High-temporal resolution probing of sol-gel transition in diffusion limited cluster aggregation process

With a tutorial on DLCA Simulations on HPC and Post processing

Report Written by GOWTHAM KUNTUMALLA in July, 2017

Contact: gowthamkuntumalla@gmail.com for any doubts and queries

THIS ARTICLE HAS NOT BEEN COMPLETELY FORMATTED FOR GOOD VISUALS. IT IS ONLY USED AS KNOWLEDGE TRANSFER MATERIAL.

Research Objectives:

The main aims behind taking up this project are as follows.

1. Performing computer simulations on Diffusion Limited Cluster- cluster Aggregation (DLCA) processes. (Stochastically modelled). Original Codes were written by Dr. Willam Heinson, (contact willbot1983@gmail.com for any doubts and errors in code)
2. Formulating kinetics of the sol to gel transition in these models along with effect of change in parameters of the system.
3. Can you predict the dependence of the kinetics and transition points for different volume fractions and relate it to a system independent parameter?

It will be helpful in chemistry, atmospheric sciences, microbe scale biology(actually any scale as far as nearly scale invariant fractal nature of particle morphology is observed.)

Introduction

In nature aggregation or agglomeration is a common phenomenon in dispersed systems. Aggregation is a simple kinetic process in which particles moving in space collide and stick to each other with a probability. If these particles do not coalesce but retain their original shape, a scale invariant fractal structure defined by a scaling dimension, D_f which is less than the spatial or Euclid dimension, D. The popular scaling law for aggregates is given below. Where N is the number of monomers per aggregate, R_g is the radius of gyration of the aggregate, and 'a' is radius of monomer (assumed spherical).

$$N = ko * \left(\frac{Rg}{a}\right)^{Df} \quad (1)$$

$$\mathrm{fv,m} = \mathrm{Nm} * \left(\frac{\frac{4\pi}{3}\mathrm{a}^3}{\mathrm{V}}\right) \quad (2)$$

In addition to above factors, Monomer volume fraction ($f_{v,m}$) is another parameter which describes the state of the system(initially). Here N_m is initial number of monomers present in the system. Note $N_{c,0} = N_m$ at time, t=0 where $N_{c,t}$ is number of clusters at time t. V is the total volume of system. There are two major motions [cite here] *viz.* Ballistic or Diffusive motion. The growth kinetics in free molecular regime is such that it starts with ballistic motion in the very dilute state and transitions into diffusive motion as the aggregate size increases.

The three common types of aggregation phenomena are (i) Diffusion limited cluster-cluster aggregation (DLCA) (ii) Reaction limited cluster-cluster aggregation (RLCA) (iii) Ballistic limited cluster-cluster aggregation (BLCA).

Most common of them, the DLCA produces fractal aggregates with a d_f=1.8. This part has been well studied in the literature{Heinson, 2017 #1;Sorensen, 2011 #2;Sorensen, 2011 #2}. Initially the process kinetics is governed by Smoluchowski equation (in the cluster dilute regime). After this point these aggregates act as super monomers to form super aggregates which have a scaling dimension of Df=2.5. In this study DLCA simulations are carried out and the sol-gel transition is probed by taking snapshots of the aggregating system during the simulation process at high temporal resolution. The high resolution is required because of the rapidly changing kinetics of the system; the sol to gel transition appears jittery and is not insightful otherwise. This warrants a deeper and elaborate study into its kinetics.

Simulation study and methods

Here in we perform off lattice montecarlo DLCA simulations. They work by finding probability of particle's motion based on its size. Larger clusters move less compared to small clusters. The simulations proceed in the following fashion and follow a canonical algorithm. Large numbers of monomers (N_m) or solid nanoparticles (N_m in the order of millions) are randomly distributed in a closed simulation box.

