Random Walks for Modelling Cell Populations
VRES 2013-14

Morgan Wall

n7532962

18 February 2014

Contents

1   Introduction

This report documents the tasks completed by Morgan Wall as part of his 2013-14 VRES scholarship with Dr. Matthew Simpson. The project covered the use of random walks in simulating cell motility, proliferation, and interaction in a variety of biological contexts. Furthermore, the derivation and analysis of partial differential equations (PDEs) to model equivalent processes was considered.

2   Simulating Random Walks in 2D

Mathematically, random walks consist of a sequence of numbers generated independently via stochastic processes [3]. Random walks can be used in modelling a wide variety of physical phenomenon. Several biological random walk models, either analsyed or proposed by Simpson, were researched. Each model accounted for cell motility and proliferation, which describes a cell's ability to move spontaneously and cell mitosis, respectively. Both of these cellular mechanisms are essential in modelling the biological processes considered [9, 10]. The models varied based on the consideration of cell interactions via volume exclusion or by the proliferation mechanism used.

2.1   Methodology

There exists several methods for visualising a random walk [3]. This research project focused on constraining the random walk to a multi-dimensional rectangular lattice. The methodology applied mimics that of Simpson's prior work [9, 10]. Considering a two-dimensional $m \times n$ lattice, for example, agents can assume positions at (i, j), where $i, j \in \mathbb{N}^+$, $i \leq m$, and $j \leq n$. Each non-boundary node (i, j) therefore has the neighbouring nodes: $(i, j+1)$, $(i, j-1)$, $(i+1, j)$, and $(i-1, j)$. Typically, a combination of periodic [4] and reflective boundary conditions [7] are imposed on the lattice.

The sequence of random numbers generated for a random walk were used in moving agents, initially positioned on the lattice, to neighbouring nodes. Furthermore, such random numbers may also have been used for placing daughter agents at an adjacent node in replicating cell mitosis. Over a period of time, such movement would form a path taken by an agent in discrete, random steps. For simulations, time was consequently discretised into fixed time steps of duration δt. Following Simpson's approach, given that N agents are present at the start of a time step, N agents were selected at random for action. Depending on the biological process being modelled, each action could include both motility and proliferation events. The exact algorithms constituting an action are defined in detail in [9, 10] and for brevity they are not redefined in this report.

2.2   Implementation

The random walk models defined in [9, 10] were reproduced using the C programming language in conjunction with Intel's Math Kernel Libraries (MKL). The results of all simulations were output to CSV files for analysis and visualisation using MALTAB. A

substantial amount of code was written to replicate the aforementioned models and was therefore omitted from this report.

2.3   Results

In Simpson's work, several PDEs were derived from agent conservation statements in describing various biological phenomenon modelled by the random walk simulations. Depending on the constraints placed on the system, the forms of the PDEs could vary quite substantially. For example, a random walk model with volume exclusion and both cell motility and proliferation, the governing PDE takes the form of Fisher's equation. In the absence of volume exclusion, an advection-diffusion equation governs the system.

To demonstrate the reproduced model, simulations were conducted on a two-dimensional lattice with periodic boundary conditions and without cell proliferation. Agent density data was obtained by averaging the occupancy of each column over multiple identically-prepared realisations. As such, a one-dimensional unbiased cell motility problem was modelled. The results of the simulations were compared against the following IVP, as derived in Simpson's prior work [10]:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} \qquad\qquad -\infty < x < \infty \qquad\qquad (1)$$

$$C(x,0) = \begin{cases} 0 & x < -h \\ C_0 - h & -h < x < h \\ 0 & x > h \end{cases} \qquad\qquad -\infty < x < \infty \qquad\qquad (2)$$

where C_0 is the initial agent density. This IVP has the analytic solution:

$$C(x,t) = \frac{C_0}{2}\left(\operatorname{erf}\left(\frac{h-x}{2\sqrt{Dt}}\right) + \operatorname{erf}\left(\frac{h+x}{2\sqrt{Dt}}\right)\right). \qquad\qquad (3)$$

To revise numerical approaches for solving PDEs, this IVP was solved numerically using a finite volume spatial discretisation with backward Euler time stepping. As seen in Figure 1, the random walk model matches the analytic and numeric solutions closely as time elapses.

Additional simulations were conducted in thoroughly examining the random walk model. As expected from (3), modifying C_0 merely scales the agent density curves (see Figure 2). Furthermore, as expected by the Law of Large Numbers, the accuracy of the reproduced model was dependent on the number of realisations averaged over [8]. As seen between Figures 1 and 3, an increase in the total realisations performed results in higher accuracy in the final solution.

3   Random Lattices in 1D

In simulating cellular phenomena via random walk models, there exists three common methods for handling the spatial domain of interest: equally-spaced lattices, random lattices,
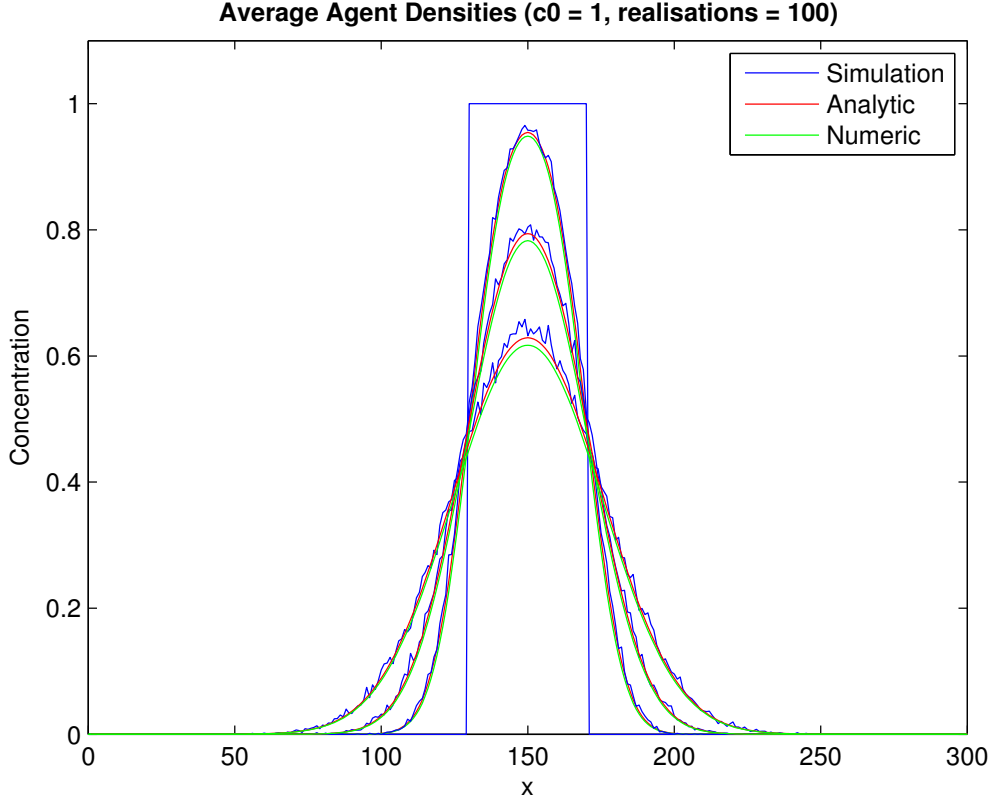
Figure 1: Comparison of agent density between a random walk model and the diffusion equation. The model had volume exclusion with unbiased motility and no proliferation. Density data was obtained by averaging the occupancy of each column over 100 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500,$ and 1000 and shows agent density diffusing across the lattice.

