Introduction to the EMAN2 GUI Tools

When possible, it is best, when starting a tutorial to go to http://eman2.org and make sure you have the most up to date version. Most of these tutorials are updated 1-3 times per year.

WE WILL GO THROUGH ONLY A FRACTION OF THIS MATERIAL AT THE WORKSHOP!

- ➡ Boxes like this one will contain additional information and tips. You can complete the tutorial without reading any of these, but they may help you better understand what you are doing.
- ➡ The main source for EMAN2 documentation is the Wiki at: http://eman2.org . There is also a Google Group for support and discussions: http://groups.google.com/group/eman2 . Anyone can read/search these sites, you only need to join if you want to post.
- ➡ **GUI Tips:** EMAN2 will work best with a 3-button scroll mouse, though there are alternatives using keyboard modifiers for one button mice or trackpads on Macs.
  - In most EMAN2 windows (2-D images, 3-D volumes, plots, etc.), the middle mouse button will open a control panel for the widget, which is different for each widget type
  - The right mouse button is used for panning in 2-D or 3-D image windows, and can be used to zoom (by shift+dragging), and to reset the zoom (clicking) in plot windows.
  - The scroll-wheel will generally act as a zoom. control-panel for more precise control
  - If you have a one button mouse or trackpad, holding down the alt/option key combined with a mouse click will serve the same role as a middle-click.
  - In the control panels, and other places in the EMAN2 interface you may encounter 'Value Sliders'. A slider is attached to a text-box with a number in it. Dragging the slider controls the number, and entering a number will change the slider. In addition, the text-box can be used to control the range of the slider and get more precise control. By typing **<value** or **>value** in the text box you can change the limits of the slider.
- ➡ Text you see in *italics* will generally refer to labels in the GUI, such as buttons to press. Text you see in **bold**, are commands to be typed in. Items like: <param> are parameters you should fill in (without the < >). Items like: [param] are optional parameters (again, fill without the [ ]).
- ➡ Check your version: The command e2version.py will tell you exactly what version of EMAN you are using. When reporting bugs or asking questions on the mailing list it is critical to include all of the lines of the output of this program with your question.
- ➡ **Mac Laptops** - If you have a Retina display (most will nowadays), then you have a lot of control over the resolution of your display. By default, your display will give you an effective resolution of 1440x900 pixels. This isn't enough space to do image processing well. For the workshop, we strongly suggest opening System Preferences-> Display and adjusting it to "More Space" which will be 1920x1200 effective resolution.
- ➡ **Windows users** - Windows **will** make your life more difficult with EMAN2. If you have no choice but Windows, please remember that programs must be run from the windows command-line, not by clicking on icons. Installation of an 'enhanced command-line' tool for windows, such as *Console 2.x*, will make your life somewhat easier.
- ➡ **Windows** likes to aggressively kill "unresponsive" programs. This means if you have asked EMAN2 to do something from the GUI, and it hasn't finished yet, simply clicking on an open window while you wait may cause the whole program to be killed. We have no solution for this at present other than "don't do that".

Why EMAN2?

- Simple pipeline for reliable refinements to near-atomic resolution
- Complete pipeline for tomography and subnanometer resolution subtomogram averaging
- Over 200 general purpose image processing algorithms (filters, masks, transformations,...)
- Multiple methods for conformational and compositional variability
- Reads and writes virtually every file format used in CryoEM
- A complete GUI workflow interface, with complete logging of all processing
- Wide range of GUI tools for 2-D plotting, image viewing and manipulation and 3-D rendering

There are now quite a few software packages available in the CryoEM community, and new users in this field are faced with a wide range choices. Larger packages like EMAN, SPIDER, XMIPP and BSOFT compete with the task-specific software like RELION, CISTEM and SIMPLE. Learning how to use the larger packages thoroughly can take a little more time than the simpler packages, however, they also offer a much wider range of analytical tools and image processing methods when you inevitably run into problems with your projects.

For some fraction of projects, you could collect your data on a Krios with a K3 camera, process with RELION and achieve a self-consistent 2 Å resolution structure. However, that fraction may not be very large. What happens when this straightforward process fails to produce the expected result, or even worse, gives you a structure which turns out to be incorrect due to model bias or other problems? How do you detect the problem and correct it?