These simulations start with a fixed number of monomers randomly distributed in spaces. Each monomer is moved randomly a distance of one monomer diameter (=2a) with a given probability ($1/N_c$) which follows Epstein diffusion dynamics. This phase of aggregation happens in a cluster dilute regime and clusters of fractal dimension D_f= 1.8 are formed under the diffusion limited cluster-cluster aggregation process. As the simulation proceeds these clusters act as super monomers with D_f = 1.8 (CITE HERE) and start agglomerating to form bigger and denser cluster with Df=2.5. This transition has been of intense study in some recent literature (CITE SOME ARTICLES HERE). Due to time restrictions and computational constraints many simulations usually stop at around 10,000 clusters or less. In this study we allow it to run till all the monomers initially present in the system aggregate (if possible) and form the dense space filling 'Gel'.

It is prudent to mention that all systems gel in finite time. But theoretically speaking any closed system containing randomly distributed moving monomers will gel at infinite time. To ensure statistical validity, a number of simulation runs are conducted in high performing computing (HPC) cluster and large amount of raw data is collected.

A normal simulation consists of following steps.

1. Define the DLCA problem. Choose the following parameters: - number of monomers, volume fraction, (box size is set by the other two). {MAKE EDITS IN APPROPRIATE EXCEL FILE} [Look at paper_runs.xlsx and new_runs.xlsx]

   NOTE: It was observed in simulations that changing the number of monomers introduces a non linear effect in the kinetics of aggregation. It is advised to use a fixed number of monomers while comparing data for different monomer volume fractions ($f_{v,m}$) to avoid discrepancies.
2. Make the appropriate folders in remote server {/research-projects/airlab/A_gowtham_simulations} on WINSCP and copy requisite files. [Look at example simulation folder "paper_run30", "paper_run55"]. Make it in any folder you wish but edit the system path in "automaton1" and "data_acq" bash files as needed.
3. FILES in a typical SIMULATION Folder:

- DLCA_kg_edit.cpp, makefile, data_for_plots(.pbs file depends on where you want to submit),
- a_import_data.m, for_proc_data folder {you will have to edit the definitions here too}

4. Edit the definitions in above files for current simulation.
5. Recheck all the edits properly.
6. Run the simulation on master node of cluster or slave nodes (.pbs scheduler required)[code for scheduler is attached]. The bash script file "automaton1" is used for scheduling multiple runs at once.
7. Use commands "top" (for master cluster submissions) or "qstat" to check the status of the runs.
8. After successful run, run "bash data_for_plots" to get proc_data folder {DO THIS IN REMOTE SERVER}
9. Save the "proc_data" folder which is formed after running data_for_plots command in any secure location. This is the important data.
10. Check the presence of "a_import_data.m" in proc_data folder. If not present adds it. Check the definitions of simulation inside these matlab files.{DO THIS ON LOCAL SERVER}

SCIENCE BEHIND DLCA AND PLANS FOR PLOTS

Two main themes have been identified based on literature survey.

i. ***Spectrum plots***. [Refer to Heinson et al. 2017 ,"Kinetic percolation" paper]
   a. Mass Spectrum plots: The wavy nature of the individual snapshot data drawn on these plots is interesting. There is mathematical constraint (a straight line) onto which all the waves 'collapse' onto. We believe this point where the waves first start forming this imaginary geometric constraint is the place where the Ideal gel point is. It can also be assumed to be the onset of gelation and at the beginning of transition regime. It is the kink in the mass spectrum graph.
   b. Number (frequency spectrum) plots: These figures enhance the visualization of the onset point discussed above. The curve becomes flat after the point in discussion. i.e. very few cluster of large size are present in the system.

Time data (corresponding to the snapshot number) is obtained by using the data_time$snapshotnumber. Note that $snapshotnumber is placeholder for snapshot number e.g. data_time264. Form all the snapshot data, by manually plotting spectra of individual snapshots and properly ***BINNING*** them the onset point is found out. [Use the mass_hist excel files for this purpose]. A simple copy paste of data in these excel files in appropriate locations will do the job.

ii. <u>1/Nc-1/Nc(0) vs. time plots</u>

a. Actual log-log plots of above. Equation proposed is as follows

$$Nm * \left(\frac{1}{Nc(t)} - \frac{1}{Nc(0)}\right) = k * \left(\frac{t}{to}\right)^{Z} \quad [3]$$

Where Nm = Number of monomers initially present in the system. Nc (t) is the number of clusters present in the system at any time t. K is a system dependant constant. The plots would be visualized in double log scale.

b. Slope (Z) vs. time: The slope of the log-log plot drawn above is Z which varies with time as Nc varies with time. Note that a polynomial fit was used to find this slope by the author of this tutorial. Readers are suggested to find alternate methods such as Numerical differentiation, piece wise spline interpolation if the result looks erroneous.

