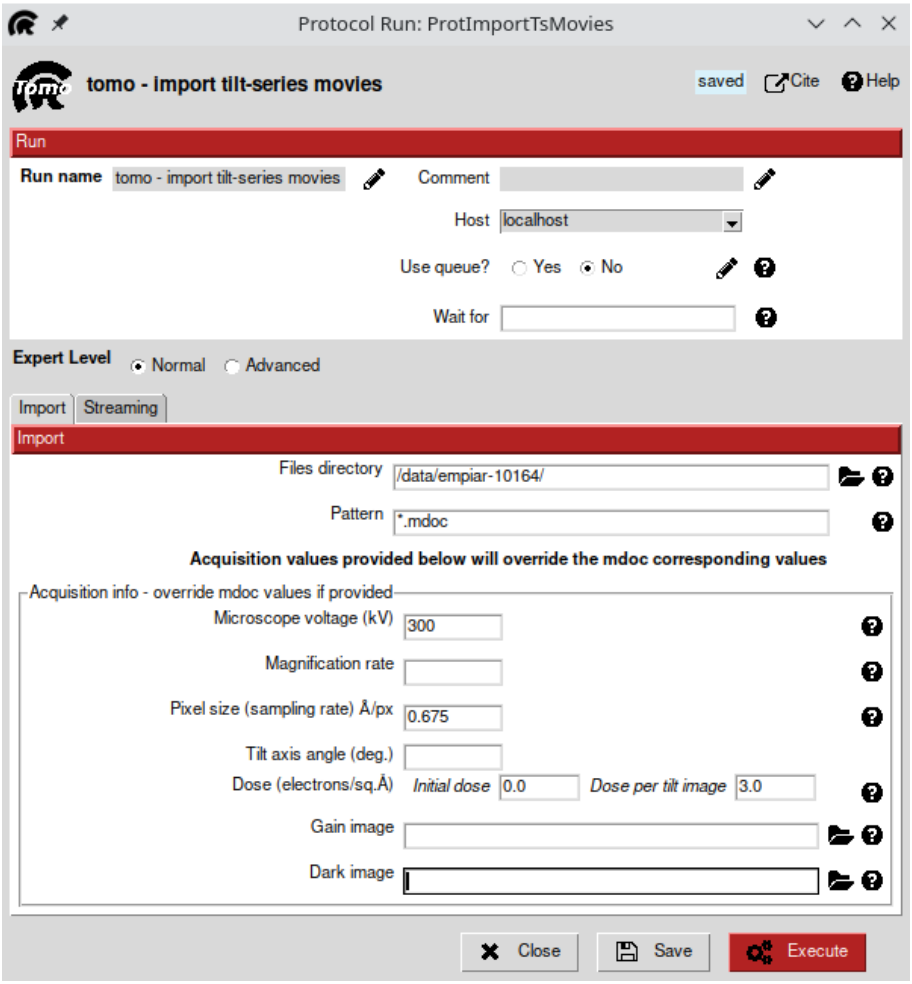

Figure 3: Import movies protocol.

This protocol (and also import tilt-series protocol) allows two import methods: via **MDOC files** or via **filename pattern**. For clarity, here we will demonstrate both approaches. MDOC is a metadata format, designed for *IMOD* (Kremer et al., 1996) and used by SerialEM (Mastronarde, 2003) data collection software. TFS Tomography software v5.6 or newer can also produce MDOC file for each tilt series. Using MDOC files is a recommended way of importing tilt-series data into Scipion project, because they contain all necessary information about both the microscope and data acquisition, including voltage, pixel size, acquisition order, tilt angles, and often dose rate.

If you open one of the downloaded MDOC files (e.g. *TS_03.mrc.mdoc*) in a text editor, you will see that it contains global metadata like pixel size, tilt axis angle, as well as information about each image of the tilt series (organized in sections with different Z-values). The values for dose rate and voltage are missing, that's why we had to input them manually above. Also, the pixel size in this MDOC file corresponds to a binned stack, not the raw movies. There are a few more important points to mention about MDOC files:

- For the import tilt-series movies protocol one MDOC file per tilt-series (per a group of movies that correspond to the same tilt-series) is expected. SerialEM can (if configured) produce MDOC file on a per movie basis, however such files cannot be used for import at the moment
- For the import tilt-series protocol also one MDOC file per tilt-series is expected, however each tilt-series must be in the $*.mrcs$ stack format
- MDOC files and binary stacks / movies should be located in the same folder
- To distinguish different tilt-series Scipion will create a unique **TS_ID** based on each MDOC file basename. This key will be used to recognize the origin of all objects created from a certain tilt-series throughout Scipion workflow
- When SerialEM creates a MDOC file for each tilt series, it is associated with the tilt-series stack that can have a different binning compared to raw movie files. In that case you need to specify the correct pixel size manually when importing movies into Scipion
- Recently, SerialEM has added a feature to rearrange tilt-series stack such that tilt images are sorted by tilt angle after the acquisition. This leads to a rearranged MDOC file. Nevertheless, Scipion is able to parse *DateTime* field from MDOC to figure out the original acquisition order
- Often (e.g. due to a missing calibration), the information about dose rate is absent from MDOC file, in such case you have to specify it manually

- SerialEM MDOC files do not contain information about spherical aberration (Cs). The default value SCIPION import protocols is set to **2.7 mm**, which is true for TFS Talos and Titan microscopes with C-Twin objective lens. Consult your microscope documentation to find out the correct value. If your data was acquired on a Cs-corrected microscope, input **0.001 mm**.

After checking that all input parameters are correct, press Execute button. The protocol should complete almost immediately as it only creates symbolic links to your raw data and registers metadata in the database. The **Summary** tab should now show that 2 tilt-series have been imported, 41 movies each, 8 frames per each movie, with the pixel size of 0.68 Å/px.

TIP

If you do not have the same number of images/movies for every tilt-series, SCIPION summary will display the number for the first imported series. Don't worry, the information is stored correctly in the database.

Now, just to practice, we will demonstrate how to import the same tilt-series movies without any MDOC files using a filename pattern. Right click on the completed protocol and make a copy. Change the input parameters as listed below and execute the protocol:

- **Files directory**: `$DATA/empiar-10164`
- **Pattern**: {TS}_{TO}_{TA}.mrc
- **Microscope voltage (kV)**: 300
- **Magnification rate**: 105000
- **Pixel size (Å/px)**: 0.675
- **Tilt axis angle (deg)**: 85.3
- **Dose per tilt image (e/Å^2)**: 3.0

As you can see, we have to input more metadata manually in this case. Moreover, the file pattern becomes more complex. It consists of 3 parts: TS to identify individual tilt-series, TO to identify acquisition order and TA to identify tilt angle from the filename. In our case, if you look for example at file TS_01_002_-3.0.mrc,

- TS = TS_01
- TO = 002

- TA = -3.0

The pattern can become more complex depending on the settings of your data collection software. If you need more information about file patterns, click on ? next to the *Pattern* parameter on the *Input* protocol tab.

Now click on Analyse Results for either import protocol and in the opened dialog verify the acquisition order, tilt angles and accumulated dose for the imported tilt-series movies. In our case, the data was collected using a dose-symmetric scheme (+/- 60 deg. starting from 0 deg.) with a 3 degrees step.

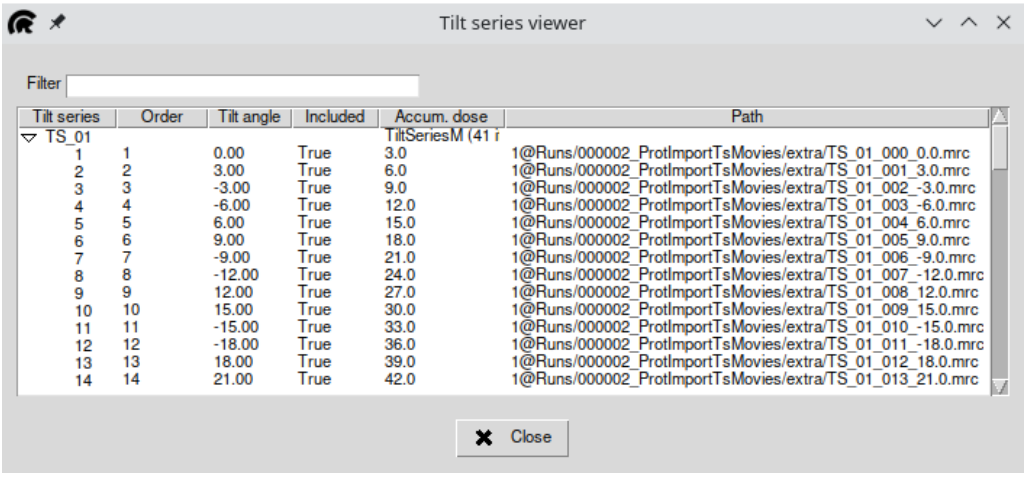

Figure 4: Tilt-series viewer. Movies are imported and sorted by acquisition order (second column).

2 Beam-induced motion correction

Aligning individual movie frames is necessary to correct for beam-induced specimen motion and restore high resolution information. In this practical we will use *motioncor2* (Zheng et al., 2017) for movie alignment. To find the corresponding protocol you can either expand the protocols tree on the left panel of the project (check the **Tilt-series movies** list) or use ctrl+F to search for motioncorr - align tilt-series movies. Select the movie set that you have just imported as input for this protocol. Set **Binning factor: 2.0** because we have imported super-resolution movies. We will not change other default options, but you are free to explore them on your own. If you have two GPUs available on your machine, set **GPU IDs** to **"0 1"** and **number of threads** to **3** - this way the protocol will run on every two movies in parallel, each on a separate GPU. The third thread is required to coordinate the jobs.

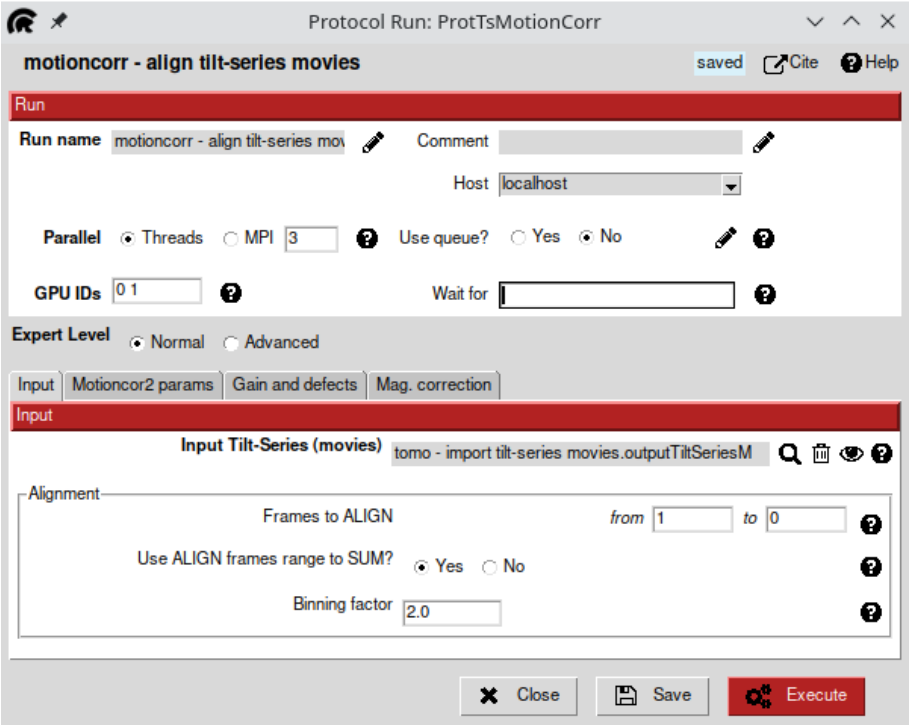

Figure 5: Movie alignment protocol for motioncor2.

After launching the protocol, we can go further and start the CTF estimation. In this case, thanks to the stream processing capability of Scipion , we don't need to wait until the previous protocol finishes to start the next one.

ALTERNATIVES

- tomo - average tilt-series movies
- xmipptomo - tiltseries FlexAlign

If you would like to first check the results of the movie alignment, wait until the first output tilt-series are produced and click on Analyze Results to open the viewer. Here you have only one option to display the output. In the opened viewer verify that the output tilt-series stacks are now sorted by tilt angle. You can also right click on the protocol output and choose *Open with ImodViewer*. Double clicking on a tilt-series name will launch *3dmod* interface to visualize the stack file.

TIP

At the moment, Scipion does not store movie frame shifts from motion correction protocols for tomography.

