The axial to radial transformation of 2D localizations for CREST

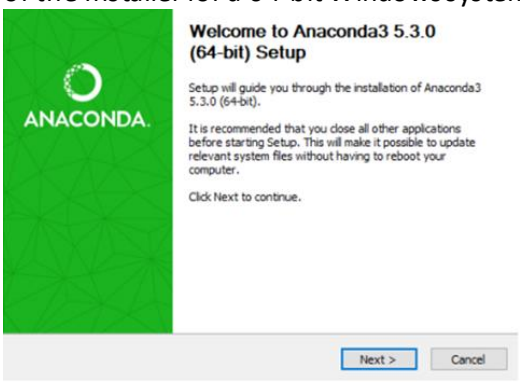
Section Outline

Section I

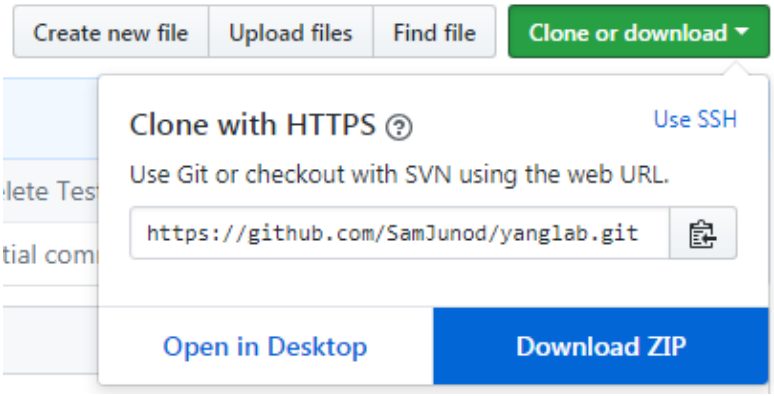
Software Download

Python 3

1) Install Python 3
    i. Since these scripts make use of several Python 3 libraries (tkinter, csv, random, os, sys, numpy, scipy, and math), the simplest way to install Python 3 with these required libraries is through the Anaconda distribution of Python 3 which can be found here (https://www.anaconda.com/download/).
    ii. Download the appropriate installer for your operating system (32-bit or 64-bit - Windows or Mac)
    iii. Once the installer has finished downloading, open it and follow the click-through instructions to install the Anaconda distribution of Python 3. An example of the first step of the installer for a 64-bit Windows system is shown below.



2) Download the script files
    i. Follow the URL link below.
        i. https://github.com/SamJunod/yanglab
    ii. Select the "Clone or Download" button and download the ZIP file. Extract the ZIP file to your preferred file location.



MatLab

1) Go to MatLab website, MATLAB - MathWorks - MATLAB & Simulink, and download the MatLab software.
    i. URL: https://www.mathworks.com/products/matlab.html

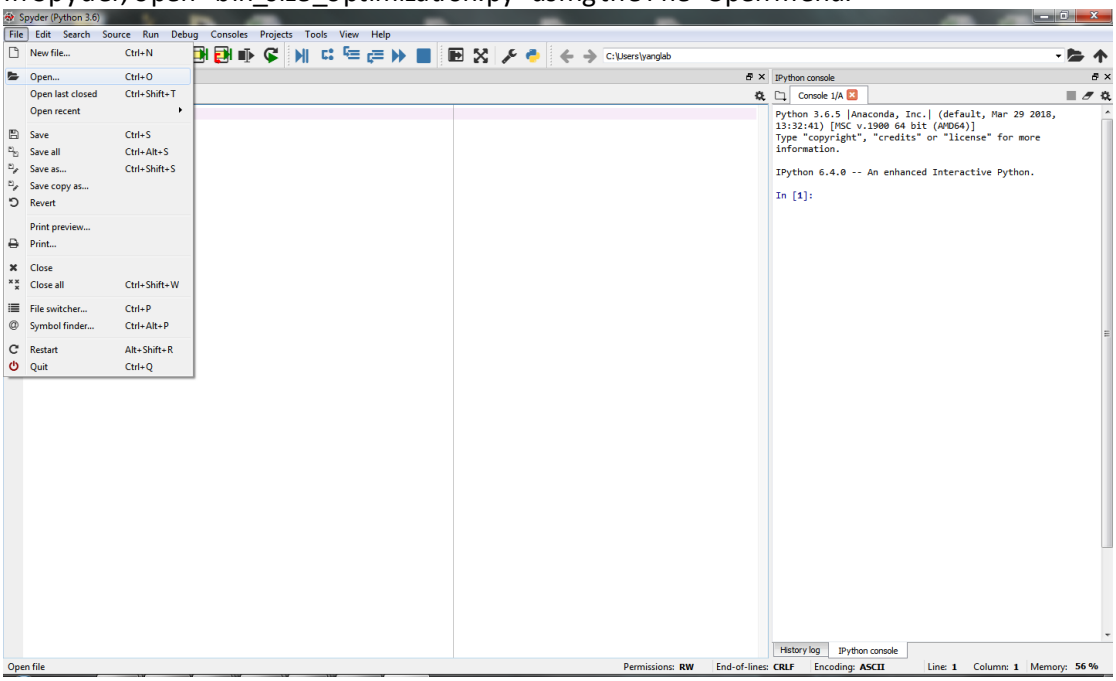
OriginLab

2) Go to OriginLab website and download the Origin software.
    i. URL: https://www.originlab.com/demodownload.aspx