<u>This part of project is a novel thing and has not been rigorously done before to the best of the author's knowledge</u>

USING MATLAB CODES AND EXCEL

A folder e.g. "fv=0.01" (Without quotes) is created for each volume fraction. This folder is present inside a folder called "data_sets". Note this has been done to ensure valid use of matlab code. If not the file paths are to be changed as per user's need. The following codes go into it. Comments present in the MATLAB file are self explanatory.

1. For Spectrum plots follow the mass_hist excel files and copy paste size (i.e. average mass) of each cluster in a given snapshot in one column of excel. Remaining columns can be used for other snapshots. **Note that each individual curve for a snapshot is a union of similar data from all the runs corresponding to one monomer volume fraction ($F_{v,m}$).** MATLAB codes are used to get this unified data as described below.
2. Following codes are used for processing data in matlab.
    a) a_get_the_full_data.m This function code is used to get the data from individual runs into a common folder (e.g. "fv=0.01", without quotes).
    b) a_Plot_mass_spectrum.m This function code is used for producing the excel file required for **spectrum** plots.

c) analyser.m This is the *main* file which uses above matlab function codes. Edit the parameter below.

NOTE: The above files need some INTELLIGENT COMMENTING and UNCOMMENTING of requisite lines of code. TAKE PROPER CARE to ensure that variable are being properly edited in the MATLAB workspace.

3. MATLAB's "cftool" can be used to get the best polynomial curve fit (Try to get a odd degree polynomial starting from cubic). Excel can be used to get the trendline for log(1/Nc-1/Nc(0)) vs log(time) plot. In Excel, Get a cubic polynomial fit for plotting slope of the above graph with time.
4. All the remaining plots are straightforward.

NOTES and CAUTION:

- The raw data sets produced are huge. Ensure that enough space is present for all the simulations. (Approximately 42GB per simulation for 1 million monomers and 418 datasets)
- Note there may be few changes in this README tutorial from real code.
- Make sure the number of simulations when run together does not cross 720 GB on AIRLAB folder space in the remote server.

Results and discussion

There are two parts in the results. Initial gelation time (to) is found out by looking manually at the kink in the mass spectrum and frequency spectrum plots. The effect of different monomer volume fractions on the kinetics of the aggregation system is studied. Slope (i.e. Z exponent) of normalized average mass vs. Time is studied.

NOTE: It will be interesting t_o see if this proposed to and t_g as proposed in literature are same (Soresnen and chakrabarti 2010).

Onset and Transition Regime of Gelation

In the literature the ideal gelation time (t_g) is assumed to be the time when nearest neighbour separation equals the average radius of gyration (FIND CITATION FROM PLOT 1 OF SORENSON REVIEW) Here it is presumed to be the kink between the dome and straight line parts of the figure. [Talk about the movement of point wrt. snapshots and talk about different volume fractions]. Sample figures are shown below.