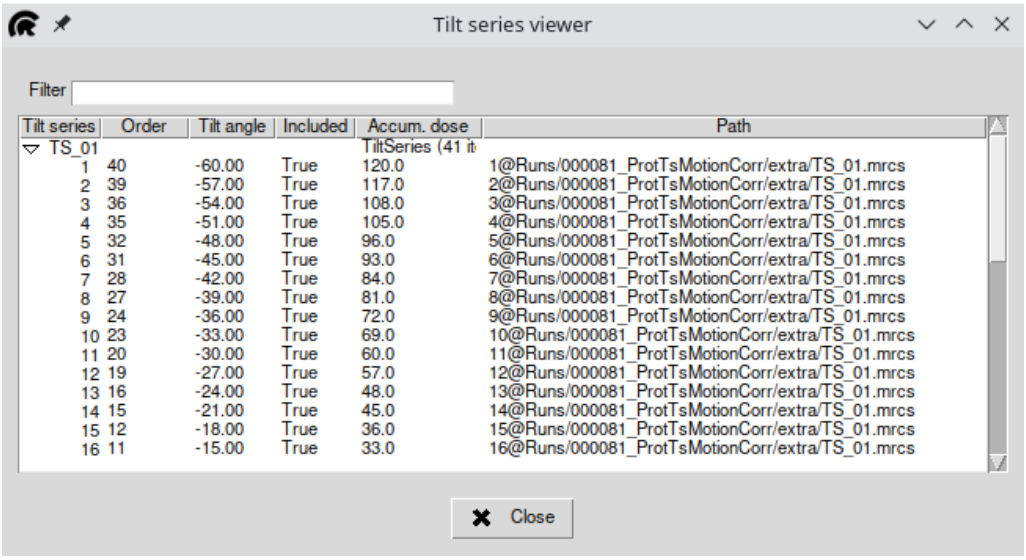

Figure 6: Tomo data viewer. Motion-corrected tilt-series are now sorted by tilt angle (third column) and assembled into individual **.mrcs* stacks.

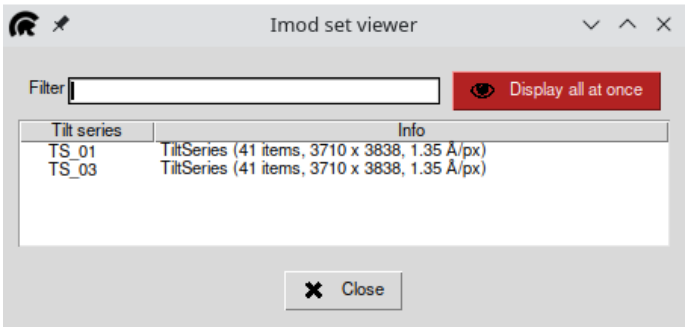

Figure 7: IMOD's tilt-series viewer. Double click on any item to launch *3dmod* interface.

If you were to use **Split and sum odd/even frames** option (visible with Expert level: Advanced) of the motioncorr protocol, you would get two more output sets. This option is commonly used to later generate two identical tomograms for the purpose of running denoising software (e.g. *cryoCARE*).

TIP

Dose-weighting option in motioncorr protocol is disabled by default. In principle, dose-weighting can be performed at many stages of the tomo pipeline. Here we will do it after CTF estimation.

3 CTF estimation

We will now estimate CTF parameters of the aligned images with *CTFFind4* (Rohou and Grigorieff, 2015). Locate cistem - tilt-series ctffind4 protocol and select tilt-series from the previous step as the input. Set number of threads equal to the number of available CPU cores. Leave all options with default values. *CTFFind4* will run on individual tilt-series **.mrcs* stacks (that were assembled by motion correction protocol) which speeds up the execution.

TIP

Make sure you always run CTF estimation on the raw data, before any dose-weighting, filtering etc.

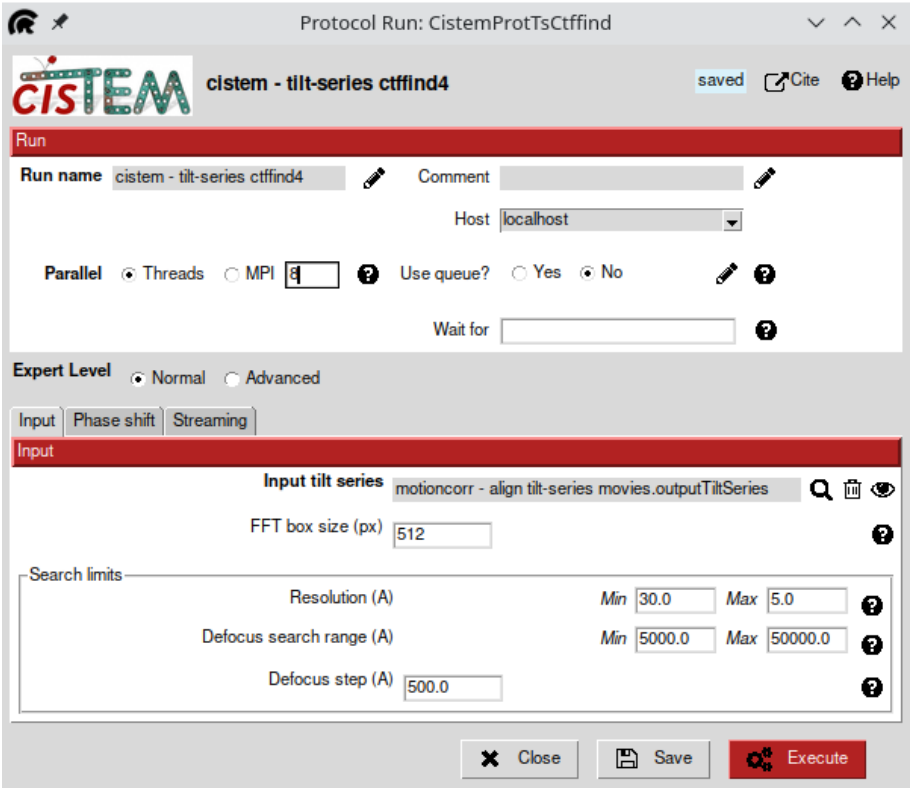

Figure 8: CTFFind4 protocol.

ALTERNATIVES

- emantomo - ctf estimation
- gctf - tilt-series gctf
- imod - CTF estimation auto/manual
- susantomo - CTF estimation

The output from *CTFFind4* (or any other CTF estimation protocol) can be

displayed upon clicking on the Analyze Results button (Figure 9). Scipion CTF viewer for tomography is split into two panels. Left mouse click on a tilt-series row on the left hand side will update the corresponding plot on the right. By default, the plot will show defocus, estimated resolution (if available) and phase shift (if estimated) as a function of tilt angle for selected tilt-series. If you click on the arrow symbol near the tilt-series name, the list will expand with all tilt images and show the table with estimated CTF parameters (defocus, astigmatism etc.) for each image. You can also left click on any tilt image and the right hand side plot will be updated with a 2D CTF fit (Figure 10). Moreover, selecting an individual tilt image and clicking 1D fit button at the top of the viewer will open a 1D CTF fit profile.

TIP

2D and 1D fit CTF profiles are only available for *CTFFind4* and *gCTF*.

If you want to discard some tilt-series you may tick the checkbox next to all tilt-series that you want to keep and click Generate subsets. This will create "good" and "bad" subsets with the corresponding items.

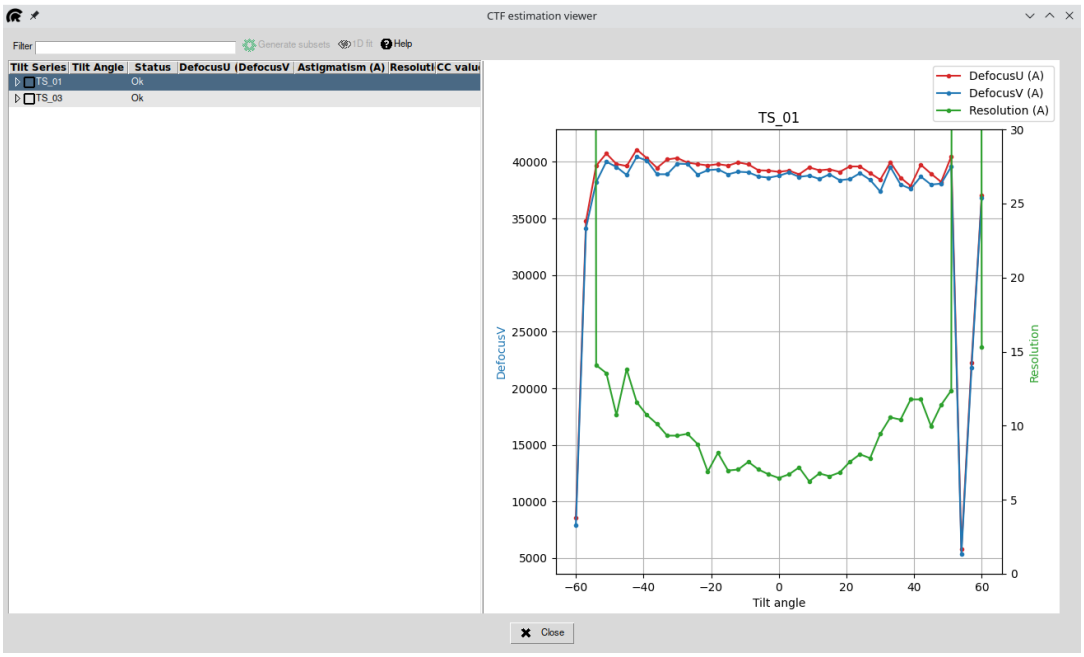

Figure 9: Visualization of CTF estimation results for each tilt-series. We can see that the defocus is consistent throughout the tilt-series, except for the high tilts where estimation has failed.

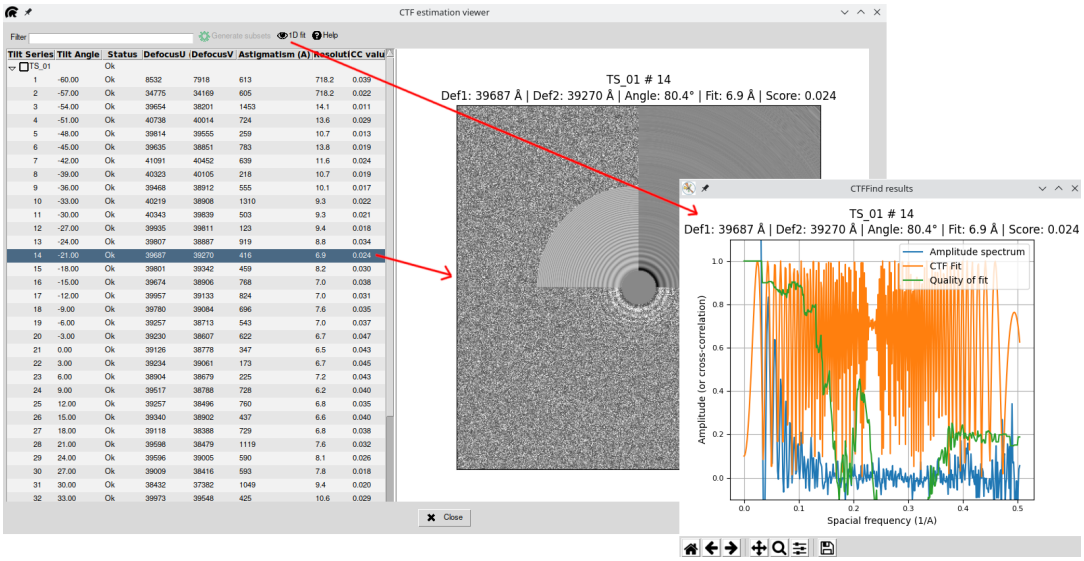

Figure 10: Individual tilt image CTF fit plots (2D and 1D) from CTFFind4.

4 Dose-weighting

As explained in (Metskas et al., 2022), dose-weighting can be done at many stages in the tomogram reconstruction pipeline, but for best effect should be done prior to tilt stack alignment and after CTF estimation. While the primary purpose of dose-weighting is to limit the contributions of Fourier space containing the most electron damage (Grant and Grigorieff, 2015), when combined with a dose-symmetric tilt scheme it effectively functions as a low-pass filter in the noise-ridden high tilts. This improves overall alignment for these tilts, along with a subtle improvement in signal to noise ratio in the reconstructed tomogram.

Now, locate imod - dose filter protocol and select tilt-series from the motion correction step as the input. Ignore other options and execute the protocol. Check the results with *3dmod* viewer. The tilt-series should now have more contrast, especially at high tilts. Note that both TS_01 and TS_03 have a completely dark image at -60 degrees (section 1). We will take this into account in the following chapter.

