10.34 Fall 2006
HW Set #11 Solution
Kinetic Monte Carlo

Notes:
The code `p1_main` will not generate all of the output for the problem, but enough for parts 1 and 2. The remaining part can be done with `p1_main_part3`. However, once these are run (or if you download the *.mat files posted), you can simply use the `p1_postprocessing` file to generate all the plots and output. If you change the name of the saved *.mat files, you will need to also change the "load" commands.

Algorithm (Gillespie):
The way to perform kinetic MC is relatively simple in theory. The idea is that you calculate a characteristic time for a reaction to occur based on the rates of all possible reaction. This characteristic time is defined as:

$$\tau = \left(\sum_{reactions\ i} rate_i \right)^{-1} \qquad units\ of\ time\ (per\ reaction\ event)$$

It is of the utmost importance to have the correct units on tau, otherwise the results will be meaningless. Tau turns out to be an extensive property and will ultimately depend on the system size, as well as the other parameters that determine the rate of a reaction (T, C_i's, parameters). You can rationalize this because if you have a larger volume (more molecules), the time between reaction events should be less. This is why you never want volume terms in you rates when computing tau. The ***units*** on the rates should be in terms of ***reactions events per time*** *for the entire system*.

Once tau is calculated, you estimate the time until the next reaction event by using a random number:

$$t_{rxn} = t_{current} - \tau \cdot \ln\left(rand\,[0,1] \right)$$

Now, you must determine which reaction will occur when t_{rxn} is reached. At this time, you have determined that some reaction will occur, and the probability of which reaction it is will be determined by the relative rates. A reaction with a fast rate will have a higher probability of occurring than one with a low rate. This is typically done by using another random number between 0 and 1, and giving each reaction a chunk of the interval [0,1] based on the value of $rate_i \cdot \tau$. Taking all reactions will span the entire interval, and reactions with larger rate will happen more often.

After this is done, it is simply a matter of keeping track of all of the appropriate variables and optimizing your code to run in an efficient manner, which is not necessarily a trivial task.

Major Issues:
There are some issues that should have been dealt with the two major ones being how to round the fractional molecules per droplet and the correct way to compute the rates. A fractional molecule should have rounded in a probabilistically-correct way depending on the fraction. If there were 0.25 molecules per droplet, you should have 0 ROOH's 75% of the time and 1 ROOH 25% of the time. A more correct way to deal with this (and droplets that had > 1 molecules) would be to use a probability distribution (e.g. Poisson distribution) for the number of molecules per droplet. Since this was not given, simple up and down rounding was acceptable. See the code for how this was done (it is easy).

The correct way/unit to compute the rate was important. The rates needed to be used with the units of reaction_events/(sec*droplet). This was accomplished by computing the rate in terms of # per second per volume, and then multiplying by the droplet volume. This will result in units of seconds for tau, which is what you want. Doing this will ensure you get the correct characteristic time. One other minor point was in dealing with the recombination rate, which was dependence on $[ROO]^2$. At the discrete level, you should actually use $N_{ROO}*(N_{ROO}-1)$ instead of N_{ROO}^2. This accounts for the fact that if one molecule is present, the reaction cannot happen. The average time between recombination events is $1/k*N*(N-1)*Vdrop$; each time a recombination event occurs N is reduced by 2.

A bit of interesting (though not required) information:
This was not required, but this shows the time progression of one trajectory for D = 250 nm. You can see the piecewise increase of ROOH (which happens in bursts with lags in between). The lags occur when 2 ROO recombine and another initiation event must happen to get things going again. In this case, you never have more than 5 ROO molecules present at any one time, and many times there are zero.
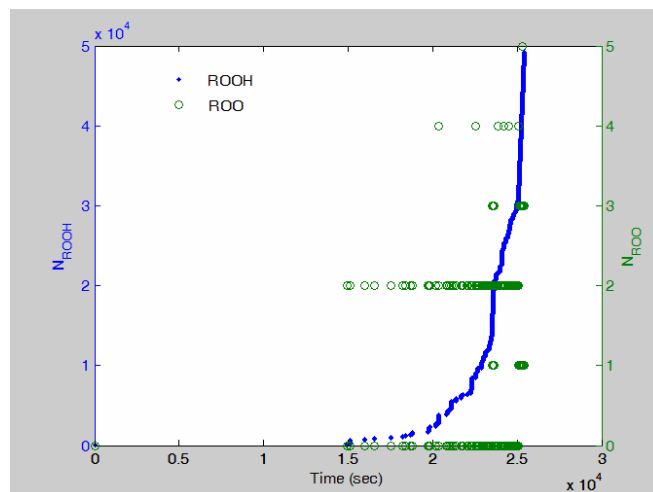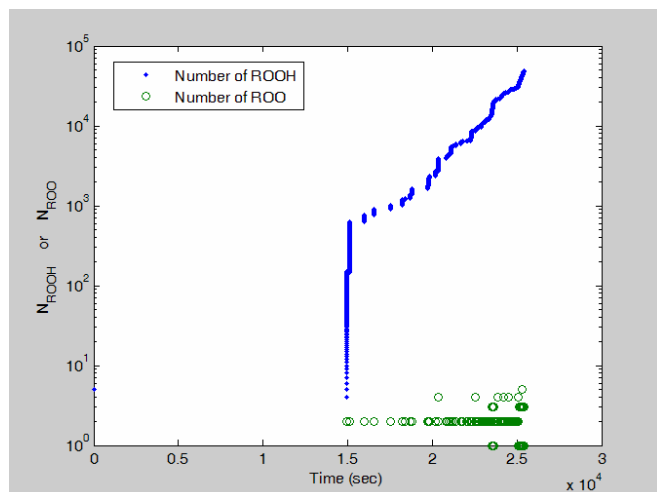
Part 1:

The following plots show the continuous solution to the problem. The three small plots show the time dependence of the concentration of ROOH, while the large figure shows the rate of failure as a function of diameter and k-value. It is easy to see that the small droplets have much more dependence on k over the ranges studied. This is due to their small size, making characteristic diffusion time to the droplet edge implied by *k* larger. For example, the rate of reaction 4 will change as follows for a small and large droplet for k from 1→ 100:

For D = 50: $1/50 = 0.02$ ➔ $100/50 = 2.0$
For D = 5000: $1/5000 = 0.0002$ ➔ $100/5000 = 0.02$

