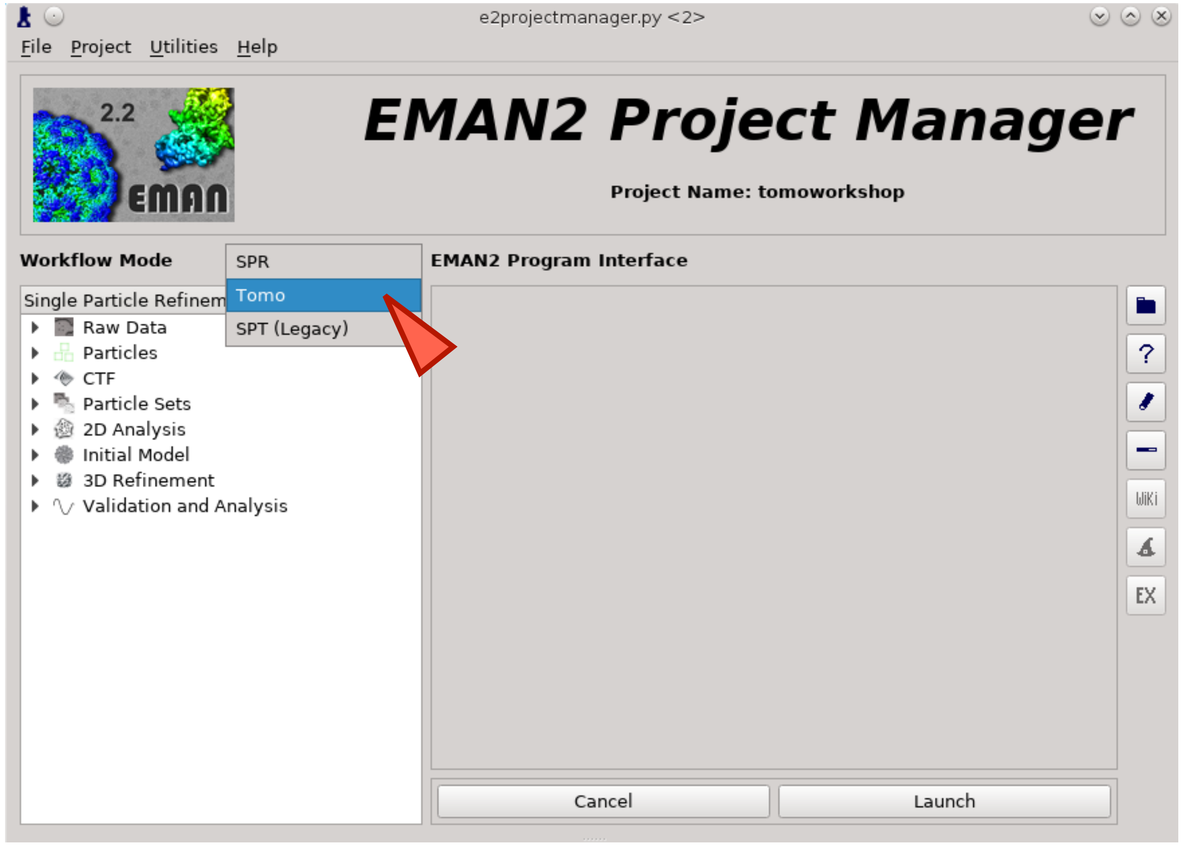
CryoET workflow

Muyuan Chen
2019-05

Prepare environment

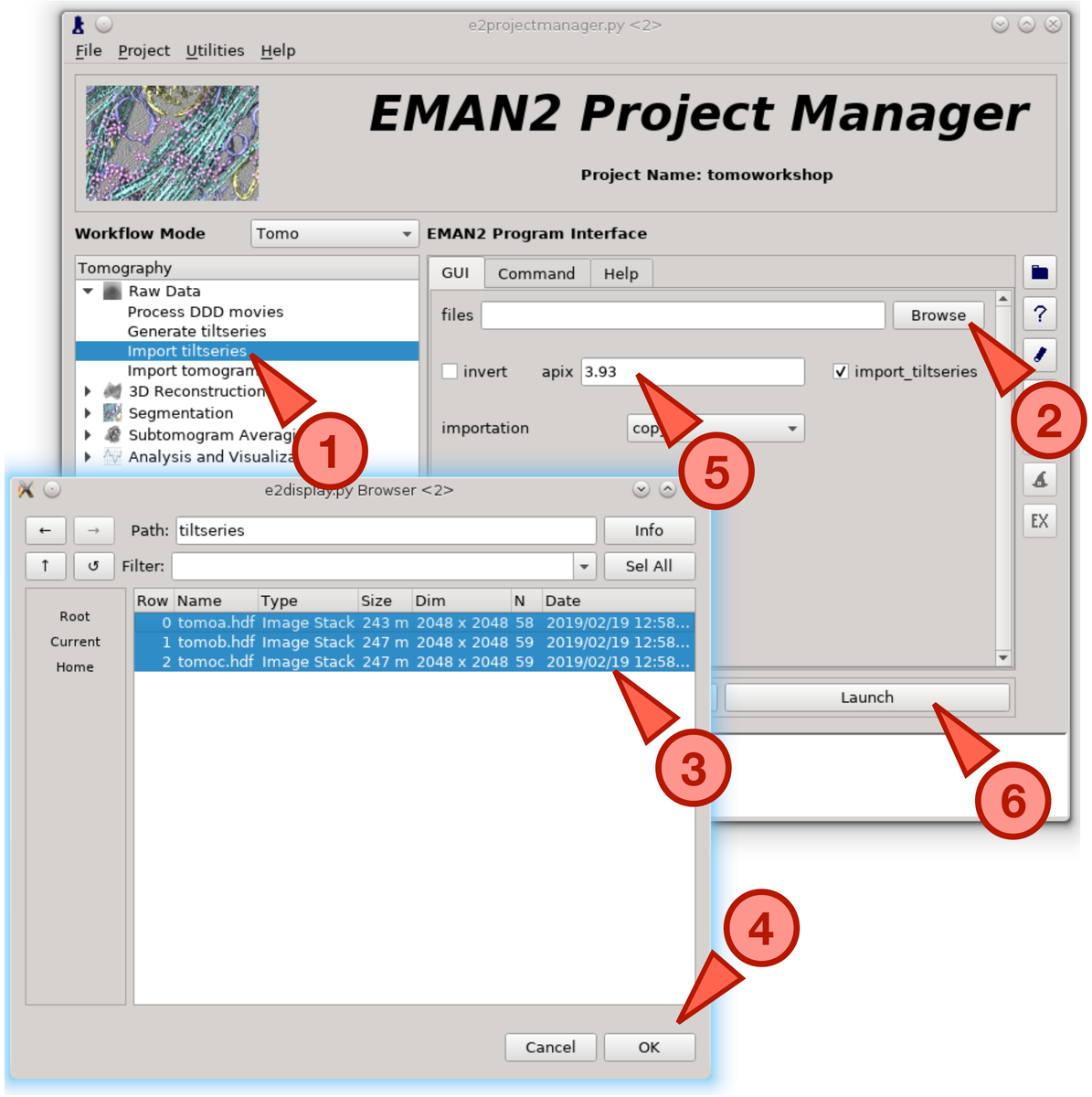
- Make a project folder
- *cd* into the project folder
- Run *e2projectmanager.py*
- Switch **Workflow Mode** *to* **Tomo**



Always run EMAN2 commands inside the project folder!