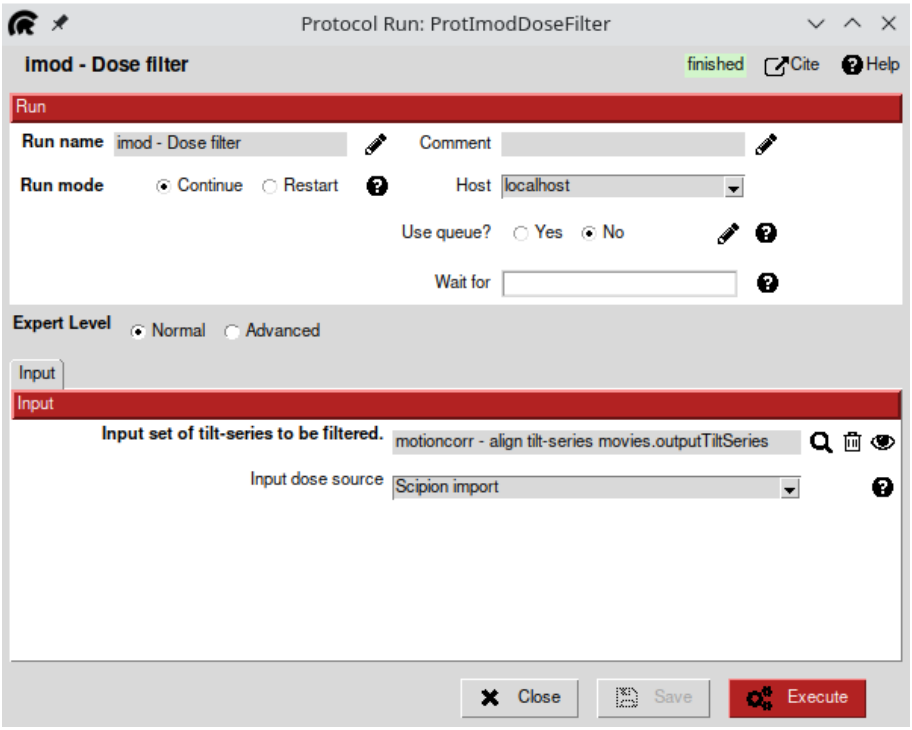

Figure 11: Dose-weighting protocol from IMOD.

5 Tilt-series alignment and reconstruction

In this chapter we describe a few common workflows that can be used to align tilt-series and reconstruct a tomogram. Users can choose which route to take depending on their experience and data complexity. The first section goes through a manual tilt-series processing via *eTomo*. Here a user is presented with full capabilities of the original *IMOD*'s graphical interface where almost any parameter can be tweaked. In this case SCIPION will simply collect and register the results. The second section describes a semi-automated approach, where the individual steps from *eTomo* interface have been wrapped into separate SCIPION protocols. Finally, the last section shows the power of *AreTomo* software that does not require fiducial markers and is able to align and reconstruct tomograms in a fully automated fashion. This tutorial dataset can be used with either of these three workflows, so you may choose any of them to follow.

5.1 Manual processing with eTomo GUI

At this point we assume that you already have some experience with *eTomo* (Mastronarde and Held, 2017) as we will not provide a complete guide for this software, but rather illustrate how to run *eTomo* interactive GUI from SCIPION and how to save the results into the project. For a more comprehensive description of every step, please refer to the official eTomo tutorials [1] and [2].

Find the imod - etomo interactive protocol and input the dose-filtered tilt-series set. Also input **fiducial marker size of 10 nm**. Set number of threads to **4**. Pixel size, tilt angles and tilt axis position will be fetched by SCIPION automatically. After executing the protocol you will see a different tilt-series dialog showing you the steps you will need to perform within *eTomo* interface. Double click on the first item in the list to launch *eTomo* GUI.

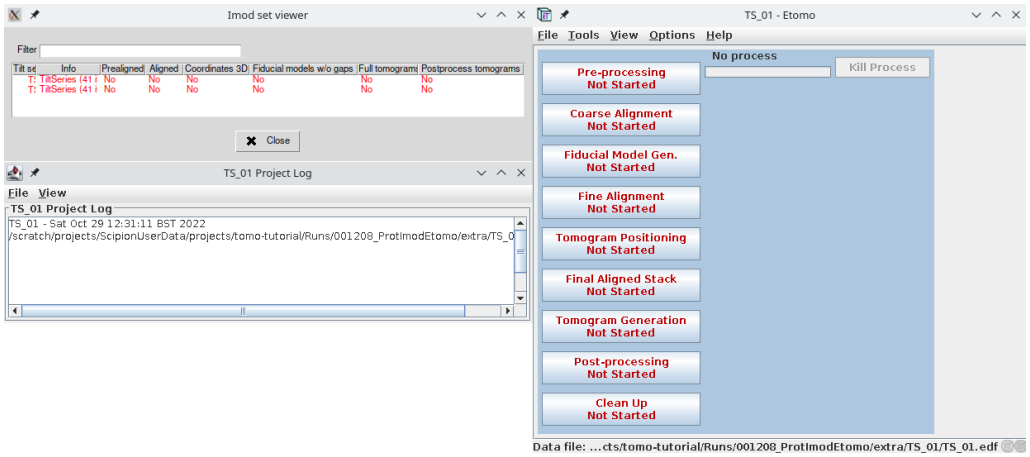

Figure 12: Tilt-series viewer of eTomo protocol (left) and eTomo interface (right).

Pre-processing

The purpose of this first step is to remove very bright or dark pixels caused by detector defects or X-rays. Such pixels can cause artifacts in a reconstructed volume so it's important to always remove them.

1. Click on Pre-processing button to open the panel for removing pixel artifacts.
2. Press Create Fixed Stack button.
3. Once completed, press on Show Min/Max for Fixed Stack and check if the min/max pixel densities distribution is uniform and there are no outliers. Deviations of 50-100 from the rest of the data will not matter in the reconstruction.
4. To proceed, press on Use Fixed Stack and then Done.

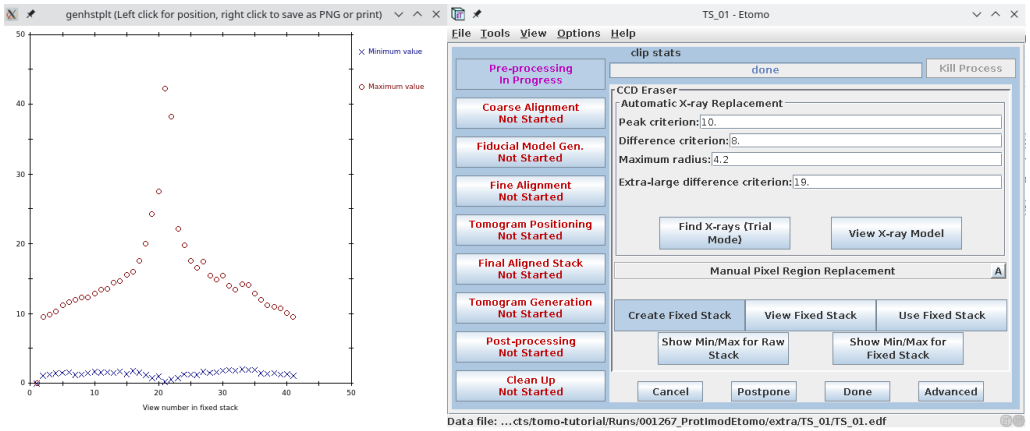

Figure 13: Removing pixel artifacts from a tilt-series stack. The plot shows minimum and maximum densities for the fixed stack. View #1 has min intensity of 0 (black image), we will get rid of it in the next steps.

Coarse alignment

The goal of coarse alignment is to calculate the cross-correlation between successive tilts and apply X and Y-shifts to the tilt-series.

1. Press the Advanced button to display more options.
2. At the end of the Tiltxcorr section enter **Views to skip: 1** to ignore the first dark image of the tilt-series.
3. Set **Coarse aligned image stack binning to 4**.
4. Check **Reduce size with antialiasing filter** checkbox.
5. Uncheck **Convert to bytes**, otherwise grayscale depth will be degraded to 8 bits.
6. Check **Float intensities to mean**.
7. Press Calculate Cross-Correlation followed by Generate Coarse Aligned stack. Tilt-series images are now aligned translationally.
8. Check the resulting aligned stack by pressing View Aligned Stack in 3dmod. Press Done.

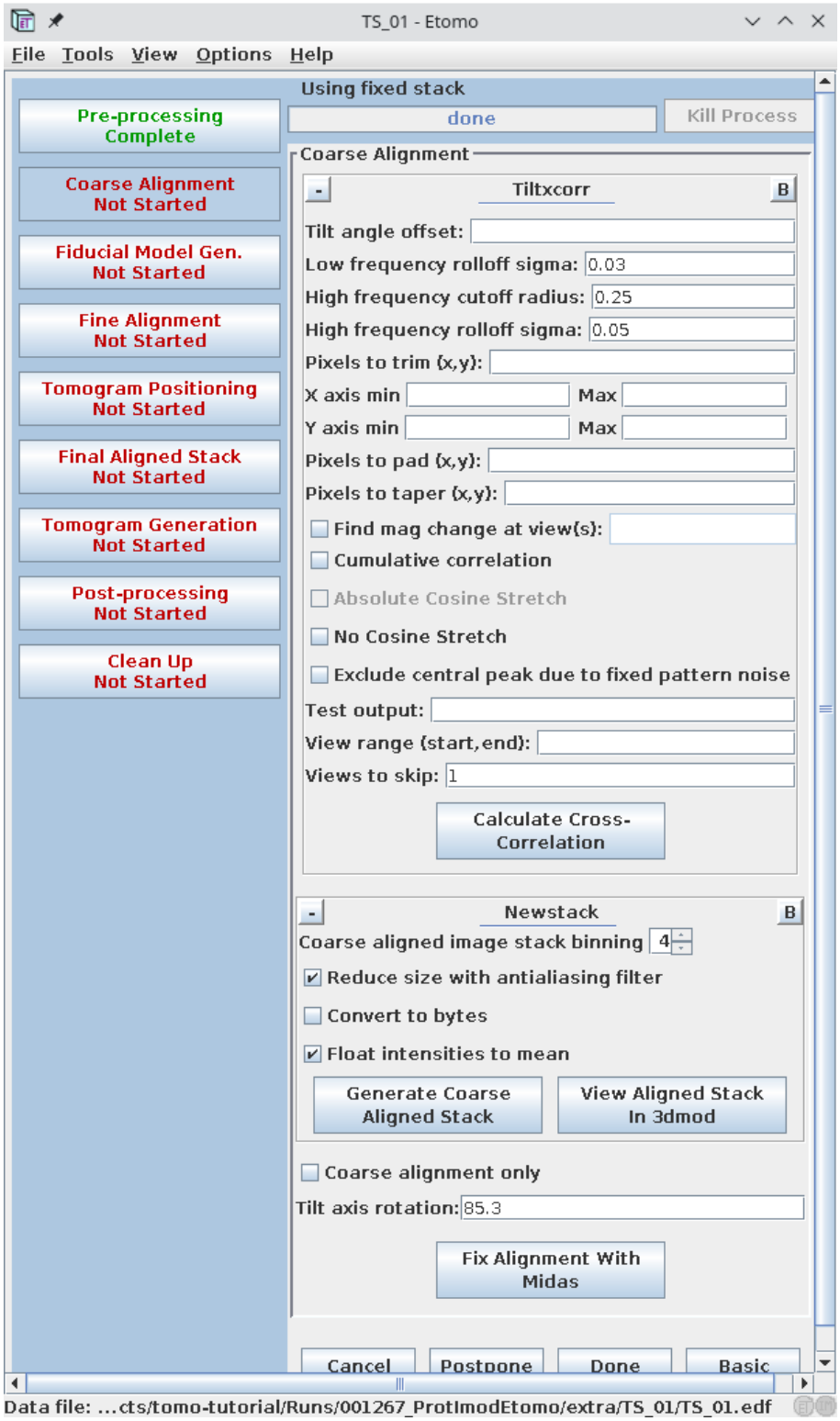

Figure 14: Coarse alignment of the tilt-series stack. We omit the first image (view) from the alignment.

Fiducial model generation

Since our data contains gold beads we will use them as fiducial markers to perform a more precise tilt-series alignment. If your data does not have any fiducials,

you may try patch tracking (not described here) or use *AreTomo*, explained in the later section.

