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!

2.2 Emai	E	EMAN2 Project Project Name: tomowork	
/orkflow Mode	SPR	EMAN2 Program Interface	
ingle Particle Refiner Raw Data Particles CTF Particle Sets 2D Analysis Initial Model 3D Refinement V Validation and	SPT (Legacy)	Cancel	Launch





Prepare dataset

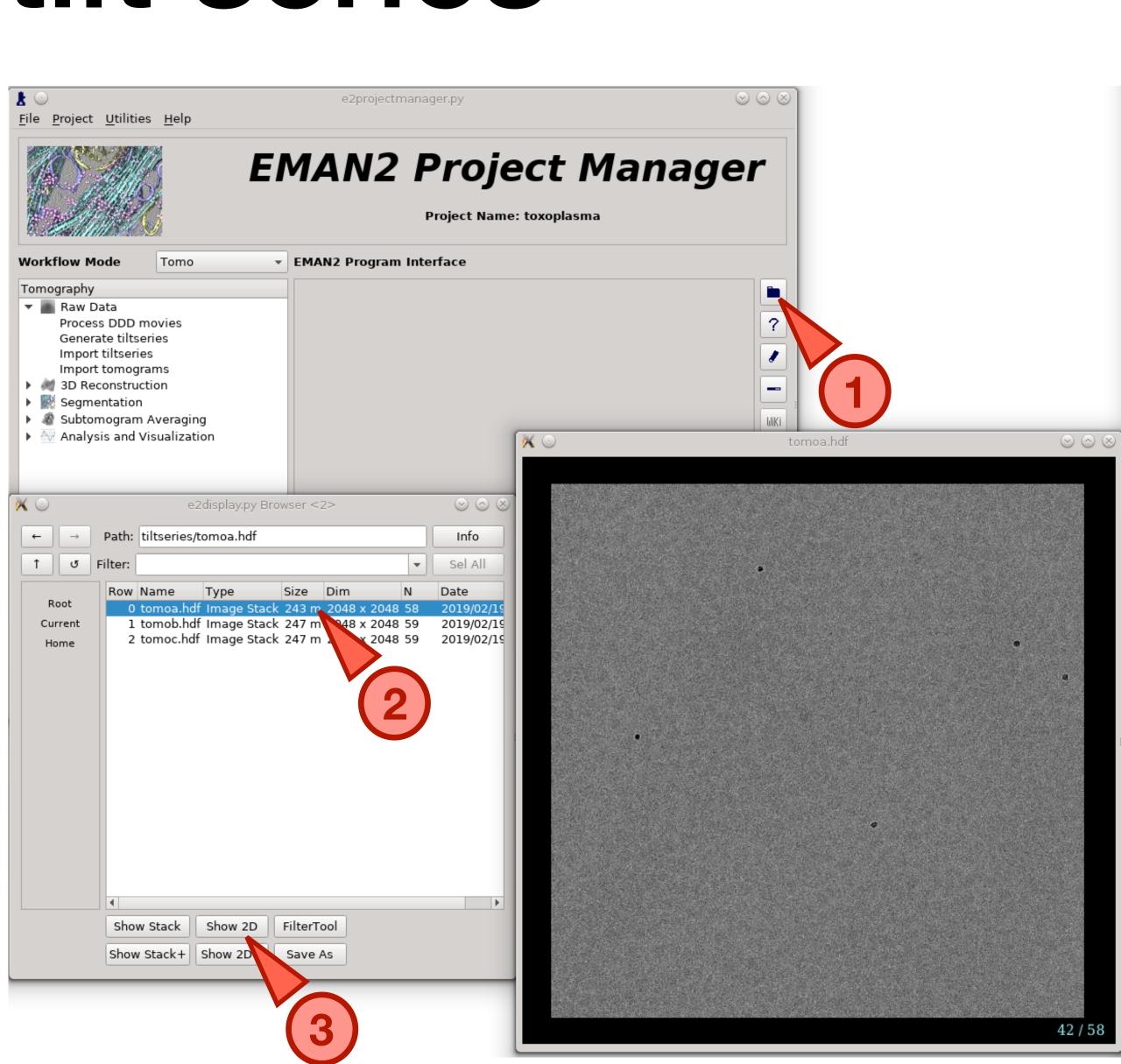
- Download from https://blake.bcm.edu/emanwiki/UTMBWorkshop2019
- Unzip the file.
- In the project manager, select Raw Data -> Import tilt series
- Click Browse next to the first box and select the unzipped files
- Set apix to 3.93
- Click Launch