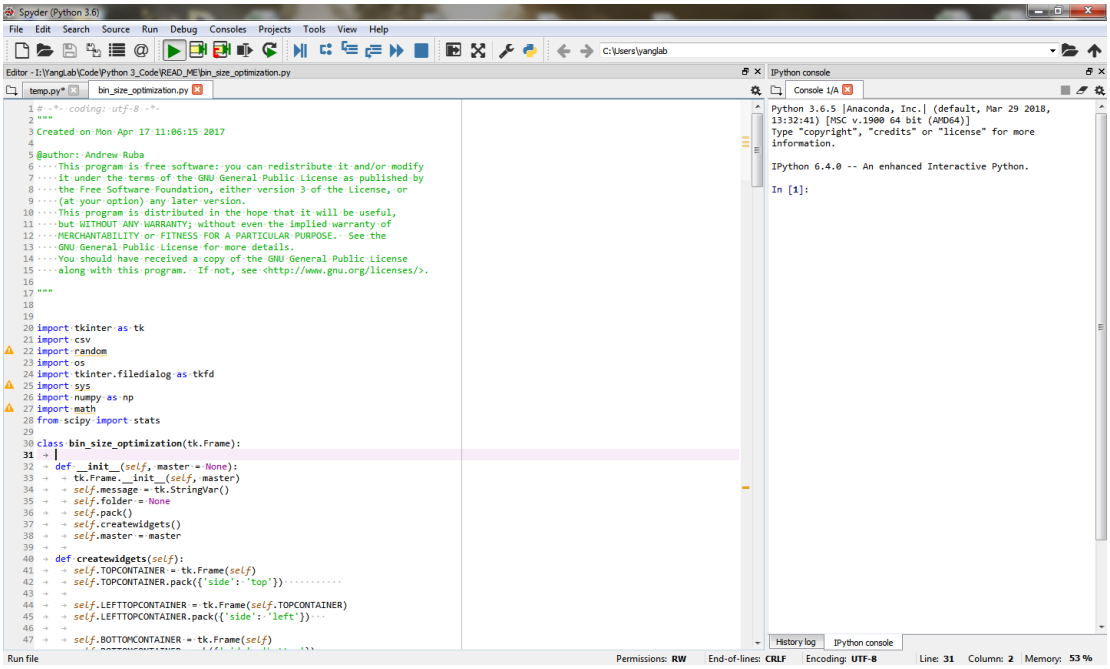
Section II
Bin Optimization

1) Packaged in the Anaconda distribution of Python 3 is an integrated development environment for Python 3 called "Spyder." Open Spyder– it should have an icon on your desktop or an icon in the program list under the start menu for Windows.

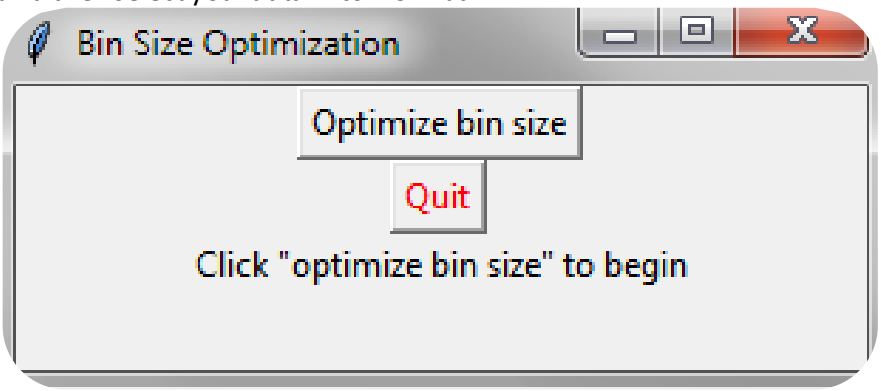
2) In Spyder, open "bin_size_optimization.py" using the File>Open menu.



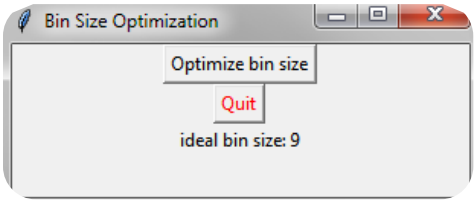
3) Then click the green "Run file" arrow in the toolbar.



4) Once the script begins to run, a pop-up window will appear. Select the Optimize bin size button and then select your data in csv format.



5) After the data is selected the script will run and output the ideal bin size. For CREST, the ideal bin size is 9.

