

Biophysics Assignment 1
Ananyapam De
18MS075

1. Answers to Ex

1.1. `input__file=raw_input('Enter input filepath and name:\n')`
`infilepath = input__file`

1.2. The number of frames is $\text{len}(\text{masterlist})/3$ because we have three coordinates to deal with: x,y,z. Hence the actual number of frames is the length of masterlist divided by three to account for the overcounting of the three coordinates.

1.3. The correlation window 'tcor' is a variable that gives a measure of the spectrum of values of time over which each averaging takes place. The correlation window moves forward with each iteration and gradually covers the entire length of the array. Too low tcor would give results in the ballistic regime where interparticle interactions become significant. On the other hand, if the correlation window is large, it takes averages for bigger sets of data hence giving poor statistics. A lower 'tcor' would give better statistics due to more averaging. To get good statistics, choose $\text{tcor} = 1/10\text{th}$ of the total data points present.

1.4. The variable norm increments the values of its elements by one in each iteration. The variable tcf increments the value by the square of the increment of length in each iteration. Then each element of tcf is divided by each element of norm to give the msd for each element. Hence the norm is used as a normalizing factor. The norm is time-dependent because it is different for each element in the array. The norm counts how many times the jth step has been taken during diffusion, so when many MSD values are obtained for a particular step in the diffusion, dividing the total sum of MSD for the frame by its norm would give the average MSD.

1.5. The modified code as below:

```
ps=raw_input("Enter the time spacing between consecutive
frames:\n")
delT=int(ps)
g=raw_input('Enter output filepath and name:\n')
out_object=open(g,'w+')
for i in range(0,tcor):
    msd = tcf[i]/norm[i]
    print 'Normalizing; time diff. = ',i
    time = delT*i
    out_object.write("%i %5.3f\n" % (time,msd))
out_object.close()
```

2. Append the following code snippet at the end of the program:

```
out_object_2=open(g,"r")
if out_object_2.mode == "r":
    contents_2=out_object_2.read()
    out_object_2.close()
old_stdout = sys.stdout
f = open("memprot_DPPC_310K_2.dat", "r")
if f.mode == "r":
    contents =f.read()
    f.close()
log_file = open("logfile.log","w")
sys.stdout = log_file
print "Input file \n"
print contents
print "\n"
print "Correlation time \n"
print tcor
print "\n"
print "Output File \n"
print contents_2
sys.stdout = old_stdout
log_file.close()
```

3. The following changes are done to the existing code in `msd_tcor_membrane.py` -
- ```
nfrm = len(masterlist)/2 #in line 20 and
for i in range(0,len(masterlist),2): #from line 44 onwards
 xi=masterlist[i]
 yi=masterlist[i+1]
 for j in range(i,len(masterlist),2):
 xj=masterlist[j]
 yj=masterlist[j+1]
 dt=(j-i)/2
 drSq=(xj-xi)**2+(yj-yi)**2
 tcf[dt] += drSq
 norm[dt] += 1
 print 'Complete frm ', i/2, '\n'
```

The following changes are done because the pollen grain data only has 2 coordinates  $x, y$  which reduces the dimension of the system to 2. Hence the code needs to be modified the following way.

4. Answers-

- 4.1. The correlation window gives better statistics on taking smaller values. This is because the correlation window is a measure of the spectrum of values it averages at a time. A smaller value of 'tcor' would imply a larger number of averages of small sets of data. A larger value of 'tcor' would mean the spectrum of values through which it averages is larger but would give poor statistics.
- 4.2. Changing 'tcor' changes the size of the window of averaging which does not affect the our results much, however the results do vary slightly due to the changing the spectrum of values of averaging.
- 4.3. Answered above in 1.4.
- 4.4. The main reasons for not getting a straight line are-
- 4.4.1. Inherent poor statistics of the system
  - 4.4.2. The finite volume of the container due to which the slope starts to decrease at the end. The interaction of the particles with the wall of the container becomes significant after a suitable time.
  - 4.4.3. Ballistic effects at small values of  $t$ . The particles are so close to each other that the interactions between them become significant and can no longer be neglected. Due to this, the slope is parabolic at the beginning.