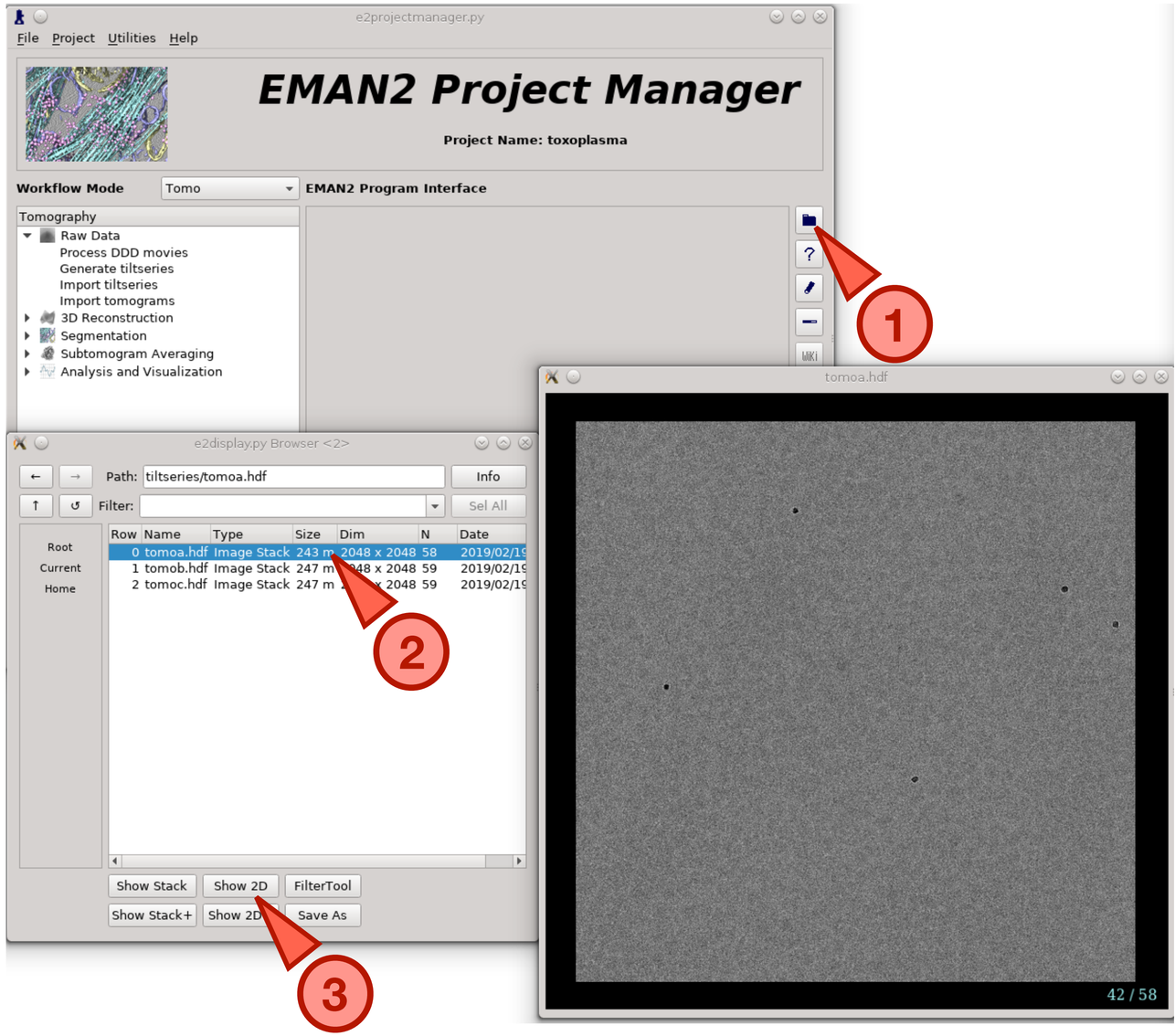
Prepare dataset

- Download from https://blake.bcm.edu/emanwiki/UTMBWorkshop2019
- Unzip the file.
- In the project manager, select **Raw Data** -> **Import tilt series**
- Click **Browse** next to the first box and select the unzipped files
- Set **apix** to 3.93
- Click **Launch**



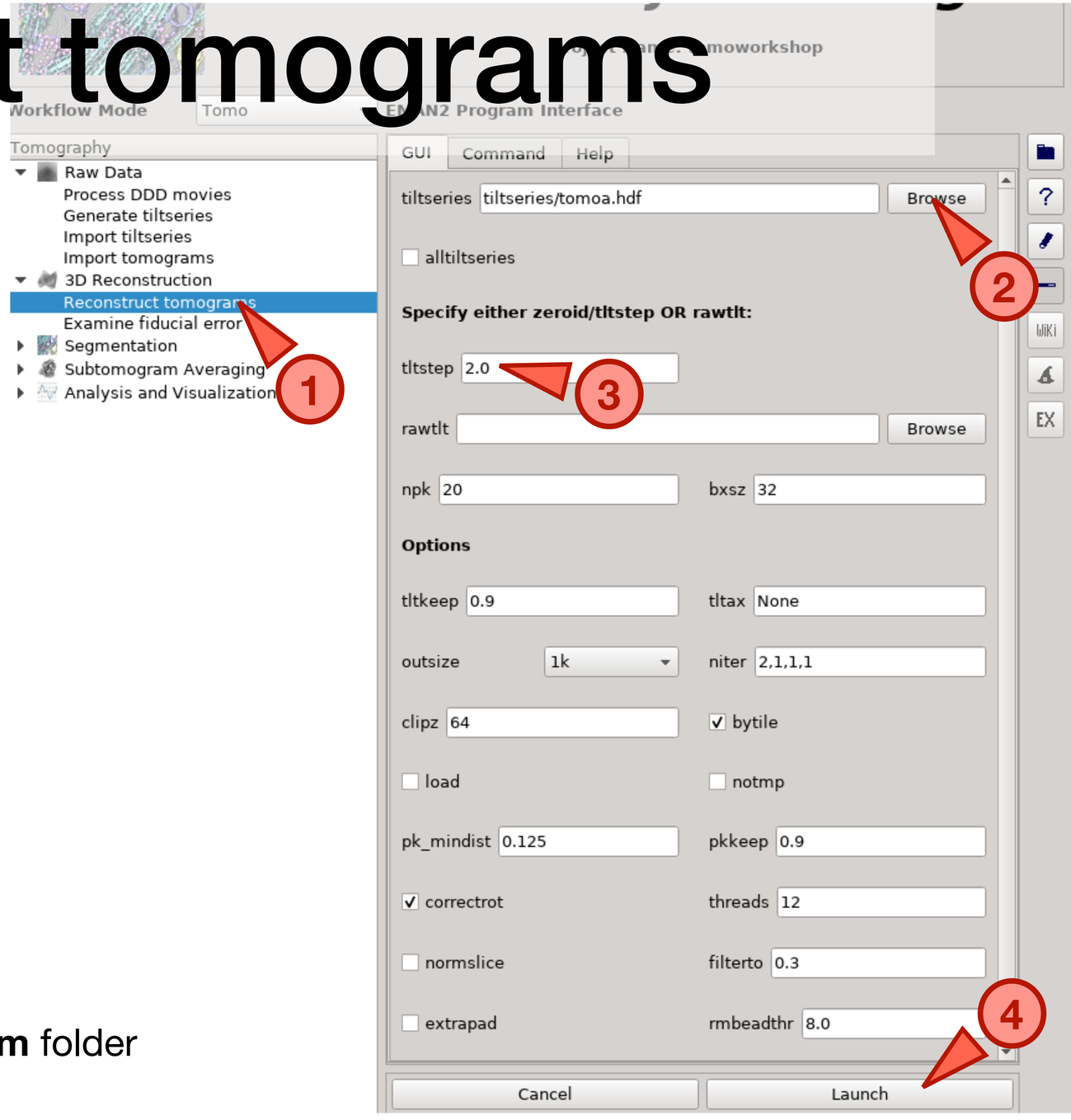
View the tilt series

- Click the Browse button in the project manager
- In the browser window, go to folder **tiltseries**
- Select a file, click **Show2D**
- Use up/down keys to move through images