Since the rates 1 and 2 are not dependent on the droplet size, it is likely that the small diameter droplet will be affected more, since the absolute rate will be larger.
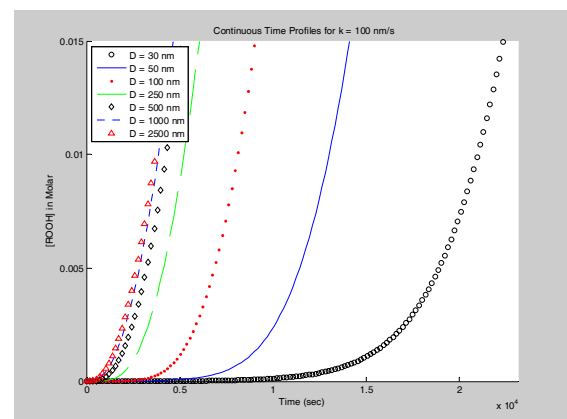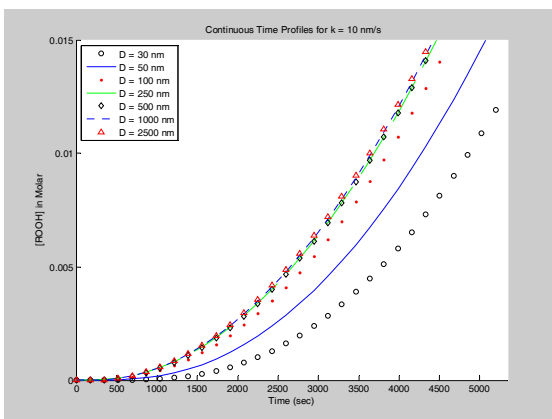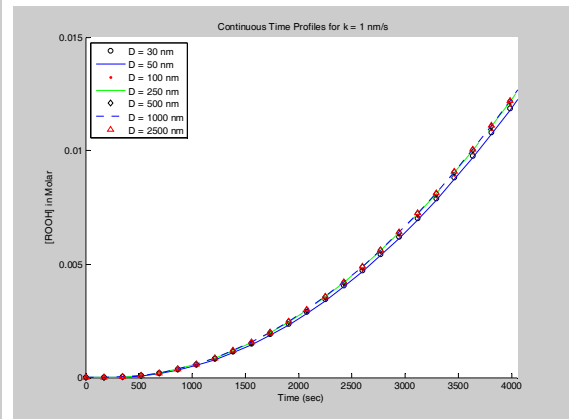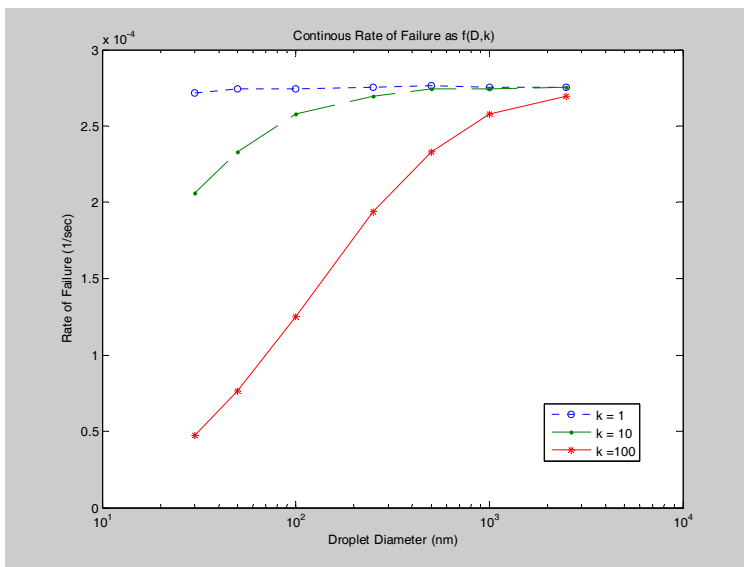
Part 2: MC solution for D = 100, 500, 1000, and 2500 nm

In this part, the solutions for all diameters are shown, even though only several diameters were asked for. The histogram of the time to failure and rate of failure are both shown, with the mean value and +/- one standard deviation shown. The mean and standard deviation only take into account the droplets that reached the critical concentration, and this will be discussed next.

When dealing with a problem like this, there is usually some degree of ambiguity in what are the appropriate quantities to measure/report. We suggested that the average rate of failure and the peak widths may give a reasonable description of the system behavior. However, even this is somewhat ambiguous, given that there are a number of ways to describe the average of something. There are two logical averages that one would think about:

1. Normal arithmetic average: $\langle r \rangle_N = \dfrac{1}{N} \sum_{i=1..N} r_i = \dfrac{1}{N} \sum_{i=1..N} \left(\dfrac{1}{t_i} \right)$

2. Log-based average: $\langle \ln(r) \rangle_N = \dfrac{1}{N} \sum_{i=1..N} \ln(r_i)$

Since the reaction times are distributed on a log scale, it seems logical that the log-based average would be more important from a statistical perspective. However, the arithmetic average is a reasonable description of the entire system, and does not run into problems when the rate of failure is zero. One other (likely industrially significant) way of looking at the problem would be to examine the point in time when the overall concentration of ROOH in the system was above some threshold value. In this case, the total system volume is defined by: $N_{drop}*V_{drop}$. The time to reach the threshold concentration could be determined by solving this implicit equation for $T_{threshold}$:

$$[ROOH]_{threshold} = \frac{1}{N \cdot V_{droplet}} \cdot \sum_{i=1}^{N} \left\{ \int_{0}^{T_{threshold}} N_{ROOH,i}(t)\, dt \right\}$$

This would require keeping track of number of ROOH's as a function of time, which would be computationally intensive. If one stored $1x10^6$ reaction events per droplet for $1x10^3$ trajectories, one would require on the order of 10 GB of memory. This limits the scale of the simulation, and nothing like this was expected in this homework assignment.

There are two general cases that we need to be concerned about. The first is when some droplets will have zero ROOH's to begin with (or start with some ROOH and reach a condition where there is no ROOH or ROO at some time > 0), and the second is when all droplets start with some ROOH and all reach the critical concentraiton. The latter is more straightforward and will be discussed first.