This is where packages like EMAN2 shine. Like RELION, EMAN2 provides a simple guided path for single particle refinement to near-atomic resolution. Indeed, for "good" data sets it is easily demonstrable that EMAN2 and RELION produce virtually identical structures. However, EMAN2 also provides an extensive set of tools for validating the accuracy of your structure, and investigating what the problem is when, for example, you have a specimen with a large degree of conformational or compositional variability. EMAN2 includes at least 6 different methods for exploring different types of specimen variability, and a range of standard tools for insuring self-consistency between raw data and reconstruction. Unfortunately packages like Appion and Scipion, which ostensibly provide access to multiple software packages through one convenient interface, generally provide access only to very basic capabilities, and are usually many years out of date in their software support.

EMAN2 also now includes a complete pipeline for subtomogram averaging using PPPT (per-particle per-tilt) methods, which can achieve subnanometer resolution. Unlike other tools which require using a sequence of different software packages, EMAN2 has a single straightforward pipeline for the entire process. EMAN2's tilt-series alignment and reconstruction methods have been shown to work in cases where other standard tools fail. If you are already familiar with single particle analysis tools in EMAN2, adding tomography to your repertoire requires very little effort.

EMAN2 GUI Tool Introduction

EMAN2 contains a wide variety of tools for visualization and image processing on images in virtually any CryoEM file format. While EMAN2 uses HDF5 for its own processing pipeline, you may freely read/write/convert to other formats. In this quick introduction, we will explore some of these tools.

For those of you who don't like to read a lot of text, the major takeaway messages of this section are: 1) press the middle mouse button (or alt-left) on just about any visualization widget in EMAN2 to get a control-panel. 2) all 3 buttons and the wheel usually have a function. 3) e2display.py can be used to visualize just about anything you encounter in the CryoEM field.

Getting Started

- Make sure your account is setup for EMAN2, by typing: **conda activate eman2**
- To test if this worked, type: **e2version.py**
- This should print information on your computer and the version of EMAN2. **If it fails, do not proceed, sort this out first!** If you ever need to report a problem or bug in EMAN2, or ask a question, it is important to report the results of this command with your question. EMAN2 uses a Google Group for this purpose:

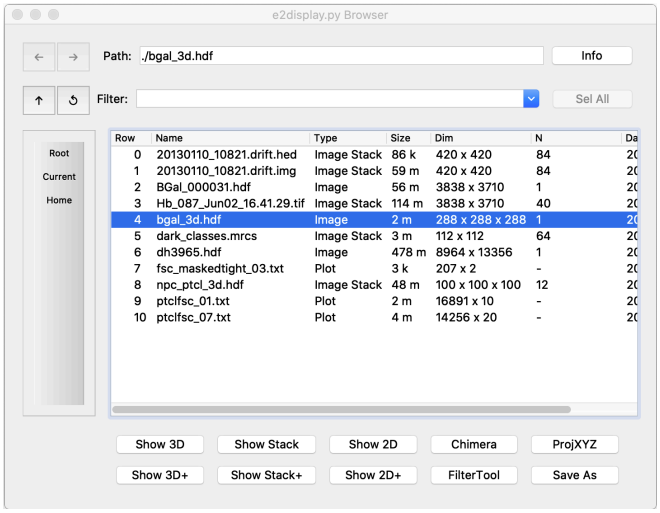
https://groups.google.com/forum/#!forum/eman2

- Next, you may consider running: **e2speedtest.py** This will give you an estimate of how fast one processor on the computer you are using is. Timings you may find later in the tutorial are based on a machine with a score of ~1.0. If you have a 4 core machine, the effective speed will be ~4x higher (if you use all of the cores).

- With the preliminaries out of the way, **cd** to the **sample_files** folder if you haven't already. We assume that you are already familiar with Linux at least enough to run commands from the command-line, and with the basic commands like: ls, cd, pwd, cp, mv and rm. If you aren't familiar with any of these, I suggest going through a basic Linux command line tutorial before proceeding. These commands are the same if you are using a Mac or a Windows machine with the embedded Ubuntu shell.
- If you now do an **ls**, you should see:

20130110_10821.drift.hed	bgal_3d.hdf	fsc_maskedtight_03.txt	ptclfsc_01.txt
20130110_10821.drift.img	dark_classes.mrcs	Hb_087_Jun02_16.41.29.tif	ptclfsc_07.txt
BGal_000031.hdf	dh3965.hdf	npc_ptcl_3d.hdf	