Reconstruct tomograms

- **3D Reconstruction -> Reconstruct tomograms**
- In the first box, select the first tilt series
- Set **tltstep** to 2 (i.e. 2 degrees between tilts)
- Set **threads** to the number of CPUs on your machine
- Optional:
  - Set **clipz** to 64
  - For the first run, uncheck **notmp**
  - Check **correctrot**
  - Change **filterto** to 0.3
  - Set **rmbeadthr** to 8.0
- Click **Launch**
- Find the reconstructed tomogram inside the **tomogram** folder



How reconstruction works

- In the e2display browser, open the **tomorecon_00** folder

- This is an iterative process, for each iteration:

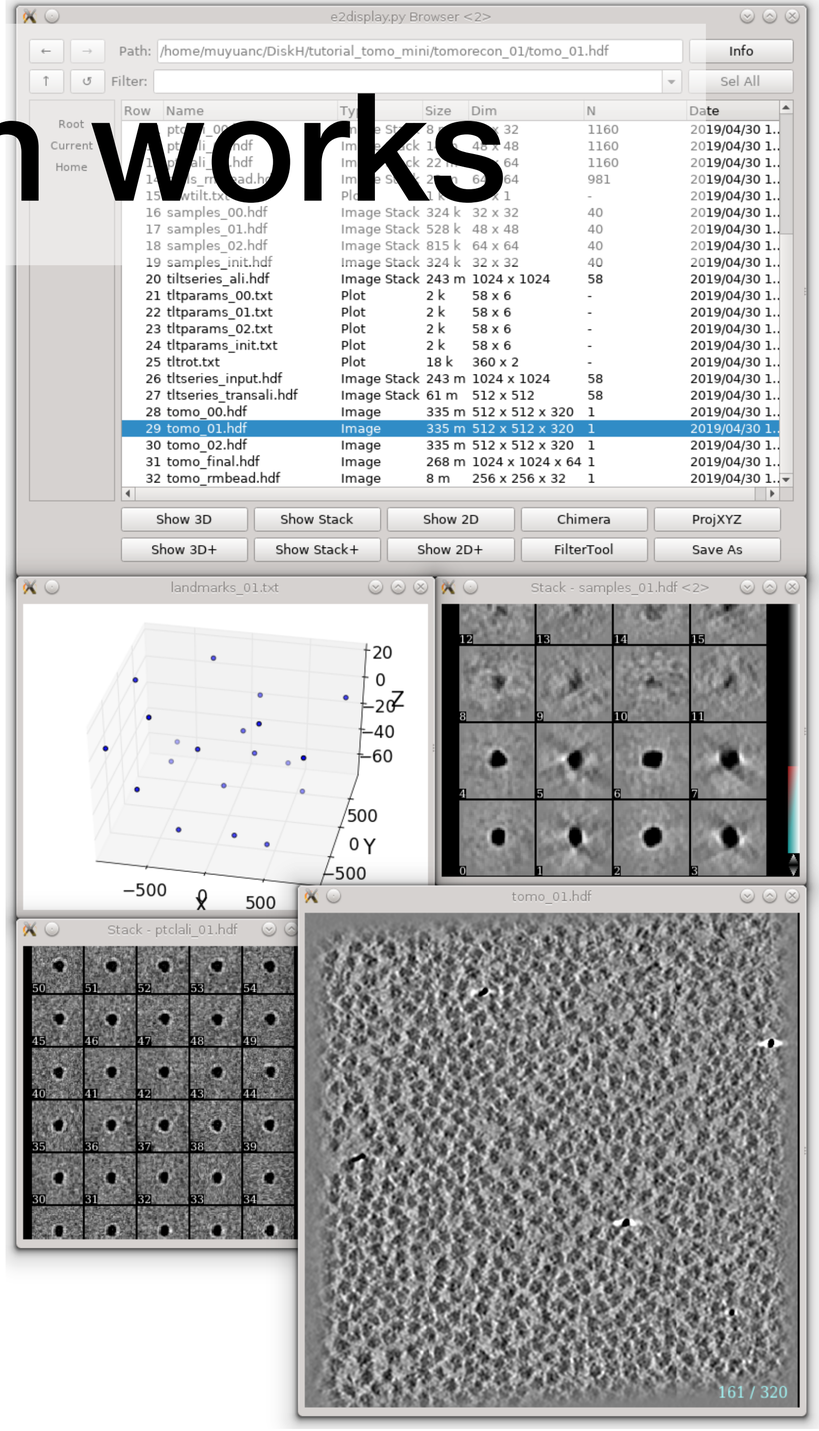
  - landmark_0X.txt : the location of the landmarks

  - samples_0X.hdf : the top and side view of those landmarks

  - ptclali_0X.hdf : the trace of each landmark throughout the tilt series

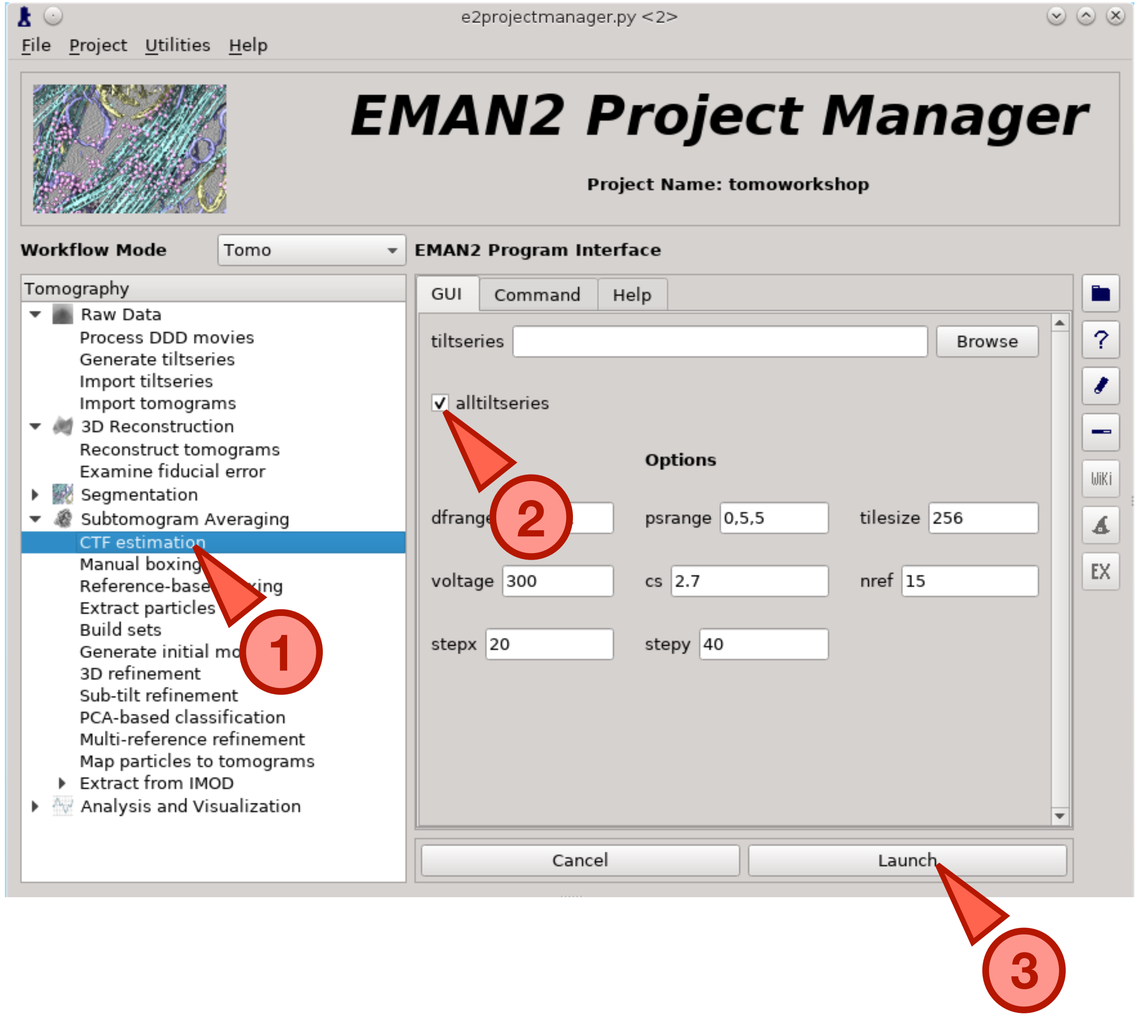
  - tomo_0X.hdf : the reconstruction after each iteration