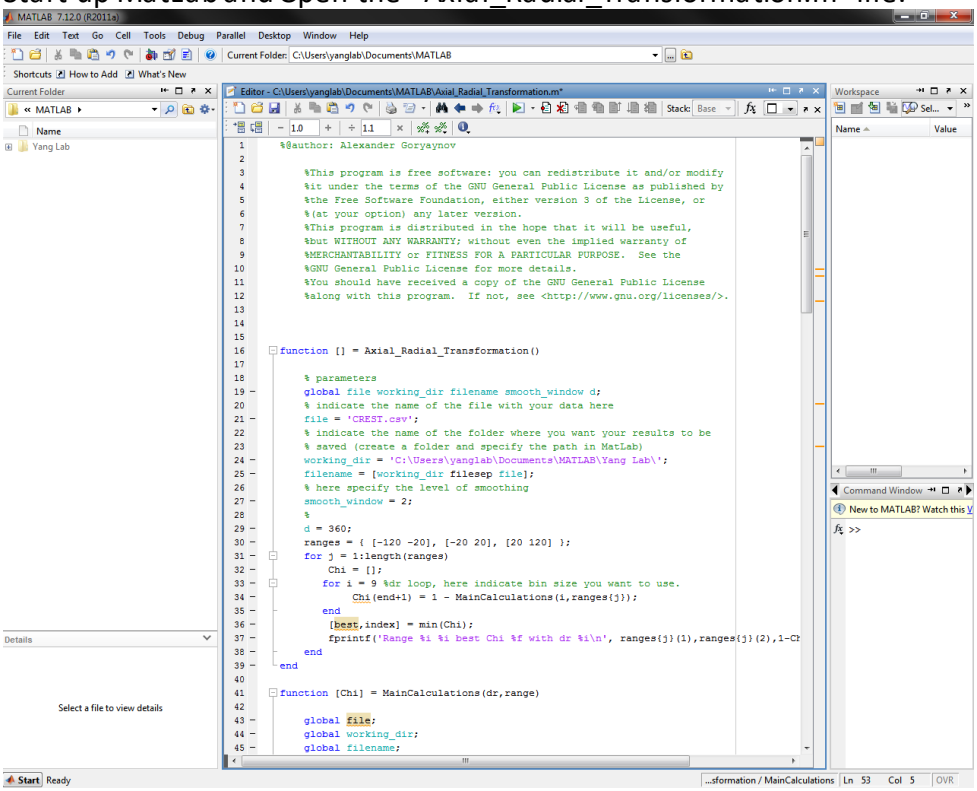


6) Select the "Quit" button to close the pop-up window.

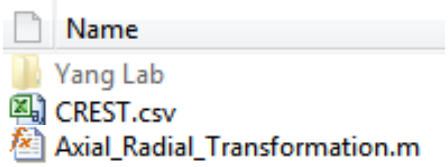
Section III

Axial to Radial Transformation

1) Start-up MatLab and Open the "Axial_Radial_Transformation.m" file.



2) The MatLab script (Axial_Radial_Transformation.m) and the 2D localizations (CREST.csv or SLD2.csv) will have to be in the same folder.



3) In the MatLab script, under the Editor box, make the following edits:

   i. File name (line 21, file = '***file name.csv***';)

```
20        % indicate the name of the file with your data here
21 -      file = 'CREST.csv';
```

   ii. Working directory(line 24, working_dir = '***folder location of data***';)

```
22        % indicate the name of the folder where you want your results to be
23        % saved (create a folder and specify the path in MatLab)
24 -      working_dir = 'C:\Users\yanglab\Documents\MATLAB\Yang Lab\';
```

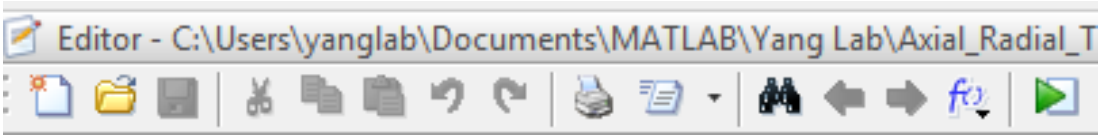
   iii. Axial ranges (line 30, ranges ={ [***axial ranges***] , [***for data***] };

```
30 -      ranges = { [-120 -20], [-20 20], [20 120] };
```

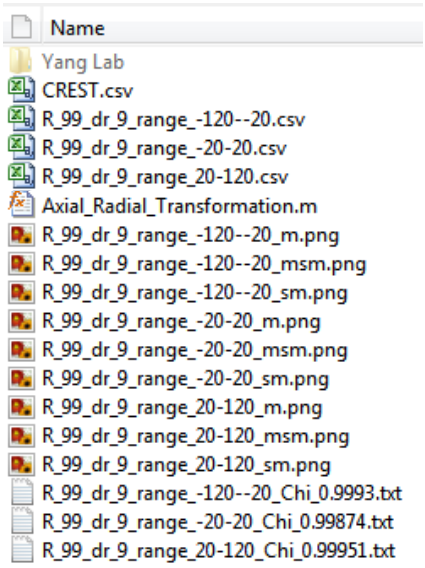
iv. Binsize (line 33, for i = ***optimal bin size***). This is the bin size determined from section II.

```
33 -        for i = 9 %dr loop, here indicate bin size you want to use.
```

4) Select "Save" and run the script by selecting the green arrow in the toolbar under the "Editor" box.



