SIMULATION OF DIFFUSION IN HETEROGENEOUS MEDIA

by

Paul Ionele

A dissertation submitted to the University of Ontario Institute of Technology in accordance with the requirements of the degree of Bachelor of Science (Hons) in the Faculty of Science.

March 20, 2016

UNIVERSITY OF ONTARIO
INSTITUTE OF TECHNOLOGY

ABSTRACT

This is a short example of an abstract.

ACKNOWLEDGEMENTS

Thank some people that you like here.

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of the University of Ontario Institute of Technology. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. Any views expressed in the dissertation are those of the author and in no way represent those of the University of Ontario Institute of Technology. The thesis has not been presented to any other University for examination either in Canada or elsewhere.

Paul Ionele
March 20, 2016

Contents

List of Figures

List of Tables

1. INTRODUCTION

What was/is the purpose of this thesis? Why did we develop these computational simulations? Really, we need a goal!

In this thesis, a lattice Monte Carlo approach was used to simulate diffusion. Also used a finite difference method to simulate diffusion. Both problems were boundary-value problems.

We also performed an analysis: MSD, mean position, etc.?? Analysis is kind of empty!

Overall overview to thesis? What kind of an overview? Or just a general background to some main concepts needed or used in this thesis? How does this compare with the abstract?

1.1 Diffusion Theory

Introduction. Theory. Supporting Figures.

Refer to page 91 in (Patton, Thibodeau , 2013). Figure 4.1 may also be used? Or is it too 'simple'?

1.2 Monte Carlo Theory

Introduction. Theory. Supporting figures.

The "Monte Carlo method" is a general probabilistic algorithm for estimating the probability of an event based on the a number of random trials. Lattice Monte Carlo

(LMC) is basically a type of discretization method. Look into: —

1.3 Master Equation Theory

Introduction. Theory. Supporting figures.

1.4 Simple Cells and Tissues

Nearly all human cells are microscopic in size; their diameters range from 7.5 µm to approximately 150 µm and a cell exhibits a particular size or shape that reflects the specific task it's designated to perform. There are many different types of cells including nerve cells, muscle cells, and gland cells, but despite their anatomical and functional differences, the cells of the human body have many similarities. It is a fact that no cell contains all cellular components found in all the cell types, so often a composite cell (Figure 1.1) is used to exhibit the most important characteristics.

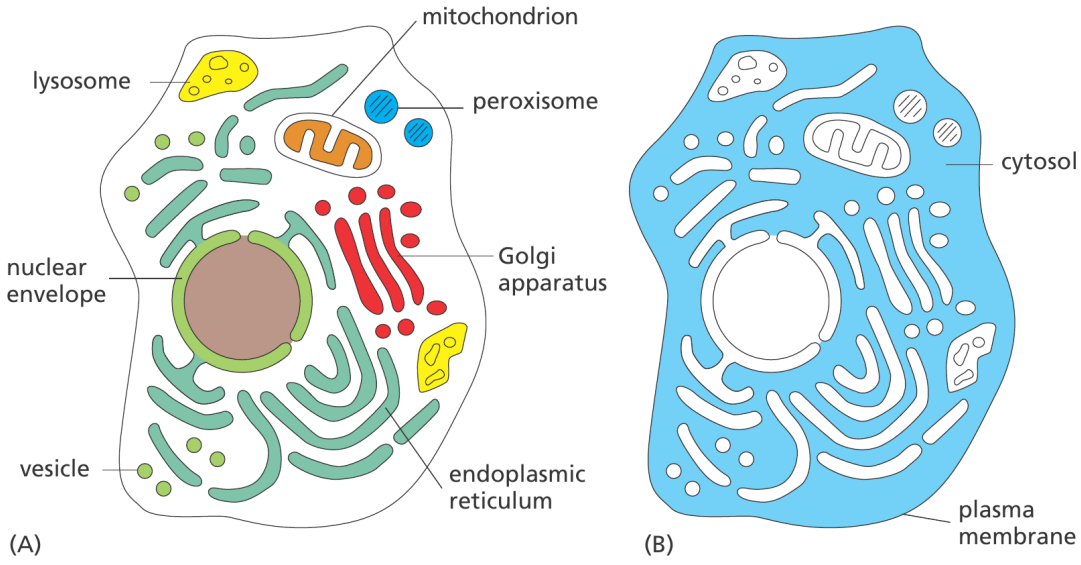

Figure 1.1: Composite cell showing various important and common internal cell structures. Most of the volume within the cell is occupied by the cytosol fluid. Figure courtesy of Essential Cell Biology, Alberts, 3rd.