- If you are satisfied with the results, go back to **Reconstruct tomograms**, check **alltiltseries** and **notmp**, to process all tilt series in the same way



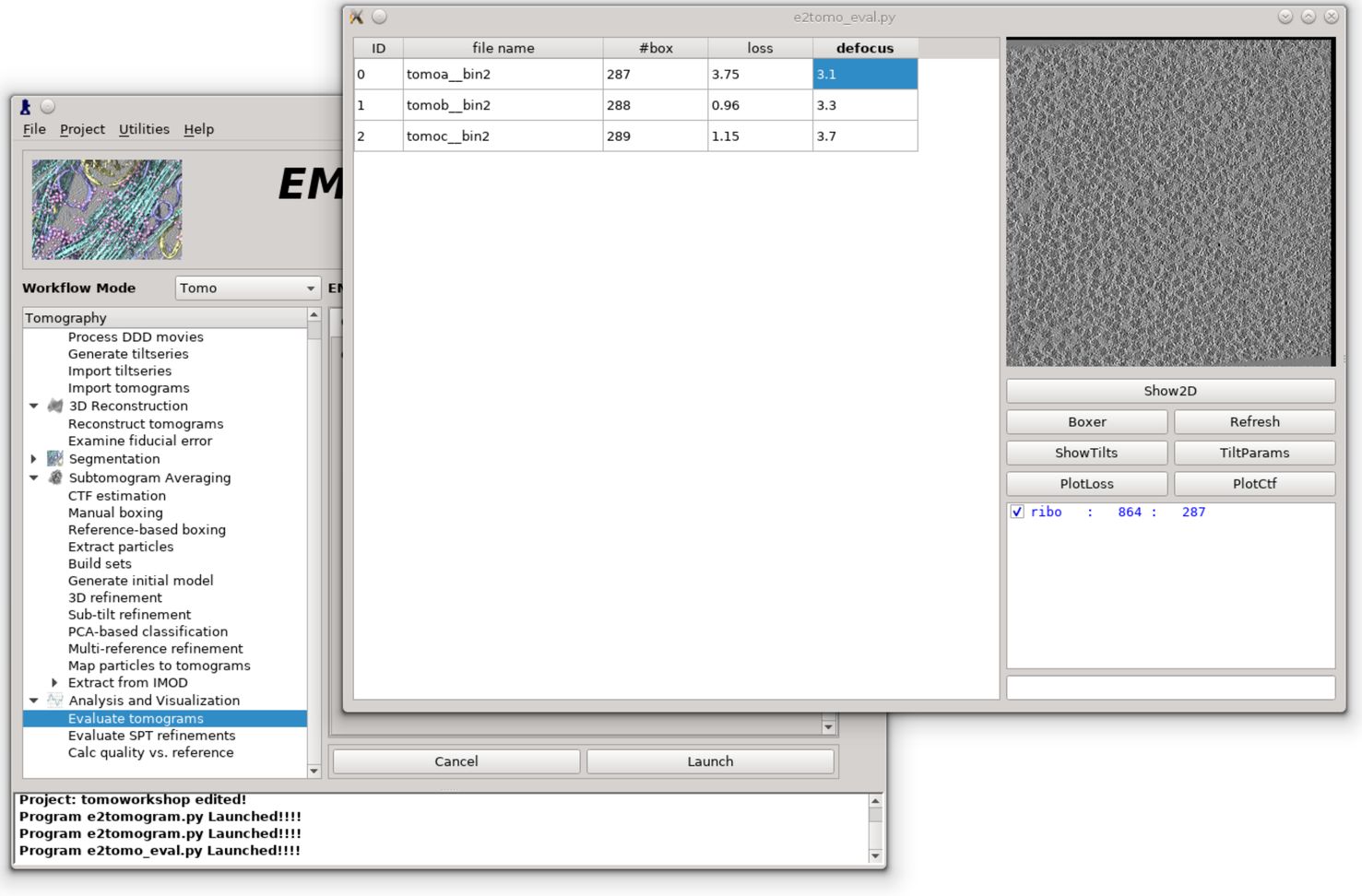
CTF estimation

- **Subtomogram averaging -> CTF estimation**
- Check **alltiltseries**
- Click **Launch**
- The theory behind this is complicated. There will be a session about CTF in more detail tomorrow…



Tomogram evaluation

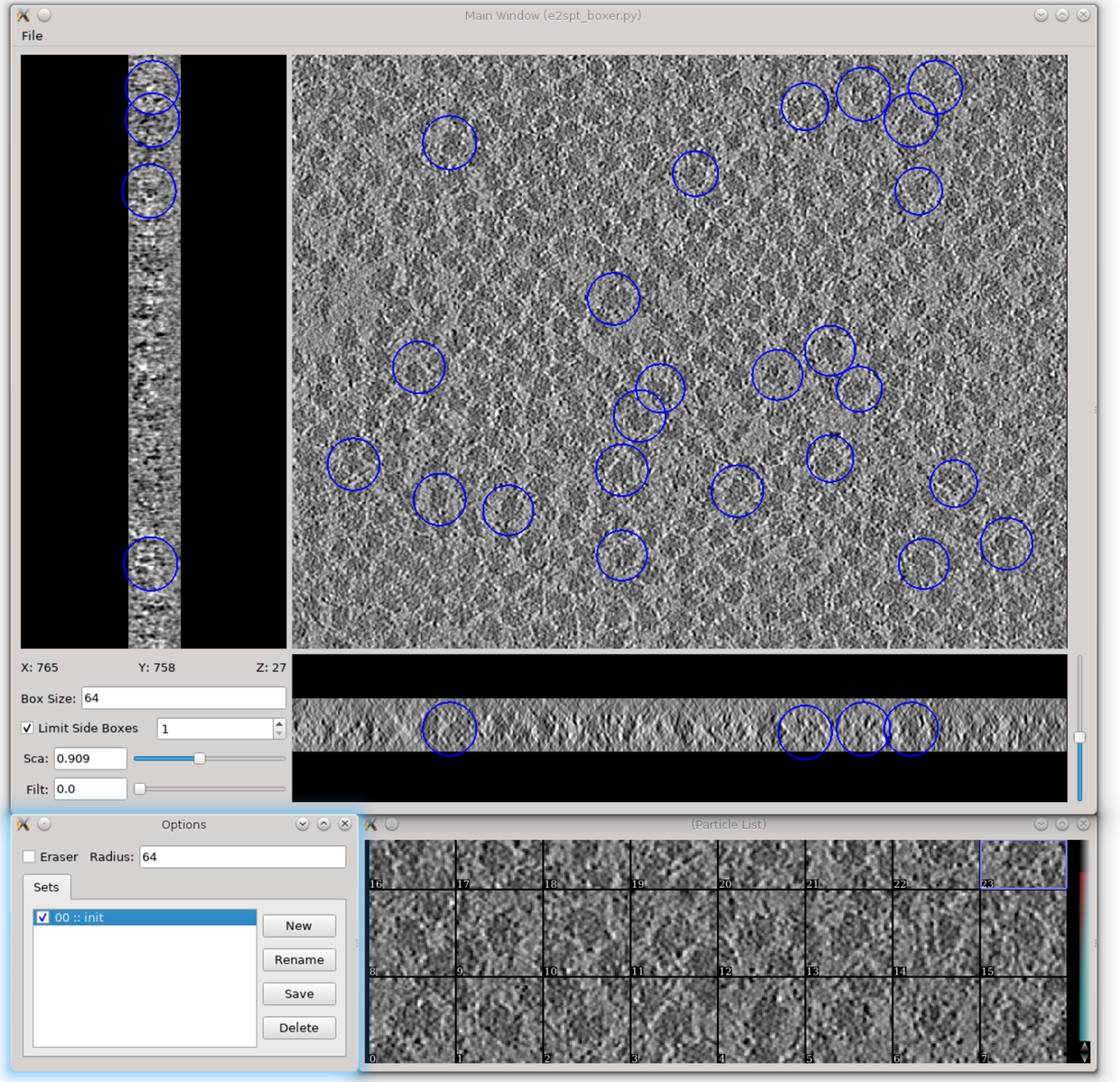
- **Analysis and Visualization -> Evaluate tomograms**
- Click **Launch**