5. Answered as below

5.1. For the membrane protein system, the system follows:

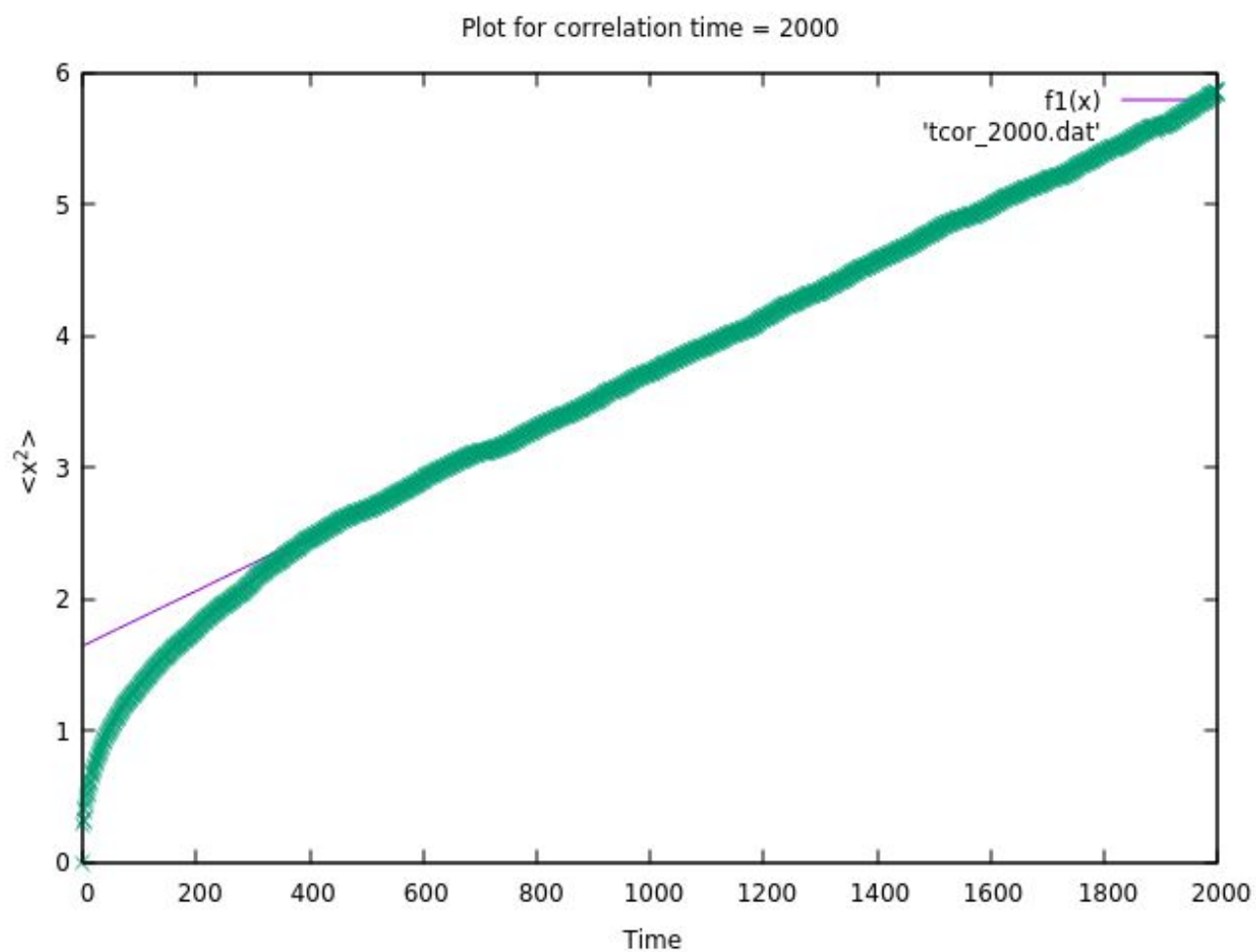
$$\langle x^2 \rangle = 6Dt$$

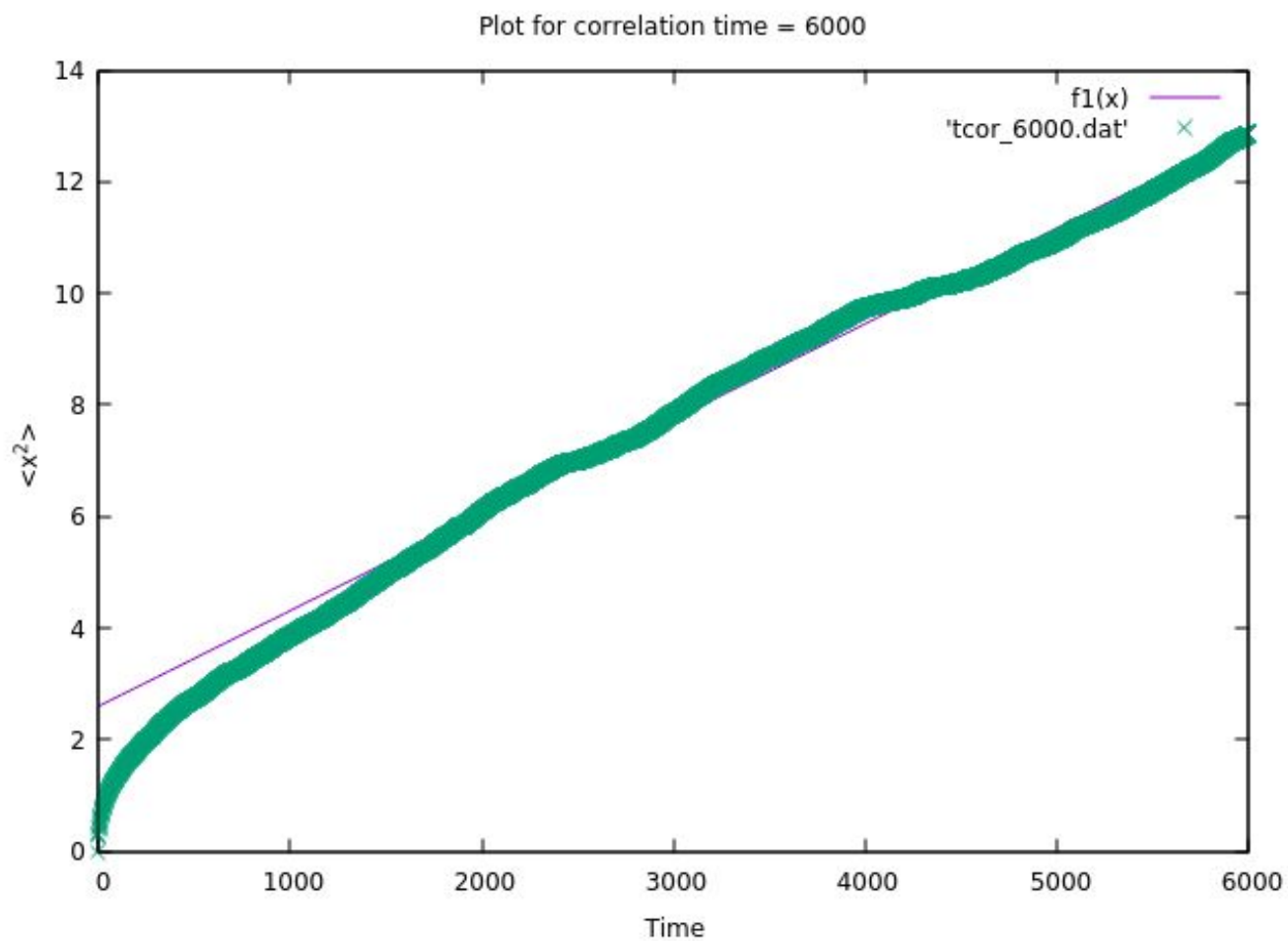
| System Name      | System Dimension | Total Number of Frames | Correlation time | Slope from fit(angstrom /10 picosec <sup>2</sup> ) | Diffusion coefficient (cm <sup>2</sup> /sec) |
|------------------|------------------|------------------------|------------------|----------------------------------------------------|----------------------------------------------|
| Membrane Protein | 3                | 50000                  | 2000             | 0.00209219                                         | $3.48 \times 10^{-9}$                        |
| Membrane Protein | 3                | 50000                  | 6000             | 0.00171614                                         | $2.86 \times 10^{-9}$                        |
| Membrane Protein | 3                | 50000                  | 8000             | 0.00172068                                         | $2.86 \times 10^{-9}$                        |