Each cell is enclosed by a plasma membrane that separates the cell contents from

the surrounding environment. The inside of the cell is mostly composed of a gel-like substance called cytoplasm that is a dense arrangement of proteins, organelles, and other molecules, suspended in a watery fluid called cytosol. The dense crowding of molecules and organelles results in frequent physical interactions which promotes high metabolic efficiency (Patton, Thibodeau , 2013). All of the fluid inside the cell may be referred to as intracellular fluid or simply, cellular fluid (CF).

The CF is separated from the extracellular fluid (ECF) by the cell/plasma membrane. This membrane is a phospholipid bilayer with various embedded macromolecular structures. Each phospholipid molecule is amphiphatic, having both a hydrophobic and hydrophilic region. A collection of phospholipid molecules will naturally arrange themselves into a bilayer that does not allow water, polar molecules, or ions to pass through easily. However, water and other molecules or ions need to traverse this membrane and the water transporting task is accomplished by aquaporin gated channel proteins. Aquaporins facilitate the passive diffusion of water through the plasma membrane, between the intracellular and extracellular regions. (Patton, Thibodeau , 2013).

1.4.1 Tissues

In a multicellular organism, there are several levels of biological organization. A cell is the lowest level of organization that is considered living; tissues are the next higher level of organization and are composed of cells similar in structure and function. This ensemble of cells resides in an extracellular matrix (ECM); a medium containing water, fibrous and adhesive proteins, glycoproteins, and other molecules. The ECM varies in composition between different tissues, but providing structural support and facilitating cell-to-cell communication are common functions of the ECM. In some cells, the cytoplasm is more viscous than the extracellular matrix (Campbell, Reece , 2008). At the cellular level some tissues are relatively organized.

In this project, a simple tissue model was constructed based on some basic defining characteristics of real cells and tissues. The simple tissue model consisted homogeneous spaces, specifically more viscous intracellular regions and less viscous extracellular regions separated by a passive semi-permeable boundary. All of the cells in a model were of the same dimensions and repeated in series.

2. MODELS AND SIMULATIONS

A simulation is intended to imitate in many cases, a real-world process or system that may be too difficult or costly to analyze directly. Before any such simulation can begin, a model of the system studied must be constructed. Models capture the characteristics and behaviours of the system they represent and in general, a model should be as simple as possible (since resources are limited) while still explaining experimental observations and making predictions with a given degree of accuracy. The simulation is the implementation of the model and can be executed on a computer to produce data for testing, analysis, and visual presentation.

In our simplified model of a biological tissue, the relatively ordered and periodic nature of cells in most simple tissues is captured as a series of repeating *unit cells*. These unit cells are the building blocks of the heterogeneous 1D and 2D models. Each unit cell is characterized by a cellular domain, separated from an extracellular domain by a semi-permeable membrane. The domains are isotropic except at the boundaries where a change in diffusivity and semi-permeable boundary exist. Within each domain, the only characteristic modelled is the diffusivity of ideal particles, and is implemented as a directional stepping probability (Section 1.1). For all of the models, the cellular domains had smaller diffusivities (diffusion coefficients) than the extracellular domains, similar to some real tissues. Regarding the boundaries, there exists two kinds in our models. The first kind is a totally-reflecting boundary; it forms the absolute boundary of the model system and represents an insurmountable physical boundary. The second kind is a semi-permeable non-active/passive boundary

and represents the selectively permeable nature of the plasma membrane. In a real biological plasma membrane, the integral membrane-bound proteins can facilitate either active or passive transport. In the simple cell model developed, the semi-permeable membranes behave in a passive transport manner and this is implemented as a boundary transition probability, a concept explained in Section 1.1.