This is more useful in large projects with many tomograms and multiple types of particle in each tomogram…

Select particles

- Select a tomogram in the **tomo_eval** window and click **Boxer**
  - You can also launch the boxer via **Subtomogram averaging -> Manual boxing** in the project manager
- In the new window, go through slices using '~' and '1' keys. Click to add a box, Shift+click to remove a box.
- Manage multiple types of particle in one tomogram using the **Sets** panel.
- At this step, we just need a few particles to make an initial model. Rename the particle set to "**init**", select 20-30 ribosome particles.
- Click **Save** in the set panel.



Generate initial model

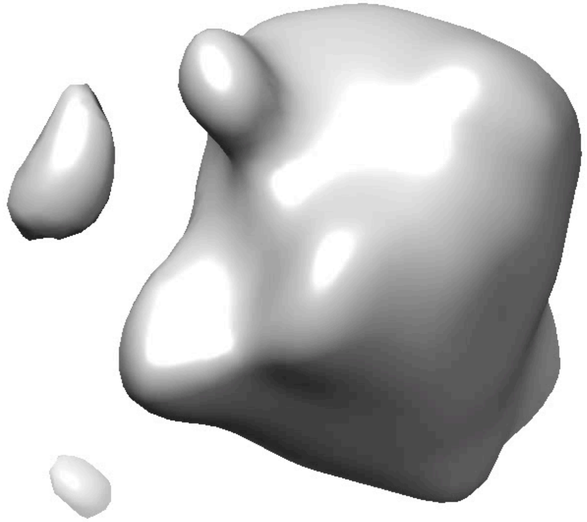
- First make sets from generated particles. In EMAN2, we combine particles of the same type from multiple tomograms into one virtual stack for further processing.
- **Subtomogram averaging -> Build sets**.
- Check **allparticles** and click **Launch**.
- **Subtomogram averaging -> Generate initial model**.
- Setting **shrink** to 2.
- Select the set we just generated (**init.lst**), and click **Launch**.

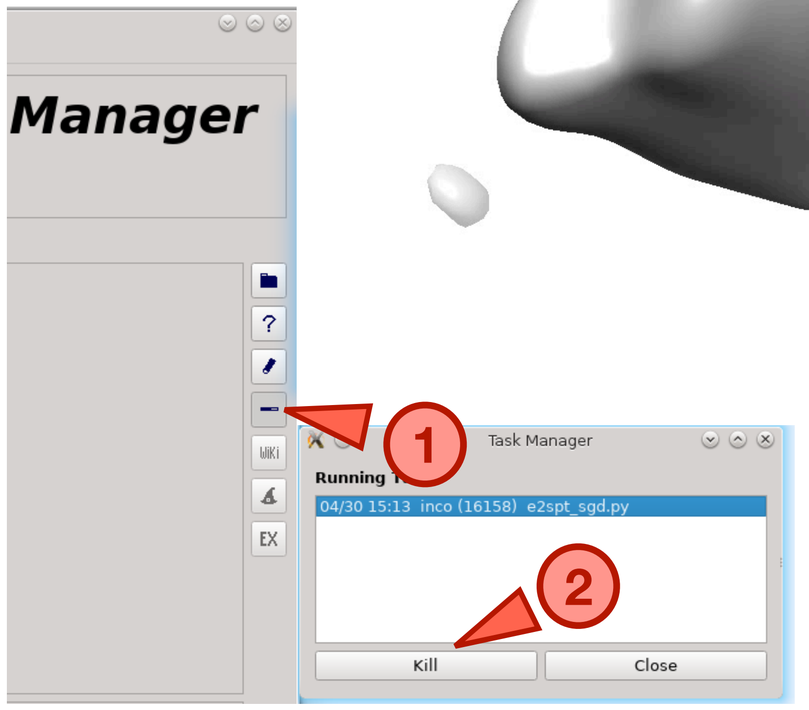
Generate initial model

- Look at the **sptsgd_XX** folder from the e2display browser.

- The file **output.hdf** will be continuously updated as the program is running. Check the file from time to time and terminate the program when the output looks reasonable (it should take 10-30 mins). Letting it run to finish does not hurt either…

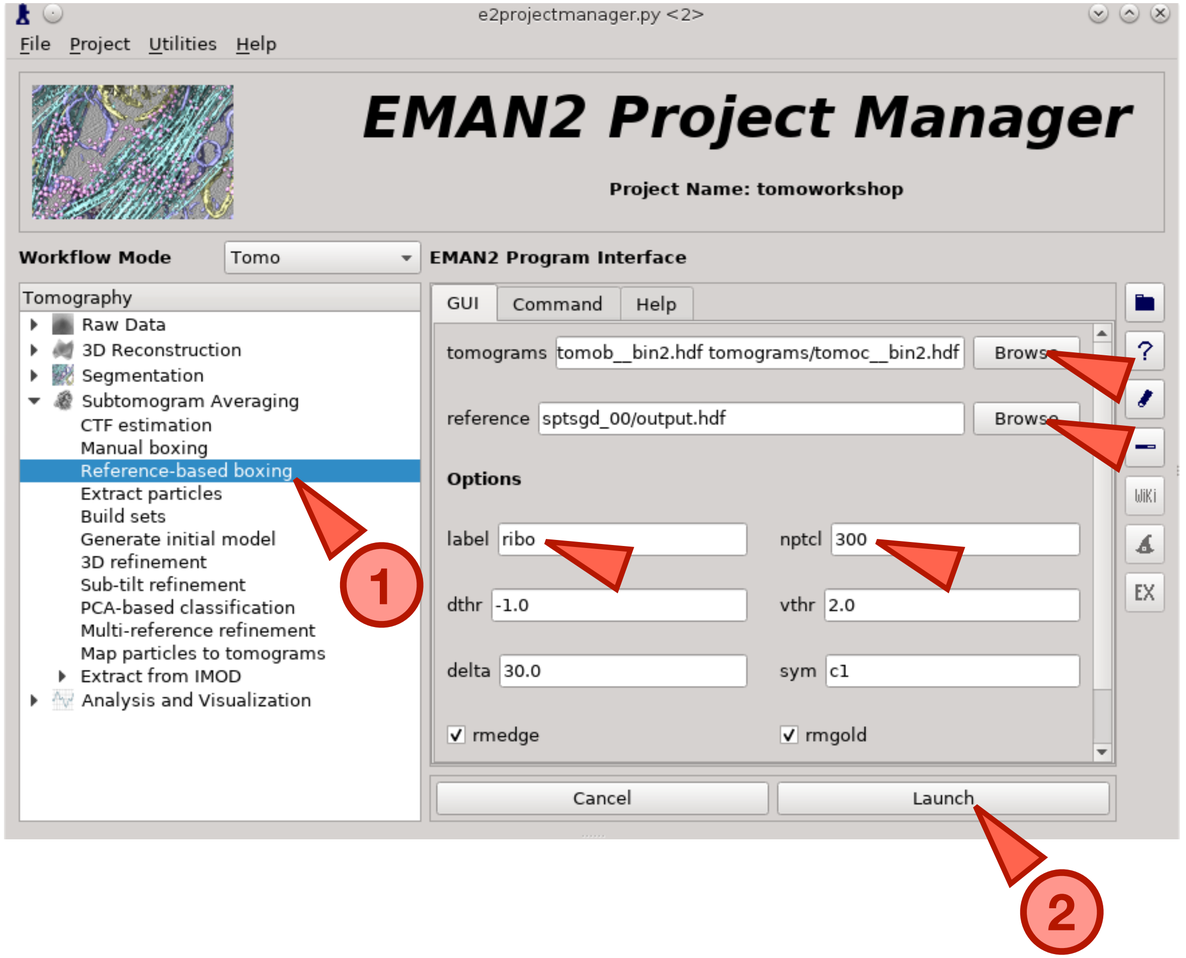
- To terminate the process, open the task manager from the project manager, select **e2spt_sgd.py** and click **Kill**.