All droplets have ROOH and reach the critical concentration
It should have been clear that the way to compute statistics (particularly distribution widths) of these cases was to use the log-scale averages and standard deviations. This can be easily seen if a histogram is generated in log-space and not linear space. The linear histograms will be greatly skewed, whereas the log histograms will be nearly normally distributed (as you will see in the ones below). This implies that

the mean and standard deviation should be computed in log-space, and then converted to a real time. If the linear average was taken, the higher time and rate values would be unfairly weighted, and the standard deviation of the unscaled points would be meaningless since the distribution is heavily skewed. The procedure to calculate the "correct" mean, standard deviation, and confidence interval are:

$$\langle \ln(r) \rangle_N = \frac{1}{N} \sum_{i=1..N} \ln(r_i) \qquad\qquad \ln(\sigma) = StDev\left[\ln(r_i)\right]$$

$$\langle r \rangle_N = \exp\left[\frac{1}{N} \sum_{i=1..N} \ln(r_i)\right] \qquad \langle r \rangle_N \pm \sigma = \exp\left[\langle \ln(r) \rangle_N \pm \ln(\sigma)\right]$$

$$\ln(CI_{99\%}) = t(0.01, N-1) \cdot \frac{\ln(\sigma)}{\sqrt{N}} \qquad where\ t(\alpha, v)\ is\ the\ student's\ t\ distribution$$

$$\langle r \rangle_N \pm CI_{99\%} = \exp\left[\langle \ln(r) \rangle_N \pm \ln(CI_{99\%})\right]$$

These averages are for the N droplets that reached the final critical concentration. As assumed here, N and the number of droplets probed are the same, but won't be below.

Some droplet start without ROOH and/or reach a condition where ROOH + ROO = 0
It should be clear in this case, the log-scale approach doesn't work, since the rate will be zero for droplets that have no ROOH (completely stable) or reach the condition where ROOH and ROO are zero (the ROOH breaks into 2 ROO, and then 2 ROO recombine before they can produce more ROOH). There is not a good universal way to deal with this case, where you will have a distribution like described above for the reacting droplets, and then a delta function at zero rate for the stable droplets. The best way to convey the information would likely be to analyze the non-zero rates as done above, and then state the fraction of droplets of that were/became stable. In this manner, you will know how many droplets would become contaminated in a certain amount of time, and can find some ensemble average rate of failure if desired (where as an average alone would not give you as much information). There seem to be two reasonable ways to compute the ensemble average rate of failure. The first, and most straightforward, would be to take a simple arithmetic average of the rates for all droplets (including the stable droplets). The other, and more appropriate, way to compute the average would be to combine the log-average and the stable droplets:

$$\langle r \rangle_{ensemble} = \alpha \cdot \frac{0}{\alpha + N} + N \cdot \frac{\langle r \rangle_N}{\alpha + N}$$

In this formulation $\langle r \rangle_N$, as shown above is the log-calculated average, α is the number of stable droplets, and N is the number of droplets that reached the critical concentration. The benefit of this approach is that the ensemble average approaches the "correct" average as $\alpha \rightarrow 0$. In these cases, the standard deviation/confidence interval is not particularly meaningful if the stable droplets are being considered. They are still calculated for the reacting droplets in the solution below, since it does provide some information on the width of the distribution.

Results:

```
Initial number of molecules of: ROO, ROOH, RH, and O2 in droplet of size D
Droplet size (nm): 30 -- ROO = 0,  ROOH = 0.0085134,  RH = 85134.0193,  O2 = 0.85134
Droplet size (nm): 50 -- ROO = 0,  ROOH = 0.039414,  RH = 394138.9783,  O2 = 3.9414
Droplet size (nm): 100 -- ROO = 0,  ROOH = 0.31531,  RH = 3153111.8267,  O2 = 31.5311
Droplet size (nm): 250 -- ROO = 0,  ROOH = 4.9267,  RH = 49267372.2915,  O2 = 492.6737
Droplet size (nm): 500 -- ROO = 0,  ROOH = 39.4139,  RH = 394138978.3316,  O2 = 3941.3898
Droplet size (nm): 1000 -- ROO = 0,  ROOH = 315.3112,  RH = 3153111826.653,  O2 = 31531.1183
Droplet size (nm): 2500 -- ROO = 0,  ROOH = 4926.7372,  RH = 49267372291.4524,  O2 = 492673.7229
```
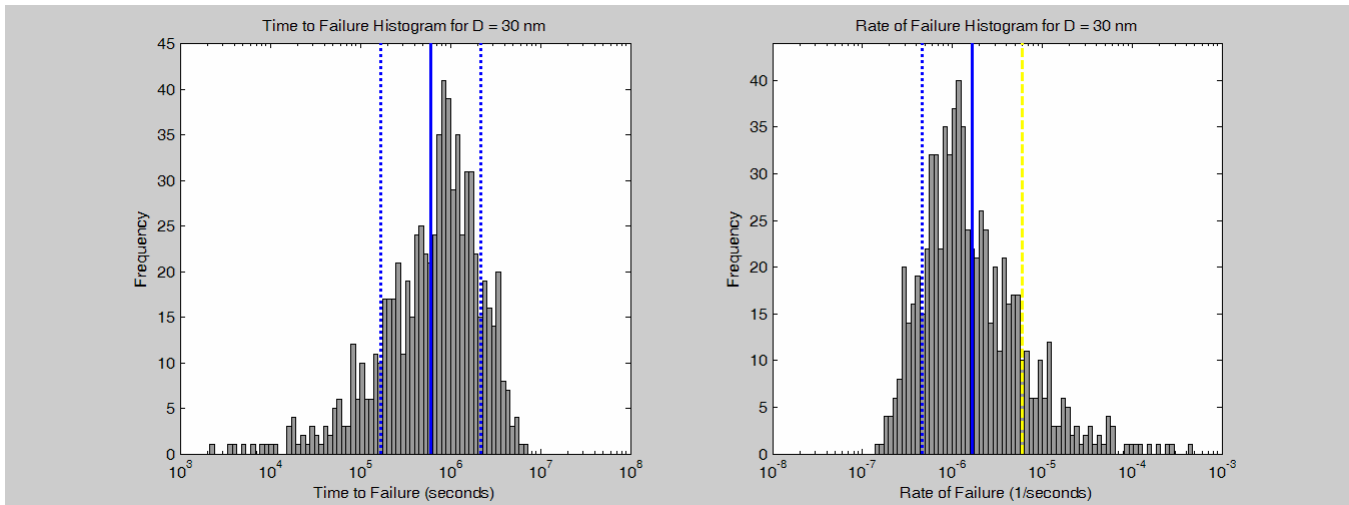
For Diameter = 30 nm

```
Number of steps that reached critical C_ROOH: 737 of 200000000 steps
Went too long (but did not become stable): 0 of 1705694 steps
Number of droplets that became stable (but reacted initially): 1704957 of 1705694
Fraction of droplets that started stable: 198294306 of 200000000 (99.1472%)

Continuum Time to Failure Stats: 1.0225 hours, or 3.68e+003 seconds
Mean Time to Failure (99% CI): 155.4258 < 167.9438 < 181.4701 hours
Time to Failure Stats (+/- StDev): 46.8414 < 167.9438 < 602.1411 hours
Time to Failure Stats (+/- StDev): 1.69e+005 < 6.05e+005 < 2.17e+006 seconds
Ensemble Avg. Time (with stable drops): 0.00061887 hours, or 2.23e+000 seconds

Continuum Rate Failure Stats: 0.97799 1/hrs, or 2.72e-004 1/sec
Arithmetic Mean Rate of Failure: 0.021593 1/hrs
Mean Rate of Failure (99% CI): 0.0055106 < 0.0059544 < 0.0064339 1/hrs
Rate of Failure Stats (+/- StDev): 0.0016607 < 0.0059544 < 0.021349 1/hrs
Rate of Failure Stats (+/- StDev): 4.61e-007 < 1.65e-006 < 5.93e-006 1/sec
```
Ensemble Avg. Rate (with stable drops): 2.1942e-008 1/hrs, or 6.09e-012 1/sec

The "mean" values and other statistics are only for the droplets that reacted, denoted by the first line of output (737 droplets in this case). An upper limit of 1×10^8 seconds was also put on the simulation, which is denoted by "Went too long". The next two lines should be self explanatory.