The simulations executed and subsequently analyzed were that of particle diffusion. More specifically, the diffusion of idealized, non-interacting particles experiencing zero net force, and adhering to boundary constraints. Particle motion was therefore undirected but occurred in only one direction (along a line) in the 1D system, and in two orthogonal directions in the 2D system. For each time step, each particle was allowed to move in only one direction or stay in its current location and since particles were non-interacting, multiple occupancy of lattice sites was permitted.

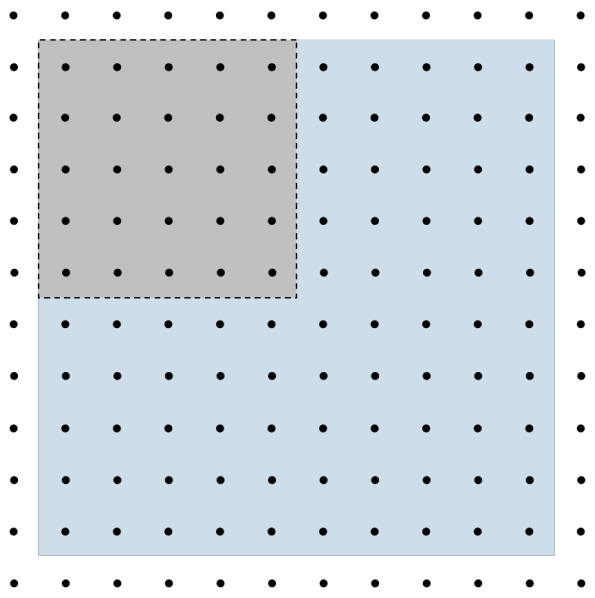

Figure 2.1: A single 2D unit cell forms the building block of the 2D model. An example lattice arrangement is overlayed on the model to show possible particle positions. The dimensions of each domain can be adjusted individually by changing the number of lattice sites used to define each dimension. Dashed lines represent semi-permeable boundaries and lattice sites outside the cell belong to adjacent cells.

It was decided from the start of the project that particles would move in the system by a constant jump/step-size, compared to a continuum step-size (ex. Gaussian).

Although this particle behaviour is generally less representative of a real diffusion process and less accurate compared to continuum Monte Carlo (MC) methods and molecular dynamics simulations, the data generated would still be sufficiently accurate for our analytical purposes and be faster to compute. Fixed step-sizes in the particle movement results in a grid or lattice-like arrangement of particle positions over time. The process simulated is therefore said to be occur on a lattice as the particles are free to move, but only to fixed lattice positions. This manner of particle movement was implemented in both the MC and master equation (ME) simulations.

Overall, the goals were to simulate the process of diffusion for homogenous and heterogeneous systems using MC and ME methods, calculate a mean-square-displacement (MSD) for every time step from the computed density distribution, and from the MSD, determine an effective diffusivity for the system. For both the MC and ME simulations, the particle density distribution data output was of the same form. MSD calculation implementations varied slightly between the MC and ME simulations, but the calculations for effective diffusivity were the same.

2.1 Monte Carlo Simulations

Using MC-based algorithms, information on the individual state of each particle including current position and path history, can be maintained. However, due to the finite number of particles used in the simulation and the stochastic nature of individual particle motion, statistical fluctuations lead to 'non-smooth' density distributions. The use of MC algorithms in simulating the process of diffusion was motivated by the random nature of the diffusion process (Section 1.1) at the level of the individual particle.

2.1.1 1D Homogenous and Heterogeneous Systems

A natural starting point for the project was 1D systems since they are more simple to model and simulate/program than those which are multidimensional. In our 1D simulations, particles move along a line and at any given lattice site, can move to either one of the two neighbour sites, unless an absolute boundary is met. A particle step-size of one lattice unit ($a = 1$) was used, meaning that the particle steps only a single lattice site in a randomly selected direction, each time step. This stepping distance is currently uncorrelated/uncalibrated with any physical or real distance. Each time step in the simulation is also uncorrelated with any characteristic or real time; in the simulation it is simply an integer used as a loop variable. Length and time-scale correlation with physical systems is possible but such an endeavour is outside the time limits of the current project.