5) The program will write several items to the folder selected in the MatLab script.



6) Open the csv files created by the program. The values within each of the csv file are the raw results from the MatLab script. The file name of the csv document corresponds to the total radius of the program, the bin size(dr), and the range of the axial segment.

7) The csv file will have 3 columns of radial information (histogram count, raw matrix equation values, and smoothed matrix equation values). Copy the third column (smoothed matrix equation values) to a separate excel file.

	A	B	C
1	37.333	-3.2487	-3.2487
2	41.333	-0.29571	-0.63033
3	44.333	1.6534	0.54286
4	45.667	0.27085	1.711
5	52	3.2086	2.3956
6	54	3.7072	4.3005
7	56.333	5.9856	5.2587
8	53.333	6.0833	6.8563
9	50	8.5	8.2973
10	33.667	10.309	6.4729
11	2	0.60999	3.6396
12	0	0	0

	A	B	C
1	(-120 to -20)	(-20 to 20)	(20 to 120)
2			
3	-3.2487		
4	-0.63033		
5	0.54286		
6	1.711		
7	2.3956		
8	4.3005		
9	5.2587		
10	6.8563		
11	8.2973		
12	6.4729		
13	3.6396		
14	0		

8) Repeat the last step for the other axial ranges from the other excel files.

	A	B	C
1	(-120 to -20)	(-20 to 20)	(20 to 120)
2			
3	-3.2487	-1.9878	-3.616
4	-0.63033	2.0531	-0.071079
5	0.54286	4.0592	2.8321
6	1.711	3.2354	3.6213
7	2.3956	2.106	5.2665
8	4.3005	0.94231	4.6083
9	5.2587	0.71181	5.7293
10	6.8563	0.61688	5.3081
11	8.2973	0.56411	6.6573
12	6.4729	0.60162	4.7512
13	3.6396	0.33024	2.7955
14	0	0	0

9) These values will have to be normalized to the number of bins in each subsection. For instance, with a bin size of 9nm the section that ranges from -120nm to -20nm is comprised of 11.11 subregions (bins). This value is calculated by taking the number the length of the axial segment (100nm) and dividing by the bin size (9nm). Likewise, the subsection that ranges from -20nm to 20nm is comprised of 4.44 subregions, and the subsection that ranges from 20nm to 120nm is comprised of 11.11 subregions. The values in the -120nm to -20nm column will be divided by 100/9. The data from -20nm to 20nm is divided by 40/9, and the data from 20nm to 120nm is divided by 100/9.

(-120 to -20)	(-20 to 20)	(20 to 120)
-0.292383	-0.447255	-0.32544
-0.0567297	0.4619475	-0.006397
0.0488574	0.91332	0.254889
0.15399	0.727965	0.325917
0.215604	0.47385	0.473985
0.387045	0.2120198	0.414747
0.473283	0.1601573	0.515637
0.617067	0.138798	0.477729
0.746757	0.1269248	0.599157
0.582561	0.1353645	0.427608
0.327564	0.074304	0.251595
0	0	0

10) Find the maximum value of the whole data set.

(-120 to -20)	(-20 to 20)	(20 to 120)
-0.292383	-0.447255	-0.32544
-0.0567297	0.4619475	-0.006397
0.0488574	0.91332	0.254889
0.15399	0.727965	0.325917
0.215604	0.47385	0.473985
0.387045	0.2120198	0.414747
0.473283	0.1601573	0.515637
0.617067	0.138798	0.477729
0.746757	0.1269248	0.599157
0.582561	0.1353645	0.427608
0.327564	0.074304	0.251595
0	0	0
	Max	**0.91332**

11) Normalize the values in each column with the maximum value by dividing each value by the maximum value. For CREST, this maximum value is 0.91332.

(-120 to -20)	(-20 to 20)	(20 to 120)
-0.32013205	-0.489702	-0.356326
-0.06211372	0.5057893	-0.007004
0.05349428	1	0.2790796
0.16860465	0.7970536	0.3568486
0.23606622	0.5188214	0.5189693
0.42377808	0.2321418	0.4541092
0.51820063	0.1753572	0.5645743
0.67563067	0.1519708	0.5230686
0.81762909	0.1389707	0.6560209
0.63784982	0.1482115	0.4681908
0.35865195	0.0813559	0.275473
0	0	0