and lattice free models. Lattices can be used under a variety of circumstances to discretise the spatial domain into a distinct set of nodes, which cells can consequently occupy. As discussed by Simpson [9], lattices can be constructed based off the expected size of a cell and the approximate distance they may move over a period of time. In constructing an equally-spaced lattice, nodes are spaced uniformly throughout a domain. Alternatively, nodes are positioned randomly or pseudorandomly within a domain in constructing a random lattice. Unfortunately, cells do not exhibit movement in a rigid, discrete manner, hence lattice-based models may not accurately simulate certain biological processes [5]. Lattice free models avoid this issue by not imposing a spatial discretisation on the simulations, which results in a major performance penalty.

Research was conducted into the application of random lattices in simulating cell motility in one dimension. Two approaches were taken in constructing a random lattice: perturbing an equally-spaced lattice and generating node locations pseudorandomly. Let x_i be the position of the ith node in the lattice. Under the perturbing approach, each node position is given by:
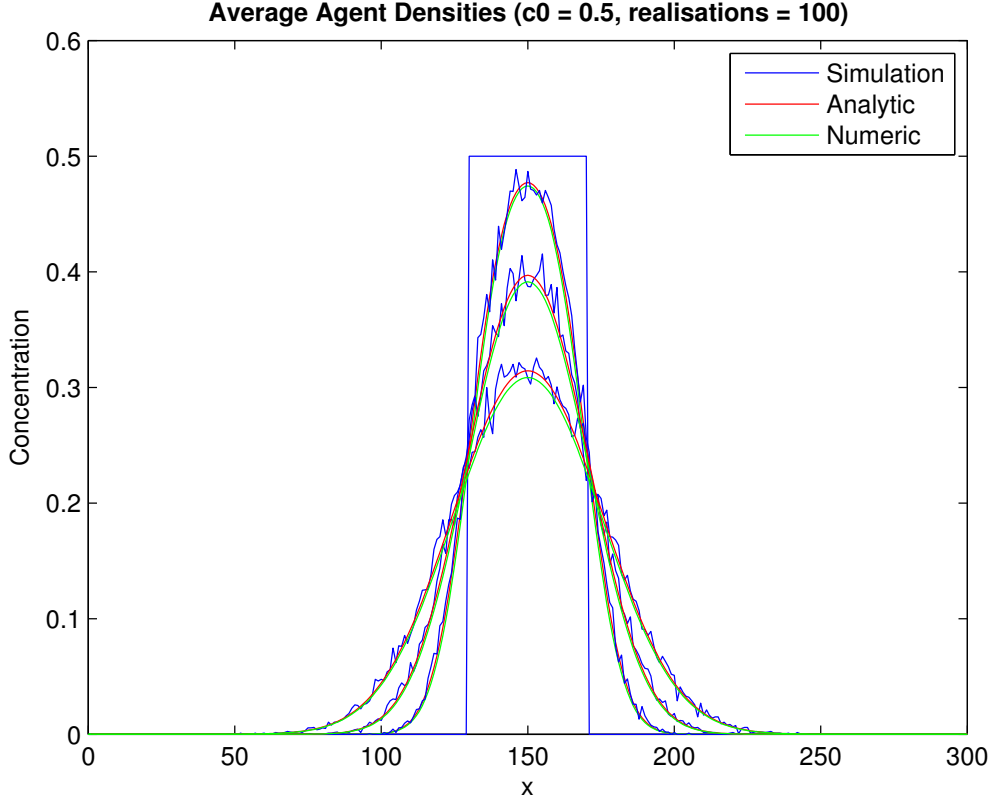
$$x_i = y_i + \Phi$$

4

Figure 2: Comparison of agent density between a random walk model and the diffusion equation. The model had volume exclusion with unbiased motility and no proliferation. Density data was obtained by averaging the occupancy of each column over 100 identically-prepared realisations satisfying (2) where $C_0 = 0.5$. Densities are given at $t = 0, 200, 500,$ and 1000 and shows agent density diffusing across the lattice.

where y_i is the location of the ith unpertubed node in the equally-spaced lattice and Φ is a random variable. Under the pseudorandom approach, each node position is given by:

$$x_i = \Theta$$

where Θ is a random variable. For all simulations, pseudorandom numbers were sourced from a variety of distributions using Intel's Vector Statistical Library (VSL).