In the homogenous system (Figure 2.2) where particle motion is unbiased, a particle that is not at the absolute lattice limits has an equal probability to step one lattice unit in either direction, we'll use the x-direction. Let those probabilities be: P_x^+ and P_x^-. The unbiased particle motion requirement is met by ensuring $P_x^+ = P_x^-$. We can introduce another physically sensible probability; the probability that a particle does not move in a given time step. Let this probability that a particle stays at its current lattice site for the time step be: $P_x^s = 1 - P_x^+ - P_x^-$. If not at an absolute boundary, the particle has three possible future states. At the absolute boundaries, the particle has only two possible future states; it may move from its current lattice site in a direction away from the boundary, or it may stay at its current lattice site. For example, with reference to Figure 2.2, if the particle is at the leftmost boundary, then it can only move to the right neighbour lattice site or stay at its current lattice site. Therefore the probabilities for the particle's motion depend on its position within the system and for this example they are: $0 < P_x^+ \leq 1$ and $P_x^s = 1 - P_x^+$. It should be noted that reflecting boundaries in the simulation model a situation where a particle attempts to

move through the boundary, but is 'reflected' back to its starting position. Therefore in the simulations, the probability of moving towards the boundary is not zero, it is the probability of crossing the boundary that is zero. Since the events of moving towards the boundary and crossing the boundary are treated as mutually exclusive, the total probability of a particle transition into a different domain is the product of the two individual probabilities. In the simulations, the position of the particle within the system is determined and P_x^s and $P_x^{-,+}$ are automatically appropriately set. At this point, it may be interesting to ask "What if the transition probability is not zero at the absolute boundaries?". If the computer simulation does not break from accessing memory space outside a predefined array, then one obtains the case for a semi-permeable boundary and this will be detailed later on.

Figure 2.2: Homogenous 1D cell model with lattice overlay. The sold lines represent the absolute physical limits of the system and behave as reflecting boundaries.

In the heterogeneous system (Figure 2.3), a particle that is not at the absolute lattice limits and is not at a lattice site next to a permeable boundary, has an equal probability to step one lattice unit in either direction, same as in the homogenous system simulation. Particle behaviour at the absolute boundaries was handled the same way as in the homogenous system. The different diffusivities of the cellular and extracellular regions were simulated by using different stepping probabilities. We introduce the subscripts i and e to differentiate between stepping probabilities in the (*intra*)cellular and extracellular domains: $P_{x,i}^{+,-}$ and $P_{x,e}^{+,-}$. Within each domain and excluding the lattice points at the boundaries, the behaviour of the particles was like that of the homogenous system. At the semi-permeable boundaries, the following condition was necessary if the long-time density distribution was to be physically

reasonable:

$$P_{\mathrm{e}\to\mathrm{i}} = \left(\frac{P_{x,\mathrm{i}}}{P_{x,\mathrm{e}}}\right) P_{\mathrm{i}\to e} \tag{2.1}$$

This equation relates the boundary transition probabilities $P_{e\to i}$ (extracellular to cellular transition) and $P_{i\to e}$ (cellular to extracellular transition) between regions of different diffusivities, characterized by the directional stepping probabilities $P_{x,i}$ and $P_{x,e}$, under the condition that $P_{x,i} < P_{x,e}$ (Section 1.1). Note that $P_{e\to i}$ *or* $P_{i\to e}$ may be set arbitrarily, but one determines the other according to Equation 2.1; they cannot be set independently if the correct density distribution in the long time is desired.