This first point is that the continuum rate of failure is much too large when compare to the either the mean value or the ensemble average. The ensemble average is much lower due to the massive number of droplets that died (only 737 out of 2×10^8 droplets reached the critical concentration). Due to the relatively small number of points, the CI in the mean of 737 points is somewhat large, on the order of 10%. The standard deviation is very significant as well. A histogram for the time and rate of failure are shown below (and for all diameters). The blue solid line is the log-scale mean, the dotted blue lines are the +/- one standard deviation, and the dashed yellow line is the arithmetic mean of the points (which happens to overlap with the std. dev. line in this case). You can see that the arithmetic average deviates significantly from the log average (half an order of magnitude). The lines will be the same for all diameters.

Time to Failure Histogram for D = 30 nm   Rate of Failure Histogram for D = 30 nm
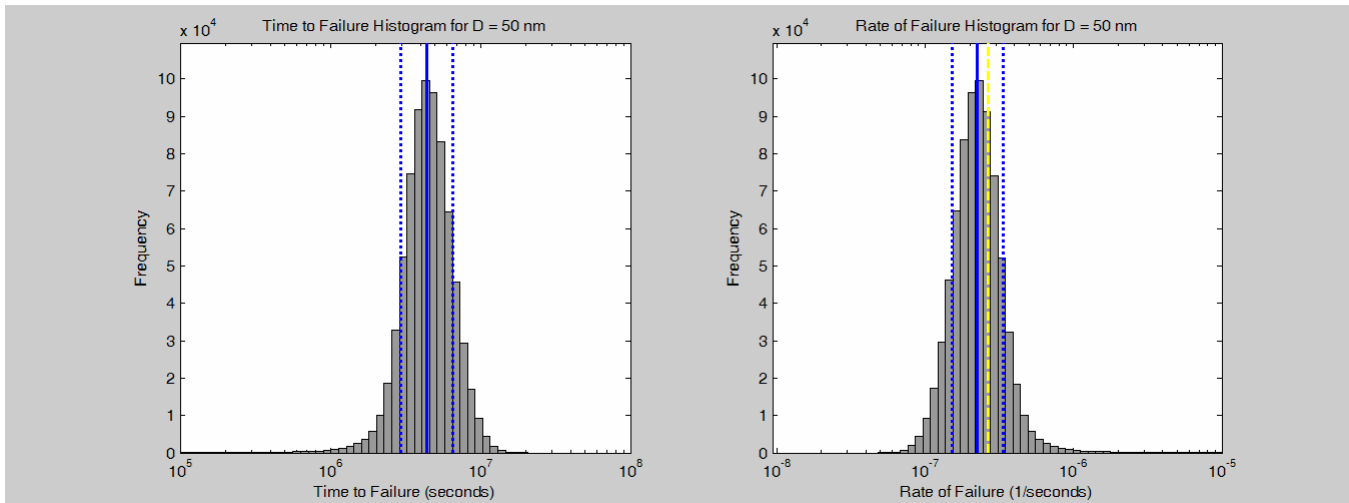
For Diameter = 50 nm

```
Number of steps that reached critical C_ROOH: 751938 of 41000000 steps
Went too long (but did not become stable): 0 of 1618626 steps
Number of droplets that became stable (but reacted initially): 866688 of 1618626
Fraction of droplets that started stable: 39381374 of 41000000 (96.0521%)

Continuum Time to Failure Stats: 1.0129 hours, or 3.65e+003 seconds
Mean Time to Failure (99% CI): 1227.6298 < 1228.5577 < 1229.4864 hours
Time to Failure Stats (+/- StDev): 824.8893 < 1228.5577 < 1829.7658 hours
Time to Failure Stats (+/- StDev): 2.97e+006 < 4.42e+006 < 6.59e+006 seconds
Ensemble Avg. Time (with stable drops): 22.5317 hours, or 8.11e+004 seconds

Continuum Rate Failure Stats: 0.98726 1/hrs, or 2.74e-004 1/sec
Arithmetic Mean Rate of Failure: 0.00095349 1/hrs
Mean Rate of Failure (99% CI): 0.00081335 < 0.00081396 < 0.00081458 1/hrs
Rate of Failure Stats (+/- StDev): 0.00054652 < 0.00081396 < 0.0012123 1/hrs
Rate of Failure Stats (+/- StDev): 1.52e-007 < 2.26e-007 < 3.37e-007 1/sec
```
Ensemble Avg. Rate (with stable drops): 1.4928e-005 1/hrs, or 4.15e-009 1/sec

The analysis of this is basically the same as 30 nm, since the number of droplets starting with an ROOH in it is small. The statistics are much better for this case, since we have 7×10^5 droplets that reached the critical concentration.

In this figure, the axes are slightly narrowed to show the histogram better. In this case, there are still values of the rate near 1×10^{-3} s^{-1}, there are just very few of them. Again, you can see that the linear average deviates from the true mean.