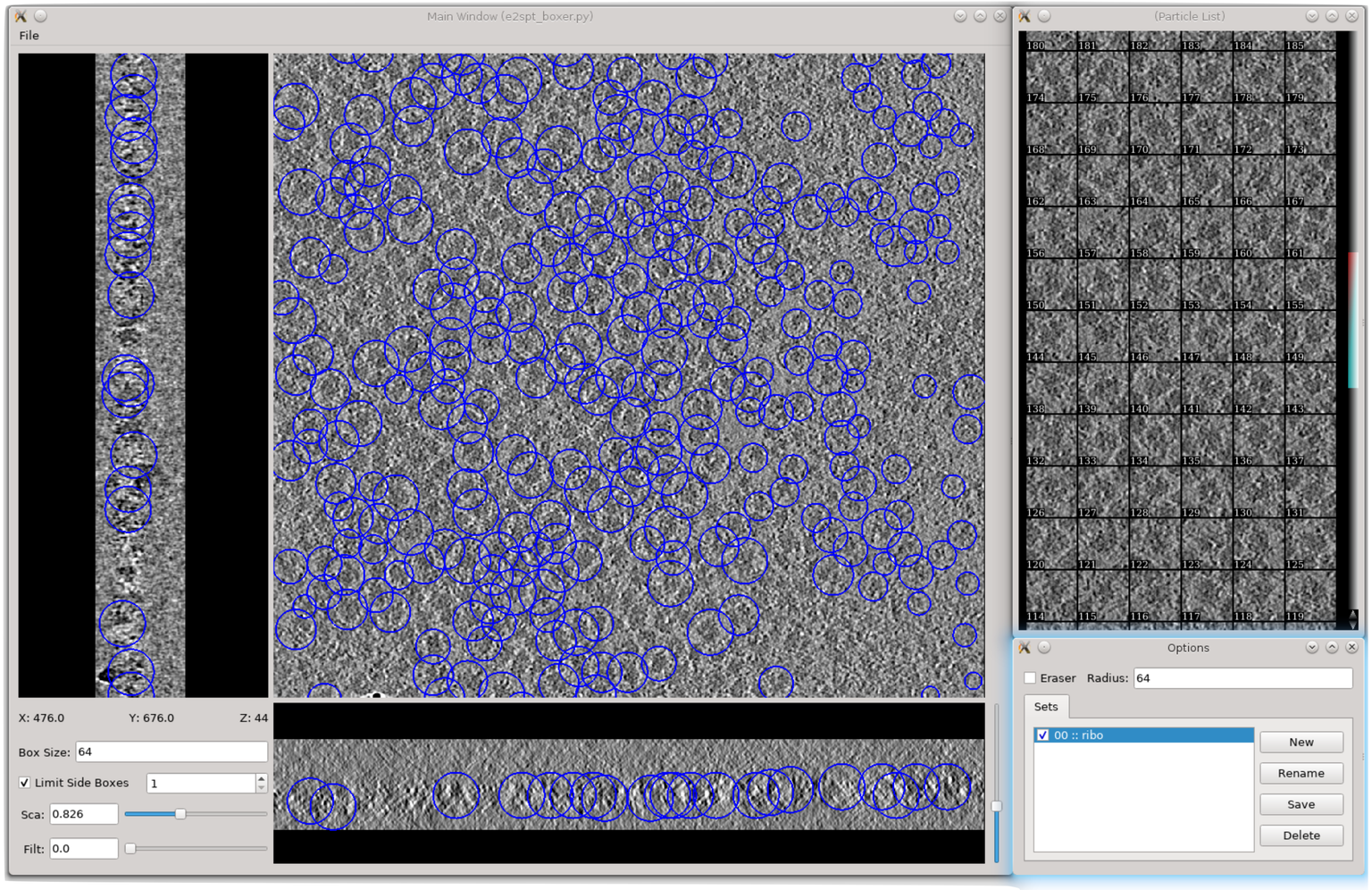
Select particles with template matching

- **Subtomogram averaging -> Reference-based boxing**.
- In **tomograms**, select all the tomograms.
  In **references**, select the output from initial model generation (**sptsgd_XX/output.hdf**).
- Provide a new label for the particles (**ribo**), set **nptcl** to 300.
- Uncheck **rmedge** and **rmgold**.
- Click **Launch**.