	e2projectmanager.py <2>	
<u>File</u> <u>P</u> ro	Ject Utilities Help EMAN2 Project Manag Project Name: tomoworkshop	e
Workflo	w Mode Tomo 👻 EMAN2 Program Interface	
Pr Ge Im Im J 31	phy GUI Command Help aw Data ocess DDD movies files Browse enerate tiltseries files Invert apix 3.93 ✓ import_tiltseries oport tomogram O Reconstruction invert apix 3.93 ✓ import_tiltseries	5
🕨 🆓 St	alysis and Visualiza	
← → Pa ↑ ত Filt	ath: tiltseries Info	
	Cancel OK	



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



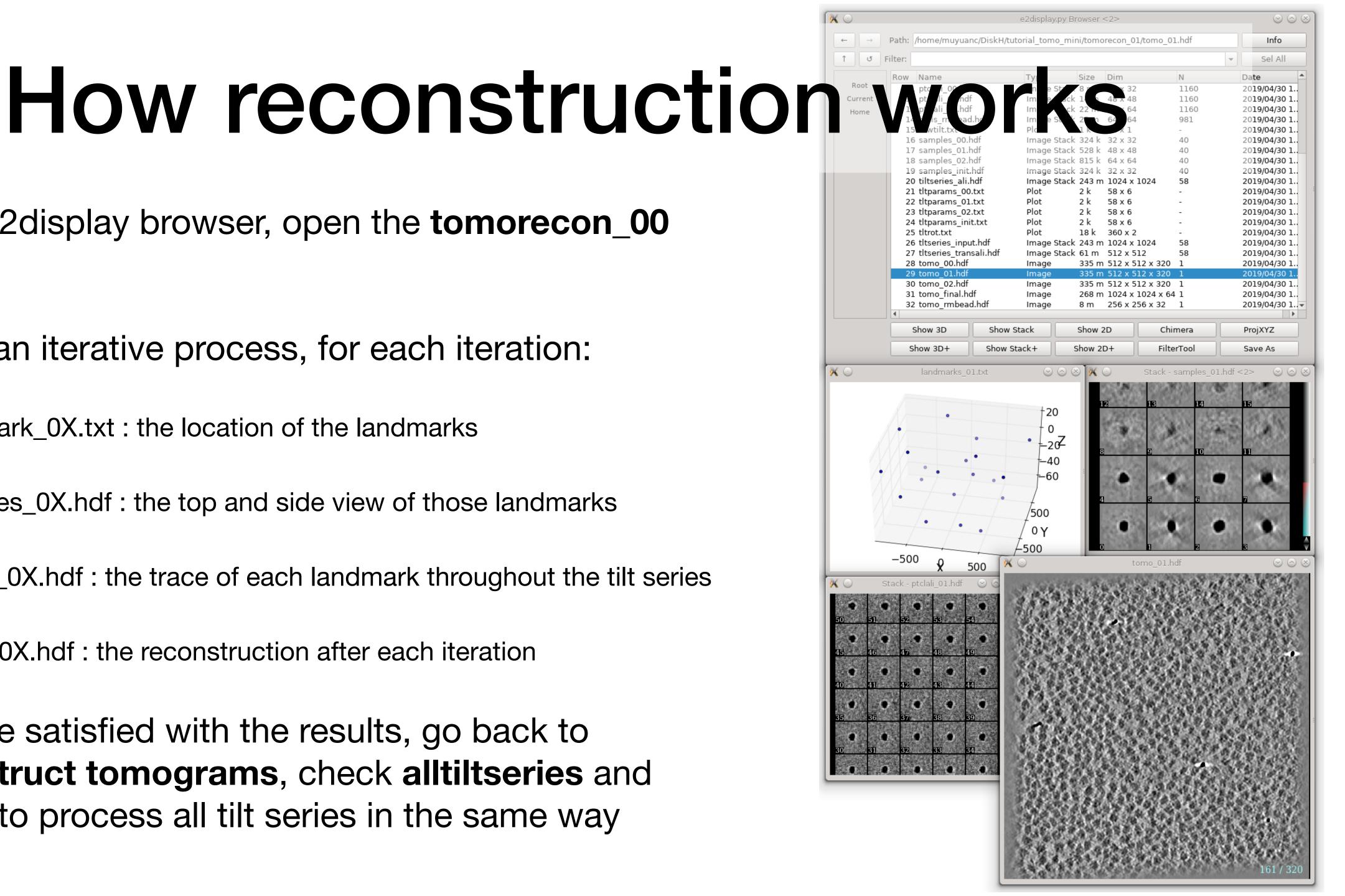
Reconstruc

- 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 tomogra

Workflow Mode	Drama EN N2 Program Interface	moworkshop
Tomography	GUI Command Help	
 Raw Data Process DDD movies Generate tiltseries Import tiltseries Import tomograms 	tiltseries tiltseries/tomoa.hdf	Brows
 3D Reconstruction Reconstruct tomograms Examine fiducial error Segmentation Subtomogram Averaging 	Specify either zeroid/tltstep OR	rawtlt:
Analysis and Visualization	rawtlt	Brows
	npk 20	bxsz 32
	Options	
	tltkeep 0.9	tltax None
	outsize 1k 👻	niter 2,1,1,1
	clipz 64	✓ bytile
	🗌 load	notmp
	pk_mindist 0.125	pkkeep 0.9
	✓ correctrot	threads 12
	normslice	filterto 0.3
am folder	extrapad	rmbeadthr 8.0
	Cancel	Launch



- 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...

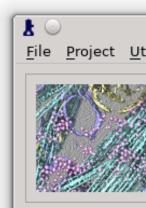
t ⊙ <u>F</u> ile <u>P</u> roject <u>U</u> tilities <u>H</u> elp	e2projectmanager.py <2>	
	EMAN2 Project Project Name: tomowork	
Workflow Mode Tomo	 EMAN2 Program Interface 	
Tomography	GUI Command Help	
 Raw Data Process DDD movies Generate tiltseries Import tiltseries Import tomograms 3D Reconstruction Reconstruct tomograms Examine fiducial error Segmentation Subtomogram Averaging CTF estimation Manual boxing Reference-base Fastract particles 	tiltseries alltiltseries options dfrange 2 psrange 0,5,5 voltage 300 cs 2.7	Bro tilesize 256 nref 15
Extract particles Build sets Generate initial mo 3D refinement Sub-tilt refinement PCA-based classification Multi-reference refinement Map particles to tomograms Extract from IMOD Analysis and Visualization	stepx 20 stepy 40	
	Cancel	Launch



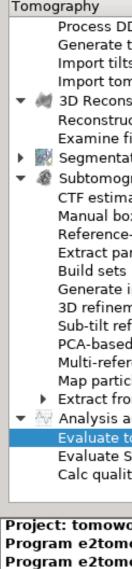
Tomogram evaluation

- 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...