- Run: **e2display.py**
- This is the EMAN2 interactive visualization tool. You should see a window appear.
- If you single click on one of the items, for example bgal_3d.hdf, you will see a set of buttons appear at the bottom of the display. These buttons will be different for each type of file, and represent various available actions.
- Note also that the columns in the browser provide information specific to image processing, like the image dimensions and number of images in the stack.
- Header information is available through the "Info"

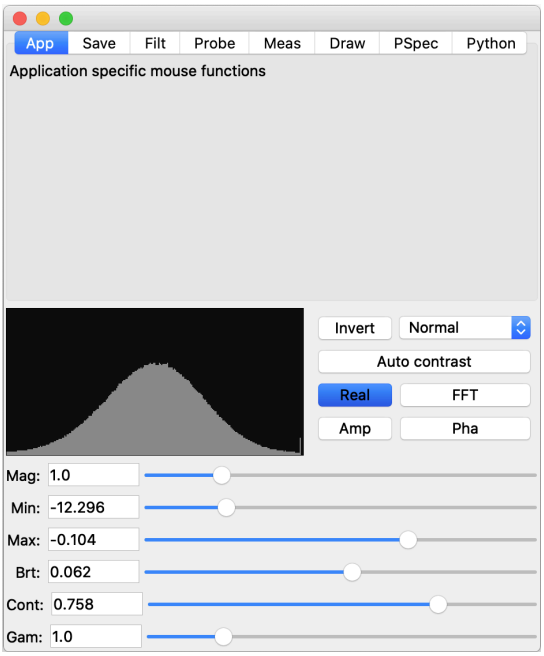


button, which is a toggle. The Info window is also context-sensitive, and will show different information for different types of files.

- If the files in the folder change (which shouldn't happen right now), you will need to press the button to update the browser list.
- Pressing the button will navigate to the parent folder. Note that subfolders will appear in the list with a "+" next to them to view the folder tree as a nested list. You can also navigate into a subfolder by double clicking. This behavior mimics a typical operating system file browser.
- The **Save As** button is available for all image files, and can be used to do file format conversions (based on the file extension). Available formats are listed: https://blake.bcm.edu/emanwiki/EMAN2ImageFormats. Note that the command line programs **e2proc2d.py** and **e2proc3d.py** are usually a better choice for file conversions.
-

2-D Single Image Display

- Let's try visualizing one of the files in the sample folder. Single click on **BGal_000031.hdf**. This file contains a single 2-D micrograph image, so there are fewer options for interacting with it. Press the **Show 2D** button. This should cause an image window to appear, displaying the micrograph. The window may be small, feel free to make it larger. You can pan by right-dragging.
- The widget being used to display the image is one of several different visualization widgets in EMAN2. To make the system easier to use, mouse buttons have similar behaviors in different tools. In any visualization widget, pressing the middle mouse button (or scroll wheel) will cause a control panel for that widget to appear. If you don't have a 3 button mouse, holding alt and left clicking should accomplish the same thing. Go ahead and open the control panel for the 2-D display widget:

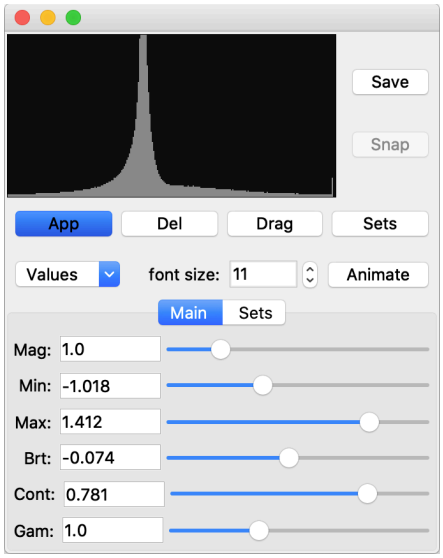


- The use of the sliders should be relatively obvious if you play with them. The histogram display, **Auto Contrast** and **Invert** buttons should be similarly obvious. Note that making adjustments in this window impact only how the image is displayed in this window. Nothing you do here will modify the image on disk.
- If you have no idea what a Fourier transform is, you can probably skip this bullet (though you will probably need to learn at some point if you are doing CryoEM). The **Real, FFT, Amp** and **Pha** buttons require a little more explanation. They allow you to visualize the Fourier Transform of the image in various ways (this is how Thon rings would be observed, among other things). This is not the best way to accomplish this task for whole micrographs, but it will at least give you an initial view of the image quality. The **Amp** button will show you the power spectrum of the image in greyscale. Similarly, the **Pha** button will show the complex phases as greyscale. The **FFT** button will show the power spectrum in greyscale and use color to represent the phases. Since cyclic, like phase, this is a more natural representation.
- The tabs along the top will allow you to perform a wide range of other operations, and will impact the behavior of the left mouse button. For example the **Save** tab will provide a range of tools for saving the image back to disk. The **Probe** and **Meas** tabs will allow you to measure distances or local density values in the image, and so on.