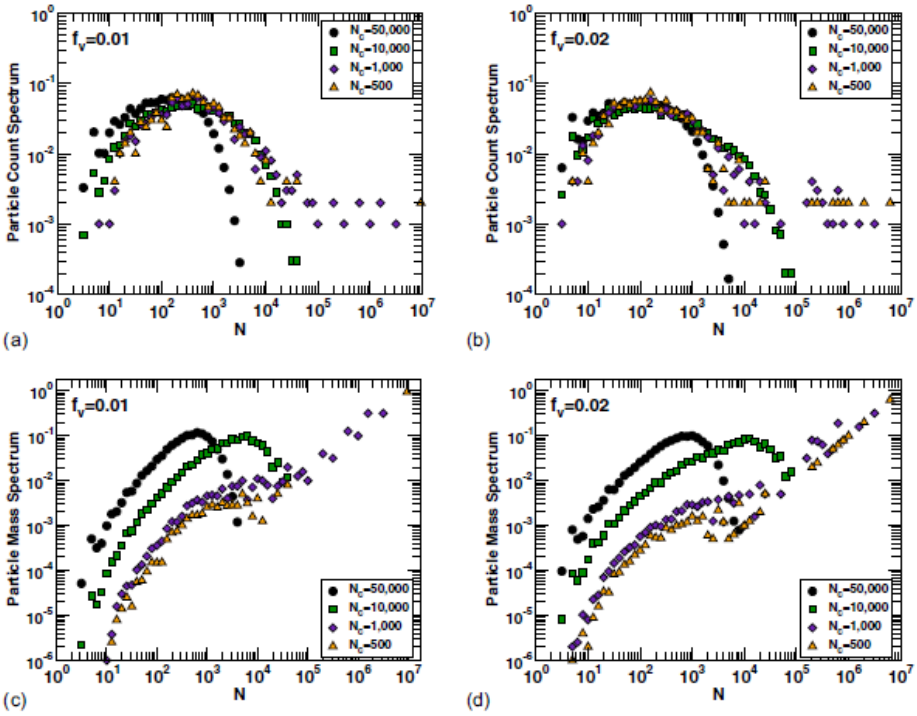


FIG. 4. Particle count spectra (upper plots) and particle mass spectra (lower plots) for two monomer volume fractions $f_v = 0.01$ and 0.02. N_c is the total number of clusters remaining in the aggregation run.

These plots are obtained in excel. It is actually a combined plot of a single volume fraction simulation conducted multiple times i.e. data from all runs at each snapshot is appended because it gives us more data from shorter simulations and error isn't too large because of high temporal resolution. Large simulations,~10 million or more monomers at once were tough to run on the LAB cluster hence they were ditched and smaller systems such as 1 million 3 million etc were used. These runs are performed multiple times everything same but with different seeds for ensuring randomness.

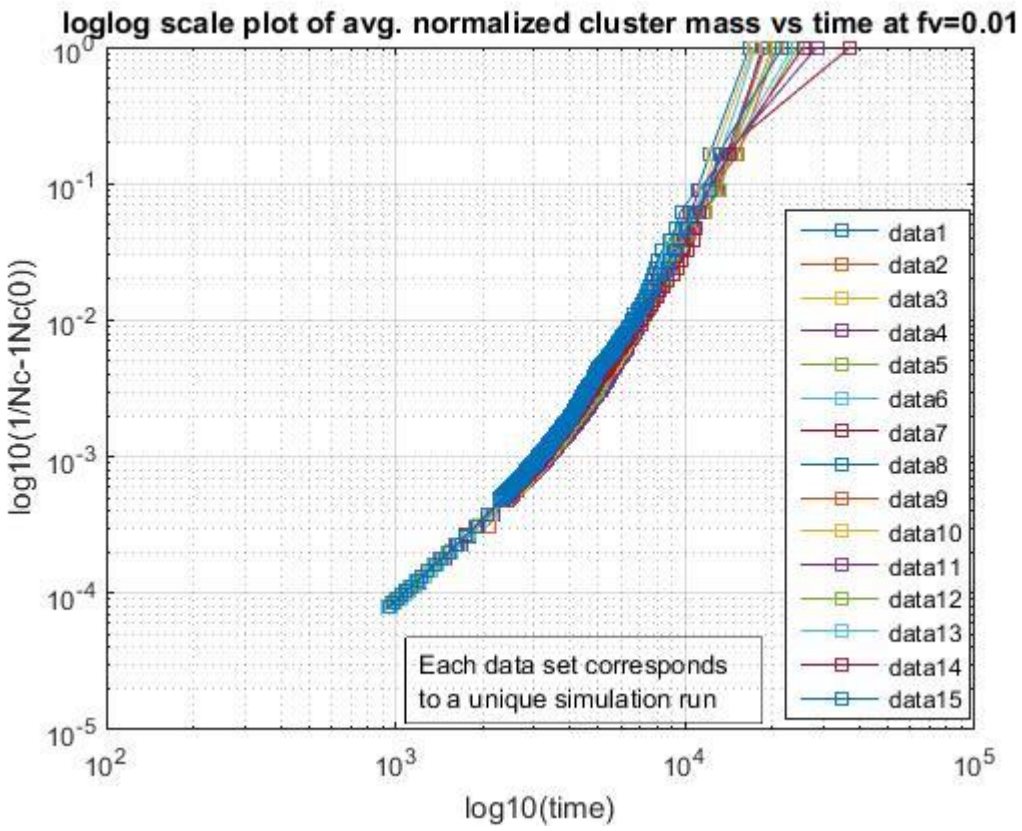
Variation of average cluster size with time