Workflow Mode



Program e2tom

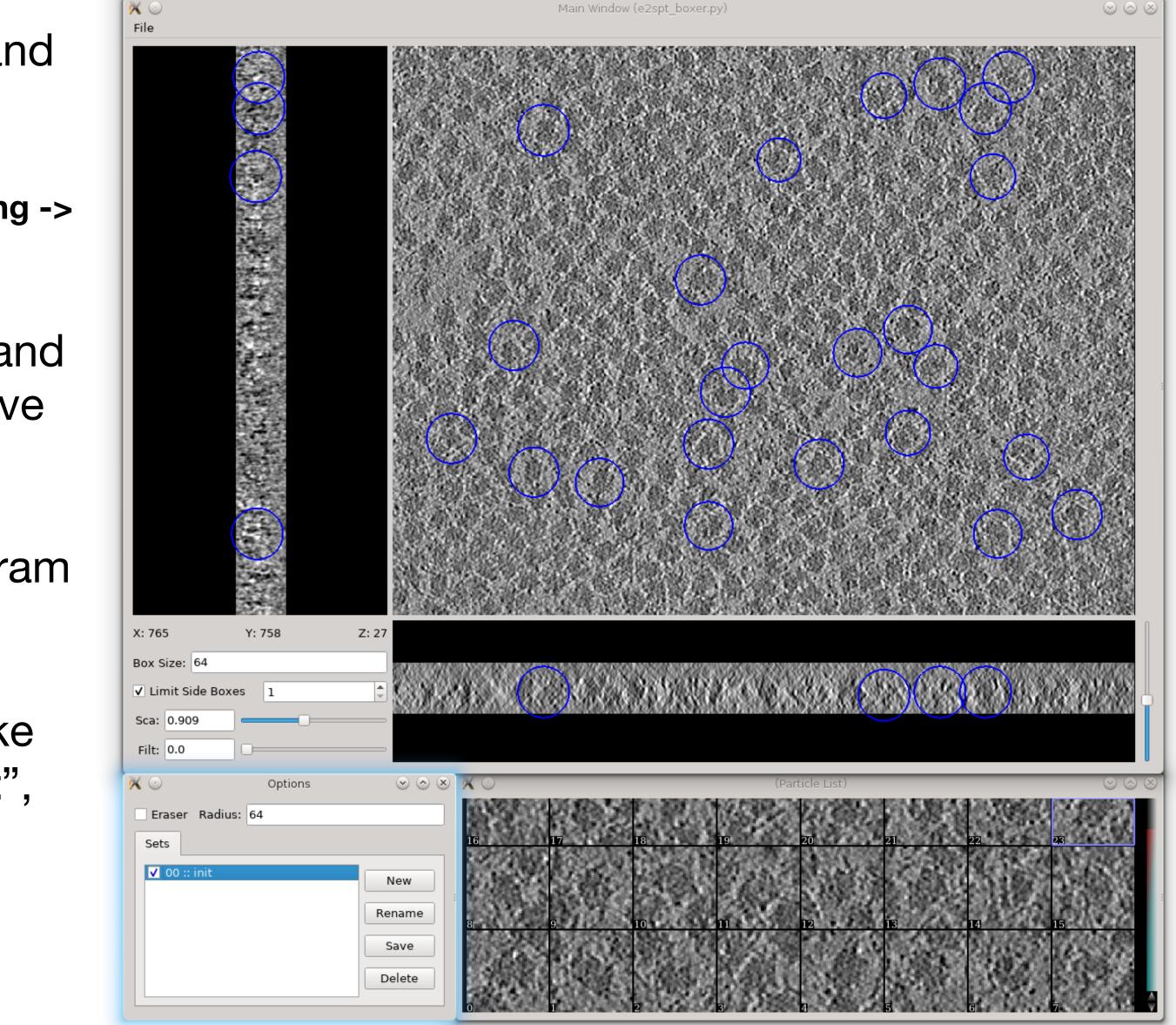
	(<u></u>							
	× 0			6	e2tomo_eval.p	Ý		
	ID	file name	#box	loss	defocus	5	Charles and the second	and the second second
	0	tomoabin2	287	3.75	3.1		14 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Sec. 19
	1	tomobbin2	288	0.96	3.3			
<u>J</u> tilities <u>H</u> elp	2	tomocbin2	289	1.15	3.7			
Le Tomo E DDD movies e tiltseries								
ltseries omograms							Show	/2D
nstruction							Boxer	Refi
ruct tomograms e fiducial error								
tation							ShowTilts	TiltPa
ogram Averaging mation							PlotLoss	Plo
ooxing ce-based boxing particles a initial model ement refinement ed classification erence refinement cicles to tomograms rom IMOD and Visualization							✓ ribo : 864 :	287
s tomograms SPT refinements	_				•			
lity vs. reference		Cancel]I	aunch				
workshop edited! mogram.py Launched!!!! mogram.py Launched!!!! mo_eval.py Launched!!!!					▲ ▼			



- 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.