1. Press Advanced to display more parameters.
2. On the *Seed model* tab in the Beadtracker section enter **View skip list: 1**, check **Refine center with Sobel filter**, input **Sobel sigma relative to bead size: 0.12** and **Overall low-pass filter cutoff: 0.3 (1/nm)**
3. In the Initial Bead Finding parameters section check **Find and adjust bead size**
4. In the Selection and Sorting parameters section input **Total number: 30** (the approx. number of beads in the tomogram).
5. Press Generate Seed Model.
6. Switch to the *Track Beads* tab and press Track Seed Model. In the *eTomo* log window check the total number of missing points. If a large number of points are missing, press Track with Fiducial Model as Seed.

TIP

You need at least 5 beads sparsely distributed in the field of view for the tracking to work correctly.

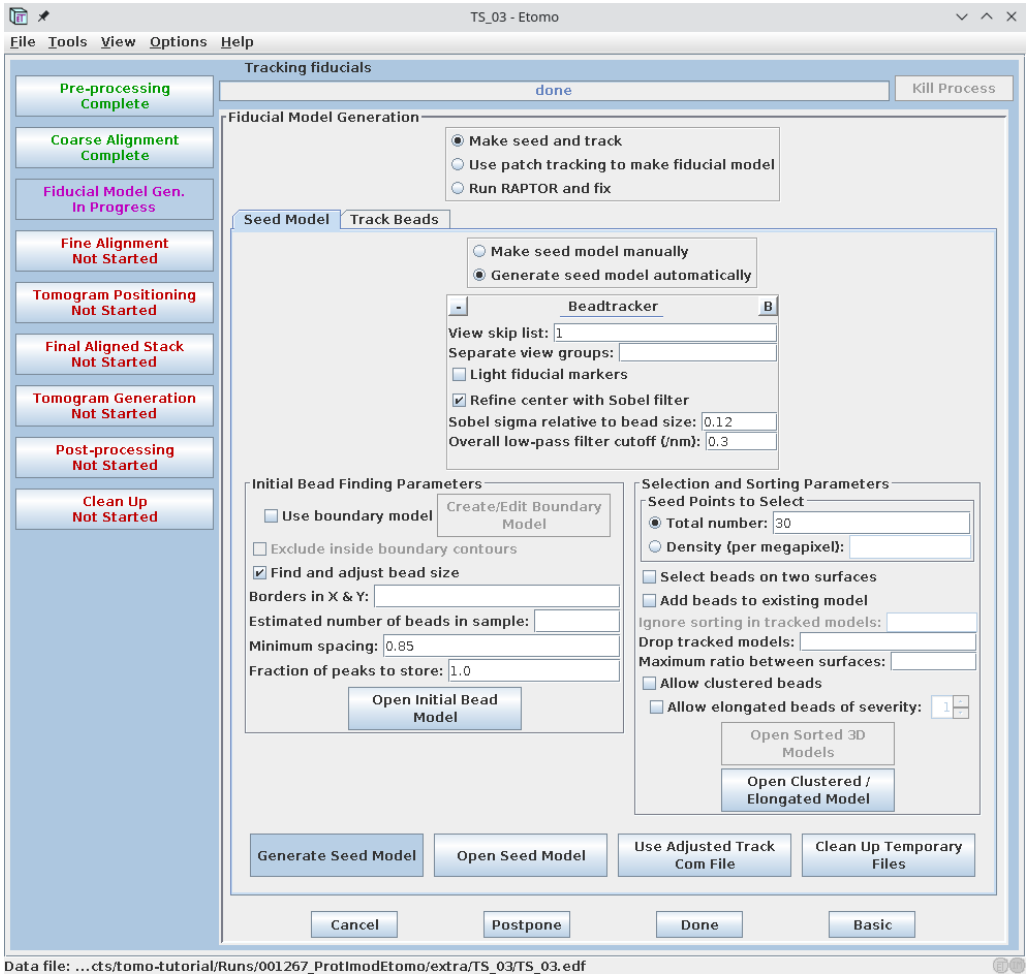

Figure 15: Fiducial model generation. Seed Model tab is displayed.

If points are still missing, the next procedure involves an iterative process to edit this fiducial model. Otherwise, just press Done to finish this step.

1. It is important to make sure that the large majority of the fiducials are tracked all the way to the two ends of the tilt-series. Press Fix fiducial model on the *Track Beads* tab, this will open the pre-aligned stack with model points overlayed along with the Bead Fixer window.

2. Click the Go to Next Gap button in the Bead Fixer window. This will highlight a point with yellow circle that has a missing model point on an adjacent section (indicated by an arrow).

3. Use the Page Up (when an "up arrow" appears) or the Page Down (see Figure 16) to switch to the image with a missing point and use the middle mouse button to add a point in the center of the gold particle (you can use right click to adjust the current point position). It is useful to increase the magnification of the image with the + key and adjust the contrast of the images, especially at high tilt.

4. Keep going to the next gap and fixing missing points until the message *No more gaps found* comes up in the main *3dmod* window.

5. Save the model file using File » Save model in the main *3dmod* window.

6. Press Track with Fiducial Model as Seed again. There should be no more missing points.

7. Press Done.

TIP

Note that it is not necessary to fill all the gaps, particularly if there is a good excess of fiducial points. Most important, you should not add a point if the bead's position is not clear or it disappears from the field of view!

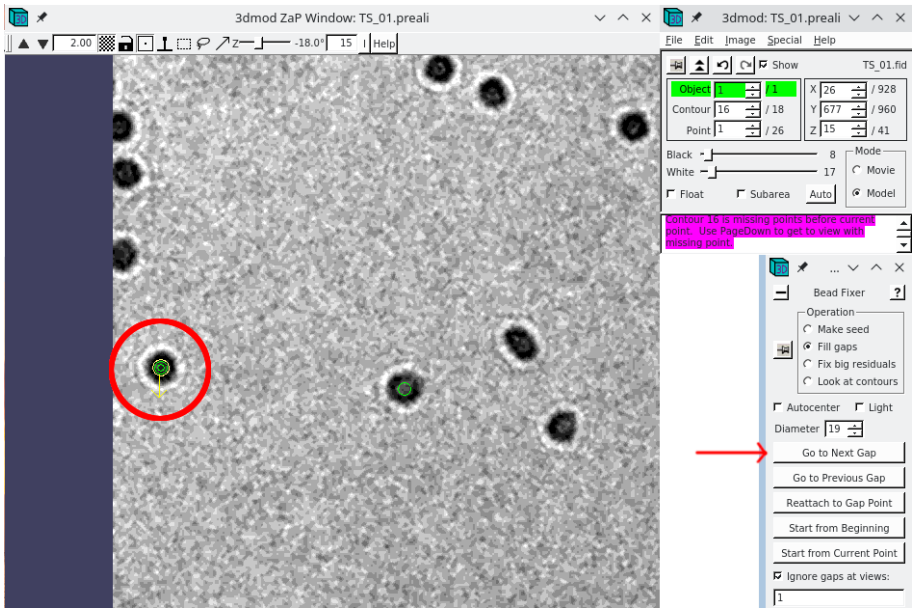

Figure 16: Filling the model gaps. Pressing Go to Next Gap (indicated by a red arrow) will highlight a bead (indicated here by a red circle) that has a missing point on the neighbouring section.

Fine alignment

The goal of this step is to solve for the displacements, rotations, tilts and magnification differences in the tilted views using the position of the gold fiducials. Usually one would aim to reduce the alignment residual error mean to 0.2 - 0.5 nm.

1. On the *General* tab set **List of views to exclude: 1** and **Threshold for residual report to 3.0**.

2. Select **Do not sort fiducials into two sufraces**
3. Check **Do robust fitting with tuning factor: 1.0**.
4. On the *Global Variables* tab set **One rotation**, **Fixed magnification at 1.0**, **Fixed tilt angles**.
5. Press Compute Alignment. The residual error is now reported in the Log window.
6. If the error is above 0.5 nm you need to optimize the alignment by removing the beads with high stdev and/or fixing bead centering manually. Depending on your microscope stage you might also adjust global variables. This is not necessary for this tutorial data, otherwise consult eTomo guide on how to fix big residuals.
7. Press Done when finished.

TIP

When playing with different parameters during final alignment always check that the residual error mean drops more significantly than the ratio of total measured values to all unknowns, otherwise you are just overfitting the model!

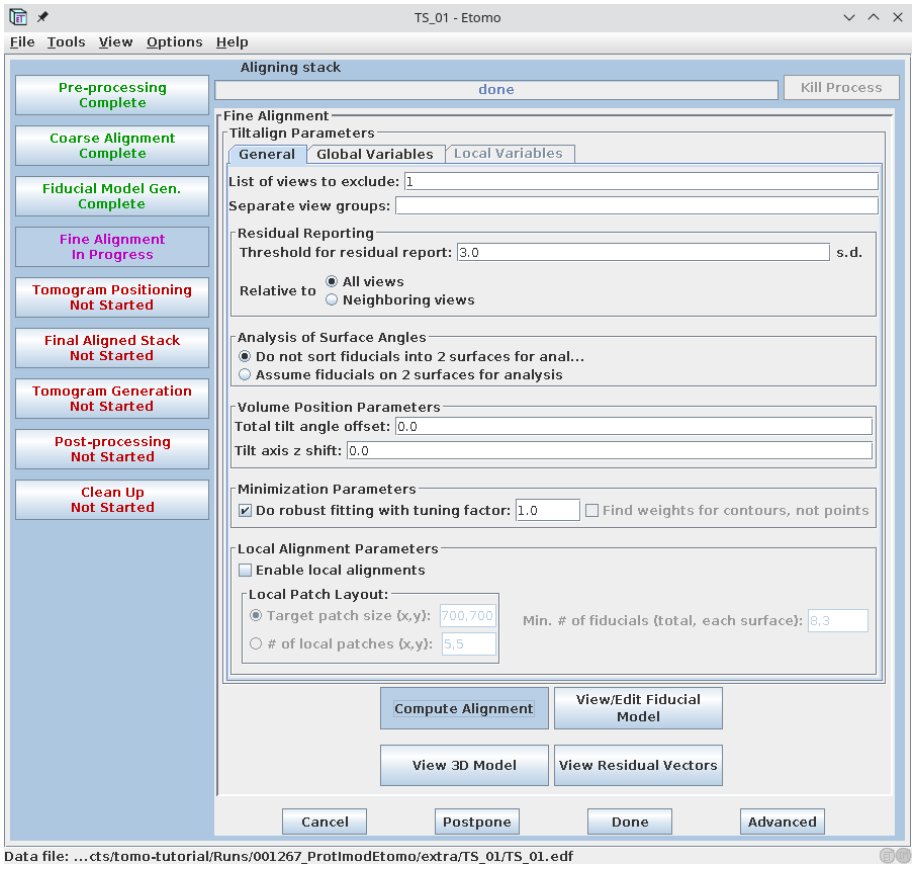

Figure 17: Fine alignment. General tab is displayed.

Tomogram positioning

The purpose of this step is to set angles and an offset in Z so that the specimen is flat and centered in the computed volume. If you plan to do subtomogram averaging it is better to skip this step (just press Done). Note that if you correct X-axis tilt at this point, the orientation of the missing wedge will be changed.

Final aligned stack generation

1. Uncheck **Use linear interpolation**.
2. Set **Aligned image stack binning to 4**. This will define the final tomogram binning (!)
3. Check **Reduce size with antialiasing filter**.
4. Press Create Full Aligned Stack followed by View Full Aligned Stack. Note that the stack has now been rotated 90 degrees so that the tilt axis became vertical. This is the default convention used by *IMOD* software. Check that the fiducials move horizontally on straight lines with no jumps when scrolling through the stack. Often it helps to draw a rectangle (use Shift + B) encompassing a few beads on its border, loop through Z and check that beads do not jump vertically. If they do – either improve the fine alignment or exclude such bad views.
5. Press Done when finished.