Time to Failure Histogram for D = 50 nm (left), Rate of Failure Histogram for D = 50 nm (right)
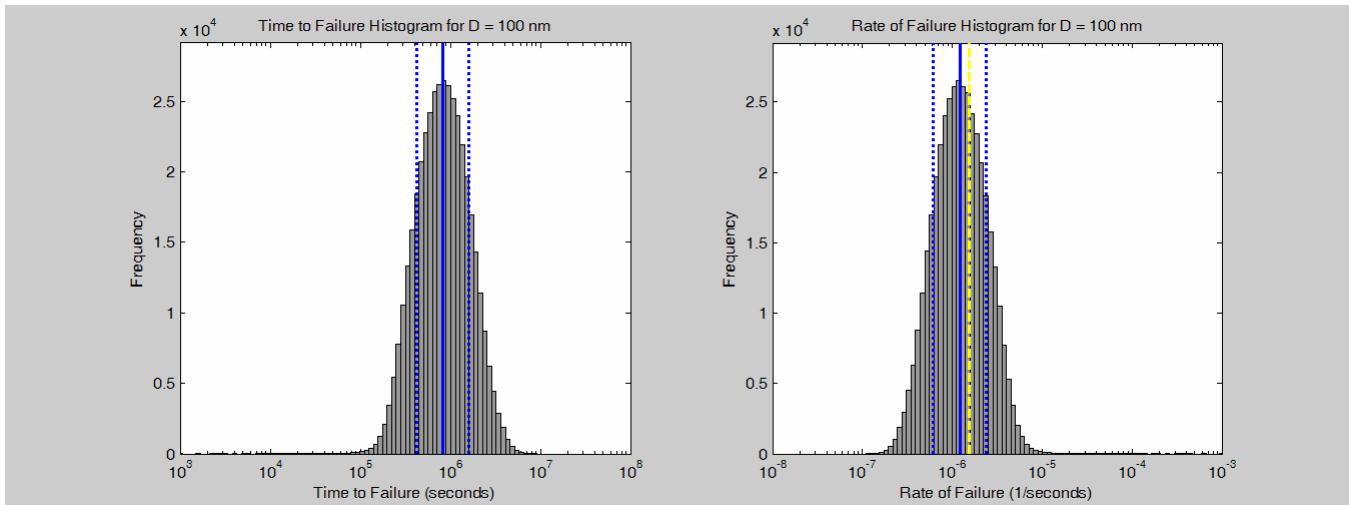
For Diameter = 100 nm

```
Number of steps that reached critical C_ROOH: 411899 of 1400000 steps
Went too long (but did not become stable): 0 of 441343 steps
Number of droplets that became stable (but reacted initially): 29444 of 441343
Fraction of droplets that started stable: 958657 of 1400000 (68.4755%)

Continuum Time to Failure Stats: 1.0129 hours, or 3.65e+003 seconds
Mean Time to Failure (99% CI): 227.9483 < 228.342 < 228.7364 hours
Time to Failure Stats (+/- StDev): 116.457 < 228.342 < 447.7195 hours
Time to Failure Stats (+/- StDev): 4.19e+005 < 8.22e+005 < 1.61e+006 seconds
Ensemble Avg. Time (with stable drops): 67.1813 hours, or 2.42e+005 seconds

Continuum Rate Failure Stats: 0.98726 1/hrs, or 2.74e-004 1/sec
Arithmetic Mean Rate of Failure: 0.0055976 1/hrs
Mean Rate of Failure (99% CI): 0.0043718 < 0.0043794 < 0.004387 1/hrs
Rate of Failure Stats (+/- StDev): 0.0022335 < 0.0043794 < 0.0085869 1/hrs
Rate of Failure Stats (+/- StDev): 6.20e-007 < 1.22e-006 < 2.39e-006 1/sec
```
Ensemble Avg. Rate (with stable drops): 0.0012885 1/hrs, or 3.58e-007 1/sec

For this case, you can see that the ensemble average is beginning to coincide with the averages for the reacting droplets because there are many droplets that started with some ROOH. This case has the most "normal" distribution (more correctly log-normal). This is likely due to the physics that arise from the problem. There seems to be two "hard" walls at times of ~1×10^7 (upper) and 3×10^3 (lower) seconds. This droplet size is such that both tails of the distribution are relatively far from these limits and can assume a nearly log-normal distribution. In this case, the mean and standard deviation do a very good job at describing the distribution of times and rate, where as the linear average would do a poor job.

The lower hard wall occurs when the initiation reaction happen very quickly, and this lower limit is the time it takes for the propagation reaction to generate enough ROOH molecules. This lower limit appears to be approximately the continuum time to failure. The upper limit is not a strict limit, but it is very rare for an initiation reaction to take longer that ~1×10^7 seconds. Once the initiation happens, it takes a relatively short time (relative to 10^7 seconds) to reach the critical concentration.
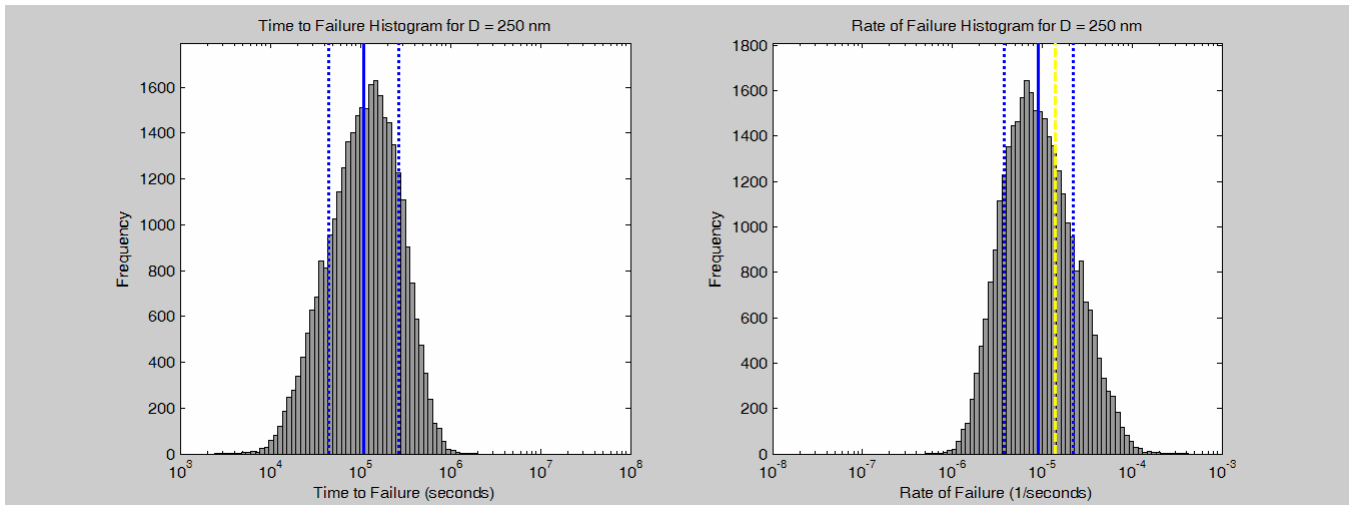
For Diameter = 250 nm

```
Number of steps that reached critical C_ROOH: 32000 of 32000 steps
Went too long (but did not become stable): 0 of 32000 steps
Number of droplets that became stable (but reacted initially): 0 of 32000
Fraction of droplets that started stable: 0 of 32000 (0%)

Continuum Time to Failure Stats: 1.0081 hours, or 3.63e+003 seconds
Mean Time to Failure (99% CI): 30.1115 < 30.3602 < 30.6109 hours
Time to Failure Stats (+/- StDev): 12.4124 < 30.3602 < 74.2596 hours
Time to Failure Stats (+/- StDev): 4.47e+004 < 1.09e+005 < 2.67e+005 seconds
Ensemble Avg. Time (with stable drops): 30.3602 hours, or 1.09e+005 seconds

Continuum Rate Failure Stats: 0.99196 1/hrs, or 2.76e-004 1/sec
Arithmetic Mean Rate of Failure: 0.050237 1/hrs
```
Mean Rate of Failure (99% CI): 0.032668 < 0.032938 < 0.03321 1/hrs
```
Rate of Failure Stats (+/- StDev): 0.013466 < 0.032938 < 0.080565 1/hrs
Rate of Failure Stats (+/- StDev): 3.74e-006 < 9.15e-006 < 2.24e-005 1/sec
```
Ensemble Avg. Rate (with stable drops): 0.032938 1/hrs, or 9.15e-006 1/sec