3.1 The Perturbing Approach

3.1.1 Lattice Properties

Initially, several key characteristics of lattices were identified in determining the preferred method for generating random lattices. The deltas between adjacent nodes and the existence of node clusters were of particular interest in generating physically realistic lattices for simulating cellular phenomenon. A series of random lattices were constructed on a dimensionless 1D region of 20 dimensionless length units in which 20 nodes would be
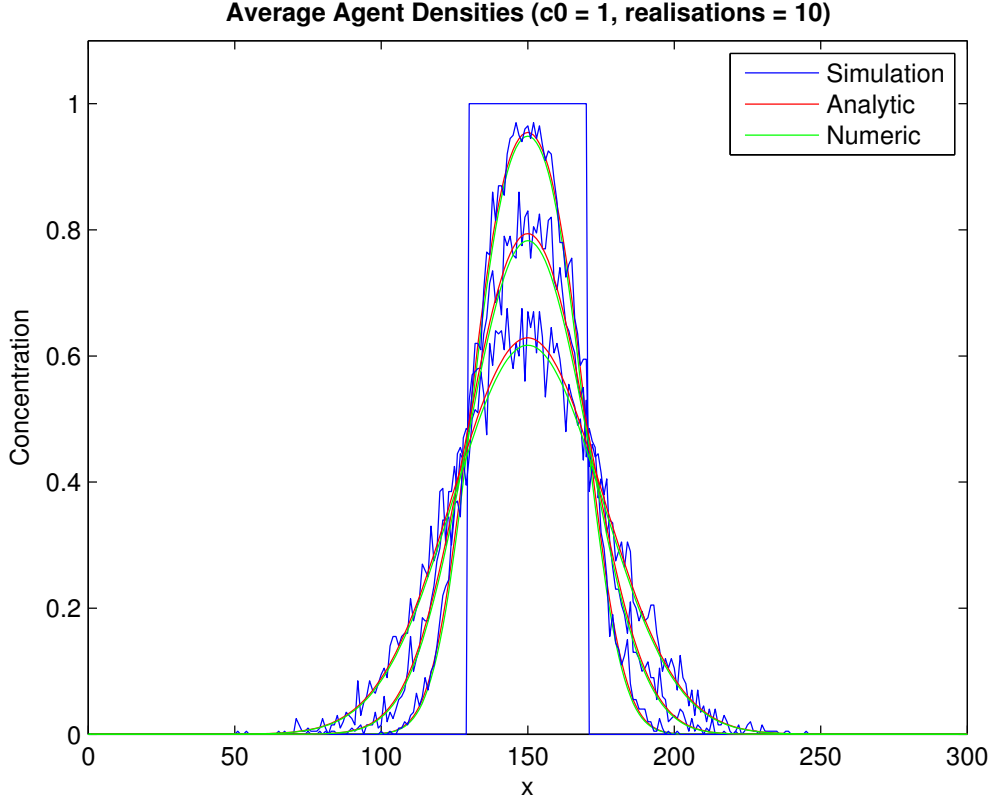
Figure 3: Comparison of agent density between a random walk model and the diffusion equation. The model had volume exclusion with unbiased motility and no proliferation. Density data was obtained by averaging the occupancy of each column over 10 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500,$ and 1000 and shows agent density diffusing across the lattice.

positioned. For the perturbation method, the equally-spaced lattice consisted of 20 nodes spaced one length unit apart with the first node residing at position zero.

Figure 4 depicts random lattices generated using the perturbation approach given by the continuous random variable

$$\Phi \sim \text{Uniform}(-\delta x, \delta x) \tag{4}$$

where δx is the maximum absolute perturbation allowed. These results quickly revealed that large perturbations are detrimental to the construction of random lattices since node clustering is evident. For example, consider the lattices where $\delta x = 0.1$ and $\delta x = 0.5$. The former closely resembles an equally-spaced lattice whereas the latter exhibits node clustering around $x = 6$ and $x = 13$ and large unoccupied regions around $x = 5$. Such disparity in lattice spacing can result in erroneous simulations since cell motility can be subject to large variations in the distance travelled per unit time. These findings are exacerbated as δx is increased further, which can be shown by considering the expected value and variance of the deltas between adjacent nodes.

Consider a random lattice generated using the perturbing approach and given the un-perturbed lattice has a fixed node spacing Δx. Let A and B be independent, identically
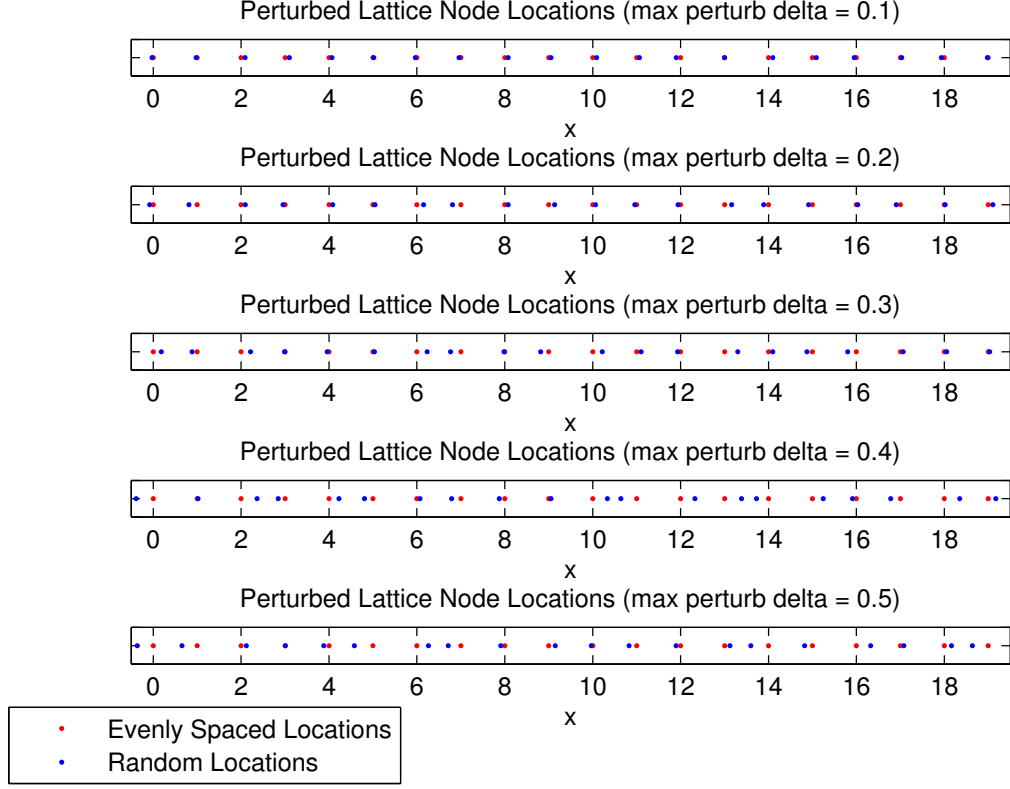
Figure 4: A series of randomly populated lattices generated by perturbing an equally-spaced lattice. Perturbations are sourced from a uniform distribution whose range is defined by a maximum allowable perturbation delta δx as given by (4). The investigated δx values range from 0.1 to 0.5.

distributed random variables. The expected node spacing is given by:

$$
\begin{aligned}
\mathrm{E}\left(x_i - x_{i-1}\right) &= \mathrm{E}\left(y_i + A - y_{i-1} - B\right) \\
&= \mathrm{E}\left(\Delta x + A - B\right) \\
&= \Delta x + \mathrm{E}\left(A\right) - \mathrm{E}\left(B\right) \\
&= \Delta x.
\end{aligned}
$$

As such, if we were to investigate the average delta between adjacent nodes over numerous meshes, then by the Law of Large Numbers the expected delta is that of the equally-spaced lattice. This finding does not explicitly describe the existence of node clustering and large deltas between nodes as seen previously. Instead, consider finding the variance of the same

statistic as follows:

$$
\begin{aligned}
\mathrm{Var}\,(x_i - x_{i-1}) &= \mathrm{Var}\,(y_i + A - y_{i-1} - B) \\
&= \mathrm{Var}\,(\Delta x + A - B) \\
&= \mathrm{Var}\,(A - B) \\
&= \mathrm{Var}\,(A) + \mathrm{Var}\,(B) - 2\mathrm{Cov}\,(A, B) \\
&= \mathrm{Var}\,(A) + \mathrm{Var}\,(B) - 2\mathrm{E}\,(AB) + 2\mathrm{E}\,(A)\,\mathrm{E}\,(B) \\
&= 2\mathrm{Var}\,(A)
\end{aligned}
$$

noting that $\mathrm{Var}\,(A) = \mathrm{Var}\,(B)$. Therefore the variability in the deltas between adjacent nodes is linearly proportional to the variance of the distribution from which perturbations are sourced from.