2-D Multi-image Display

- Next, let's look at the multi-image 2-D display widget. Close the 2-D image display from the first example, and single click on **dark_classes.mrcs**, an MRC stackfile containing 64 images. You will now see an additional set of 2 buttons: **Show Stack** and **Show Stack+**.
- Before moving to Show Stack, let's quickly try the **Show 2D** button. When the image opens, open the control-panel again with the middle button as well.
- Note 2 changes as compared to the last time we did this: there is a *1/64* in the lower right corner of the image window, and there is now a new *N#* slider in the control panel. This approach allows you to see all of the images in the stack, one at a time, which is good for comparing one image to another in sequence, and gives you all of the measurement tools we looked at above.
- Once you've played with that, close the single image display window and press the **Show Stack** button. When the image opens, go ahead and open its control panel with the middle mouse button. You will see that it has some similar sliders, but the feature set in this control panel is noticeably different.
- The 4 radio buttons: **App, Del, Drag** and **Sets** control the action of the left mouse button, and give you various strategies for manipulating the set of images manually. The Sets method is discussed as part of the single particle analysis tutorial, so we will skip that.
- Go ahead and select the **Del** button. If you now click on individual images, they will toggle to being a solid color. Again, note that this is not modifying the file on disk, this is purely happening in the display at the moment. Once you have marked some of the images for deletion, you will need to press the **Save** button to create a new file which omits the marked images.
- Note that there is a number appearing in the lower left corner of each image. As should be obvious, this number is the image number within the file. Note that in EMAN the first image in the file is numbered 0, not 1.
- It is possible to display other per-image values in this way as well. If you click on the **Values** pop-up menu, you will see a list of all header values available for this set of images. Select **sigma** from this menu, and you will see the standard deviation of each image displayed in the lower left corner.
- Now bring up this menu again, and this time select **MRC.minimum**. You will see that a second value is now displayed as well. Confusingly, however, you will see that this value is exactly the same for all of the images!? This is the reason EMAN2 uses *HDF5* rather than *MRC* as our default file format. When you save images in *MRC* stacks, there is a single image header for the entire stack of images, so there is no way for the file to contain per-image information. In HDF5, you can associate as much information as you like with each image (http://eman2.org/Eman2Metadata).
- This has some real consequences. The images you are looking at are class-averages (averages of many aligned particles). In *HDF5*, you would be able to see how many particles went into each

This is a good time to point out that historically MRC files had no way of storing sets of particle images. Each file could hold a single 2-D image or a single 3-D volume. This led to the idea of 'stacking' a bunch of 2-D images in 3-D. The downside to that approach is that, until very recently there was no way to identify which 3-D volumes were real volumes and which were sets of 2-D images. EMAN2 uses the same convention used in several other packages. If a 3-D volume has the *.mrcs* file extension, it will be treated as a stack of 2-D images. If it has any other extension, it will be treated as a single volume.



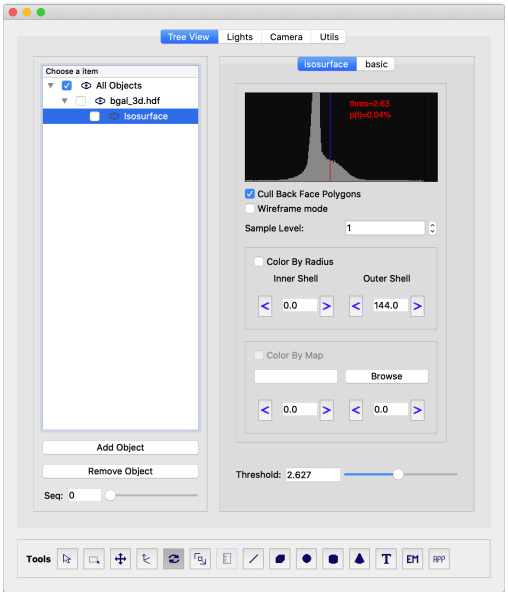
average, and even identify the specific particles used in each average. By converting from *HDF5* to *MRC*, all of this information has been lost.
- Close the multi-image display when you are done.