As an example, consider a particle at a lattice site adjacent to semi-permeable boundary. If the particle is in a cellular region, the total transition probability for the particle to the extracellular region is the product of the transition probability from the cellular to extracellular region and the *cellular* directional stepping probability.

$$P_{\mathrm{total},\,\mathrm{i}\to\mathrm{e}} = \left(\frac{P_{x,\mathrm{e}}}{P_{x,\mathrm{i}}}\right) P_{\mathrm{e}\to i} \cdot P_{x,i} \tag{2.2}$$

Using Equation 2.1 for the transition probabilities, the total transition probability across any semi-permeable boundary can be determined. In the case of absolute reflecting boundaries, the transition probabilities in Equation 2.1 are zero, and hence the total transition probability is zero.

Figure 2.3: Heterogeneous 2D unit cells with cellular and extracellular regions. Dashed lines indicate semi-permeable boundary, separating regions characterized by different diffusivities. Dots outside coloured regions indicate continuity of system.

All directional stepping probabilities, boundary transition probabilities, region dimensions, number of particles used, and time step limit are initialized prior to running the simulation. More specifically, the region dimensions are defined by a number of lattice sites used for that region (i.e. the length of a cellular space could be set as n lattice sites) and the size of a unit cell is the sum of lattice sites in the cellular and extracellular regions. Therefore, the density of lattice sites within any region is always the same, regardless of the size of that region. The total length of the system is the product between the number of lattice sites used in a unit cell and the number of unit cells.

Our MC-based simulations of diffusion require a 'random' number generator. Pseudo-random numbers, drawn from a uniform distribution, were generated during the execution of the simulation using the `rand()` C-library function and `RAND_MAX` built-in constant. It was desired that the random numbers (rnd) be uniformly distributed over the interval $0 \leq \text{rnd} < 1$, so all random numbers returned by the `rand()` call were normalized by `RAND_MAX`.

The simulations produced a particle density distribution at every time step. One of the goals of this project was to compute the MSD of the particles at every time step and collect this data for further analysis. Since the individual position of each particle was tracked during the simulation, it was possible to calculate the MSD of all particles for a time step. Let x_i be the position of the i^{th} particle at t_n, the MSD:

$$\langle \Delta x^2 \rangle = \langle x_i^2 \rangle + \langle x_i \rangle^2 \tag{2.3}$$

The angle brackets indicates the sum of the positions divided by the number of particles:

$$\langle x_i^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} x_i^2 \tag{2.4}$$

$$\langle x_i \rangle = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{2.5}$$

Since the particles experienced no external force, it was expected and shown that for every time step, $\langle x_i \rangle \approx x_0$, where x_0 is the initial starting lattice site of all the particles. The mean-squared-position $\langle x_i^2 \rangle$ was not constant for every time step, it increased with time reflecting the 'spreading' of the particles outwards from their origin.

The output of our simulations was two *.txt files. One file contained the density distribution data computed at every time step. The second file contained the simulation analytics: $\langle x_i \rangle$, $\langle x_i^2 \rangle$, and MSD data computed from the density distribution for every time step. These quantities were all that was needed for further calculations and the creation of various plots for analysis purposes. From the MSD, effective particle diffusivities could be computed for various cellular and extracellular diffusivities, semi-permeable boundary transition probabilities, and geometrical variations of the model. The density distribution data was processed by a separate program to create plots for every time step; the result an animation of the particle diffusion process in the system.

2.1.2 2D Heterogeneous Systems

Simulations of particle diffusion were also performed for 2D systems. The 2D model developed and simulated was heterogeneous; no homogenous model was developed. The MC algorithm for the simulation of the 2D system was in principle the same as for the 1D heterogeneous system, however, there are a few important differences to note. For every given time step, a particle at some lattice site that is not at an absolute boundary has four directional steps available and particle movement in the orthogonal directions are independent but cannot both occur in a single time step.

Similar to the 1D simulation, the 2D unit cell is the building block of the 2D model, the particle moves 'on a lattice', and a step-size of one lattice unit is permitted. Each unit cell consists a cellular region separated from an extracellular region by a semi-permeable boundary. A key characteristic of the unit cell is that, when linked in series, an extracellular channel is formed providing particles the option to move freely without obstruction. In Figure 2.4, this is visible as an uninterrupted region or channel in the lower half of the figure. It is possible to arrange more than one series of unit cells so that a larger $(m \times n)$ unit cell configuration is obtained, but this was not studied. Only $(1 \times n)$ unit cell models were simulated.