Select particles with template matching

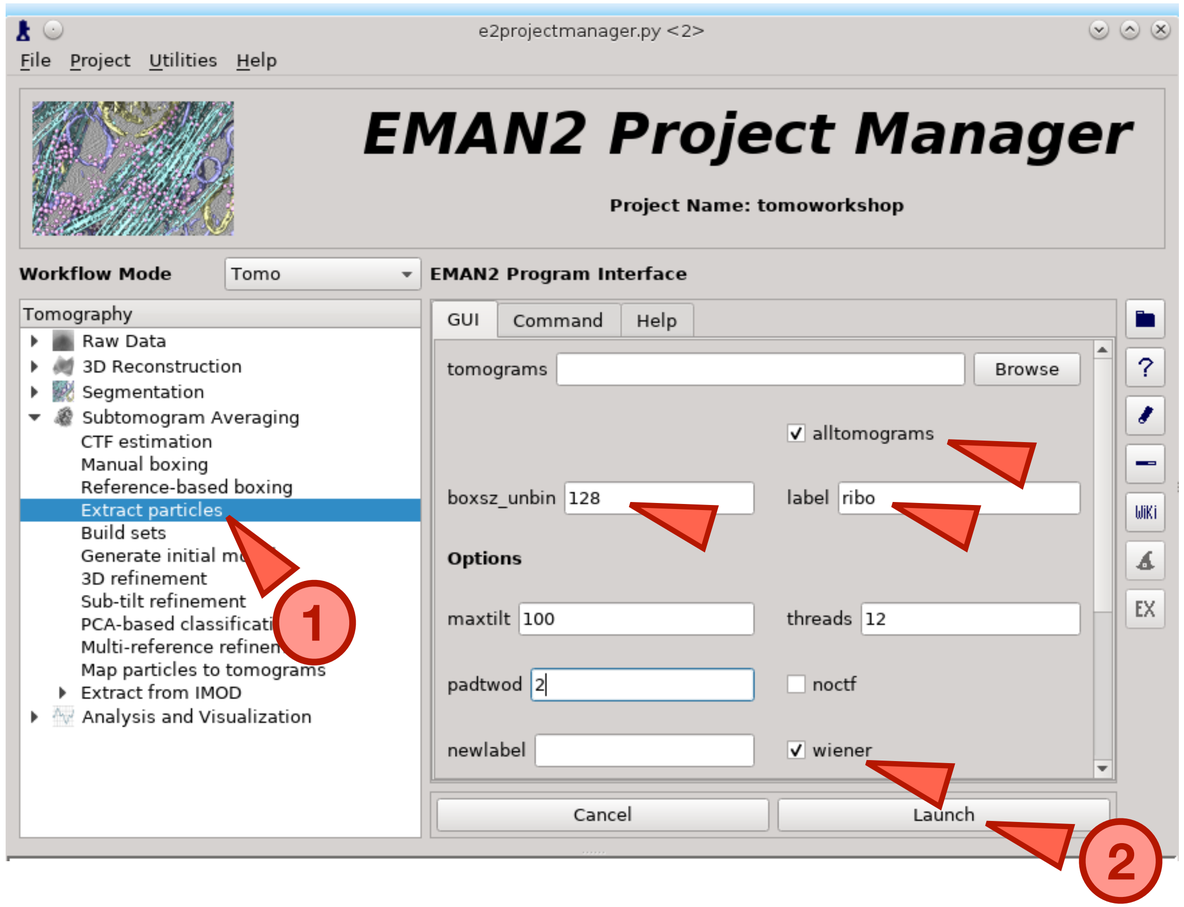
Take a look at the particles selected by template matching from the boxer window.



- You can remove obvious bad particles or add new particles manually.
- Particles inconsistent with others will be excluded during refinement, so having a small amount of bad particles here is fine.

Extract unbinned particles from tilt series

- The tomogram we reconstructed is binned by 2 by default, so any particles directly extracted from the tomogram are downsampled.
- To get full sized particles, we extract 2D particles from raw tilt series and reconstruct them into 3D particles.
- Per-particle-per-tilt CTF correction is performed internally at this step.
- **Subtomogram Averaging -> Extract particles**.
- Check **alltomograms**, set **boxsz_unbin** to 128, set **label** to **ribo**, check **wiener**, and click **Launch**.



3D refinement

- First combine all particles from all tomograms into a set. In **Build** sets, check **allparticles** and click **Launch**.

- If the previous initial model is not satisfying, consider re-run the initial model generation with reference and use the full particle set before 3D refinement.

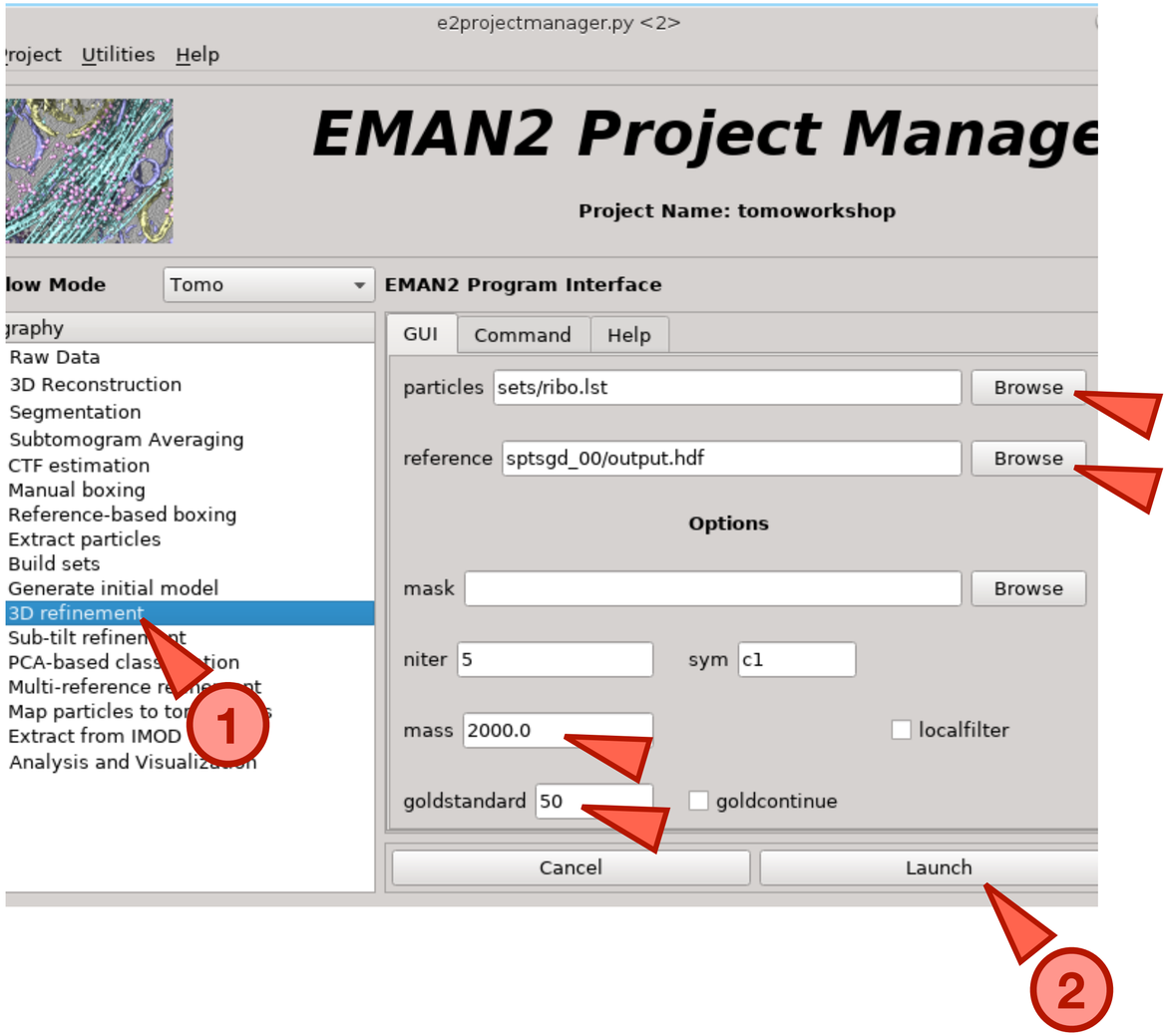
- Depending on the computer used, it can take 0.5 to 2 hours to finish one iteration with 900 particles.