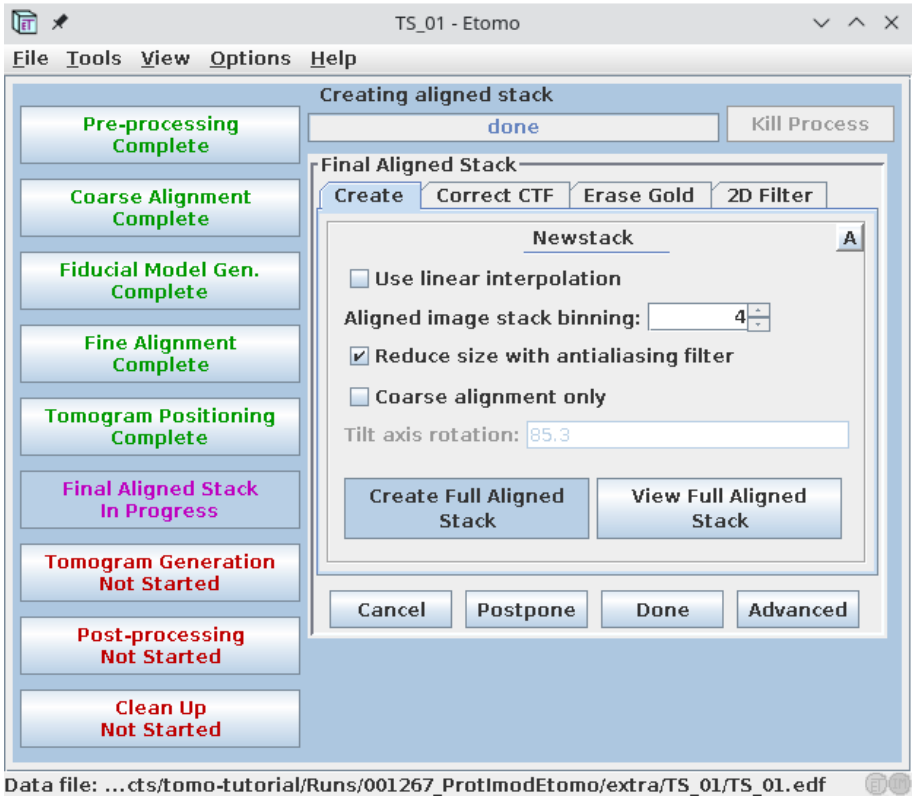

Figure 18: Making a final aligned stack. Create tab is displayed.

TIP

If you plan to proceed later with subtomogram averaging, you would usually generate a bin 1 tomogram with WBP reconstruction method (see below) and a binned SIRT-like filtered tomogram with high contrast for particle picking. Remember, the Final Aligned stack step is used to set the desired binning of the tomogram.

Tomogram generation

Note that all views excluded at the **Fine Alignment** step will be automatically omitted from the final reconstruction.

1. Select **Back projection** method.
2. Check **Parallel processing** and set 4 CPUs to use in the parallel section above.
3. Uncheck **Take logarithm of densities**
4. Set **Tomogram thickness in Z to 2000**. This is in unbinned pixels (!)
5. Set **Use SIRT-like filter equivalent to 8 iterations**.

6. Press Generate Tomogram and wait until it's completed. Then press View Tomogram In 3dmod and open Slicer (press \). In the new Slicer window set **X rotation to 90** and **Img to 30** to increase the slicing thickness. This will display x-z slice. Scroll through the tomogram using **View axis position** slider. Verify that (a) the thickness is correct and (b) the specimen is positioned in the middle of the slice, otherwise change **Z-shift** value (use positive unbinned pixels value to shift it upwards) and recalculate the tomogram. This step is equivalent to the *Tomogram positioning*. In the case of TS_01 you need to recalculate the tomogram using 1360 px final thickness and 300 px Z-shift.

7. Press Done when finished.

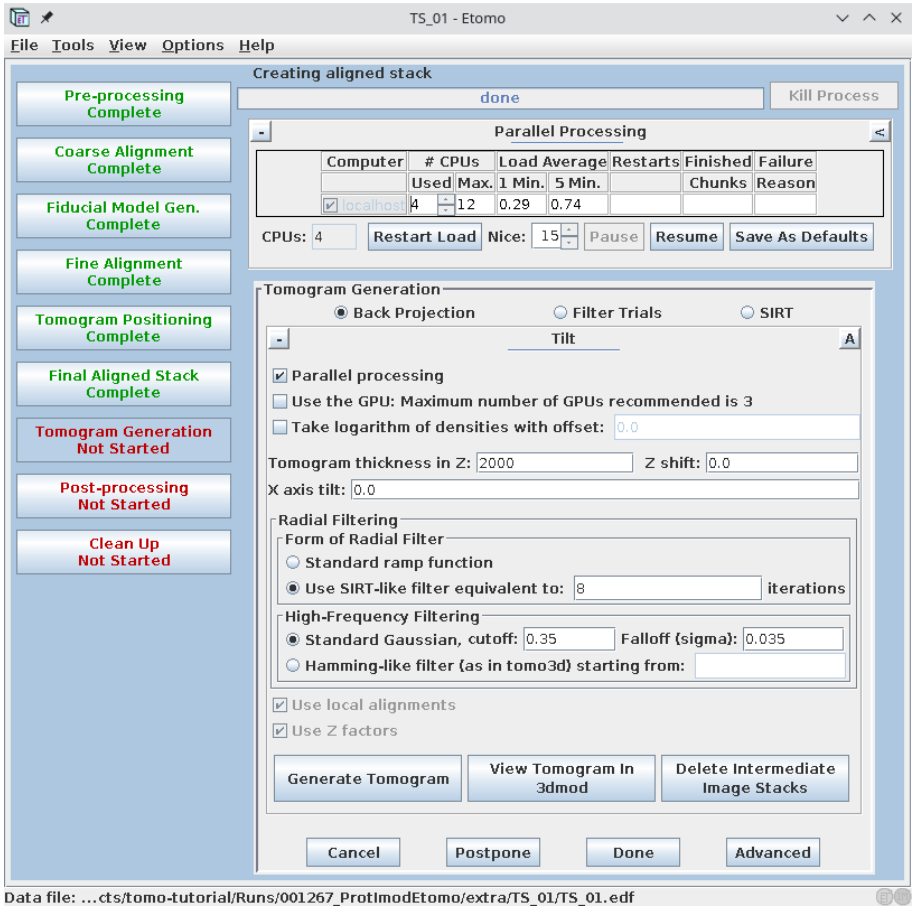

Figure 19: Creating a bin 4 tomogram with the SIRT-like filter.

Post-processing

1. Uncheck **Convert to bytes**.

2. Select **Rotate around X axis**. This will orient the tilt axis vertically while preserving the handedness.

3. Press Trim Volume. Display the trimmed volume if you like.

4. Press Done when finished.

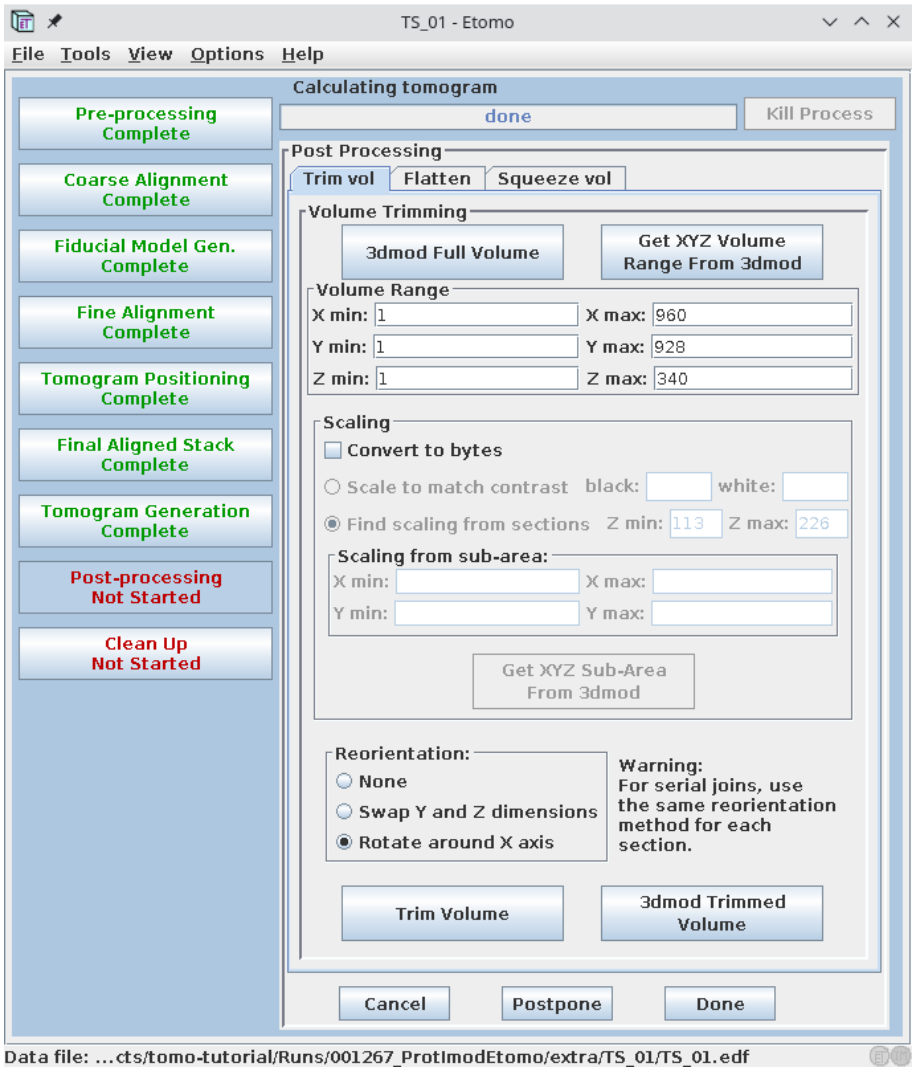

Figure 20: Post-processing the reconstructed tomogram.

Clean-up

Since deleting some intermediate files might make it impossible for Scipion to generate some of the outputs, we will skip this step. Just press Done and close *eTomo* window.

If you now look at the Scipion viewer dialog that stayed open while you used *eTomo*, you will notice that all steps for the first tilt-series have been completed and the first row has become green.

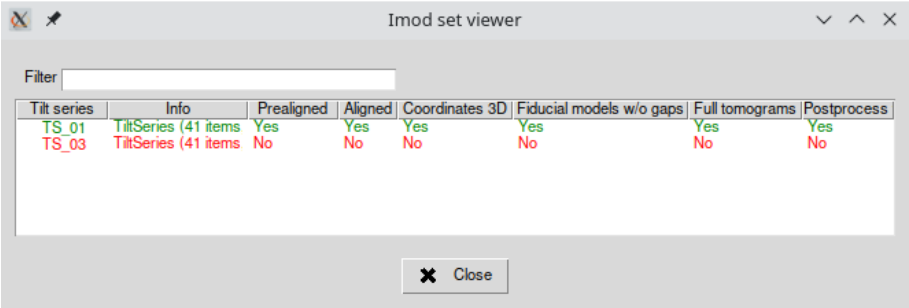

Figure 21: The SCIPION tilt-series viewer registers all steps performed with *eTomo* GUI.

EXERCISE

Repeat the eTomo workflow for the second tilt-series. Don't forget to exclude the dark view #1 during the processing.

Now, close both *eTomo* GUI and the tilt-series dialog. SCIPION will now register the following outputs for each processed tilt-series:

- Prealigned tilt-series (coarse aligned stack)
- Aligned tilt-series (final aligned and rotated stack)
- Set of 3D coordinates of the fiducials
- Set of landmark models that include fiducials positions and residuals
- Raw tomogram
- Post-processed (rotated) tomogram

All these outputs may be now visualized with the installed viewers and/or connected to other SCIPION protocols continuing the workflow.

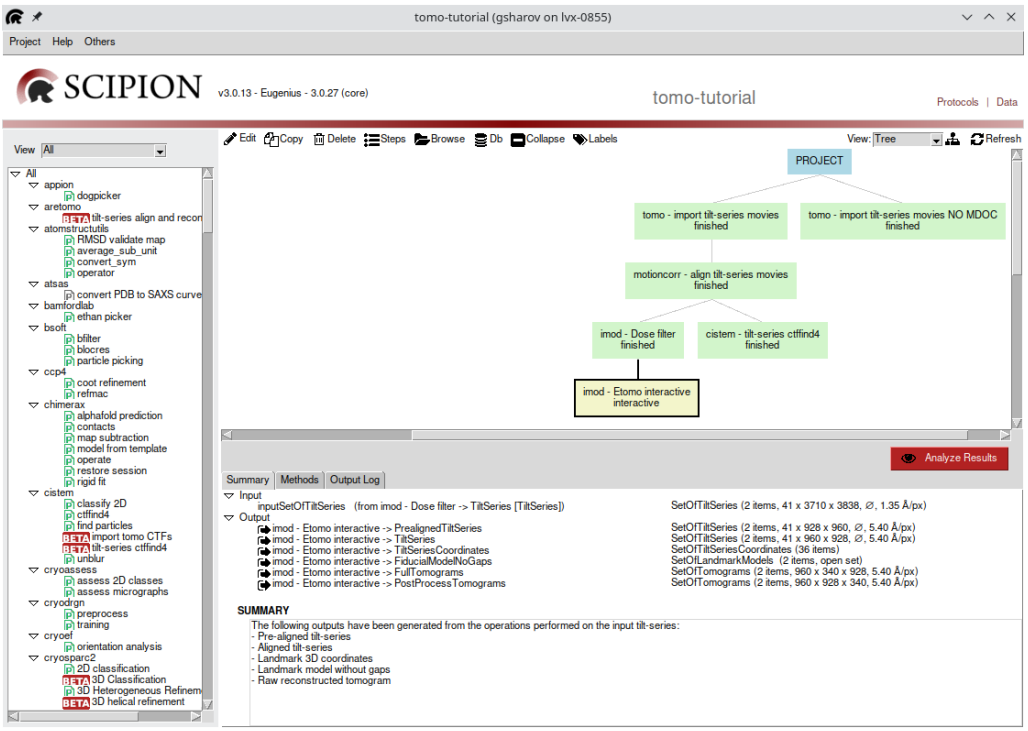

Figure 22: Outputs from *eTomo* interactive protocol.