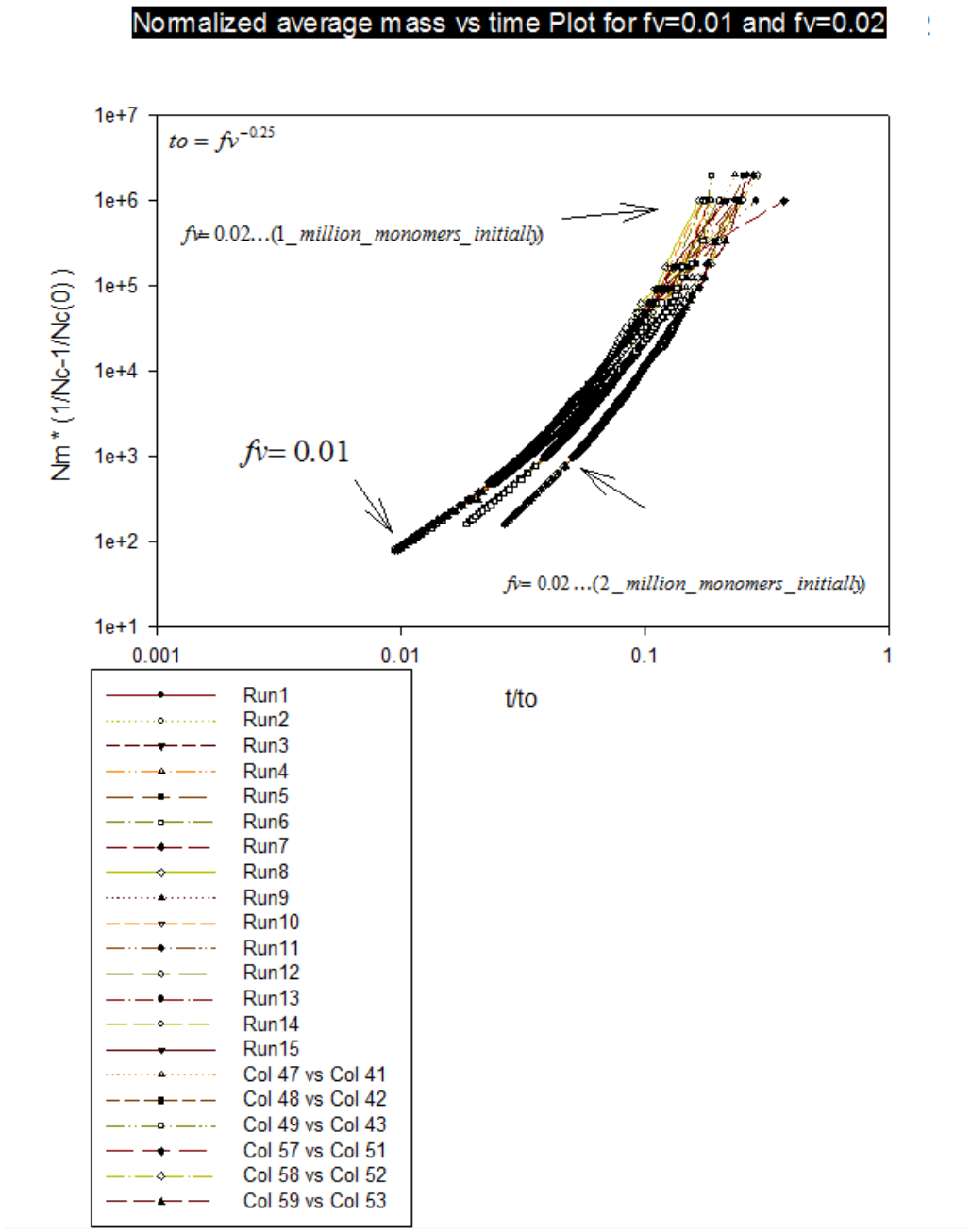
Normalized average cluster size or mass ($1/N_c$) in the system is plotted against time to look at the aggregation kinetics.