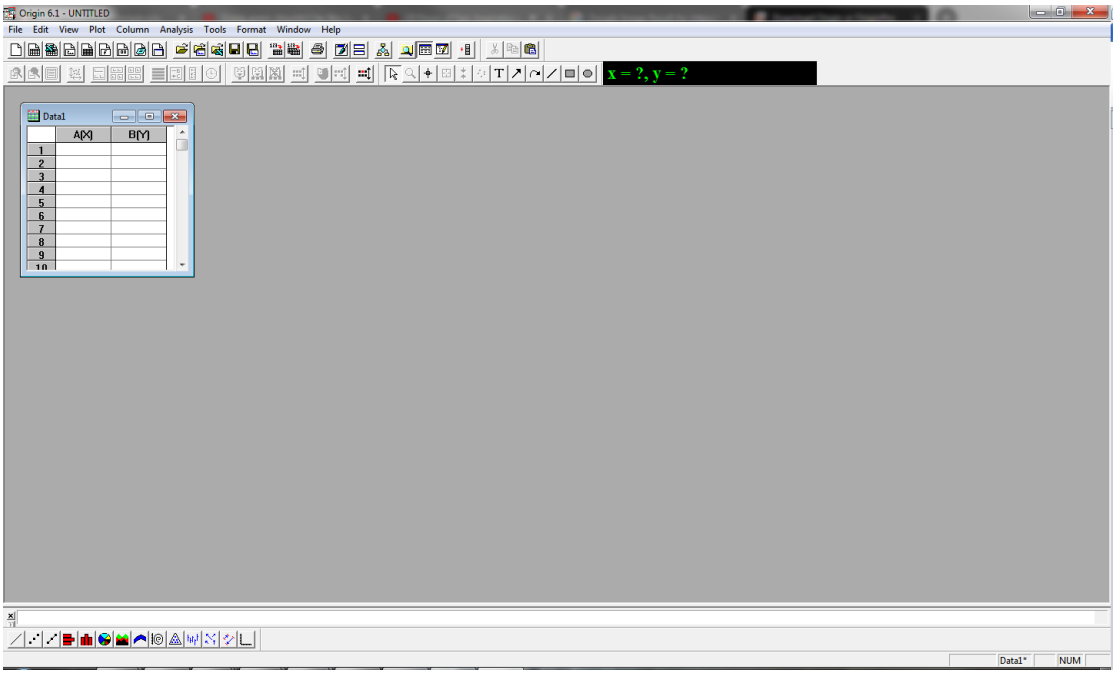
12) Even after normalization some negative values may still be present. Replace these values with 0.

(-120 to -20)	(-20 to 20)	(20 to 120)
0	0	0
0	0.505789	0
0.05349428	1	0.2790796
0.16860465	0.797054	0.3568486
0.23606622	0.518821	0.5189693
0.42377808	0.232142	0.4541092
0.51820063	0.175357	0.5645743
0.67563067	0.151971	0.5230686
0.81762909	0.138971	0.6560209
0.63784982	0.148211	0.4681908
0.35865195	0.081356	0.275473
0	0	0

13) Then add a radial dimension column "R (nm)". This column will start at half the bin size then increase by the bin size. For CREST, with an optimal bin size of 9, the column will start at 4.5 nm then increase by 9 nm for each density measurement.

R(nm)	(-120 to -20)	(-20 to 20)	(20 to 120)
4.5	0	0	0
13.5	0	0.505789	0
22.5	0.05349428	1	0.2790796
31.5	0.16860465	0.797054	0.3568486
40.5	0.23606622	0.518821	0.5189693
49.5	0.42377808	0.232142	0.4541092
58.5	0.51820063	0.175357	0.5645743
67.5	0.67563067	0.151971	0.5230686
76.5	0.81762909	0.138971	0.6560209
85.5	0.63784982	0.148211	0.4681908
94.5	0.35865195	0.081356	0.275473
103.5	0	0	0

14) To determine the peak density at different radial distances, you will then transfer the data set into Origin.



15) Copy the dataset into a new worksheet. With the radial dimension values as (X) and the density values as (Y).

Data1

	A[X]	B[Y]	C[Y]	D[Y]
1	4.5	-0.32013	-0.4897	-0.35633
2	13.5	-0.06211	0.50579	-0.007
3	22.5	0.05349	1	0.27908
4	31.5	0.1686	0.79705	0.35685
5	40.5	0.23607	0.51882	0.51897
6	49.5	0.42378	0.23214	0.45411
7	58.5	0.5182	0.17536	0.56457
8	67.5	0.67563	0.15197	0.52307
9	76.5	0.81763	0.13897	0.65602
10	85.5	0.63785	0.14821	0.46819
11	94.5	0.35865	0.08136	0.27547
12	103.5	0	0	0
13				

Data1

	A[X]	B[Y]	C[Y]	D[Y]
1	4.5	0	0	0
2	13.5	0	0.50579	0
3	22.5	0.05349	1	0.27908
4	31.5	0.1686	0.79705	0.35685
5	40.5	0.23607	0.51882	0.51897
6	49.5	0.42378	0.23214	0.45411
7	58.5	0.5182	0.17536	0.56457
8	67.5	0.67563	0.15197	0.52307
9	76.5	0.81763	0.13897	0.65602
10	85.5	0.63785	0.14821	0.46819
11	94.5	0.35865	0.08136	0.27547
12	103.5	0	0	0
13				

16) Select the column containing the radial dimension values (For this example, A(X)) and then while holding the "Ctrl" button on the keyboard select one of the density value columns (For this example, B(Y), or C(Y), or D(Y)).