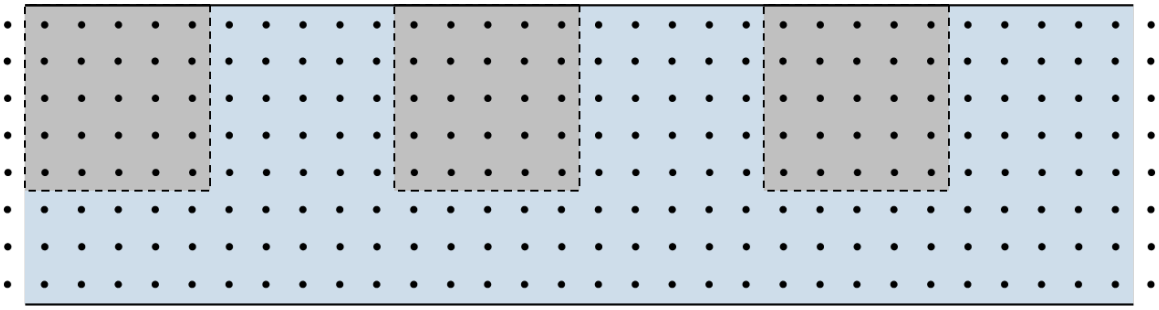

Figure 2.4: A $(1 \times n)$ series configuration of 2D unit cells with cellular and extracellular regions. Dashed lines indicate semi-permeable boundary and solid lines indicate absolute boundary. Dots outside the coloured region indicate continuity of the system.

In 2D, there are an additional two neighbour lattice sites the particle can move to. If not at an absolute boundary, the particle has five possible future states. At the absolute boundaries, the particle has only 4 possible future states; it may move from its current lattice site in a direction away from the boundary, along the boundary, or it may stay at its current lattice site. We introduce an additional set of directional stepping probabilities in $\hat{y}$: $P_{y,i}^{+,-}$ and $P_{y,e}^{+,-}$. The directional stepping probabilities in $\hat{x}$ and $\hat{y}$ have the conditions: $0 < P_{i,e}^{+,-} \leq 0.25$. For unbiased motion (excluding boundary lattice sites) all directional stepping probabilities must be equal. Therefore, the probability that the particle y-coordinate does not change for a given time step is $P_y^s = 1 - P_y^+ - P_y^-$. Similarly for the x-coordinate $P_x^s = 1 - P_x^+ - P_x^-$. So the total

probability that a particle remains at its current lattice site is $P^s = 1 - P^s_x - P^s_y$. When a random number is generated, its value is tested for membership in intervals defined by the stepping probabilities, and this determines the motion of the particle.

Handling of the boundary transition probabilities was done in the same manner as the 1D simulation, in accordance with Equation 2.1. In the program developed, for every time step, the current particle position and possible future particle position are compared. If it is found that the future possible position is in a different region than the current position, then the probability to cross into the new region is automatically set using Equation 2.2 solved for $P_{\text{total, i}\to\text{e}}$ or $P_{\text{total, e}\to\text{i}}$. Absolute boundaries were implemented as reflecting boundaries.

Similar to the 1D simulations, density distribution data and other analytical data was computed and written to *.txt files. The density distribution data output was in a matrix form; each line of data represented a different row of lattice sites in the system. For each time step, all rows were written to the file and an empty line inserted before writing data for the next time step, for parsing purposes. Producing 2D animations of the diffusion process required a program different than the one used to create the 1D animations; however, similar programs were used to simulation analytics.