This data is normalized to show a general behaviour.

$$<N>=Nm/Nc \;\&\; 1/Nc = <N>/Nm \qquad (4)$$

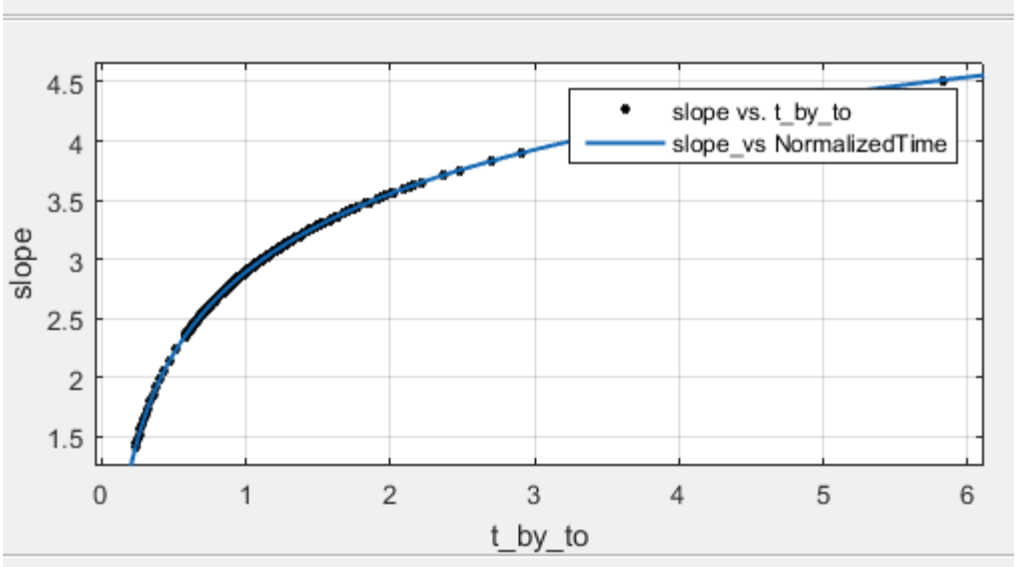
As expected in this irreversible system(Sorensen and Chakrabarti 2011 and Fry Sorensen paper 2002) the number of clusters in the system decrease with passage of time and become more bulky. This power law kinetics with exponent Z has been of great interest because of the apparent change in rate of aggregation with time.


loglog scale plot of avg. normalized cluster mass vs time at fv=0.01
log10(1/Nc-1Nc(0))
log10(time)
Each data set corresponds
to a unique simulation run
data1
data2
data3
data4
data5
data6
data7
data8
data9
data10
data11
data12
data13
data14
data15



In the above plot 3 kinds of data are present. (i) fv= 0.01, 1million monomers. 15 runs (ii) fv= 0.01, 1million monomers. 3 runs (col 47, 48, 49) (iii) fv= 0.02, 2 million monomers. 3 runs (col 57, 58, 59)

Results show that the kinetics becomes faster as the number of clusters at time t decreases because of agglomeration (Note $N_{c,0} = N_m$ at time, t=0). Due to resource constraints, only a small number of simulations could be carried out. Because of this the exact point of transition (i.e. the kink mentioned earlier) becomes difficult to identify from the plots. {Talk about the curve fits}.