Data1

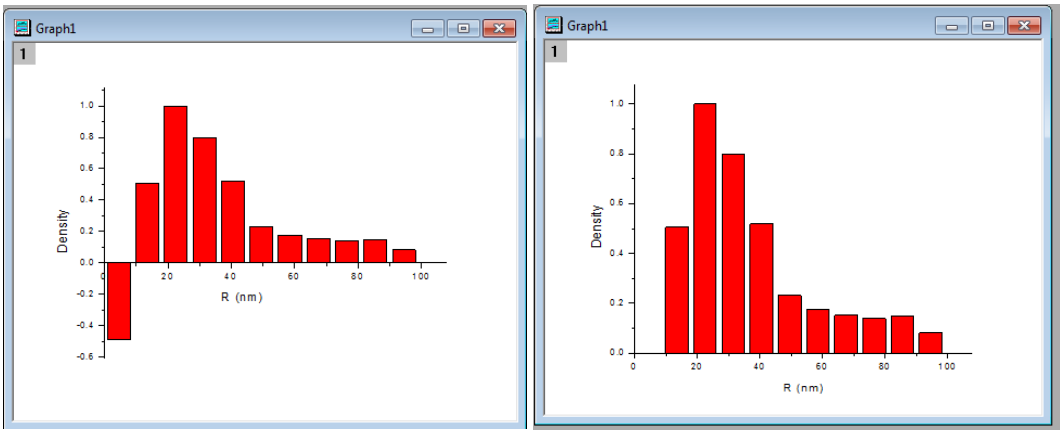
	A[X]	B[Y]	C[Y]	D[Y]
1	4.5	-0.32013	-0.4897	-0.35633
2	13.5	-0.06211	0.50579	-0.007
3	22.5	0.05349	1	0.27908
4	31.5	0.1686	0.79705	0.35685
5	40.5	0.23607	0.51882	0.51897
6	49.5	0.42378	0.23214	0.45411
7	58.5	0.5182	0.17536	0.56457
8	67.5	0.67563	0.15197	0.52307
9	76.5	0.81763	0.13897	0.65602
10	85.5	0.63785	0.14821	0.46819
11	94.5	0.35865	0.08136	0.27547
12	103.5	0	0	0
13				

Data1

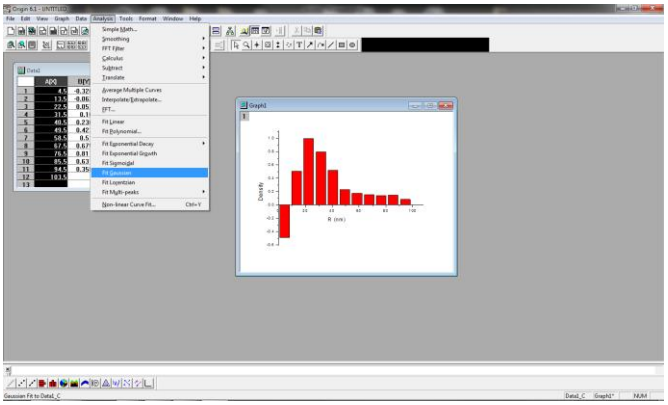
	A[X]	B[Y]	C[Y]	D[Y]
1	4.5	0	0	0
2	13.5	0	0.50579	0
3	22.5	0.05349	1	0.27908
4	31.5	0.1686	0.79705	0.35685
5	40.5	0.23607	0.51882	0.51897
6	49.5	0.42378	0.23214	0.45411
7	58.5	0.5182	0.17536	0.56457
8	67.5	0.67563	0.15197	0.52307
9	76.5	0.81763	0.13897	0.65602
10	85.5	0.63785	0.14821	0.46819
11	94.5	0.35865	0.08136	0.27547
12	103.5	0	0	0
13				

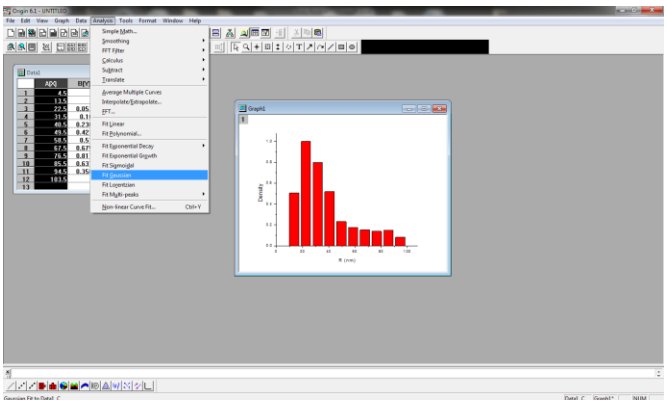
17) Then select the red vertical column graph button on the toolbar at the bottom of the Origin program.

18) After selecting the vertical graph column a graph will appear.