Select particles





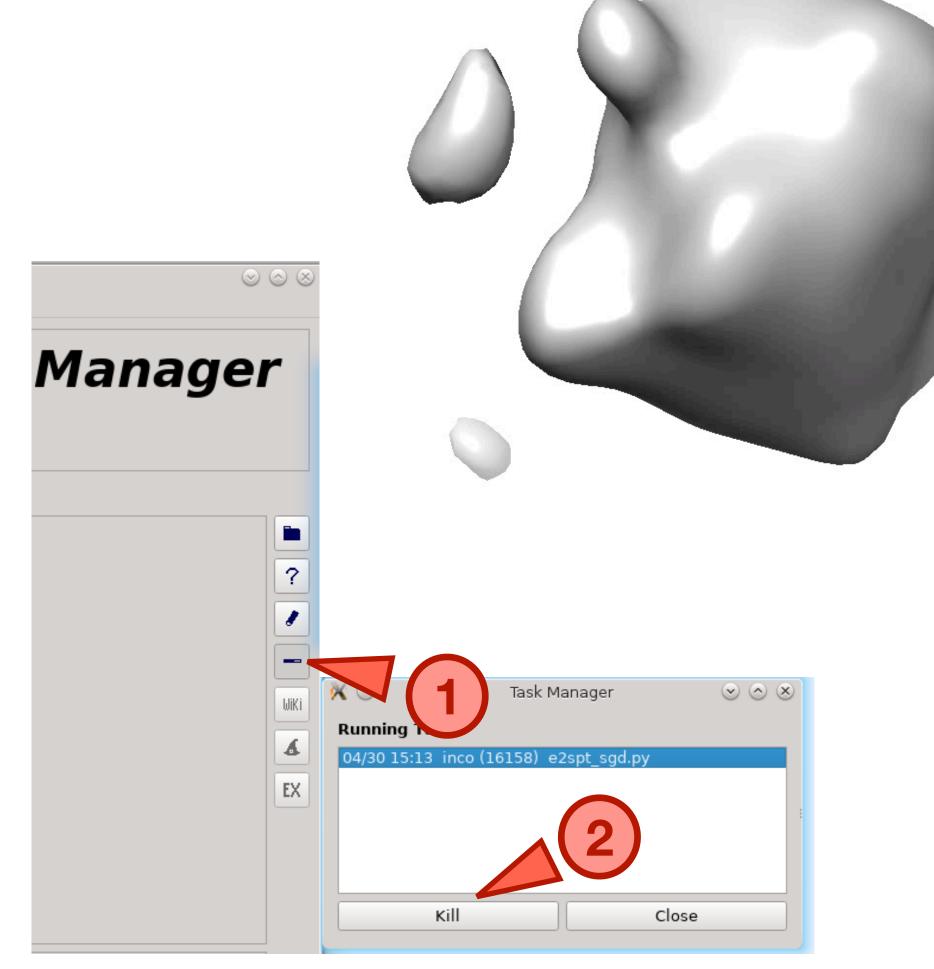
- processing.
- Subtomogram averaging -> Build sets.
- Check allparticles and click Launch.
- Subtomogram averaging -> Generate initial model.
- Setting **shrink** to 2.
- Select the set we just generated (init.lst), and click Launch.

Generate initial model

• First make sets from generated particles. In EMAN2, we combine particles of the same type from multiple tomograms into one virtual stack for further