The above plot is one of Z exponent vs time on normal linear scales. Not that this variation before this curve would be a near constant straight line of Z exponent of 1~1.2. Initiation of gelation is assumed to occur at t=to where Z~2.5 which is a little more than the value in literature (Z=2).

Conclusions and Future work

It is observed that kinetics of gelation is a fast phenomenon at higher stages of agglomeration. The kinetic Z exponent varies with time in a non uniform fashion. It is not shown here but previous studies {CITE HERE} have shown that z exponent is close to 1 or 1.2 before the onset of gelation. In that regime (before gelation starts) mean field kinetics hold good and Smoluchowski equation is valid.

Further studies may include calculation of variation of Z exponent, Structure factor S(q) [Cite Sorensen et.al review paper here] during the gelation transition period. The variation of kinetics kernel is also an interesting case to investigate.

Acknowledgments

G.K gratefully acknowledges Prof Rajan Chakrabarty, Dr. William R. Heinson, Pai Liu for mentoring and guidance and EECE Department of WUSTL, USA for scholarship.

References

[1] “Kinetic Percolation”, Heinson et al. 2017

[2] “SIMULATION STUDIES ON SHAPE AND GROWTH KINETICS FOR FRACTAL AGGREGATES IN AEROSOL AND COLLOIDAL SYSTEMS “PhD dissertation, W.R.Heinson 2011

[3] “The sol to gel transition in irreversible particulate systems”; Sorensen, Chakrabarti 2010

[4] “AGGREGATION IN COLLOIDS AND AEROSOLS” PhD Dissertation Flint G.Pierce, 2007

[5] “Statistical Thermodynamics of Irreversible Aggregation: The Sol-Gel Transition” Matsoukas 2015

[6] “Fractal Geometry: What is it, and what does it do?” Mandelbrot 1989

[7] “Scaling theory and exactly solved models in the kinetics of irreversible aggregation” Levyraz 2013

[8] Bernoulli 5(1), 1999, 3±48 DAVID J. ALDOUS

[9] Rajan Dhaubhadel , Corey S. Gerving , Amitabha Chakrabarti & Christopher M. Sorensen (2007) Aerosol Gelation: Synthesis of a Novel, Lightweight, High Specific Surface Area Material, Aerosol Science and Technology, 41:8, 804-810, DOI: 10.1080/02786820701466291

[10] Pai Liu, William R. Heinson & Rajan K. Chakrabarty (2017) Fractal scaling of soot packing density across five size decades, Aerosol Science and Technology, 51:7, 879-886, DOI: 10.1080/02786826.2017.1316359

[11] W. R. Heinson, F. Pierce, C. M. Sorensen & A. Chakrabarti (2014) Crossover from Ballistic to Epstein Diffusion in the Free-Molecular Regime, Aerosol Science and Technology, 48:7, 738-746, DOI: 10.1080/02786826.2014.922677

[12] “Off-lattice Monte Carlo simulations of irreversible and reversible aggregation processes” DOI: 10.1039/b510449a

[13] Soft Matter, 2013, 9, 4437

[14] [IMPORTANT PAPER] DOI: 10.1103/PhysRevLett.89.148301

[15] Journal of Colloid and Interface Science 375 (2012) 65–69

Glossary

A run is a single simulation like paper_run14, new_run22 etc

Snapshot = an instant of time during simulation

1. Number of monomers initially placed, N_m

2. Volume fraction (or initial monomer packing density in the simulation box), f_v

3. Number of clusters at a given time instant, N_c

4. Number of monomers in a given cluster, N

5. Average number of monomers (or mass) per cluster at a give time instant in the system, <N>

6. Radius of gyration of a given cluster, R_g

7. Average radius of gyration in a snapshot, $<R_g>$ - This gives an idea about how the system is evolving

8. t_g or to= time at onset of Gelation (i.e. the kink in the mass spectrum graph)

REMOTE SERVER = WHERE YOU WORK ON SHELL ONLINE (PUTTY SSH client is used to connect to it on Windows OS, WinSCP is used as file transfer wizard)

LOCAL SERVER = Personal Computer or Workstation