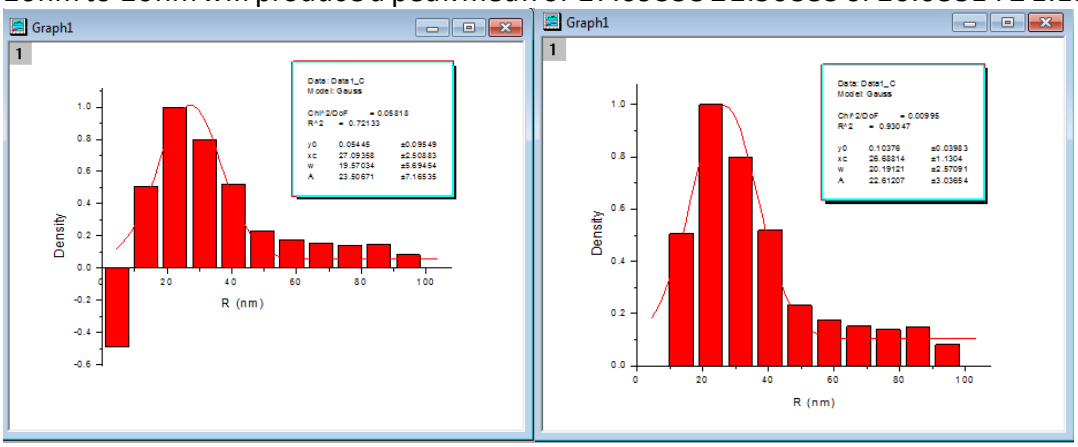


19) Select the graph, then on the toolbar Analysis>Fit Gaussian





20) Origin will fit a Gaussian to the selected graph. The xc will be the mean of the peak density distribution. This value is the transport route with largest density distribution within that axial range. The density distribution of CREST, within the axial range of -20nm to 20nm will produce a peak mean of 27.09358 ± 2.50883 or 26.68814 ± 1.1304.

