KMC Protein Clustering Notes

by Carlos Alas
September 23, 2021

I. Problem Description

Here we present a description of the problem, along with the given variables, and the transition rates W for our KMC simulations. We follow the notation of reference [1], with simplifications made to deal with only one nanodomain rather than several.

We will look at the phenomenon of monomer emerin proteins freely diffusing through an inhomogenous nuclear membrane surface with spacial varying diffusion rates. Our system will consist of $M = 1600$ total monomer emerin proteins. We wish to track the trajectories of the emerin proteins. The motion of the emerin proteins is governed by diffusion through the membrane lipids. Within the nanodomain, the diffusion coefficient is about an order of magnitude slower ($D_{in} = 250$ nm^2/s) than that outside of it ($D_{out} = 3500$ nm^2/s).

We will use a nanodomain in the shape of a square and of area $A_{in} = 12100$ nm^2. The domain of our calculations will be a square patch of area $A = 160000$ nm^2, within which the nanodomain of monomers is contained. Therefore the area outside the nanodomain is $A_{out} = A - A_{in}$. The domain patch will be split into $K = 1600$ uniform lattice square cells with side length $a = 10$ nm, since $a^2 = A/K = 100$ nm. At lattice cell i, for $0 \leq i \leq K-1$, there are N_i emerin proteins. On the four sides of our domain patch, we will enforce periodic boundary conditions, so

$$\sum_{i=0}^{K-1} N_i = M \tag{1}$$

at all times.

We will start the simulations with a uniform distribution of monomers on the domain patch, so $N_0 = N_1 = ... = N_i = ... = N_{K-2} = N_{K-1} = M/K = 1$ at $t = 0$. We will model the monomer protein trajectories as a series of hops as they diffuse through the media. In our KMC simulations, we sample the times a protein will diffuse from one lattice cell to another from a Poisson distribution fitted to the mean hopping time– τ_{in} or τ_{out}– which we divide by the number of proteins occupying the lattice cell, since a higher occupancy would imply a higher chance of having a protein diffuse out and so the event times in lattice cells of higher occupancy should have more occurances compared to other lattice cells of similar diffusion rates. We simulate, by use of a priority queue (binary min heap), where the lattice cells are arranged by their event times, one monomer emerin protein hop at a time. A monomer emerin protein can hop up, down, left, or right to a neighboring lattice cell at each hop. We use a K by 4 array nn to store the nearest neighbors of each lattice cell i in $0 \leq i \leq K-1$, where the 1st column stores the neighbor index above and the remaining neighbors are stored in clockwise order, so up, right, down, and then left, for $j = 0, 1, 2, 3$, respectively. If a protein is inside the nanodomain, the time it must wait to hop is dependent on the hopping rate determined from the diffusion rate and how many other protein are in the same lattice cell. The average hopping time to another lattice cell is $\tau_{in} = a^2/D_{in}$. Outside the nanodomain, it must wait on average, to hop, a time $\tau_{out} = a^2/D_{out}$.

Since we are looking at diffusion in an inhomogenous media, we will reference the paper [1] for the general form of the transition rates W. In this paper, the notation $\boldsymbol{N} = \{N_i\}$ and $\boldsymbol{m} = \{m_i\}$ represent the state of the system at a time t by its spacial distribution of monomer emerin proteins and the random incremental changes to the spacial distribution due to hopping at each site m_i, respectively. Since we simulate only one emerin protein hopping per iteration, m_i is -1 for a hop out of site i and m_j is $+1$ for the site j it hops into.

For the kinetic Monte Carlo simulations we calculate the transitions rates in each lattice cell;

$$W(\boldsymbol{N};\boldsymbol{m}) = \frac{1}{4}\sum_{i=1}^{K} \frac{N_i}{\tau_i}\delta(m_i + 1) \sum_{jnnofi}^{4} \delta(m_j - 1) \prod_{k\neq i,j}^{K-2} \delta(m_k), \tag{2}$$

where j nn of i are of the four neighboring lattice sites at lattice site i. The first summation over i only gives the term for the protein hopping out, and then the summation over j that remains gives only the term for where the protein hops into. The product series evaluates to 1 always since $m_k = 0$ for $k \neq i, j$. Note that in the subvolume reference paper [2] they use s in place of W as their notation. So $W_i = (1/4)N_i/\tau_i$. The transition rate is simply the number of proteins in the lattice cell divided by the average hopping time and the number of directions it can hop to. With the transition rates we can sample the event times in each cell as

$$t = -\tau_i \ln(u)/N_i = -\ln(u)/W_i \tag{3}$$

for i in $0 \leq i \leq K-1$, where u is a random number between 0 and 1. At each step we must update the lattice cells which have changed in occupancy number N_i or transitions τ_i, though in our case the transition rates are constant through time, in each cell, so we need only mind changes is occupancy; the priority queue is re-orded each iteration using min heapify up/down percolation functions.

For the simple case of free diffusion, the analytical solution is known to give the distribution of proteins inside the nanodomain, ρ, as

$$\rho = \frac{\tau_{in}}{\tau_{in} + \tau_{out}}, \tag{4}$$

at steady state.

If we plug in the values of the hopping times we use in our simulations, we find that the fraction of proteins in the nanodomain at the systems steady state is $\rho \approx 53\%$.