- 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.

Generate initial model





Select particles with template matching

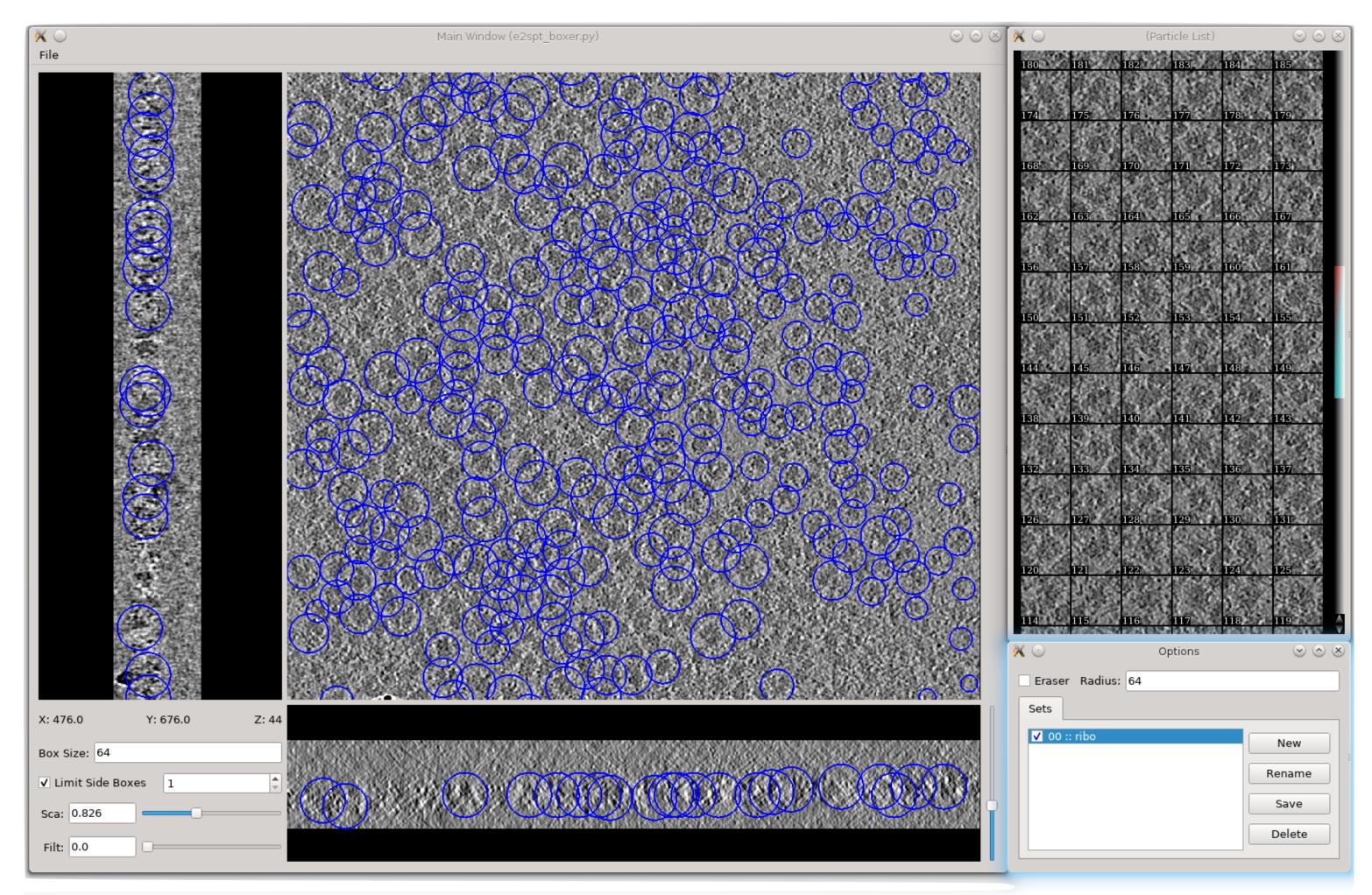
- Subtomogram averaging -> Reference based boxing.
- In tomograms, select all the tomogram
 In references, select the output from in model generation (sptsgd_XX/output.)
- Provide a new label for the particles (rib set nptcl to 300.
- Uncheck **rmedge** and **rmgold**.
- Click Launch.