2.2 Master Equation Simulations

Using the ME methods and evolving the particle density distribution in time, discontinuities in the distributions due to statistical fluctuations are no longer an issue; however, at the expense of information on individual particle state. Using this method, it is not possible to know individual particle position history which restricts the use of certain analyses. For example, it is not possible to reconstruct the paths of particles through the system which can be used to generate a frequency histogram to score the number of times a particle crosses a semi-permeable boundary. The main reason

for use of the ME methods was that most of our intended analysis only required knowledge of $\langle x_i \rangle$, $\langle x_i^2 \rangle$, and MSD. These quantities could be calculated, although in a slightly different way than implemented in the MC methods. In addition, for the basic models simulated, ME-based simulations were much faster than MC-based simulations.

Consider the situation of a single particle at a lattice site that, at a given time step, has a directional stepping probability of $P_x^{+,-} = 0.5$ (similar to a coin-toss experiment). In the MC simulations, a random number is drawn from uniform distribution and tested for membership in the intervals defined by the directional stepping probabilities.

$$\begin{cases} x_0 - 1 & P_x^- : 0.0 \leq \text{rnd} < 0.5 \\ x_0 + 1 & P_x^+ : 0.5 \leq \text{rnd} < 1.0 \end{cases}$$

For a small, even number N of particles, it may not be reasonable to expect that the number of particles that step -1 or $+1$ in $\hat{x}$ is exactly N/2. In terms of particle positions, even though the particles have an equal probability of stepping in either direction, the mean position of the particles $\langle x \rangle$ will likely not equal x_0, the expected value of the mean position. However, the law of large numbers guarantees that the arithmetic mean of the position of the particles practically converges to the expected value in the limit of very many particles. In the ME simulations, it is assumed that there are very many particles, so many that the fluctuation in the mean position of the particles is also assumed zero. Instead of starting with some number of particles at x_0, we start with a particle density ρ_0 at x_0 and evolve the particle density distribution in time. As an example, if the density is $\rho(x_0) = \rho_0$ with the stepping probabilities $P_x^{+,-} = 0.2$ and $P_x^s = 0.6$, then at the next time we have the density distribution $\rho(x_0 - 1) = \rho(x_0 + 1) = 0.2$ and $\rho(x_0) = 0.6$.

ME simulations were also performed on a lattice; that is, the step-size in 'par-

ticle' movement was fixed and for each time step, only neighbour lattice sites were accessible. Within a cellular or extracellular region, directional stepping probabilities were equal for unbiased motion. At absolute and semi-permeable boundaries, the density distribution was, in principle, treated similar to how individual particles were treated in MC, although the implementation was different. Instead of the value of a random number determining the movement of a particle, the evolution of the density distribution to the next time step depended on the probability distribution of possible particle movements at that lattice site. Clarifying with an example, consider a lattice site with particle density ρ_0 in a cellular region and adjacent $(-\hat{x})$ to the site is a semi-permeable boundary. The boundary transition probability follows Equation 2.1 solved for $P_{i\to e}$ and the total transition is probability (Equation 2.2):

$$P_{\text{total, i}\to\text{e}} = \left(\frac{P_{x,\text{e}}}{P_{x,\text{i}}}\right) P_{\text{e}\to i} \cdot P_{x,i} \tag{2.6}$$

Therefore, the density distribution at the next time step is:

$$\begin{cases} \rho(x_0 - 1) = \rho_0(x_0) \cdot P_{\text{total, i}\to\text{e}} \\ \rho(x_0 + 1) = \rho_0(x_0) \cdot P_{x,i} \\ \rho(x_0) = 1 - \rho_0(x_0) \cdot P_{\text{total, i}\to\text{e}} - \rho_0(x_0) \cdot P_{x,i} \end{cases}$$

3. RESULTS AND ANALYSIS

4. FUTURE WORK AND CONCLUSION

May just put this into the results section near the end as conclusions and future work.

REFERENCES

Navarro, J.F., Frenk, C.S., & White, S.D.M. 1995, MNRAS, 275, 720

Campbell, N.A., Reece, J.B., et al. 2008. Biology. 8th ed. Pearson Benjamin Cummings.

Patton K.T., Thibodeau G.A.. 2013, 8th ed. Elsevier.

A. APPENDIX

This is just an example of an appendix.