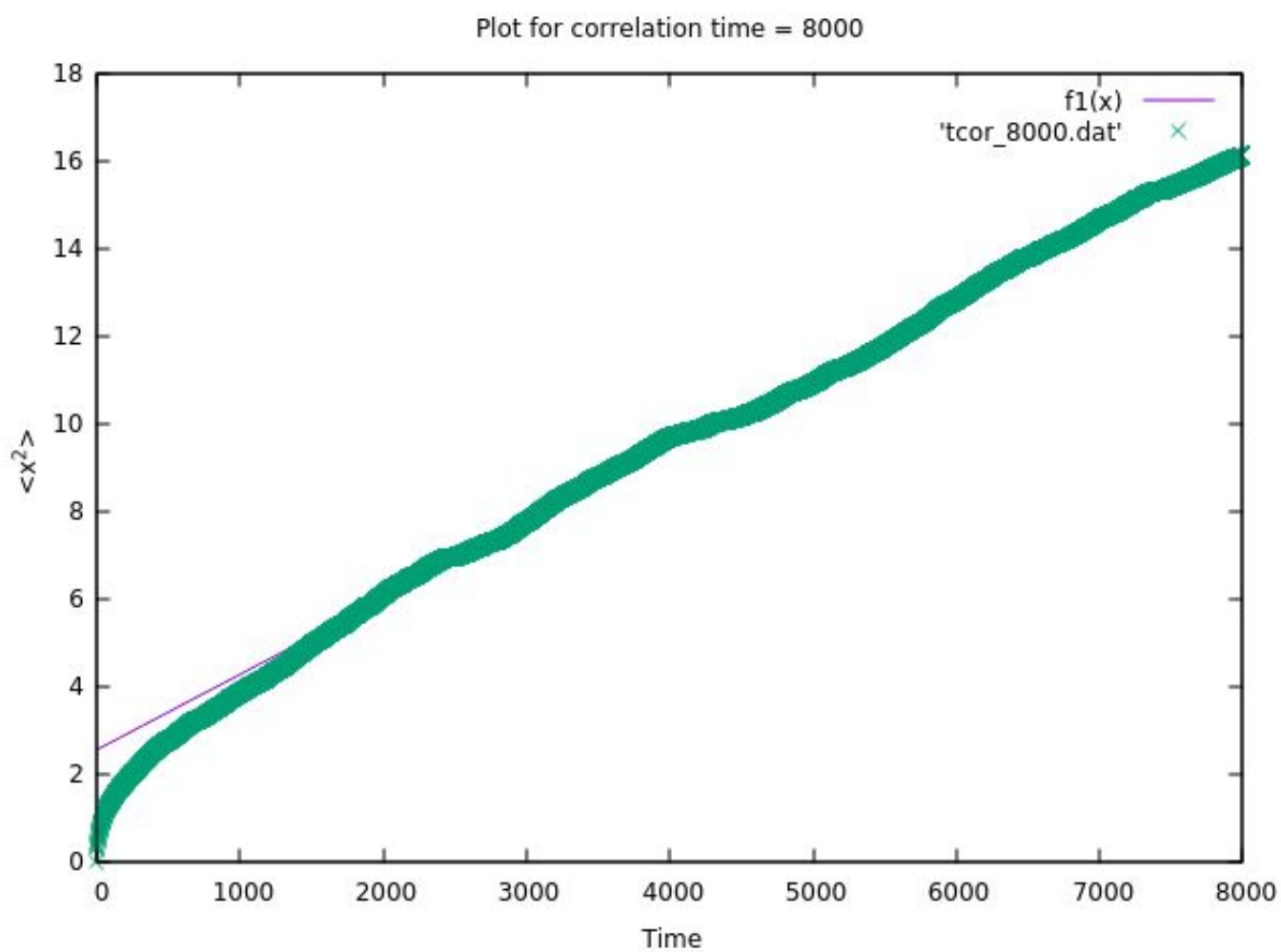
Note for all the three graphs below, the scales are as follows-