			ΜΑ		oject N Name: tomoworksho	_
S.	Workflow Mode	Tomo	- EMAN2	2 Program Interface	3	
	Tomography		GUI	Command Help		
itial	Raw Data					
Incial	Segmentation	lon	tomo	grams tomob_bin2.r	ndf tomograms/tomoc	bin2.hdf Brows
ndf).	CTF estimation	veraging	refere	ence sptsgd_00/outpu	ıt.hdf	Browse
	Reference-base Extract particles	~	Optic	ons		
)) .	Generate initial 3D refinement		label	ribo	nptcl 300	
-	PCA-based clas Multi-reference	sification refinement	dthr	-1.0	vthr 2.0	
	Extract from IM	DD	delta	30.0	sym cl	
			✓ rm	ledge	✓ rmgold	
		w Mode Tomo Apply Tomo Apply Tom Ap		Cancel		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 \bullet internally at this step.
- Subtomogram Averaging -> Extract particles.
- Check alltomograms, set boxsz_unbin to 128, set label to ribo, check wiener, and click Launch.

e2projectmanager.py <2> File Project Utilities Help EMAN2 Project Manager Project Name: tomoworkshop EMAN2 Program Interface Workflow Mode Tomo Tomography Command Help GUI Raw Data 3D Reconstruction tomograms Browse Segmentation 🔻 🆓 Subtomogram Averaging ✓ alltomograms CTF estimation Manual boxing Reference-based boxing boxsz unbin 128 label ribo Extract particles Build sets Generate initial m Options 3D refinement Sub-tilt refinement maxtilt 100 threads 12 PCA-based classificat Multi-reference refine Map particles to tomograms padtwod 2 noctf Extract from IMOD Analysis and Visualization newlabel ✓ wiener Cancel 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.
- 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

Depending on the computer used, it can take 0.5 to 2 hours to finish one

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.

roject <u>U</u> tilities	<u>H</u> elp	E۸	1Α	N2		Oje Name: to			าล
low Mode	Tomo	•	EMAN2	Program In	terface				
graphy			GUI	Command	Help]			
Raw Data 3D Reconstruct Segmentation	ion		partic	les sets/ribo.					Brow
Subtomogram / CTF estimation Manual boxing	Averaging		refere	ence sptsgd_0	0/output	.hdf			Brow
Reference-base Extract particles Build sets	-					Option	IS		
Generate initial	model		mask						Brow
3D refinement Sub-tilt refinen PCA-based clas Multi-reference	e tion		niter	5		sym c	1		
Map particles to Extract from IM Analysis and Vi	OD		mass	2000.0				🗌 local	lfilter
			golds	tandard 50		gold	continue	•	
			1	Canc			[Launc	h