For this diameter, we have finally moved into the regime where all droplets react and reach the critical concentration. You will also see that the ensemble average has converged to the mean value, and the statistics are valid for the standard deviation and CI's in the mean. However, we are still quite far from the continuum value of ~1 hr for the time to failure, with the MC value estimated to be ~30 times longer, +/- about 1%.

The histogram story is much the same, with the linear mean doing a poor job at predicting the peak in the distribution.
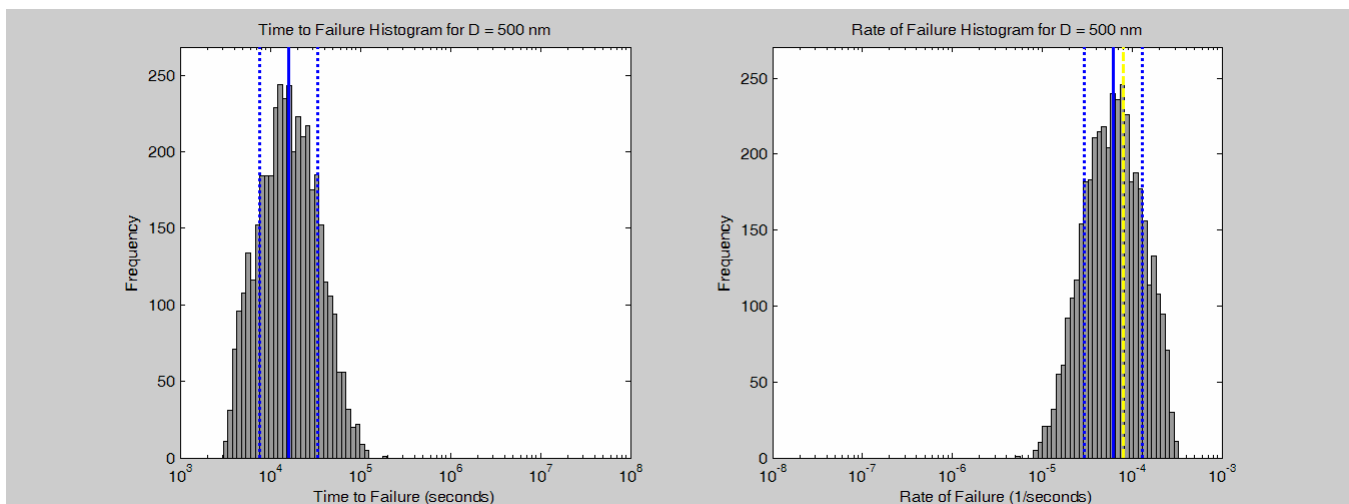
For Diameter = 500 nm

```
Number of steps that reached critical C_ROOH: 4100 of 4100 steps
Went too long (but did not become stable): 0 of 4100 steps
Number of droplets that became stable (but reacted initially): 0 of 4100
Fraction of droplets that started stable: 0 of 4100 (0%)

Continuum Time to Failure Stats: 1.0057 hours, or 3.62e+003 seconds
Mean Time to Failure (99% CI): 4.4083 < 4.4931 < 4.5794 hours
Time to Failure Stats (+/- StDev): 2.1414 < 4.4931 < 9.4273 hours
Time to Failure Stats (+/- StDev): 7.71e+003 < 1.62e+004 < 3.39e+004 seconds
Ensemble Avg. Time (with stable drops): 4.4931 hours, or 1.62e+004 seconds

Continuum Rate Failure Stats: 0.99433 1/hrs, or 2.76e-004 1/sec
Arithmetic Mean Rate of Failure: 0.28905 1/hrs
```
Mean Rate of Failure (99% CI): 0.21837 < 0.22257 < 0.22684 1/hrs
```
Rate of Failure Stats (+/- StDev): 0.10607 < 0.22257 < 0.46698 1/hrs
Rate of Failure Stats (+/- StDev): 2.95e-005 < 6.18e-005 < 1.30e-004 1/sec
```
Ensemble Avg. Rate (with stable drops): 0.22257 1/hrs, or 6.18e-005 1/sec