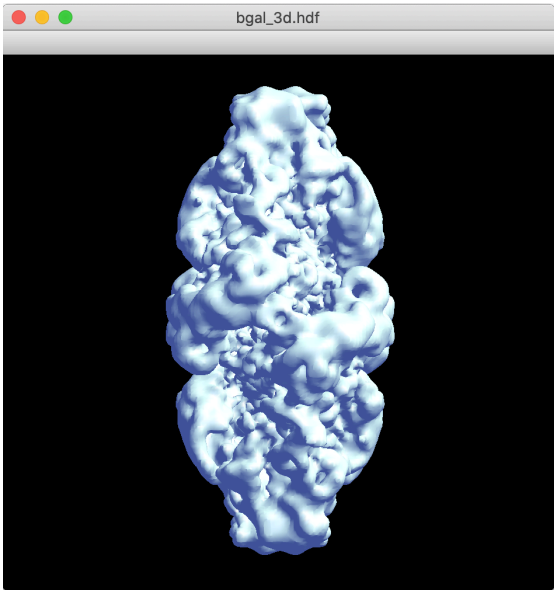
3-D Volume display

- Next we will take a look at a 3-D volume. Single-click on **bgal_3d.hdf**. You will see no fewer than 10 different buttons appear!
- The **Chimera** button is a convenience function, and if properly configured, will open the volume in UCSDF Chimera for visualization. This is no different than launching Chimera and opening the associated file.
- The **ProjXYZ** is also convenient, particularly in the context of subtomogram averaging. Go ahead and press this button. You should see a 2-D single image display open with X, Y and Z projections of the volume. Close the window once you've taken a look.
- Similar to the case for image stacks, the **Show Stack** and **Show 2D** buttons will allow you to visualize the volume as a set of slices along the Z axis. If you need to look along a different direction with this tool, you will need to rotate the volume first, so Z is in the desired direction. Try these if you like, and close them when done.

- What we want to focus on next is the **Show 3D** widget. Press this button and you should see a window appear with a 3-D isosurface view of your 3-D map.
- While this tool does not attempt to compete with a full-featured visualization program like Chimera, it does have fairly extensive capabilities, and a few tools, like the ability to easily open several volumes and quickly cycle through them are much easier to do in this tool than Chimera.
- Open the control panel with the middle mouse button.
- When building scenes of objects, it is traditional to arrange things in a 'tree' structure. In this case, you will see *All Objects*, which contains a single data set, *bgal_3d.hdf*. This data set in turn contains a single visualization: *Isosurface*, which is currently highlighted in the list. The controllers in the right side of the control panel are for the selected *Isosurface* element. You can adjust the isosurface threshold either by dragging in the histogram or by adjusting the **Threshold** slider. Go ahead and try doing that.

- If you see a window appear, but it doesn't contain a 3-D model, you may need to read http://eman2.org/Remote.
- If you get a segmentation fault or other crash, try opening e2display.py again and see if that fixes the problem. If it doesn't please seek assistance on the Google Group as you may have a software installation problem.