X-axis: 1 Unit = 10 picoseconds

Y-axis: 1 Unit = 1 Angstrom<sup>2</sup>







Average Diffusion Coefficient =  $3.17 \times 10^{-9} \text{ cm}^2/\text{sec}$

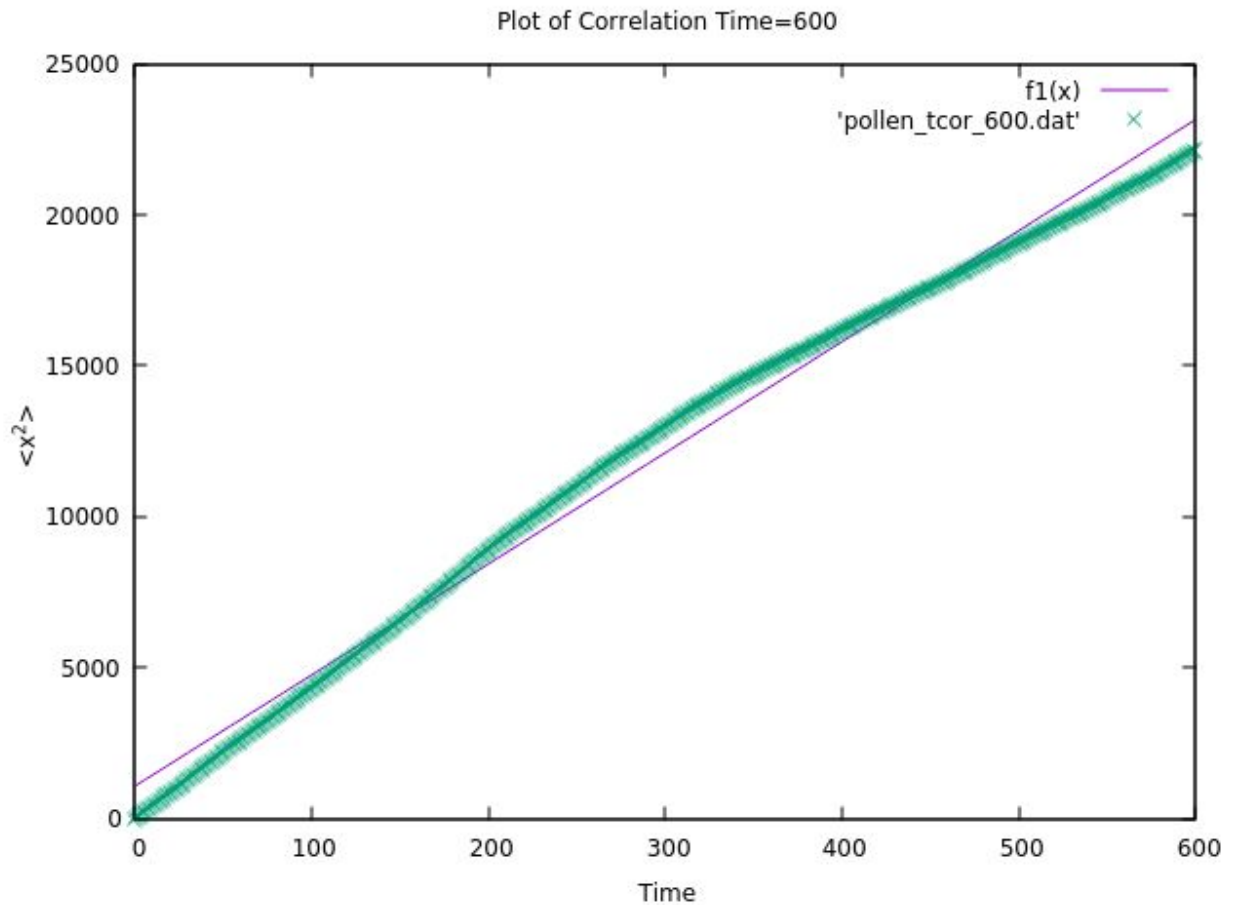
5.2. For the pollen grain system, the system follows:  $\langle x^2 \rangle = 4Dt$