These same results were exhibited by averaging adjacent node spacings over a series of random lattices. For differing values of δx, 100 random lattices were generated in which 40 nodes were placed over a 1D region consisting of 40 dimensionless length units. As seen in Figure 5, the variability in minimum and maximum node spacings tends to increase linearly with δx. Furthermore, the mean node spacing (not shown) consistently converges to the expected value of $\Delta x = 1$ by the Law of Large Numbers. Additional analysis was conducted into the identification of node clustering in the absence of averaging, however, the results were somewhat inconclusive due to randomness and are not discussed here.

3.1.2 Simulating Random Walks

Random walks were conducted on perturbed lattices to obtain information regarding how such perturbations may impact cell motility. As per Simpson's analysis [10], the net agent displacement,

$$
X_t - X_0 = \sum_{i=1}^{t} (X_i - X_{i-1}), \tag{5}
$$

and the sum of squared displacements,

$$
S_t = \sum_{i=1}^{t} (X_i - X_{i-1})^2, \tag{6}
$$

were investigated for different reasons. Note carefully the distinction between X_i and x_i: X_i is the location of a cell at the ith time step whereas x_i is the location of the ith node. The net displacement gives information on how a cell displaces itself from its starting position. Given the random walk method described section 2, it is expected that the net displacement should fluctuate about zero. On the other hand, the sum of square displacements is used to describe how the cell moves over the course of the simulation. As such, fluctuations can be observed as a cell travels over large distances and within node clusters, which is not adequately represented by (5). The expected sum of squared displacements is

$$
\mathrm{E}\,(S_t) = \sum_{i=1}^{t} \mathrm{E}\left((X_i - X_{i-1})^2\right) \tag{7}
$$

by the linearity of the expected value operator. Let A, B, and C be independent, identically distributed random variables and Δx be the spacing between nodes on the equally-spaced
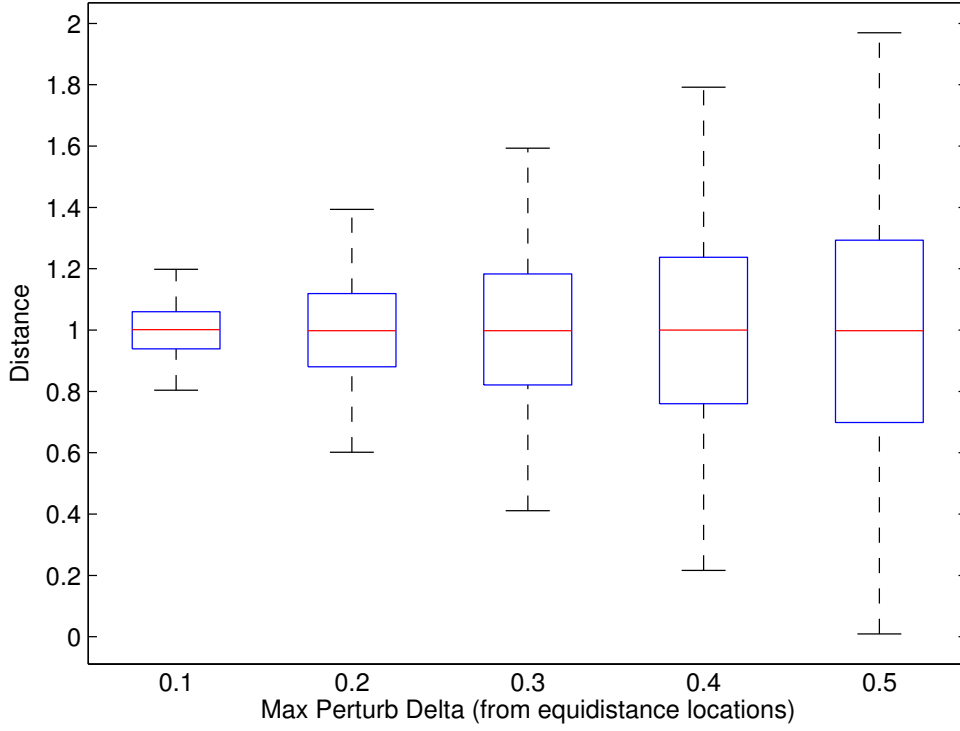
8

Figure 5: Boxplots of the spacing between adjacent nodes as collected over 100 random lattices for varying values of δx. Each random lattice consisted of 40 nodes spread over 40 dimensionless length units. The investigated δx values range from 0.1 to 0.5.

lattice. Additionally, suppose that at the ith time step the cell is located at x_j. Now consider

$$
\begin{aligned}
\mathrm{E}\left((X_i - X_{i-1})^2\right) &= \mathrm{E}\left(\frac{1}{2}(x_{j+1} - x_j)^2 + \frac{1}{2}(x_{j-1} - x_j)^2\right) \\
&= \frac{1}{2}\mathrm{E}\left((x_{j+1} - x_j)^2\right) + \frac{1}{2}\mathrm{E}\left((x_{j-1} - x_j)^2\right) \\
&= \frac{1}{2}\mathrm{E}\left((\Delta x + A - B)^2\right) + \frac{1}{2}\mathrm{E}\left((\Delta x + C - B)^2\right).
\end{aligned}
$$

For brevity, consider

$$
\begin{aligned}
\frac{1}{2}\mathrm{E}\left((\Delta x + A - B)^2\right) &= \frac{1}{2}\Delta x^2 + \Delta x \mathrm{E}(A - B) + \frac{1}{2}\mathrm{E}\left((A - B)^2\right) \\
&= \frac{1}{2}\Delta x^2 + \frac{1}{2}\left(\mathrm{E}(A^2) - 2\mathrm{E}(AB) + \mathrm{E}(B^2)\right).
\end{aligned}
$$

9

Noting that A, B, and C and independent and $\mathrm{Var}\,(A) = \mathrm{Var}\,(B) = \mathrm{Var}\,(C)$,

$$\frac{1}{2}\mathrm{E}\left((\Delta x + A - B)^2\right) = \frac{1}{2}\Delta x^2 + \frac{1}{2}\left(\mathrm{E}\left(A^2\right) - \mathrm{E}\left(A\right)^2 + \mathrm{E}\left(B^2\right) - \mathrm{E}\left(B\right)^2\right)$$
$$= \frac{1}{2}\Delta x^2 + \frac{1}{2}\left(\mathrm{Var}\,(A) + \mathrm{Var}\,(B)\right)$$
$$= \frac{1}{2}\Delta x^2 + \mathrm{Var}\,(B).$$