This size droplet has gotten us even closer to the continuum, with an error of ~4.5 times the continuum.
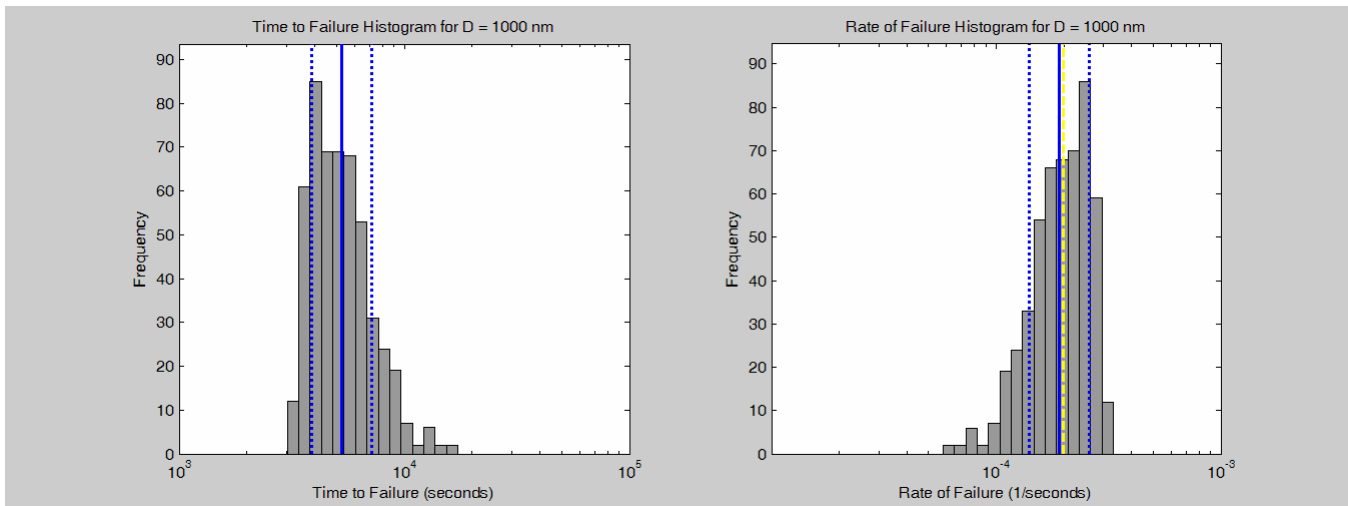
For Diameter = 1000 nm

```
Number of steps that reached critical C_ROOH: 510 of 510 steps
Went too long (but did not become stable): 0 of 510 steps
Number of droplets that became stable (but reacted initially): 0 of 510
Fraction of droplets that started stable: 0 of 510 (0%)

Continuum Time to Failure Stats: 1.0081 hours, or 3.63e+003 seconds
Mean Time to Failure (99% CI): 1.4285 < 1.461 < 1.4942 hours
Time to Failure Stats (+/- StDev): 1.0735 < 1.461 < 1.9882 hours
Time to Failure Stats (+/- StDev): 3.86e+003 < 5.26e+003 < 7.16e+003 seconds
Ensemble Avg. Time (with stable drops): 1.461 hours, or 5.26e+003 seconds

Continuum Rate Failure Stats: 0.99196 1/hrs, or 2.76e-004 1/sec
Arithmetic Mean Rate of Failure: 0.71491 1/hrs
```
Mean Rate of Failure (99% CI): 0.66926 < 0.68448 < 0.70004 1/hrs
```
Rate of Failure Stats (+/- StDev): 0.50296 < 0.68448 < 0.93151 1/hrs
Rate of Failure Stats (+/- StDev): 1.40e-004 < 1.90e-004 < 2.59e-004 1/sec
```
Ensemble Avg. Rate (with stable drops): 0.68448 1/hrs, or 1.90e-004 1/sec

We have now nearly converged to the continuum value, being only about 50% too large in terms of time to critical concentration. The distribution has also become narrower and more skewed, since it is now abutting the "hard" lower limit of ~3000 seconds for time. The linear average has also started to converge to the true mean since the distribution has become narrower.

For Diameter = 2500 nm

```
Number of steps that reached critical C_ROOH: 40 of 40 steps
Went too long (but did not become stable): 0 of 40 steps
Number of droplets that became stable (but reacted initially): 0 of 40
Fraction of droplets that started stable: 0 of 40 (0%)

Continuum Time to Failure Stats: 1.0081 hours, or 3.63e+003 seconds
Mean Time to Failure (99% CI): 1.0203 < 1.0308 < 1.0414 hours
Time to Failure Stats (+/- StDev): 0.99208 < 1.0308 < 1.071 hours
```

```
Time to Failure Stats (+/- StDev): 3.57e+003 < 3.71e+003 < 3.86e+003 seconds
Ensemble Avg. Time (with stable drops): 1.0308 hours, or 3.71e+003 seconds

Continuum Rate Failure Stats: 0.99196 1/hrs, or 2.76e-004 1/sec
Arithmetic Mean Rate of Failure: 0.9708 1/hrs
Mean Rate of Failure (99% CI): 0.96027 < 0.97012 < 0.98007 1/hrs
Rate of Failure Stats (+/- StDev): 0.93367 < 0.97012 < 1.008 1/hrs
Rate of Failure Stats (+/- StDev): 2.59e-004 < 2.69e-004 < 2.80e-004 1/sec
Ensemble Avg. Rate (with stable drops): 0.97012 1/hrs, or 2.69e-004 1/sec
```
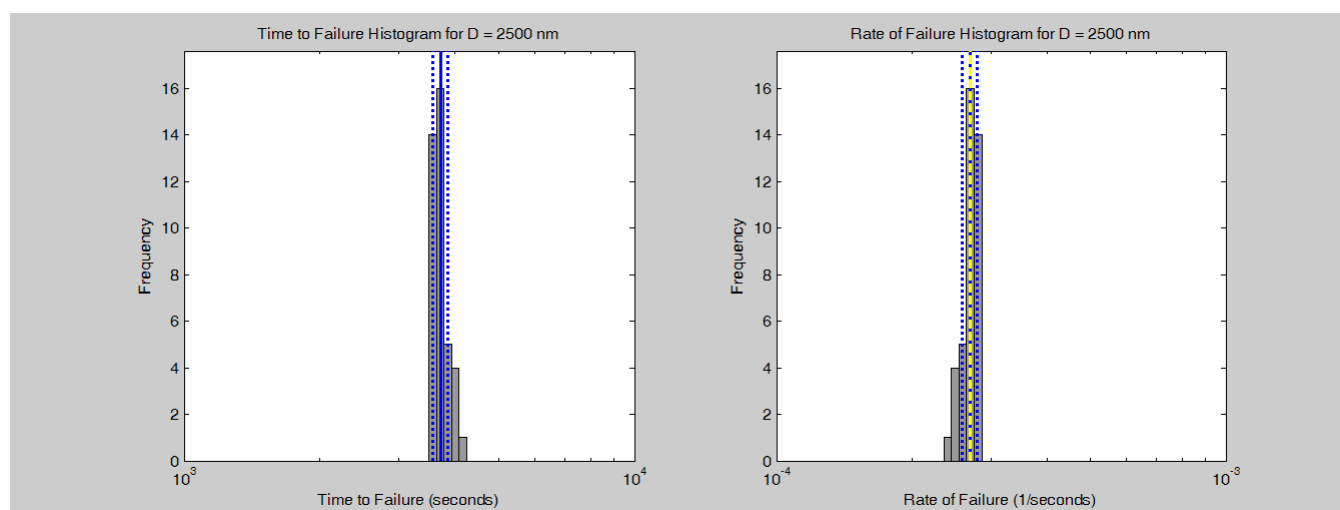
For this large diameter (and many ROOH's per droplet), we have achieved about 97% of the continuum value, with the confidence interval for the mean lying well within the 10% error that we desire. This set of data points (although small, 40), has a very narrow distribution, and the arithmetic average is equivalent to the log-scale average.