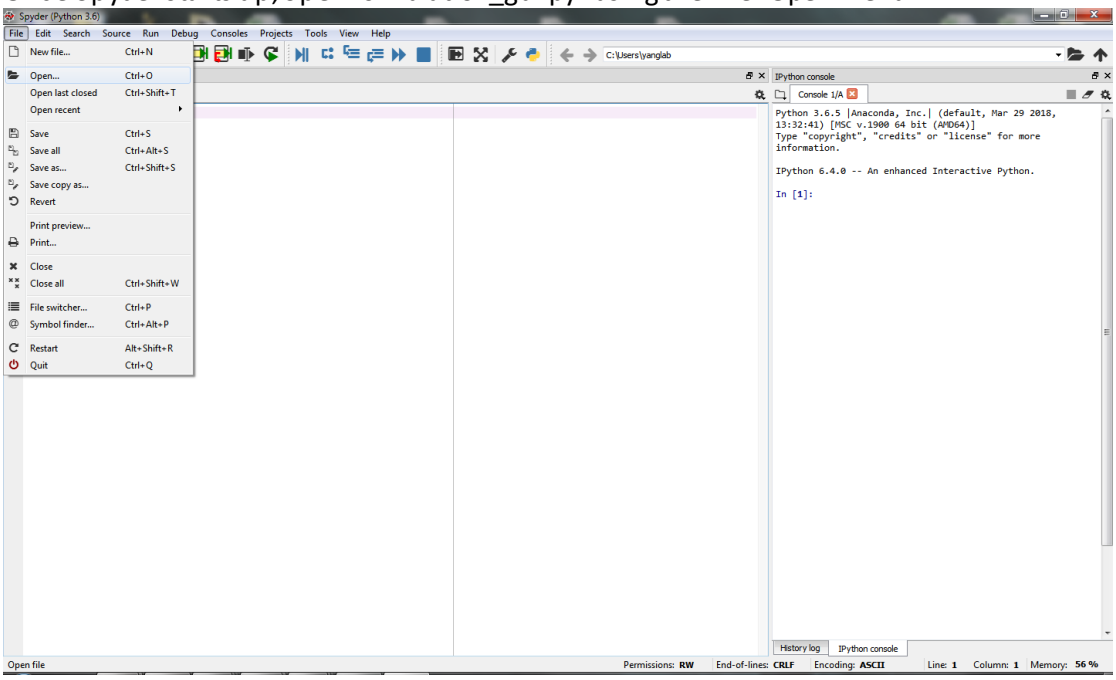


Section IV

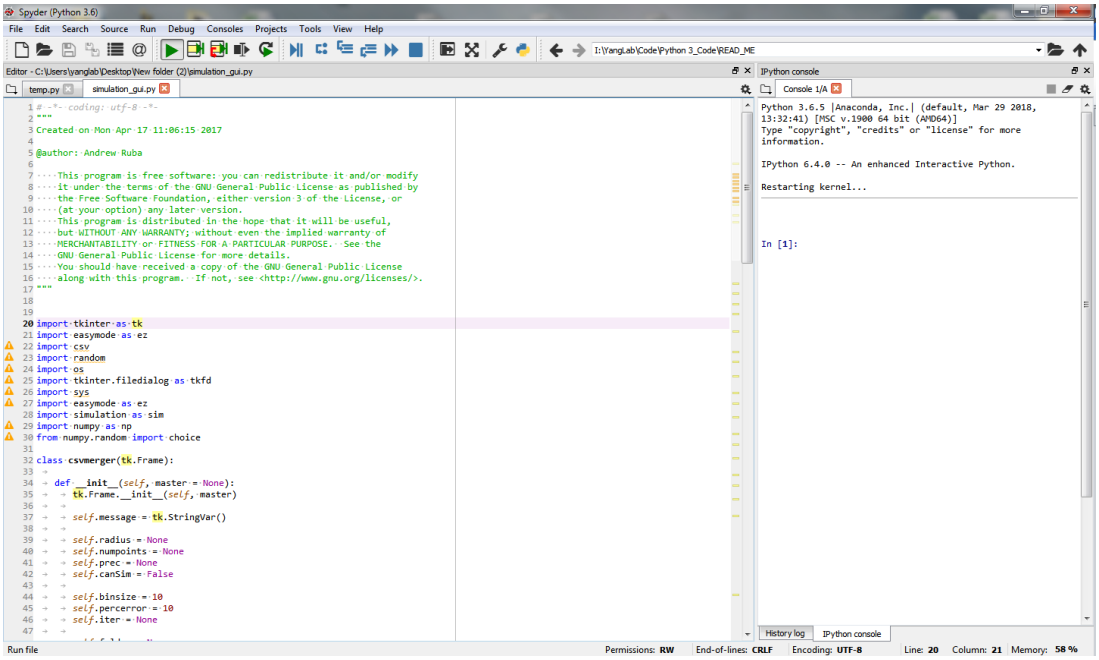
Point Number Simulation

1) Similarly to section II, Open Spyder – it should have an icon on your desktop or an icon in the program list under the start menu for Windows.

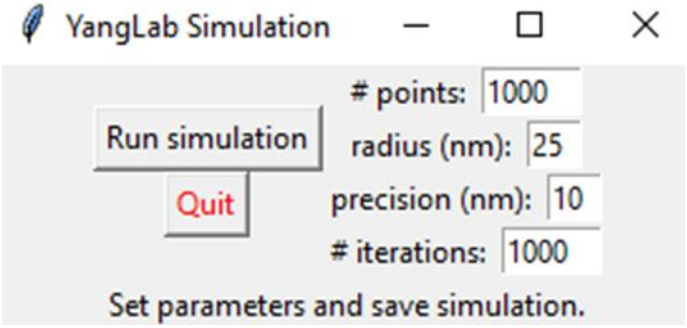
2) Once Spyder starts up, open "simulation_gui.py" using the File>Open menu.



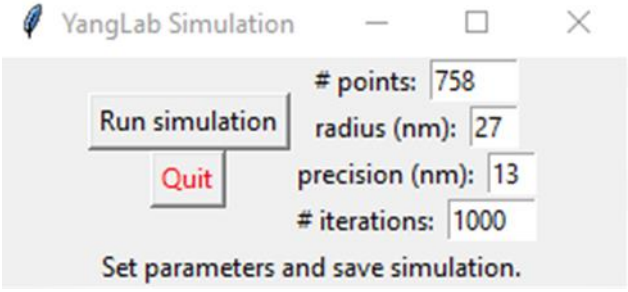
3) Once "simulation_gui.py" is open, click the green "Run file" arrow in the toolbar.



4) Once the script begins to run, a pop-up window will appear.

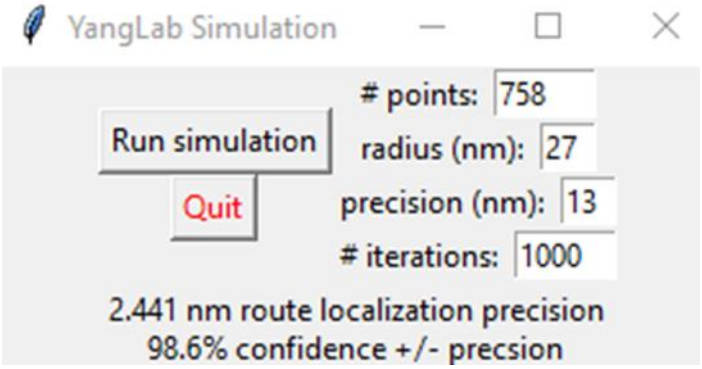


5) Input the following values;
   i. The number of points (***# points:***).
   ii. The peak radius found from Origin as explained in section III (***radius (nm):***).
   iii. The precision of the your single-molecule localizations (**precision (nm):**).
   iv. The number iterations you would like the simulation to run (**# iterations:**).



6) Click "Run simulation" after entering the proper integer values for the simulation parameters you would like to run. The optimal bin size will be dynamically calculated according to the

parameters and the reproducibility rate and route localization precision will be written into the gui message area.



7) Select the "Quit" button to close the pop-up window.