Applying this result in simplifying (7), gives

$$\mathrm{E}\left((X_i - X_{i-1})^2\right) = \frac{1}{2}\Delta x^2 + \mathrm{Var}\,(B) + \frac{1}{2}\Delta x^2 + \mathrm{Var}\,(B)$$
$$= \Delta x^2 + 2\mathrm{Var}\,(B),$$

therefore,

$$\mathrm{E}\left(S_t\right) = t\left(\Delta x^2 + 2\mathrm{Var}\,(B)\right). \tag{8}$$

Using a similar approach as above, in the absence of perturbation it can be shown that,

$$\mathrm{E}\left(S_t\right) = t\Delta x^2. \tag{9}$$

Alternatively, consider the case where there is no variance in the perturbations, which is equivalent to translating the domain of interest. In this case, (8) clearly reduces to (9). Therefore, as the variance in perturbations increases so does $\mathrm{E}\left(S_t\right)$.

Random walk simulations of a single agent (i.e. cell) were conducted over a series of lattices in ensuring the averaged results matched (8) and (9). Initially, equally-spaced lattices were considered for 500 time steps. Periodic boundary conditions were applied but never encountered by the agent by making the ratio of Δx and the length of the domain sufficiently small given the number of time steps. Linear least-squares fits were made to both the net displacement and the sum of squared displacements in comparing the results to the expected solutions. As seen in Figure 6, given that $\Delta x = 1$, the results match the expected solutions well with a fitted gradient for S_t of approximately 1.0.

As increasily large perturbations are introduced, the value of $\mathrm{E}\left(S_t\right)$ increased. For example, individual walks were conducted on 100 perturbed lattices where perturbations were sourced from (4) for $\delta x = 0.5$ and $\Delta x = 1$. As seen in Figure 7, the expected value of S_t has increased with δx compared to Figure 6. Under such a perturbation scheme, the expected gradient of S_t is

$$\mathrm{m} = \Delta x^2 + 2\mathrm{Var}\,(B) = 1 + \frac{1}{6} = \frac{7}{6} = 1.1\overline{6}.$$

The fitted gradients of 1.16 for S_t and 0.0 for $X_t - X_0$ match the expected values quite well. The expected result was reproduced for a variety of values for δx.

The expected solutions derived earlier are general enough to apply to a variety of distributions. For example, perturbations were sourced from a Gaussian distribution with a mean of zero and a standard deviation of 0.2887. This standard deviation provides a variance equivalent to that of (4) with $\delta x = 0.5$, whose results are seen in Figure 7. Unsurprisingly, the same results are generated by sourcing from the aforementioned Gaussian distribution (Figure 8).

Figure 6: The progression of S_t and $X_t - X_0$ for a single agent over time on an equally-spaced lattice with $\Delta x = 1$. Linear least-squares fits were made from this data where the resultant functions were constrained to pass through $(0, 0)$. The fitted gradient for S_t was 1.0 and for $X_t - X_0$ it was 0.0 (rounded to one decimal place).

3.1.3 Pertubing and Diffusivity

As outlined in section 2.3, Simpson showed the diffusion equation can be used to describe certain cellular processes when simulated on an equally-spaced lattice. It was consequently investigated whether such an analytic solution also applies to simulations conducted on a perturbed lattice.

Initially, simulations were designed and conducted to compare (3) with (2)) (i.e. the diffusion equation) against agent density data averaged over multiple realisations. A series of simulations were performed for a one-dimensional equivalent of the model detailed in section 2.3. Instead of visualising the random walks on an equally-spaced lattice, however, a perturbed lattice was used instead. As seen in Figure 9, the averaged simulation data fits the diffusion equation quite well. Similar results were obtained for $\delta x < \Delta x/2$.