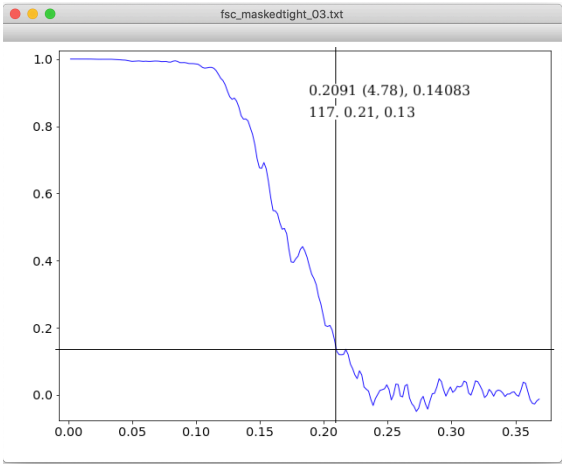




- The *Tools* shown at the bottom control what the left mouse button will do. By default the left mouse button will rotate the object, and the wheel allows you to zoom in and out. Try rotating the object a bit.
- We can also add other visualizations to the display:
  - Select **bgal_3d.hdf** instead of Isosurface
  - Press **Add Object**. A new window will appear.
  - In the new window, select **Slice** from the drop-down menu, and press **Add Node**.
- You should see that a slice through the middle of the map has now been added to the display. Try zooming out and rotating to better understand what you are seeing. Note that there is now a *Slice* element below the *Isosurface* element in the tree list. If you select **Slice**, you can adjust the parameters of the slice to reorient or reposition it.
- You and add additional nodes to the scene in this way and adjust each independently. We won't take the time to go through all of the possibilities here as most of the options should be fairly straightforward to understand once you know they exist, but there are a few more points to raise.
- Note that next to each item in the tree is a small icon:  , indeed, this is supposed to represent an eye. This controls whether the object in question is visible on the screen. To toggle this you must use the right mouse button on the eye. Try right clicking the Slice eye, and you should see the slice disappear from the display.
- We will only look at one other feature of the 3-D widget for now. Go ahead and close the 3-D display window.
- Go back to the browser, make sure **bgal_3d.hdf** is selected, and press **Show 3D** again.
- At first everything will seem normal, just like the first time you did this. However, if you open the control panel, you will see that there are now 2 copies of bgal_3d in the list. What's going on? When you press Show 3D, Show 2D or similar buttons, a new display window is only created if an existing window of that type doesn't exist. If a window already exists, it will be re-used. When you close the display window, it doesn't actually get deleted, it is just hidden. When you display a new image in a 2-D window it replaces the original, but when you do it in a 3-D window it gets added to the list of objects in the window!
- You have probably been wondering why there are Show 3D+, Show 2d+, etc. buttons. When you press one of these buttons, rather than re-using an existing window, it will create a new window of the appropriate type, and this new window will be come the target of any new operations. Any other open windows will remain open with their current contents. If you wish to compare 2 3-D objects in the same window, then it makes sense to just press Show 3D multiple times. However, if you wish to do side-by-side comparisons, then the "+" buttons allow you to do that. The display windows will remain in existence until you close/exit the file browser window, which will trigger all of the other windows to close as well.

2-D Plotting/Graphing

- Next, let's take a look at 2-D plotting. Close any windows you still have open, except for the browser, and select **fsc_maskedtight_03.txt**. Press **Plot 2D**.
- You should now see a 2-D plot of a Fourier shell correlation (FSC) curve. These are typically used to estimate resolution in single particle analysis. To estimate the resolution, we would normally look for the resolution at which this curve crosses 0.143. Use the left mouse button and position the crosshairs at this point. You will note that a display of the coordinates of the crosshairs is



displayed in the upper right corner of the plot. You may also see another line or two like the "117." line in the example above. Another number is shown in parentheses as well. The numbers shown are: X (1/X), Y. Since the FSC is computed in Fourier, also known as reciprocal, space, the X-axis is actually the reciprocal of the number usually cited as the resolution. In my estimate, this curve shows a resolution of ~4.78 Å. The reciprocal is computed automatically for convenience. If another line is shown, that means your cursor was within a pixel or 2 of an actual data point in the file. and the exact data point is then shown.

- What if you want to estimate the number more precisely? You can zoom in on the plot by dragging a box with the right mouse button. This will give more precision in your estimate, though in this case that would be of questionable value. Once you have 'zoomed in' if you wish to rescale to the original scaling, simply click the right mouse button without dragging.
- As usual, you can open a control panel with the middle button. Most of the items there should be fairly obvious and straightforward. some of the more advanced features of this interface can be explored with the multi-column ptclfsc_01.txt file, but that is beyond the scope of this tutorial. Some of those capabilities are discussed in the single particle reconstruction tutorial.

FilterTool

One button we have not explored is FilterTool. This is an extremely useful and capable utility for image processing in both 2-D and 3-D, and has its own tutorial(s):
http://eman2.org/Programs/e2filtertool
the linked video tutorial is a great way to see how this tool works.

Conclusion

In this short tutorial, we have covered only the basics of GUI interaction in EMAN2. Just remember that when using any program in EMAN2, if you see an image/volume/plot display, that you can open a control-panel by middle-clicking, even if that display is embedded in a window with lots of other things. If you encounter other aspects of the GUI you find confusing, please consult the EMAN2 Mailing list:
https://groups.google.com/forum/#!forum/eman2
(which also has a fully searchable history)