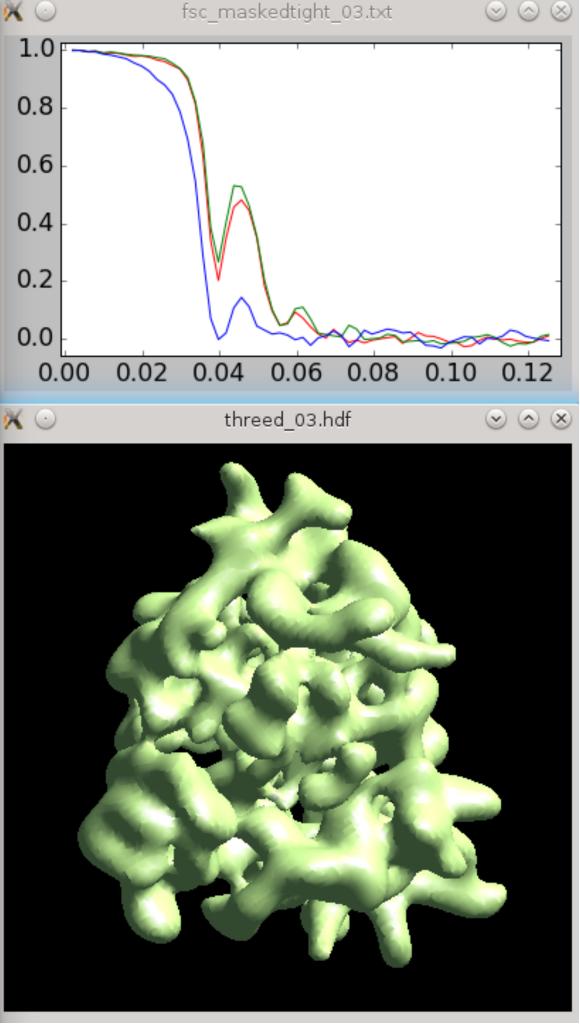
e2projectmanager.pv <2>



3D refinement

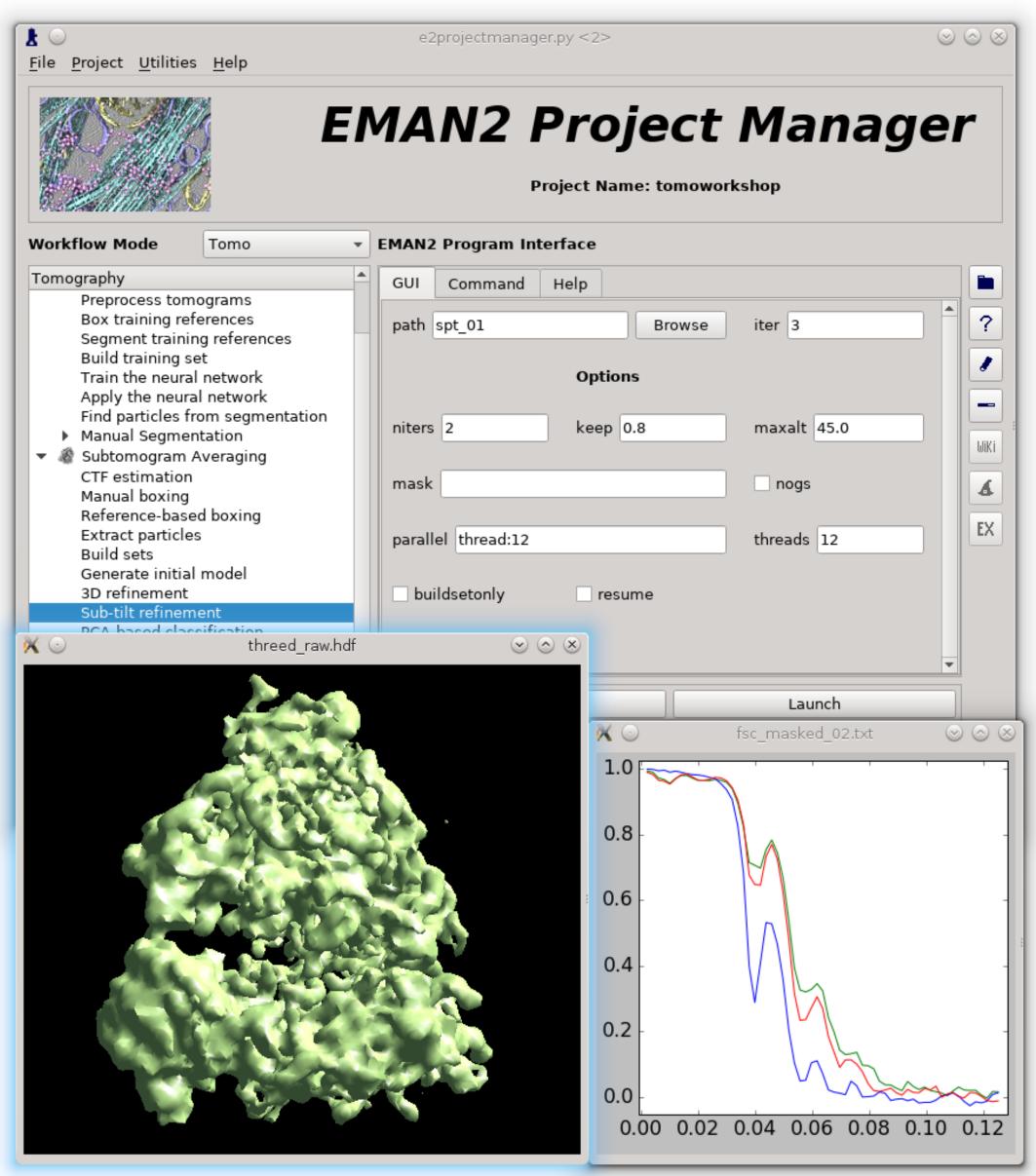
- Take a look at the refinement results in the **spt_XX** folder.
- threed_XX.hdf contains the averaged structure and fsc_xxx.txt is the Fourier shell correlation under different masking after each iteration.