5.2 Semi-automated processing with IMOD protocols

In this section we describe an alternative workflow that can replace *eTomo* interface with a group of separate Scipion protocols. Note that not all functions available from *eTomo* GUI have been wrapped into individual protocols.

1. Find and execute imod - X-rays eraser protocol on the dose-weighted tilt-series.

2. If you now display both output tilt-series with *3dmod*, you may notice that the first tilt image (-60 deg.) is completely dark. Thus we need to exclude it from further processing. Find imod - exclude views protocol and input the tilt-series from the previous step. Select Expert level: Advanced. Now, create a text file containing two lines:

   ```
   TS_01 1
   TS_03 1
   ```

   The first column is tilt-series ID, while the second contains a list of views to be removed. Save and input this file as **Exclude views file** into the protocol and then execute. The output tilt-series should now contain 40 images each.

3. Next, let's downsample our tilt-series to speed up the processing. Find the imod - tilt-series preprocess protocol and input the tilt-series from the previous step. Use a **binning of 4.0**.

4. Afterwards, run imod - Xcorr prealignment protocol using the previous output. Set **Generate interpolated tilt-series? Yes** and input **Binning of 1**. This protocol will generate two output tilt-series: non-interpolated (with assigned alignment) and interpolated (aligned). You can display the interpolated output with *3dmod* viewer to check if the alignment was successful.

5. Next, run the imod - generate fiducial model protocol using the non-interpolated tilt-series from the step above, with options **Find on two surfaces? No** and **Fiducial radius: 10 nm**. Select Expert level: Advanced and also input the **Number of fiducials: 30**. After execution, the protocol will create a *SetOfLandmarkModels* which corresponds to a model containing fiducials positions and their residuals. You can display the output fiducial model with *3dmod*.

6. Continue the workflow with the imod - fiducial alignment protocol, using the set of landmark models from the previous step as input. On the *Input* tab also set **Find on two surfaces? No**, **Generate interpolated tilt-series? Yes** and **Binning of 1**. On the *Global variables* tab choose **Solve for one rotation**, leaving other parameters fixed. SCIPION will produce several outputs here, but you might want to check only the interpolated (aligned) tilt-series.

7. Finally, locate the imod - tomo reconstruction protocol. Input a set of non-interpolated tilt-series from the step above and set **Tomogram thickness to 350 px** and **Z-shift to 75 px**. Select Expert level: Advanced and set **Iterations of a SIRT-like equivalent filter: 8**. If you have gone through the *eTomo* section above, you know that the optimal tomogram thickness for this data is about 1400 px. Here we use $1400/4 = 350$ because we binned our tilt-series by a factor of 4. If you do not have any GPUs, set GPU IDs to **No**.

In the end, your workflow should look something like this:

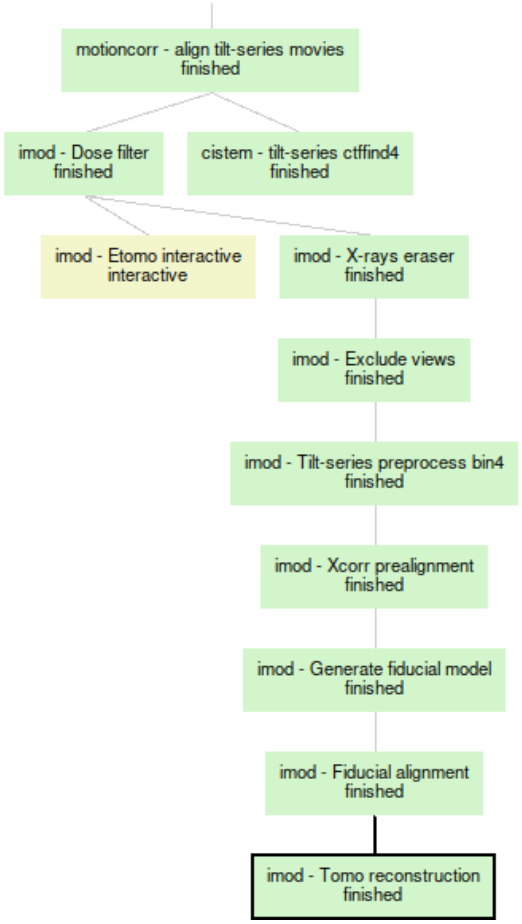

Figure 23: Project workflow for semi-automated processing with *IMOD*.

5.3 Automated processing with AreTomo

AreTomo (Zheng et al., 2022) offers an automated marker-free tilt-series alignment and reconstruction. The program is accelerated on GPU and can produce aligned tilt-series and/or a reconstructed volume. Users can choose either weighted back projection (WBP) or simultaneous algebraic reconstruction technique (SART) to reconstruct their tomograms. *AreTomo* implements both global and local alignments. The global alignment determines the tilt angle offset, translations of the tilt images, and orientations of the tilt axis as it varies throughout the tilt series. Local alignment can be used to correct for local motions due to the progressive sample deformation under repeated beam exposures. Moreover, the software can apply dose-weighting filter to aligned tilt-series or reconstructed tomogram.

Find `aretomo - tilt-series align and reconstruct` protocol and fill in the parameters listed below:

- **Input set of tilt-series**: provide the set of dose-filtered tilt-series
- **Reconstruct tomogram?** Yes

- **Binning**: 8
- **Volume height (voxels)**: 1800. This is in unbinned voxels (!)
- **Tomogram thickness (voxels)**: 2000. This is in unbinned voxels (!)

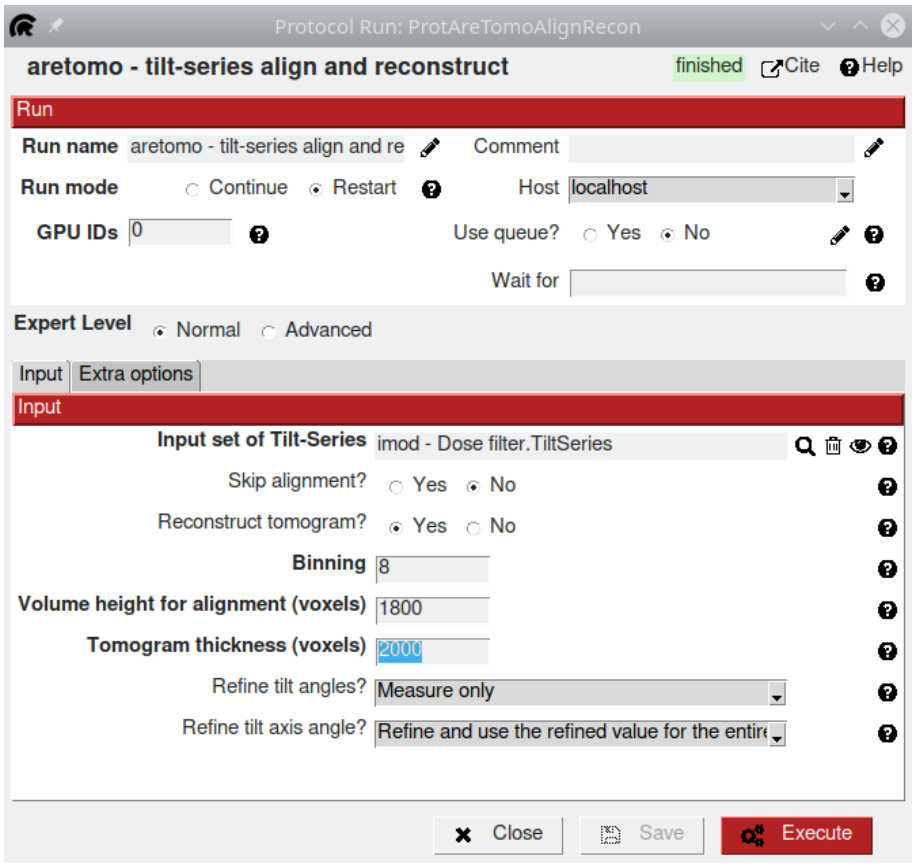

Figure 24: AreTomo protocol is used to align and reconstruct dose-filtered tilt-series.

Do not change other options and execute the protocol. At this point we intentionally chose a large thickness value and high binning factor in order to estimate the correct tomogram thickness. Upon completion, the protocol has produced two outputs: non-interpolated tilt-series (the original ones with alignment metadata) and a set of tomograms. Right click on the *outputSetOfTomograms* output and choose *Open with ImodViewer*. Double click on the first tomogram (TS_01) to launch *3dmod* interface.

To check the tomogram thickness you need to run 3dmod's Slicer: press \ key or choose Image > Slicer in the *3dmod* menu. In the new Slicer window set **X rotation to 90** and **Img to 30** to increase the slicing thickness. This will display x-z slice. Scroll through the tomogram using **View axis position** slider. The current binned tomogram thickness is 2000/8=250 px. Draw a rectangle (use Shift + B) encompassing sample and all beads to estimate the specimen thickness. The rectangle dimensions will be displayed instead of Lo/Hi buttons in the Slicer window

when drawing. Remember that the pixels are 8x binned! Our estimate is 140 px * 8 = 1120 px.

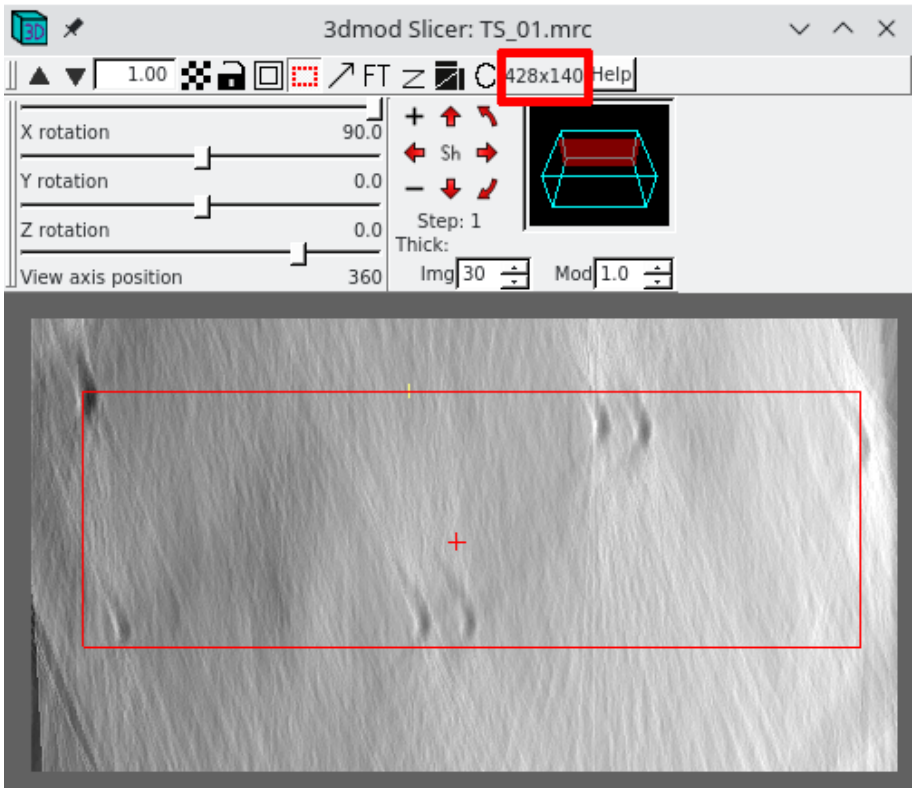

Figure 25: Estimating tomogram thickness using Slicer in *3dmod*. Rectangle dimensions are indicated with a red box.