II. Subvolume KMC method

The subvolume KMC method allows only one protein to hop at a time. The transition rates are assigned respecting the diffustion rates inside the nanodomain and outside of it. We begin with a uniform distribution of monomer emerin proteins. We assign the hopping events in lattice cells event times sampled from a Poisson distribution fitted to $1/W_i$ and sort them with a min binary heap. We begin the simulation by having the protein in the lattice cell of the event time at the top of the min binary heap hop randomly to one of its neighbors, using a uniform distribution. We then update the transition rates of the lattice cell from which the protein hops out of and in the lattice cell it hops into. Then we sample new subsequent event times to both of these lattice cell and re-sort the min binary heap. We do this over and over again until the simulated time reaches a little over 4 hours.

III. Results

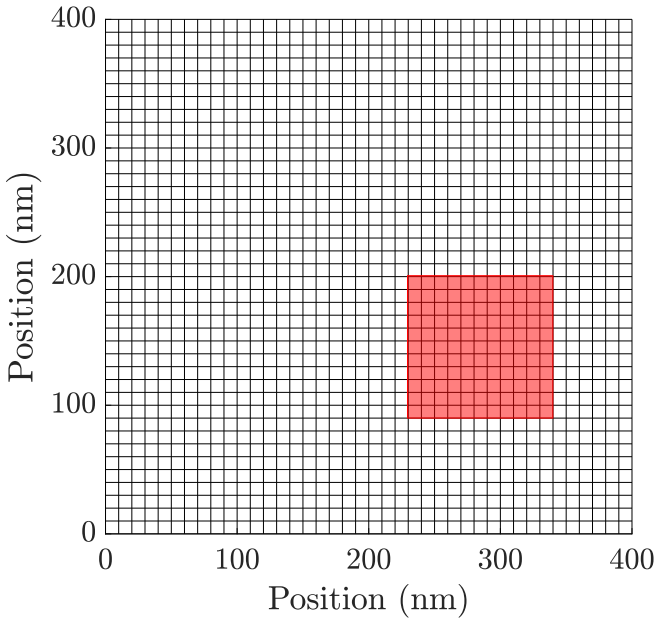

Figure 1. We ran the subvolume KMC method of reference paper [2] using the parameters discussed in the main text of these notes above. Here we show the lattice decomposition in our system domain. The red shaded area is the nanodomain area.

In our simulations, we observed that the system for about 4 hours. At the steady state, the averaged concentration of emerin proteins is about 53% and agrees to within 1% of the analytical solution. In figure 1, we show the domain patch decomposed into lattice cells with the nanodomain lattice cells highlighted in red. In figure 2, we show the monomer protein emerin concentration as a function of KMC simulation time in red and the analytical solution in blue; the percent error was averaged over the entire KMC simulation. It appears that different diffusion rates, in an inhomogeneous media, lead to the clustering of monomer emerin proteins. Since monomer emerin proteins cluster in response to mechanical stimuli from the surrounding environment, our simulations seem to suggest that the clustering is a result of the mechanical stimuli altering the diffusion rates on a domain patch. Due to the clustering of proteins,

the nanodomains are able to maintain a lower diffusion rate for periods of time longer than the mechanical stimuli events.

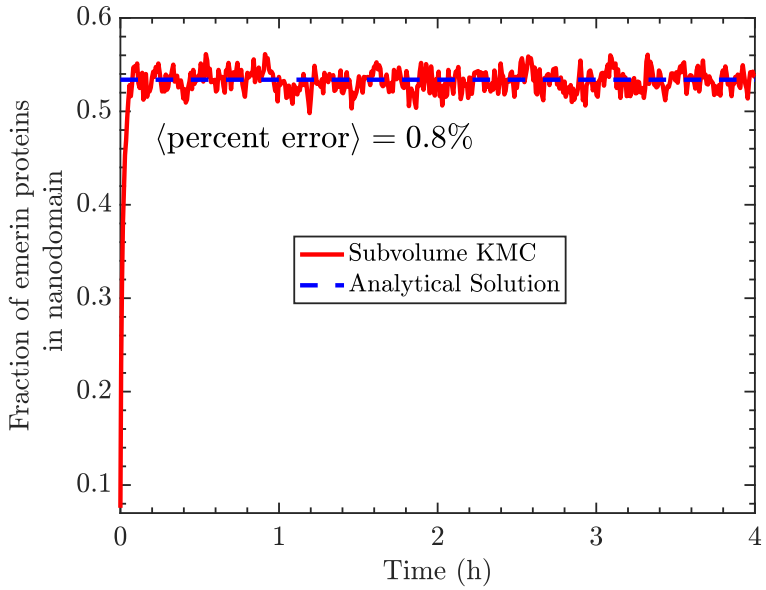

Figure 2. We ran the subvolume KMC method of reference paper [2] using the parameters discussed in the main text of these notes above. The red line is the KMC simulation of the concentration of monomoer emerin proteins in the nanodomain. The blue dashed line is the analytical solution of eqn. (4). The average percent error of the KMC simulation from the analytical solution throughout the entire run is about 0.8%

.

.

Bibliography

[1] Yiwei Li, Osman Kahraman, and Christoph A. Haselwandter. Distribution of randomly diffusing particles in inhomogeneous media. *Physical Review E*, 96(3), Sep 2017.

[2] Johan Elf, Andreas Doncic, and Mans Ehrenberg. Mesoscopic reaction-diffusion in intracellular signaling. In Sergey M. Bezrukov, Hans Frauenfelder, and Frank Moss, editors, *Fluctuations and Noise in Biological, Biophysical, and Biomedical Systems*. SPIE, May 2003.