Originally, Simpson considered a one-dimensional agent conservation statement in showing that the diffusion equation can be used to describe certain cellular phenomenon. Let $C_i(t)$ denote averaged agent occupancy (i.e. concentration) at the ith node at time t. Consider a random walk model with volume exclusion and unbiased cell motility. If an agent is chosen for action, suppose that there is probability P that the agent attempts to move. As such,
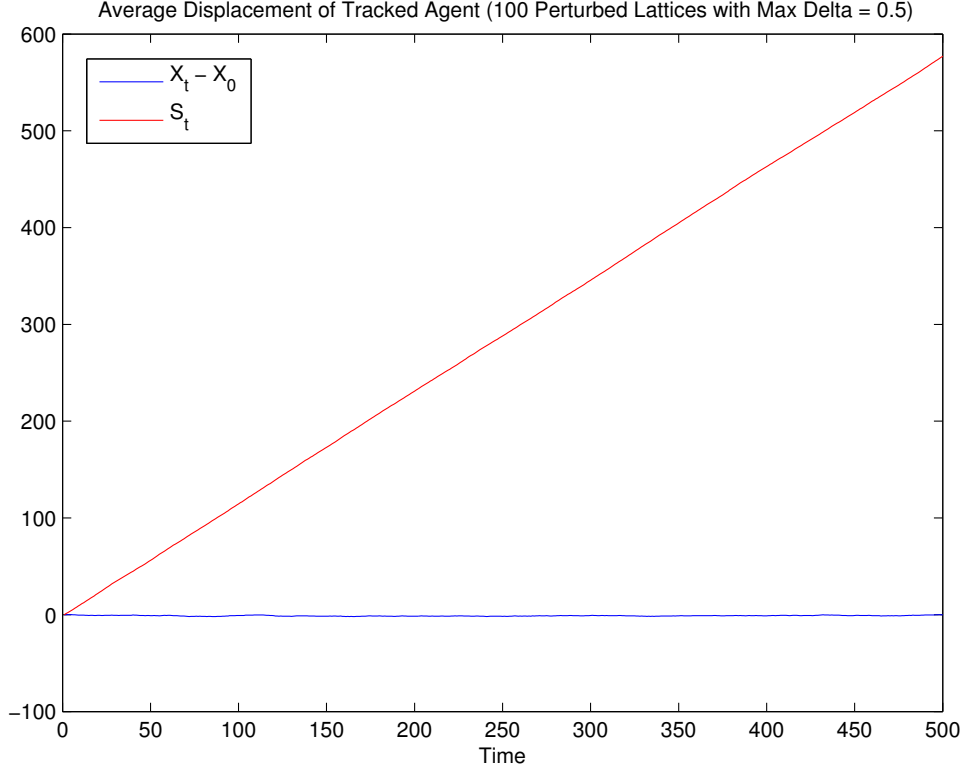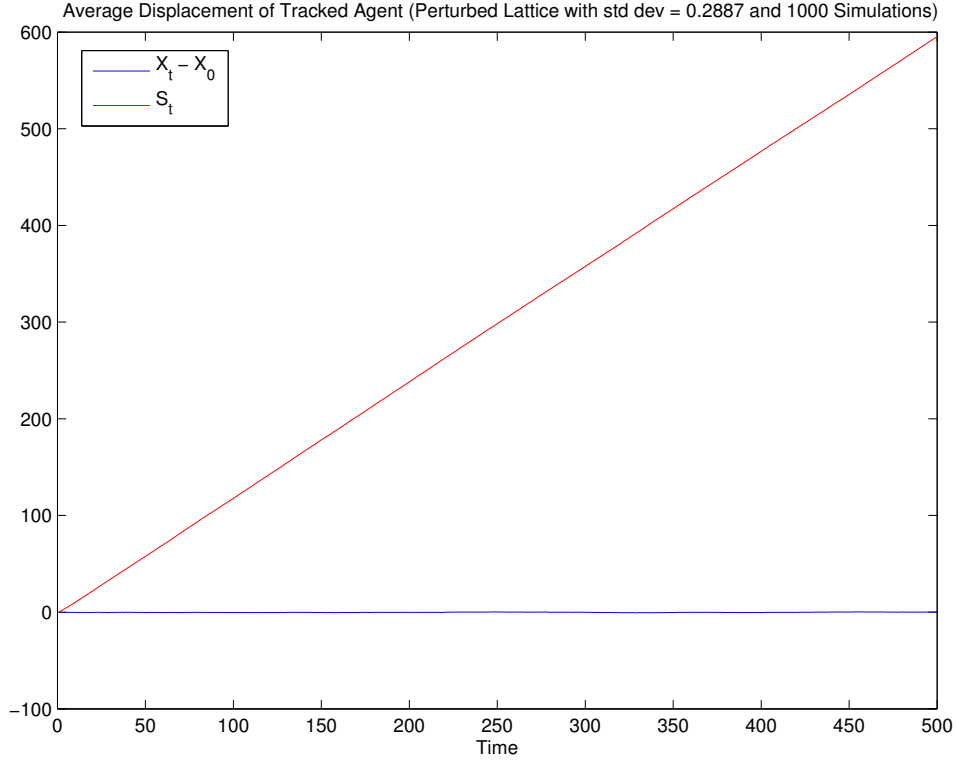
Figure 7: The progression of S_t and $X_t - X_0$ for a single agent over time averaged over 1000 perturbed lattices with $\Delta x = 1$ and perturbations sourced from (4). Linear least-squares fits were made from this data where the resultant functions were constrained to pass through $(0,0)$. The fitted gradient for S_t was 1.16 and for $X_t - X_0$ it was 0.0 (rounded to two decimal places).

the conservation of agent statement at x_i is:

$$C_i(t + \delta t) - C_i(t) = C_{i-1}(t)\frac{P}{2}\left(1 - C_i(t)\right) + C_{i+1}(t)\frac{P}{2}\left(1 - C_i(t)\right)$$
$$- C_i(t)\frac{P}{2}\left(1 - C_{i-1}(t)\right) - C_i(t)\frac{P}{2}\left(1 - C_{i+1}(t)\right) \tag{10}$$

where δt is the time step size. This statement is a simplification of Simpson's agent conservation statement in two-dimensions with support for biased motility [10]. Note carefully that due to volume exclusion $0 \leq C_i(t) \leq 1$, hence agent motility events can only occur if the designated move location is vacant. Expanding (10) gives

$$C_i(t + \delta t) - C_i(t) = \frac{P}{2}\left(C_{i-1}(t) - C_{i-1}(t)C_i(t)\right) + \frac{P}{2}\left(C_{i+1}(t) - C_{i+1}(t)C_i(t)\right)$$
$$- \frac{P}{2}\left(C_i(t) - C_i(t)C_{i-1}(t)\right) - \frac{P}{2}\left(C_i(t) - C_i(t)C_{i+1}(t)\right)$$
$$= \frac{P}{2}\left(C_{i-1}(t) - 2C_i(t) + C_{i+1}(t)\right).$$

Dividing by δt gives the standard finite difference approximation on the right-hand

Figure 8: The progression of S_t and $X_t - X_0$ for a single agent over time averaged over 1000 perturbed lattices with $\Delta x = 1$ and perturbations sourced from a Gaussian distribution with mean zero and a standard deviation of 0.2887. Linear least-squares fits were made from this data where the resultant functions were constrained to pass through $(0,0)$. The fitted gradient for S_t was 1.16 and for $X_t - X_0$ it was 0.0 (rounded to two decimal places).

side:

$$\frac{C_i(t + \delta t) - C_i(t)}{\delta t} = \frac{P}{2\delta t} \left(C_{i-1}(t) - 2C_i(t) + C_{i+1}(t) \right). \tag{11}$$

Let α_i be the difference between node i and $i - 1$ as determined by the perturbation approach. Therefore,

$$\alpha_i = \Delta x - A_i + B_i$$

where A_i and B_i independent, identically distributed random variables with an expected value of zero as defined by (4). Alternatively,

$$\alpha_i = \Delta x + r_i$$

where r_i is a random variable given by $r_i = B_i - A_i$. Let x_i be the position of ith node, now $C_{i-1}(t)$ (i.e. $C(x_{i-1}, t)$) and $C_{i+1}(t)$ (i.e. $C(x_{i+1}, t)$) can be approximated by a second Taylor series centred at x_i as follows:

$$C(x_{i-1}, t) = C_i(t) - \left. \frac{\partial C}{\partial x} \right|_{x = x_i} (\Delta x + r_{i-1}) + \frac{1}{2} \left. \frac{\partial^2 C}{\partial x^2} \right|_{x = x_i} (\Delta x + r_{i-1})^2$$
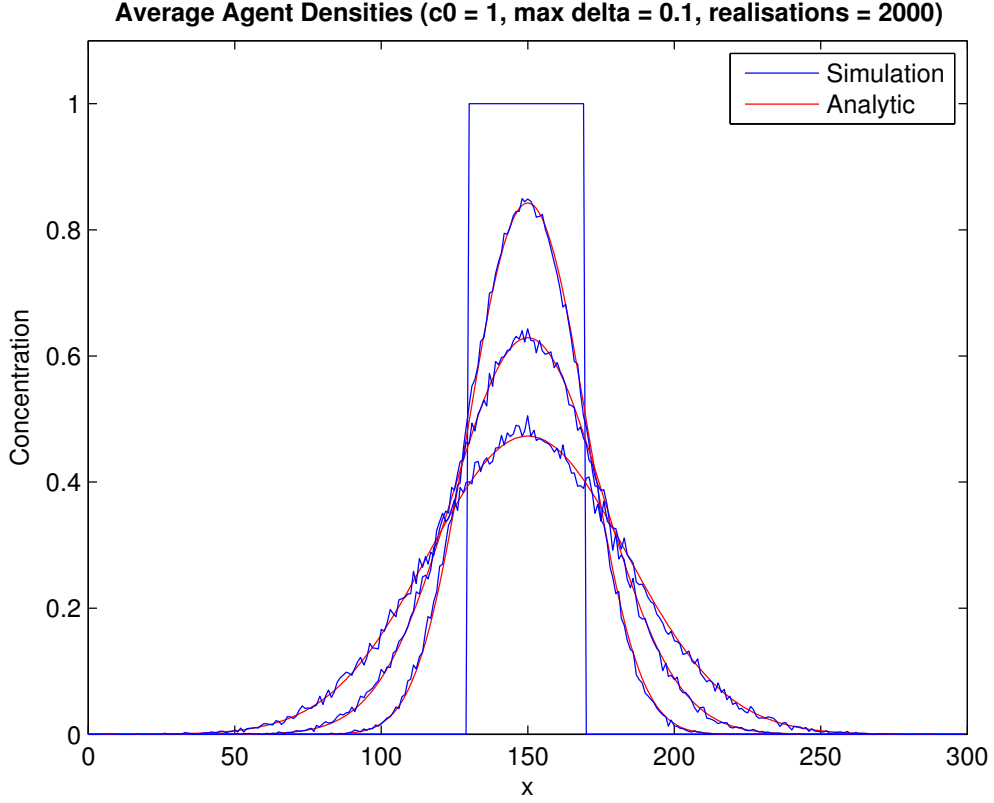$$- \mathrm{O}\left((\Delta x + r_{i-1})^3 \right)$$

13

Figure 9: Comparison of agent density between a random walk model and the diffusion equation. The model had volume exclusion with unbiased motility and no proliferation. The random walks were visualised on a random lattice generated using the perturbation approach with $\delta x = 0.1$. Density data was obtained by averaging the occupancy of each column over 2000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500,$ and 1000 and shows agent density diffusing across the lattice.

$$C(x_{i+1}, t) = C_i(t) + \left.\frac{\partial C}{\partial x}\right|_{x=x_i} (\Delta x + r_{i+1}) + \frac{1}{2}\left.\frac{\partial^2 C}{\partial x^2}\right|_{x=x_i} (\Delta x + r_{i+1})^2$$
$$+ \mathrm{O}\left((\Delta x + r_{i+1})^3\right).$$

Substituting these equations into (11) gives,

$$\frac{C_i(t+\delta t) - C_i(t)}{\delta t}\frac{2\delta t}{P} = (\Delta x + r_{i+1})\left.\frac{\partial C}{\partial x}\right|_{x=x_i} - (\Delta x + r_{i-1})\left.\frac{\partial C}{\partial x}\right|_{x=x_i}$$
$$+ \frac{1}{2}\left.\frac{\partial^2 C}{\partial x^2}\right|_{x=x_i}(\Delta x + r_{i+1})^2$$
$$+ \frac{1}{2}\left.\frac{\partial^2 C}{\partial x^2}\right|_{x=x_i}(\Delta x + r_{i-1})^2$$
$$+ \mathrm{O}\left((\Delta x + r_{i+1})^3\right) + \mathrm{O}\left((\Delta x + r_{i-1})^3\right)$$

14

Given $r_i \ll \Delta x$,

$$\frac{C_i(t+\delta t) - C_i(t)}{\delta t}\frac{2\delta t}{P} = \Delta x^2 \frac{\partial^2 C}{\partial x^2}\bigg|_{x=x_i} + \mathrm{O}\left(2\Delta x^3\right) + \mathrm{O}\left(\mathrm{Max}\left(r_i{}^2\right)\right).$$

Now, taking the limit as $\Delta x \to 0$ and $\delta t \to 0$ whilst keeping $\Delta x^2/\delta t$ constant [2], the error terms are removed and the discrete derivative in time is replaced by it's continuous analogue giving

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2}$$

where

$$D = \lim_{\Delta x, \delta t \to 0}\left(\frac{P\Delta x^2}{2\delta t}\right).$$

This governing PDE is equivalent to that derived by Simpson for equally-spaced lattices [10]. This finding shows that a random lattice does not influence simulation results provided the perturbations are sufficiently small compared to Δx. This is interesting considering that the diffusivity of a single cell is influenced by the distribution perturbations are sourced from. Any influence on the diffusivity of a single cell could be negated in the presence of volume exclusion, which was not considered in the derivation of (8). Further investigation into this finding was planned for future work.

3.2 The Random Approach

3.2.1 Lattice Properties

Equivalent simulations were conducted for the random and perturbing approaches to facilitate comparison. As such, a series of random lattices were constructed on a dimensionless 1D region of 20 dimensionless length units consisting of 20 nodes. Figure 10 depicts a series of random lattices generated by the random approach where

$$\Theta \sim \mathrm{Uniform}\left(0, 20\right). \tag{12}$$

In comparison to Figure 10, the random approach exhibits node clustering quite clearly. This is expected given that the random approach is greatly influenced by stochastic processes compared to the perturbed approach where $\Delta x \gg \mathrm{Max}\left(\Phi\right)$. Therefore, any node clustering and large deltas between adjacent nodes can be avoided by sourcing a large number of node positions in creating a fine lattice.