Summary of Part 2:

When the continuum and Monte Carlo solution disagree, which is more accurate?
In general, the MC simulations will give you a much better idea of what is going on in the system and yields a more accurate characteristic time to failure than a continuum method, when the number of molecules per droplet is small. They should converge (as was seen) when the number of ROOH's get large. This is especially true when you have fractional numbers of molecules per droplet, and the continuum method will have reaction occurring, even when there should be zero molecules in some droplets base on probabilities.

Comment on the width of the distributions, and on how many trajectories need to be run to achieve good statistics on the mean and width of the distribution.
There is not one correct answer for this question…
The widths (on a log scale… it doesn't make sense to take about them on a linear scale) depended mainly on how close the mean was to the "hard" walls at $3x10^3$ and $1x10^7$ seconds. If the mean was close to one of these walls (especially the lower wall), the distribution would tend to be relatively narrow and be skewed away from the "wall." When the mean was near the upper "wall," the tail tended

to be longer, making the distribution wider. When the mean value was located away from both walls, the width of the distribution tended to be about 2 orders-of-magnitude.

The number of trajectories needed to achieve good statistics on the mean can be fairly well determine by using the confidence interval discussed earlier. There are certain assumptions about the distribution being Gaussian (which are not strictly true for this system), but for a first-order estimate, it is sufficient. You can see that for all cases (expect for D = 30 nm), the confidence interval in the mean value was relatively small (less than 5%). This indicates that the number of steps run was sufficient to achieve good statistics on the mean. The width is somewhat more difficult to characterize. One simple way is to examine the distribution graphically, and if it looks relatively smooth, then you can be confident that the width is pretty well converged. A more rigorous way would be try to find the confidence interval for the standard deviation using higher-order moments of the distribution.

How large does the diameter need to be in order to reach within 10% of the continuum solution?
The figure below shows the convergence of the MC simulation with the continuum value for varying diameters. Although data for diameters of 1500 and 2000 nm would be useful, we can fairly confidently say that the MC solution will reach within 10% of the continuum somewhere between 1250 and 2000 nm. Given the shape of the curve, it is likely that the value is closer to 1250 nm.
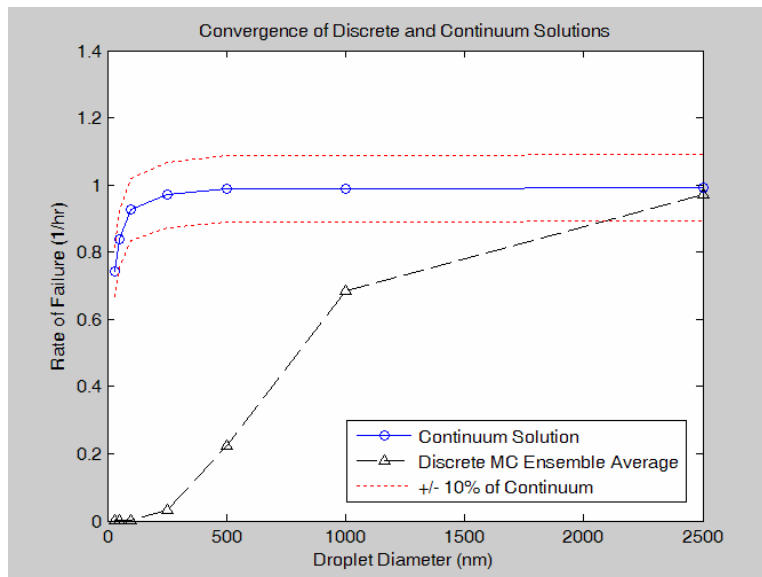
Part 3: 250 nm droplet – dependence on *k*

The figure below shows the dependence of the mean rate of failure (dots) on *k*, for *k* = 1→1000. The 99% confidence intervals on the mean values are shown as well with triangles. It shows a pretty interesting trend, with an increase in the rate from small to intermediate values of *k*, and then a sharp decrease as *k* increases further.

The sharp decrease at large values is easy to explain. The k-value can essentially be thought of as the diffusivity of the ROO molecule, and as it increases, the molecule has a large chance of reaching the edge of the droplet and reacting via reaction 4 (ROO → ROOH). When ROO get to the edge and converts to ROOH, then the very slow initiation reaction must happen again to regenerate the catalyst. For a droplet that only has a few molecules in it, this is important and can significantly affect the rate. This serves to slow the propagation reaction (ROO → ROO + ROOH) by taking some ROO molecules out of circulation. This, on average, slows the rate of failure.

The initial increase is likely due to a more subtle effect cause by the edge. This may be due to the competition between the recombination pathway (2 ROO → inert) and the edge pathway. Essentially, increasing the likelihood of having the edge reaction occur will decrease the likelihood of the recombination happening. Since the recombination is relatively worse (since it eliminates 2 ROO's instead of 1), the net result is a benefit. The easiest way to understand this is with an example:

Assume you have a droplet with 4 ROO's, and *k* is such that: $r_{prop} = 0.989$, $r_{recomb} = 0.01$, $r_{edge} = 0.001$. In this case, you have a 1% chance of killing 2 ROO's, and a 0.1% chance of converting 1 ROO into an ROOH. If one of these undesirable side reactions happens, it will likely be the recombination. The result will be 2 ROO's left in the droplet. Suppose *k* increases by a factor of 10, so that now $r_{edge} = 0.01$. Now you have an equal chance of killing 2 ROO's, or convert 1 into an ROOH. The latter is a much better result since you only lose 1 ROO and add to the pool of ROOH as well (it will also create an odd number of ROO's in the droplet, which may be beneficial, as discussed below). Also, with this increase in *k*, you have only had a minor effect the propagation rate. If *k* increases by another factor of 10, then nearly 10% of the time the edge reaction will happen, which will results in a significant decrease in the overall production of ROOH (i.e. the beginning of the sharp decrease at large k-values).

Another reason for the initial increase could have to do with the differences created when you have an odd or even number of ROO's in the droplet. Since ROO's are always created in pairs, and pairs must be present to convert 2 ROO's into an inert, it would be beneficial to have an odd number of ROO's to ensure all ROO's cannot be converted into inerts. When diffusion is slow, it is difficult to reach the odd ROO state, since you would more often have two ROO recombine than have one reach the edge. When two are present, you have the benefit of having 2x the propagation rate, but face the chance of recombining and having to undergo the slow initiation reaction again. If you have one diffuse to the edge, you have only one ROO left to run the propagation reaction, but it can work carefree without worrying about being annihilated. However, if the diffusion gets too fast, then the last ROO can diffuse to the edge and the chain reaction will be quenched (causing the sharp decline at large values of *k*).

Both of these scenarios will contribute to the behavior seen in the plot below.

Part 3 - Dependence of Rate of Failure on k with 99% CIs