Now let's re-run the protocol using the newly estimated thickness value and then verify tilt-series alignment. Right click on the completed protocol box and select **Edit**. This time input **binning 4**, volume height of **1120 px** and tomogram thickness of **1300 px**. Once the run is finished, double check the output tomogram again with *3dmod*.

To verify the tilt-series alignment we can use the non-interpolated tilt-series output (bin 1) that has alignment metadata. Find imod - apply transformation matrix protocol and input the set of tilt-series from the *AreTomo* protocol. Set **Binning to 4** and press Execute.

TIP

AreTomo protocol can generate aligned interpolated tilt-series when you choose not to reconstruct a tomogram. This can also be used to quickly check the accuracy of alignment. You might notice that interpolated tilt-series have less tilt images than the input, this is due to the fact that *AreTomo* automatically finds and discards dark tilt images. The parameter that controls this behaviour (**Dark tolerance**) can be found on the *Extra options* tab of the protocol.

Once the *IMOD*'s protocol run is completed, let's check the tilt-series alignment. Right click on the *InterpolatedTiltSeries* output and choose *Open with ImodViewer*. Double click on the first item (TS_01) to launch *3dmod* interface. Middle click on the image (ZaP) window to scroll though aligned tilt-series. The aligned stack is displayed as x-y view with the vertical tilt axis. If the alignment has been successful, the gold beads should only move horizontally (perpendicular to the tilt axis) without any vertical jumps. You might also notice that the first dark tilt image has been removed from these tilt-series, since it was detected by *AreTomo* and marked as disabled in the non-interpolated set.

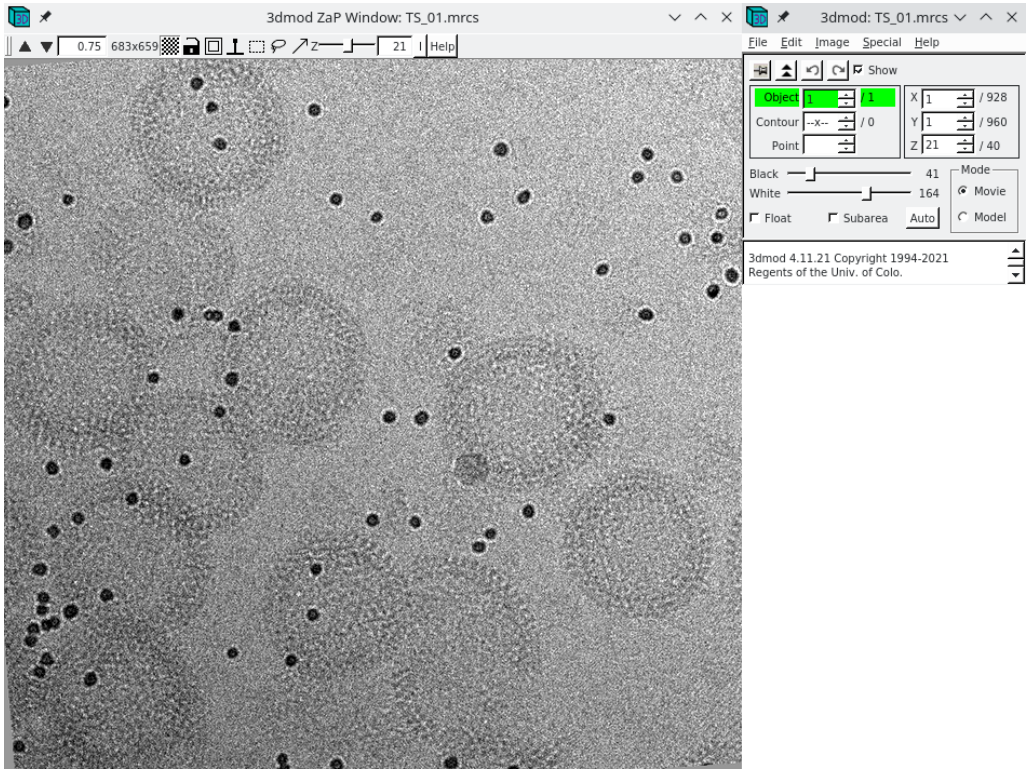

Figure 26: Displaying aligned tilt-series with *3dmod* interface.

ALTERNATIVES

- emantomo - align tilt series
- emantomo - tomo reconstruction
- imod - tomogram reconstruction
- tomo3d - tomo3D

6 CTF correction

CTF correction in cryo-ET is an active research area. Conventionally, CTF estimation and correction is performed on tilt-series before tomogram reconstruction. In the **2D strip-based** approaches, implemented e.g. in *IMOD*, each tilt image is divided into strips parallel to the tilt axis. Defocus for the central strip is estimated by averaging over overlapping tiles and then extrapolated to the other strips using a linear relationship depending on the distance from the tilt axis and the tilt angle. Each strip is then CTF-corrected by phase flipping. Corrected aligned stack of tilt-series is then used for tomogram reconstruction. This method only considers the defocus gradient caused by tilting the sample.

3D CTF-correction approaches additionally consider the sample thickness, that creates another defocus gradient along the direction of electron beam. In the case of *NovaCTF* (Turoňová et al., 2017) during the reconstruction of a tomogram by back-projection, each voxel is calculated from tilt images that were CTF-corrected with defocus values corresponding to the position (given by x and z coordinates) of that voxel at each tilt. To achieve this, each image in the tilt-series is CTF-corrected multiple times with different defocus value. The number of different CTF corrections performed per image depends upon how finely the defocus gradient should be sampled and is a user-defined parameter.

CTF correction can also be performed at **subtomograms level**, taking into account their known 3D positions inside a tomogram volume. Such methods have been implemented in *Relion* and *EMAN2*. However, this imposes additional challenges of dealing with missing wedge, extra interpolations etc. *Relion* 4.0 has now implemented a different concept of pseudo-subtomograms (Zivanov et al., 2022) comprised of 3D arrays of CTF pre-multiplied 2D tilt-series images and other auxillary data.

More recently, several high-resolution subtomogram averaging pipelines have been derived using **per-particle per-tilt averaging** algorithms (e.g. *EMAN2*, *emClarity* or *M*). In case of *EMAN2* the tomograms are only used for particle picking and once the coordinates are known, subtilt series are extracted from the original tilt series, one for each particle. These subtilt series come with CTF information based on where each particle is located in a 3D volume, independent for each tilt in the subtilt series. This information is used for CTF correction during subtomogram

averaging which essentially becomes subtilt averaging.

It's important to note that all CTF correction methods mentioned above depend on the accuracy of the CTF estimation. In many cases *CTFFind4 or gCTF* are used, that were originally designed for single-particle analysis. The low signal-to-noise ratio of tomographic tilt images makes estimation of the CTF less robust and precise than with single particle data, as the power spectra of higher tilt images show fewer Thon rings for fitting. An error in the defocus estimation of 250 nm limits the resolution to 10 Å, and an error of 63 nm limits the resolution to 5 Å (Kudryashev, 2017).

6.1 2D CTF correction with IMOD

In this tutorial we will explore both 2D and 3D CTF correction methods. Let's start with the first one. If we look at the *CTFFind4* estimation that we executed at the beginning of the analysis, we can notice two issues: a) high tilts have wrong defocus estimates, b) we did not remove the first dark image of each tilt-series before running the estimation. In principle, one could use imod - exclude views protocol to remove bad images and then re-run CTF estimation. Let's do this and also explore *IMOD* protocols for CTF estimation.

1. Find and execute imod - exclude views protocol. Input motion-corrected tilt-series from *motioncor2* and the text file with excluded views numbers that we have already used before.

2. Create a text file containing two lines as below. We will use this as defoci list for the two tilt-series:

```
TS_01 4000.0
TS_03 1500.0
```

3. Locate imod - Automatic CTF estimation protocol and change its input parameters: a) input motion-corrected tilt-series with dark views excluded, b) **Expected defocus file: the file created above**, c) autorefinement **Angle step: 0**, d) **Search astigmatism? Yes**. Using a value of zero for the angular step will force *IMOD* to fit CTF to each single image separately.

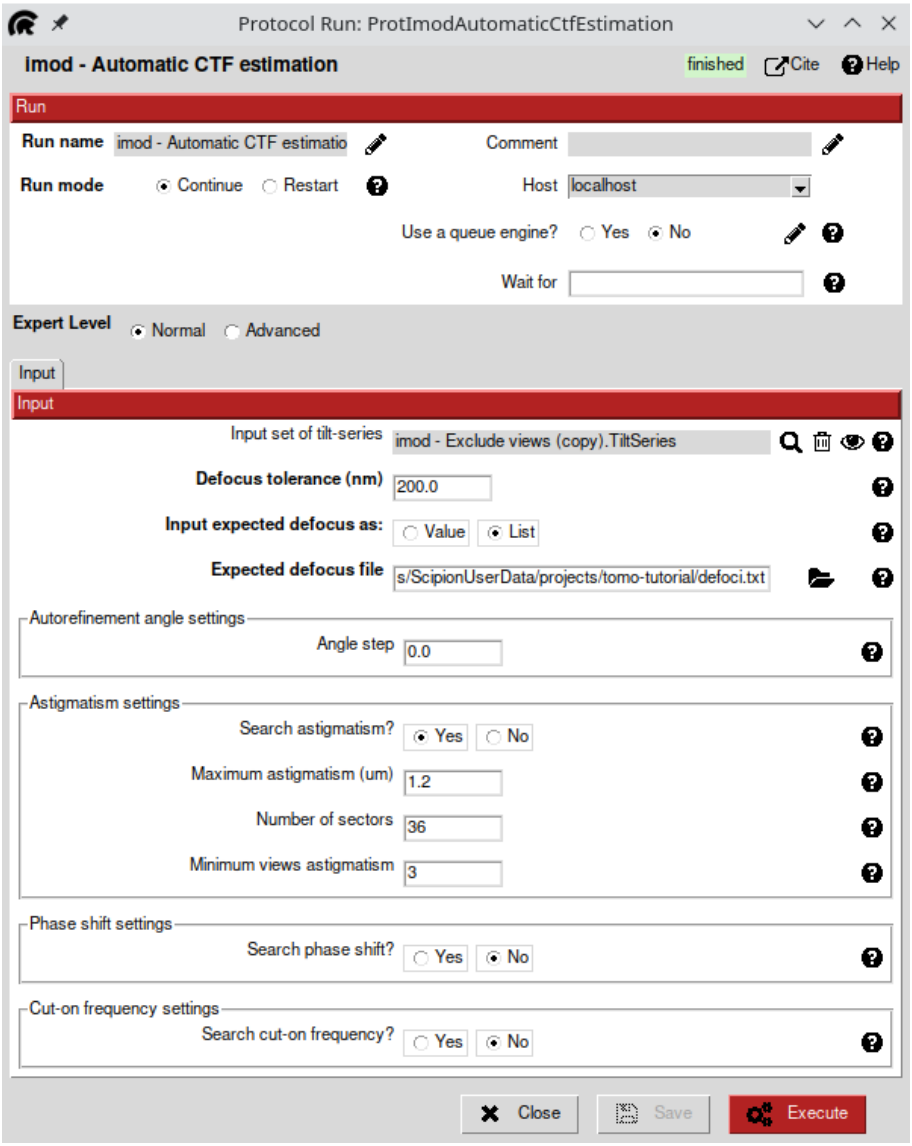

Figure 27: Automatic CTF estimation with IMOD.

4. If we now analyse the results of this CTF estimation, we can notice that high tilt views no longer deviate from the rest of tilt-series. We will now continue with these values, though you could run manual CTF estimation with *IMOD* instead to verify the 2D CTF fit for each image.

5. Next, we will execute imod - CTF correction protocol using the CTF estimation above and non-interpolated tilt-series (bin 4) from imod - fiducial alignment protocol. Non-interpolated tilt-series have alignment information and the estimated astigmatism will be rotated to match the tilt axis alignment automatically by *IMOD* protocol. We also made sure that in our workflow the excluded views are the same so that the tilt-series can be matched with the CTF estimation.