		e2display.py	/ Brows	er			$\odot \odot \otimes$	
Path:	spt_01	/threed_03.hdf					Info	
Filter:						•	Sel All	
Row	Name	,		Туре	Size	Dim	1	
0		_params.json		JSON	668	-		
1		_ref.hdf		Image			x 128 x 128	
2	_	nasked_01.txt		Plot	1 k	63 x		
3	_	nasked_02.txt		Plot		63 x		
4	_	nasked_03.txt		Plot	1 k	63 x		
5	_	naskedtight_01.t		Plot		63 x		
6	_	naskedtight_02.t		Plot	1 k	63 x		
7	_	haskedtight_03.t		Plot		63 x		
8	_	nmasked_01.txt		Plot	1 k	63 x		
9	_	nmasked_02.txt		Plot	1 k	63 x		
	_	nmasked_03.txt	-	Plot	1 k	63 x	2	
	_	core.pdf		-	1 k	-	1	
	_	core.txt		Plot	22 k	864 >		
	3 mask.			Image			x 128 x 128	
		_tight.hdf		Image			x 128 x 128	÷
		el_input.hdf		Image			x 128 x 128	
	-	le_parms_01.jsc		JSON	325 k			
		le_parms_02.jsc		JSON	324 k			
	-	le_parms_03.jsc	on	JSON	324 k	-		
		d_01.hdf		Image	8 m	128;	x 128 x 128	
		d_01_even.hdf		Image	8 m	128;	x 128 x 128	
		d_01_odd.hdf		Image	8 m		x 128 x 128	
		d_02.hdf		Image	8 m	128;	x 128 x 128	
		d_02_even.hdf		Image	8 m	128;	x 128 x 128	
2	4 three	d_02_odd.hdf		Image	8 m	128 :	x 128 x 128	
2	5 three	d_03.hdf		Image	8 m	128 :	x 128 x 128	
2	6 three	d_03_even.hdf		Image	8 m	128 :	x 128 x 128	
2	7 three	d_03_odd.hdf		Image	8 m	128 :	x 128 x 128	
2	8 three	d_even_unmask	ed.hdf	Image	8 m	128;	x 128 x 128	
29	9 three	d_odd_unmaske	d.hdf	Image	8 m	128;	x 128 x 128	
•							Þ	
Sho	w 3D	Show Stack	Show	2D (Chimera	a	ProjXYZ	
Show	w 3D+	Show Stack+	Show	2D+ F	ilterToo	ы [Save As	



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 lacksquareand we cannot finish one iteration on a laptop by the end of day...



Thank you