- Build a subset of 200 particles for testing using:

```
e2proclst.py sets/ribo.lst —create sets/ribo_subset.lst —range 0,800,4
```

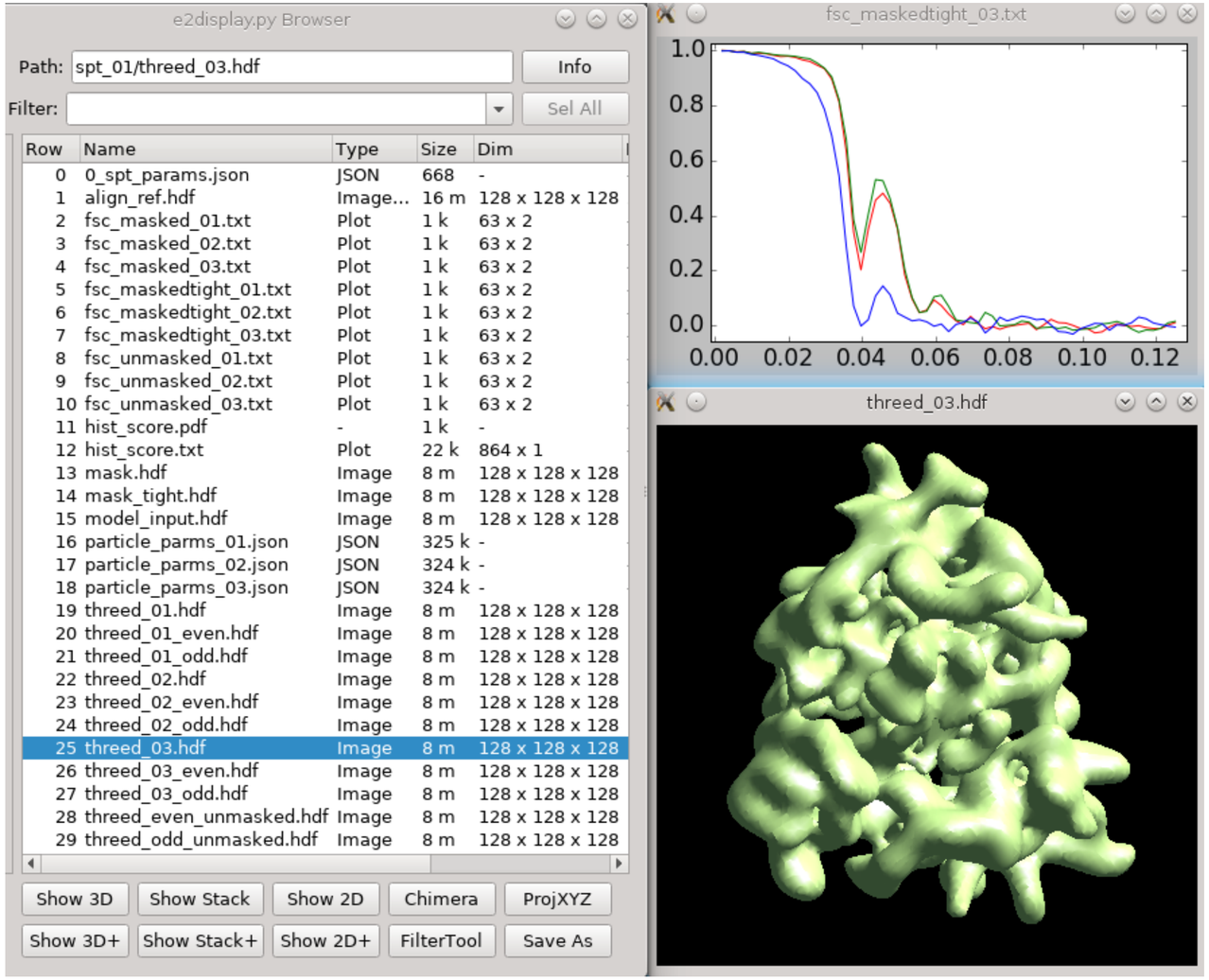
3D refinement

- In **Subtomogram averaging -> 3D refinement**, select **sets/ribo.lst** as particles, and **sptsgd_XX/output.hdf** as reference.
- Set **mass** to 2000 and **goldstandard** to 50.
- Click **Launch**.



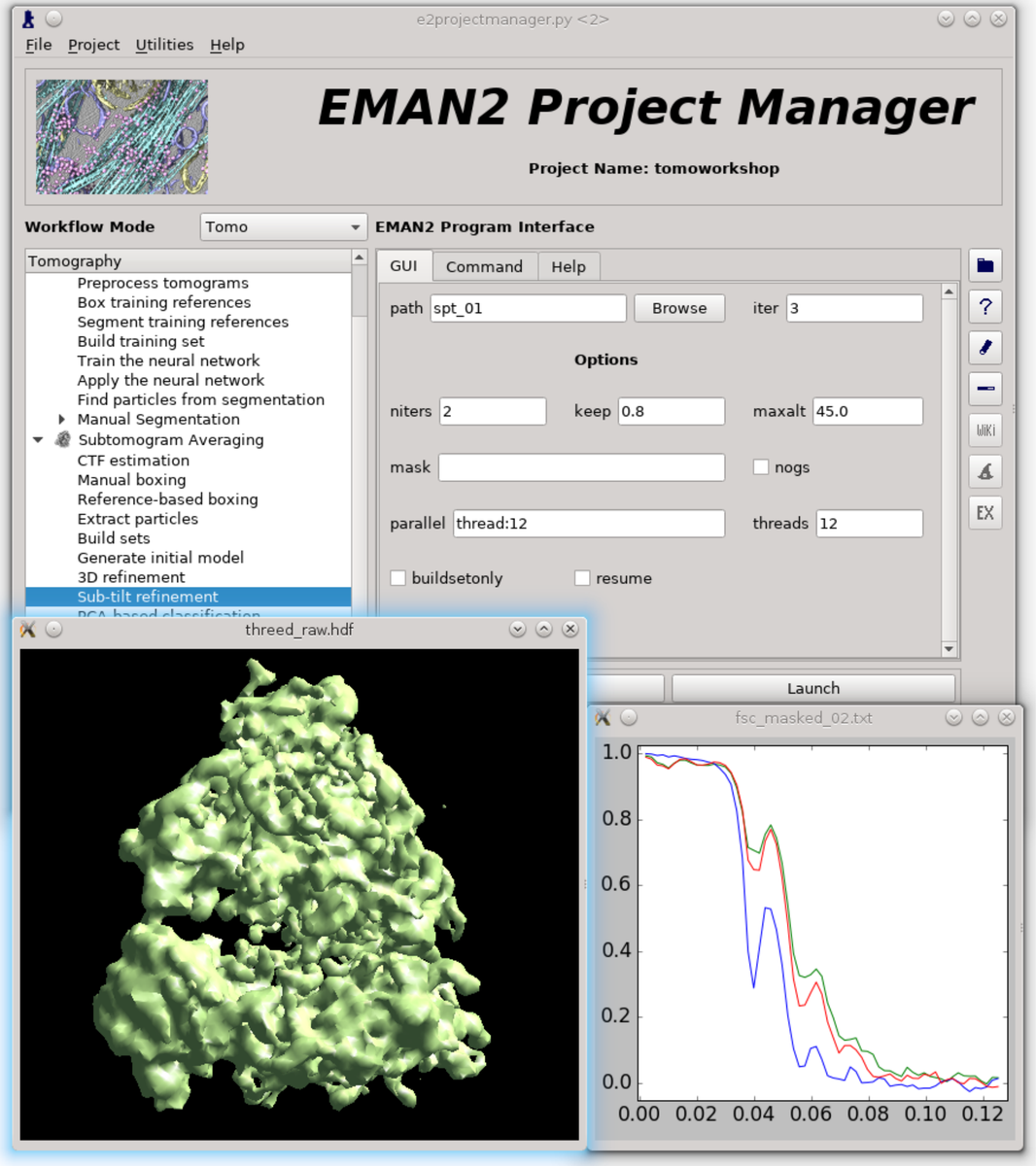
3D refinement

- Take a look at the refinement results in the **spt_XX** folder.
- **threed_XX.hdf** contains the averaged structure and **fsc_xxxx.txt** is the Fourier shell correlation under different masking after each iteration.



Subtilt refinement

- After the subtomogram refinement, select **Subtomogram Averaging -> Sub-tilt refinement**.
- In **path**, select the previous subtomogram refinement folder (**spt_XX**). **iter** should be the last iteration in the **spt_XX** folder.
- **niters** is the number of iterations subtilt refinement will run. **maxalt** excludes 2D particles from higher tilt images with more radiation damage.
- Do **NOT** click Launch. This process is very slow and we cannot finish one iteration on a laptop by the end of day…



Thank you