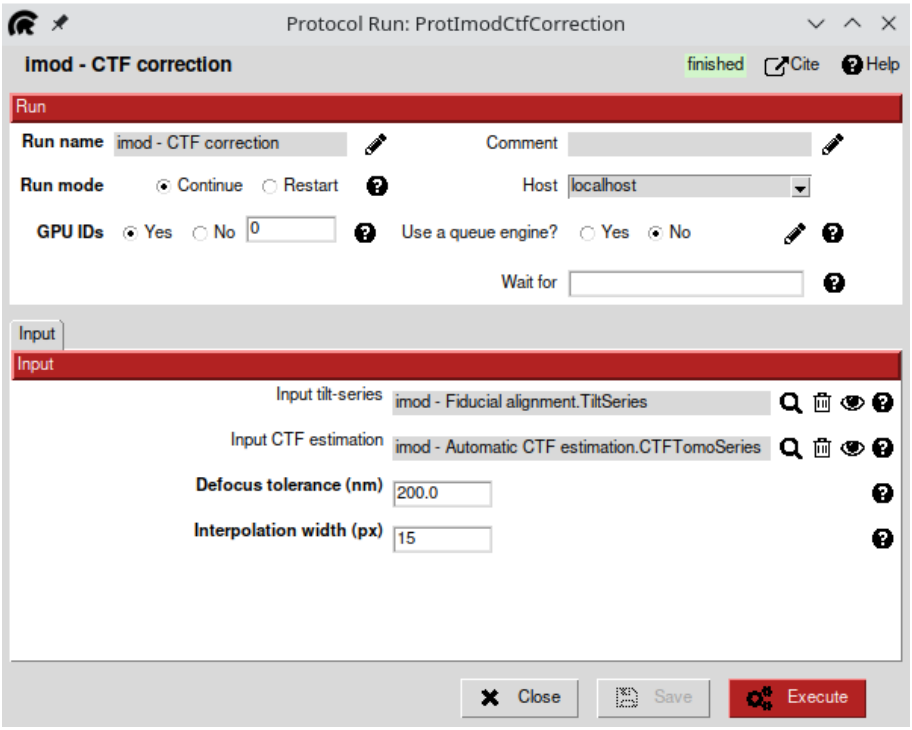

Figure 28: 2D strip-based CTF correction in IMOD.

6. Finally, CTF corrected and aligned tilt-series can be reconstructed with imod - tomo reconstruction protocol. In the end, you project workflow should look something like below.

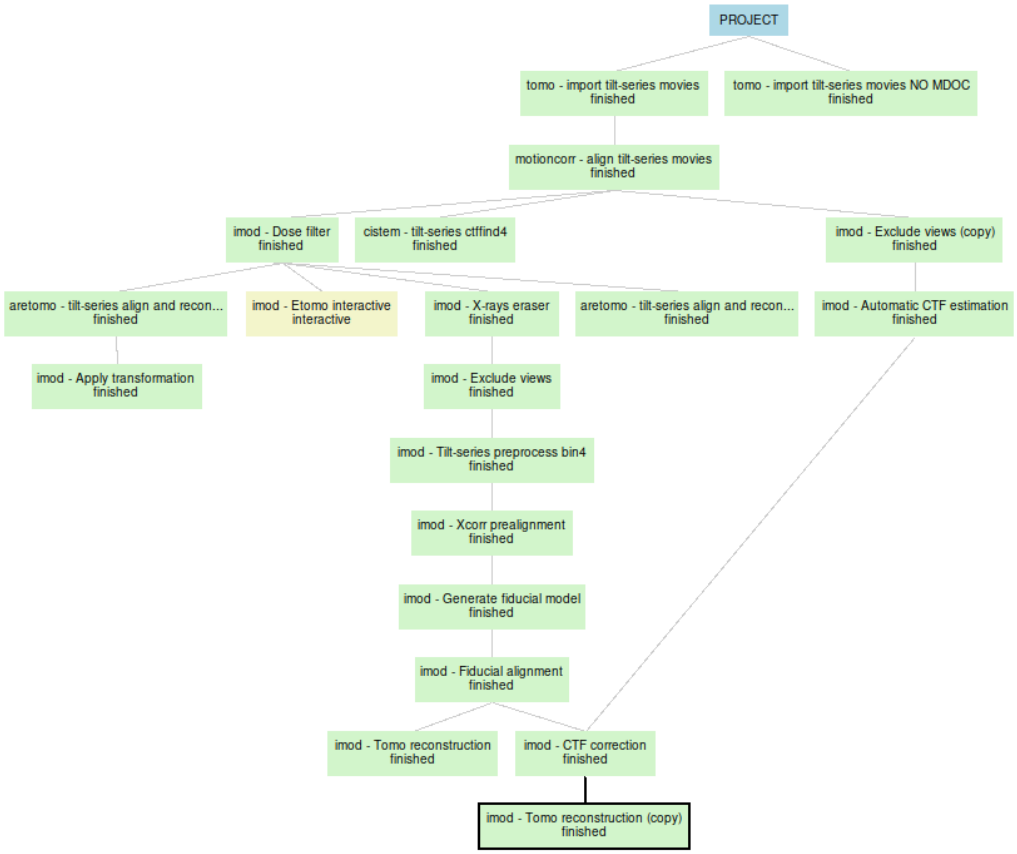

Figure 29: Project workflow after CTF correction.

6.2 3D CTF correction with NovaCTF

3D CTF correction workflow with *NovaCTF* consists of two Scipion protocols: compute defocus array and 3D CTF correction and reconstruction. Find the first one and input there the non-interpolated tilt-series (bin 4) from imod - fiducial alignment protocol. Also input CTF estimation from the last *IMOD* run. For tomogram thickness use **350 px** with Z-shift of **75 px**. Instead of default recommended defocus step of 15 nm you can use **30 nm** to speed up processing. We will use phase flipping as CTF correction method, though multiplication can be used as well, albeit at a cost of reducing the tomogram visual quality. The protocol should finish very quickly as it only generates multiple defocus files used for CTF correction later.

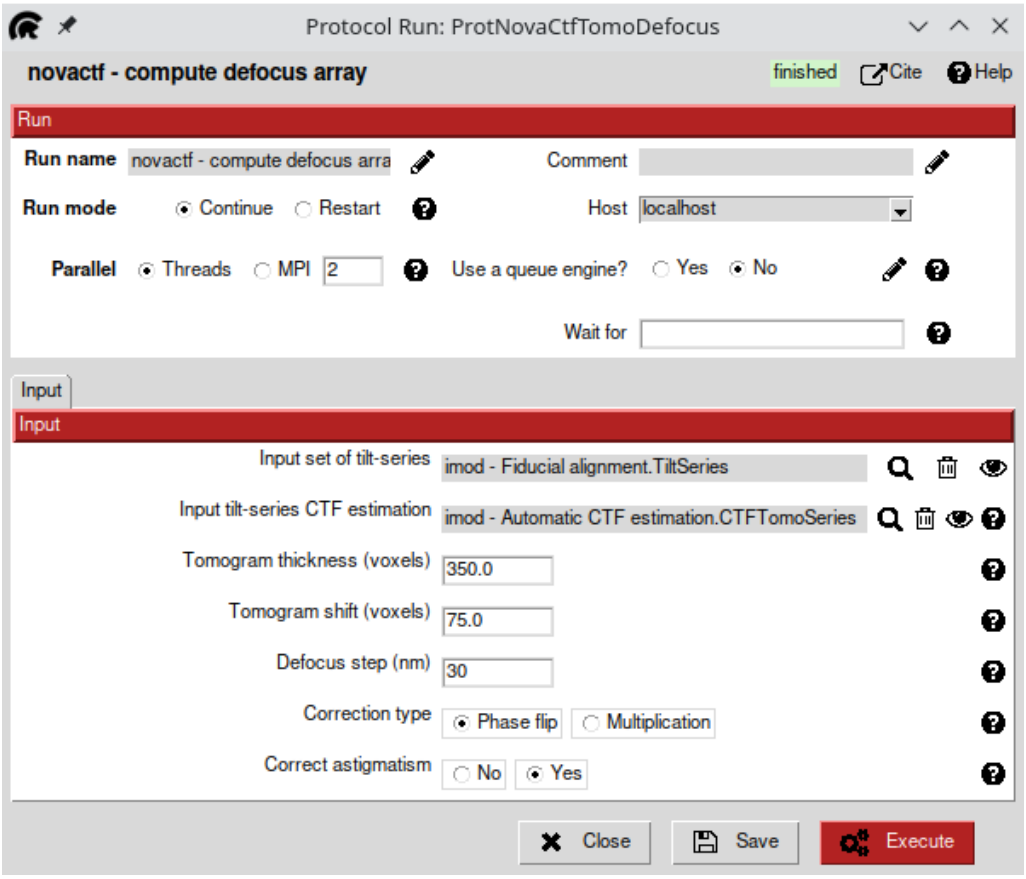

Figure 30: NovaCTF - compute defocus array protocol.

Now locate novactf - 3D CTF correction and reconstruction protocol and use the previous protocol as input. Set **Apply tilt-series alignment? Yes** and leave the other options with default values. Set **Number of threads** equal to the number of available CPU cores to parallelize steps execution. This protocol will perform CTF correction, apply alignment and filter input tilt-series N times, where N is the number of defocus files generated in the previous protocol. Afterwards, it will combine all stacks into a final CTF-corrected tomogram.

7 Wrapping up

That's it! If you have completed the tutorial and read and understood the guide up to this point, you should be ready for some serious tomo data processing using Scipion framework! If you found any mistakes in this guide or have more questions, get in touch with us. Below are the resources for further reading:

- Our paper on using Scipion for tomography (Jiménez de la Morena et al., 2022).
- Other Scipion tutorials for tomography

References

de la Rosa-Trevín, J., Quintana, A., del Cano, L., Zaldívar, A., Foche, I., Gutiérrez, J., Gómez-Blanco, J., Burguet-Castell, J., Cuenca-Alba, J., Abrishami, V., Vargas, J., Otón, J., Sharov, G., Vilas, J., Navas, J., Conesa, P., Kazemi, M., Marabini, R., Sorzano, C., and Carazo, J. (2016). Scipion: A software framework toward integration, reproducibility and validation in 3d electron microscopy. *J. Struc. Biol.*, 195(1):93 – 99.

Grant, T. and Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-em using a 2.6 Å reconstruction of rotavirus vp6. *eLife*, 4:e06980.

Jiménez de la Morena, J., Conesa, P., Fonseca, Y., de Isidro-Gómez, F., Herreros, D., Fernández-Giménez, E., Strelak, D., Moebel, E., Buchholz, T., Jug, F., Martinez-Sanchez, A., Harastani, M., Jonic, S., Conesa, J., Cuervo, A., Losana, P., Sánchez, I., Iceta, M., del Cano, L., Gragera, M., Melero, R., Sharov, G., Castaño-Díez, D., Koster, A., Piccirillo, J., Vilas, J., Otón, J., Marabini, R., Sorzano, C., and Carazo, J. (2022). Scipiontomo: Towards cryo-electron tomography software integration, reproducibility, and validation. *Journal of Structural Biology*, 214(3):107872.

Kremer, J. R., Mastronarde, D. N., and McIntosh, J. (1996). Computer visualization of three-dimensional image data using imod. *Journal of Structural Biology*, 116(1):71–76.

Kudryashev, M. (2017). Resolution in electron tomography. In *Biological and Medical Physics, Biomedical Engineering*, pages 261–282. Springer International Publishing.

Mastronarde, D. N. (2003). Serialem: A program for automated tilt series acquisition on tecnai microscopes using prediction of specimen position. *Microscopy and Microanalysis*, 9(S02):1182–1183.

Mastronarde, D. N. and Held, S. R. (2017). Automated tilt series alignment and tomographic reconstruction in imod. *Journal of Structural Biology*, 197(2):102–113. Electron Tomography.

Metskas, L. A., Wilfong, R., and Jensen, G. J. (2022). Subtomogram averaging for biophysical analysis and supramolecular context. *Journal of Structural Biology: X*, 6:100076.

Rohou, A. and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struc. Biol.*, 192(2):216–221.

Turoňová, B., Schur, F. K., Wan, W., and Briggs, J. A. (2017). Efficient 3d-ctf correction for cryo-electron tomography using novactf improves subtomogram averaging resolution to 3.4Å. *Journal of Structural Biology*, 199(3):187–195.

Zheng, S., Palovcak, E., Armache, J.-P., Verba, K., Cheng, Y., and Agard, D. (2017). Motioncor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods*, 14(4):331–332.

Zheng, S., Wolff, G., Greenan, G., Chen, Z., Faas, F. G., Bárcena, M., Koster, A. J., Cheng, Y., and Agard, D. A. (2022). Aretomo: An integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction. *Journal of Structural Biology: X*, 6:100068.

Zivanov, J., Otón, J., Ke, Z., von Kügelgen, A., Pyle, E., Qu, K., Morado, D., Castaño-Díez, D., Zanetti, G., Bharat, T. A. M., Briggs, J. A. G., and Scheres, S. H. W. (2022). A bayesian approach to single-particle electron cryo-tomography in RELION-4.0.