| System Name  | System Dimension | Total Number of Frames | Correlation time | Slope from fit(micron <sup>2</sup> /0.1 second) | Diffusion coefficient (cm <sup>2</sup> /sec) |
|--------------|------------------|------------------------|------------------|-------------------------------------------------|----------------------------------------------|
| Pollen Grain | 2                | 8000                   | 600              | 36.8464                                         | $3.68 \times 10^{-6}$                        |
| Pollen Grain | 2                | 8000                   | 1000             | 37.0245                                         | $3.70 \times 10^{-6}$                        |
| Pollen Grain | 2                | 8000                   | 2000             | 48.3628                                         | $4.83 \times 10^{-6}$                        |

Note for all the three graphs below, the scales are as follows-

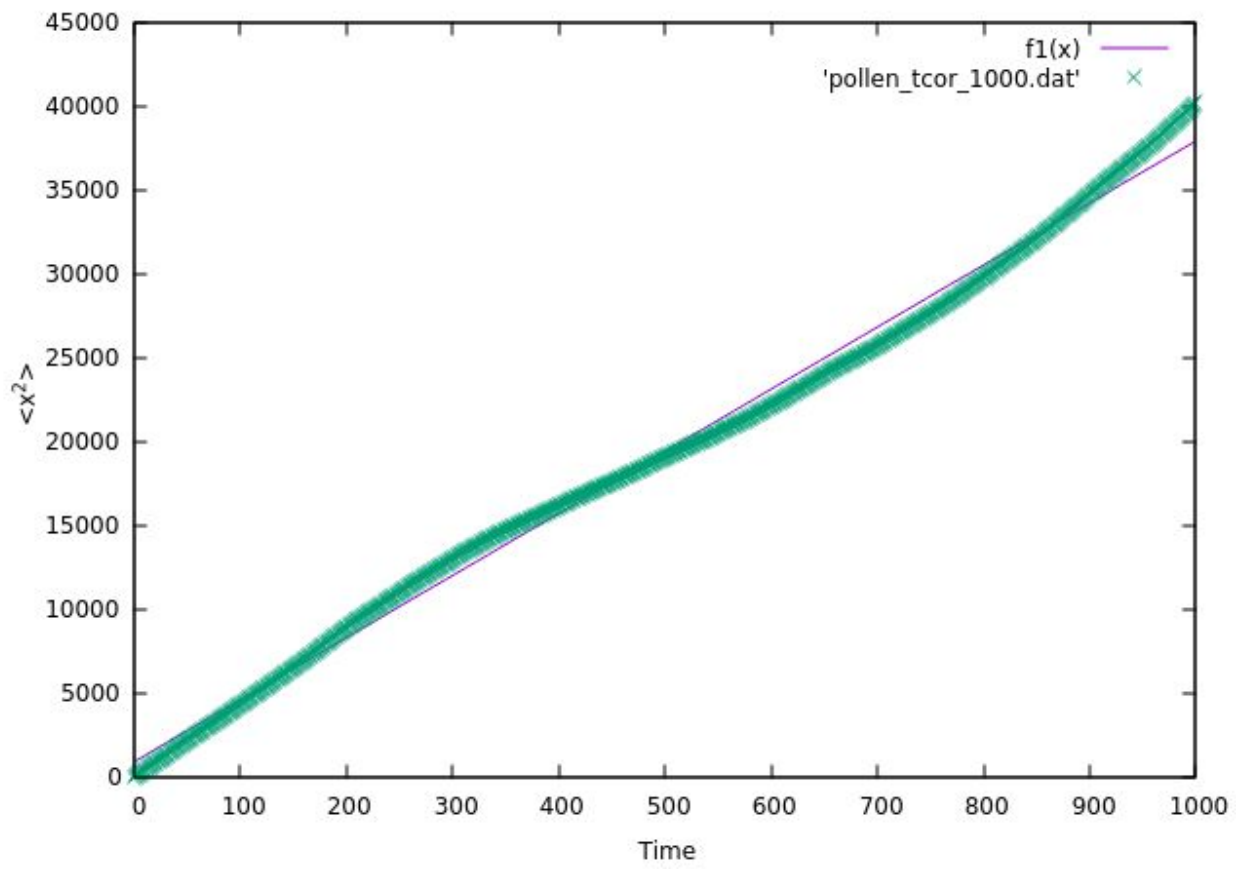
X-axis: 1 Unit = 0.1 second

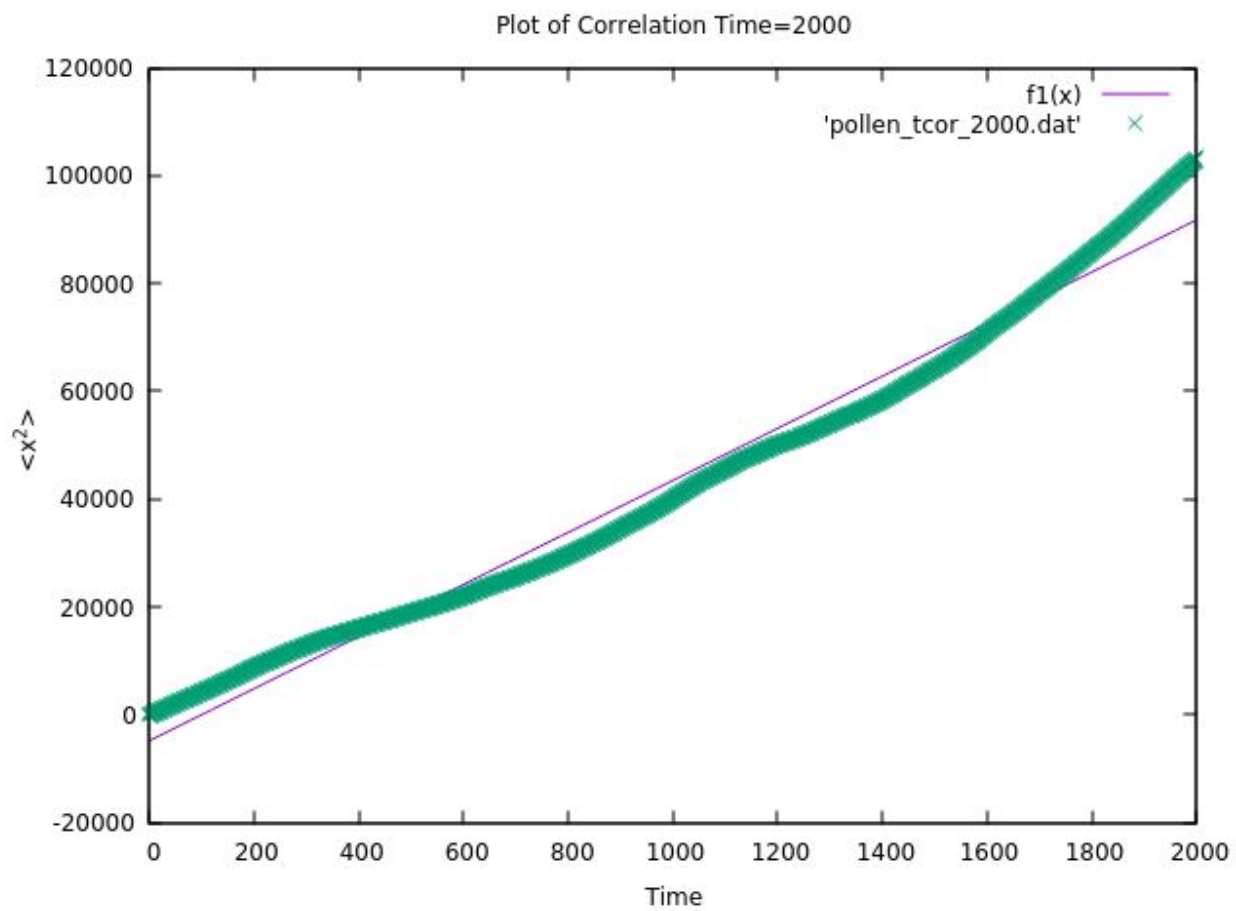
Y-axis: 1 Unit = 1 Micron<sup>2</sup>





Plot of Correlation Time=1000





Average value for Diffusion Coefficient =  $4.07 \times 10^{-6} \text{ cm}^2/\text{sec}$