3.2.2 Simulating Random Walks

Using approaches similar to those applied in section 3.1.2, it can be shown that

$$\mathrm{E}\left(X_t - X_0\right) = 0$$

and

$$\mathrm{E}\left(S_t\right) = 2t\mathrm{Var}\left(\Theta\right)$$

where Θ is the random variable defining the locations the nodes. Several sets of simulations were designed and executed in an attempt to reproduce the expected results. Due to
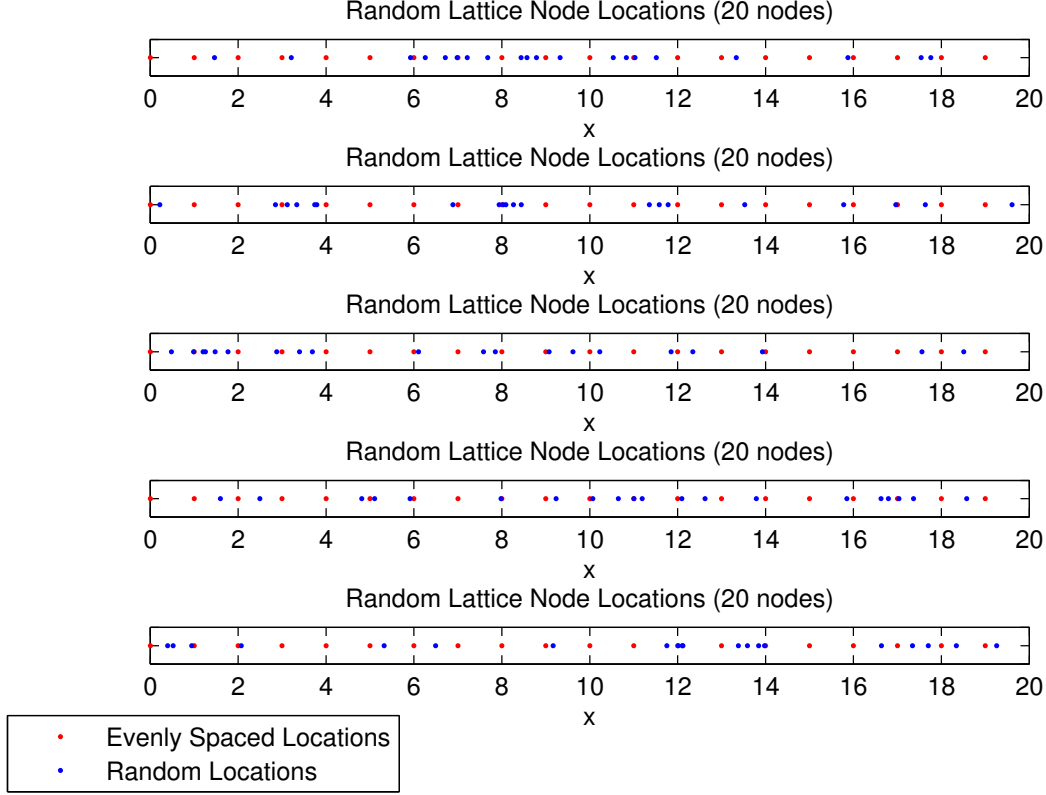
Figure 10: A series of randomly populated lattices generated by sourcing node positions from a uniform distribution spanning the region of interest as given by (12). Multiple lattices are displayed to illustrate the variability between lattices generated using this approach.

the large impact that stochastic processes have on this method it was computationally infeasible to construct simulations which would be impacted by the Law of Large Numbers. Meshes would need to be extremely fine or a large number of simulations would need to be run to compensate for the variability. As such, generating random meshes using a purely pseudorandom method is arguably unsuitable for large scale simulations or applications which do not require a fine mesh.

4 Investigating Travelling Wave Solutions

4.1 Phase Plane Analysis of Fisher's Equation

As stated in section 2.3, several PDEs were derived in Simpson's prior work in describing cellular phenomenon. For certain biological processes, variants of Fisher's equation can be used as a governing equation [10]. As such, the travelling wave solutions of Fisher's equation for population growth,

$$\frac{\partial U}{\partial t} = \nu \frac{\partial^2 U}{\partial y^2} + kU(1 - U), \tag{13}$$

16

was thoroughly analysed via [1]. To ease subsequent analysis, (13) can be nondimensionalised via the unit-reducing substitutions:

$$\tau = \alpha t \qquad \text{and} \tag{14}$$

$$x = \beta y \tag{15}$$

where α and β have units t^{-1} and m^{-1}, respectively. Applying the substitution in (14) and the chain rule, we obtain:

$$\frac{\partial U}{\partial \tau}\frac{d\tau}{dt} = \nu\frac{\partial^2 U}{\partial y^2} + kU(1-U)$$

$$\alpha\frac{\partial U}{\partial \tau} = \nu\frac{\partial^2 U}{\partial y^2} + kU(1-U)$$

By inspection, it is clear that (13) can be nondimensionalised in time and simplified when $\alpha = k$. Hence,

$$\frac{\partial U}{\partial \tau} = \frac{\nu}{k}\frac{\partial^2 U}{\partial y^2} + U(1-U) \tag{16}$$

Applying (15) in a similar manner gives

$$\frac{\partial U}{\partial \tau} = \frac{\beta^2\nu}{k}\frac{\partial^2 U}{\partial x^2} + U(1-U).$$

As such, (16) can be nondimensionalised in space and simplified when $\beta = \sqrt{k/\nu}$, which results in

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2} + U(1-U). \tag{17}$$

As discussed by Canosa [1], this nondimensionalised equation describes nonlinear population growth in one spatial dimension. Consequently, the modelled habitat (i.e. space) can only feasibly support a maximum population per unit length. Given the form of the logistic source term in (17), the maximum population is unity. The following initial condition can therefore be applied:

$$0 \leq U(x,0) \leq 1, \qquad -\infty < x < \infty. \tag{18}$$

Canosa [1] was specifically interested in solving (17) subject to (18) such that all the derivatives in x tended to zero as $x \to \pm\infty$ coupled with the following far-field conditions:

$$\lim_{x\to-\infty} U(x,\tau) = 1 \qquad \lim_{x\to\infty} U(x,\tau) = 0 \qquad t \geq 0. \tag{19}$$

Any one-dimensional PDE in space and time, $U(x,\tau)$, exhibiting travelling wave solutions supports the transformation

$$U(x,t) = U(x - c\tau) \equiv U(s) \tag{20}$$

which merges both time and space variables to provide a "standing wave" solution where the spatial origin is locked at the peak of the travelling wave moving at speed c. As such, the spatial coordinate system can be seen to translate in time in moving with the peak of the wave. Applying (20) with the chain rule gives

$$-c\frac{dU}{ds} = \frac{d^2 U}{ds^2} + U(1-U)$$

17

which has reduced Fisher's equation into a second order non-linear PDE. This equation can be further simplified into the following system of first order differential equations:

$$\frac{dU}{ds} = V(s) \tag{21}$$

$$\frac{dV}{ds} = U^2 - U - cV. \tag{22}$$

Phase plane analysis was conducted on (21) and (22) using pplane8 for MATLAB [6].

Fisher originally discovered that (13) has travelling wave solutions where $c \geq 2$ [1]. This fact can be verified by solving the eigenvalue problem associated with (21) and (22) to determine the nature of any critical points in the phase plane. Given $c \geq 2$ it can be shown that, $(U, V) = (0, 0)$ is a stable node and $(U, V) = (1, 0)$ is a saddle point. Conversely, given $c < 2$, $(U, V) = (0, 0)$ is a stable spiral. The phase planes seen in Figures 11 and 12 illustrate these results for specific values of c. Figure 11 exhibits the physically impossible solutions for $c < 2$. Note how certain phase trajectories can lead to negative values of U, which is physically impractical in modelling population growth.
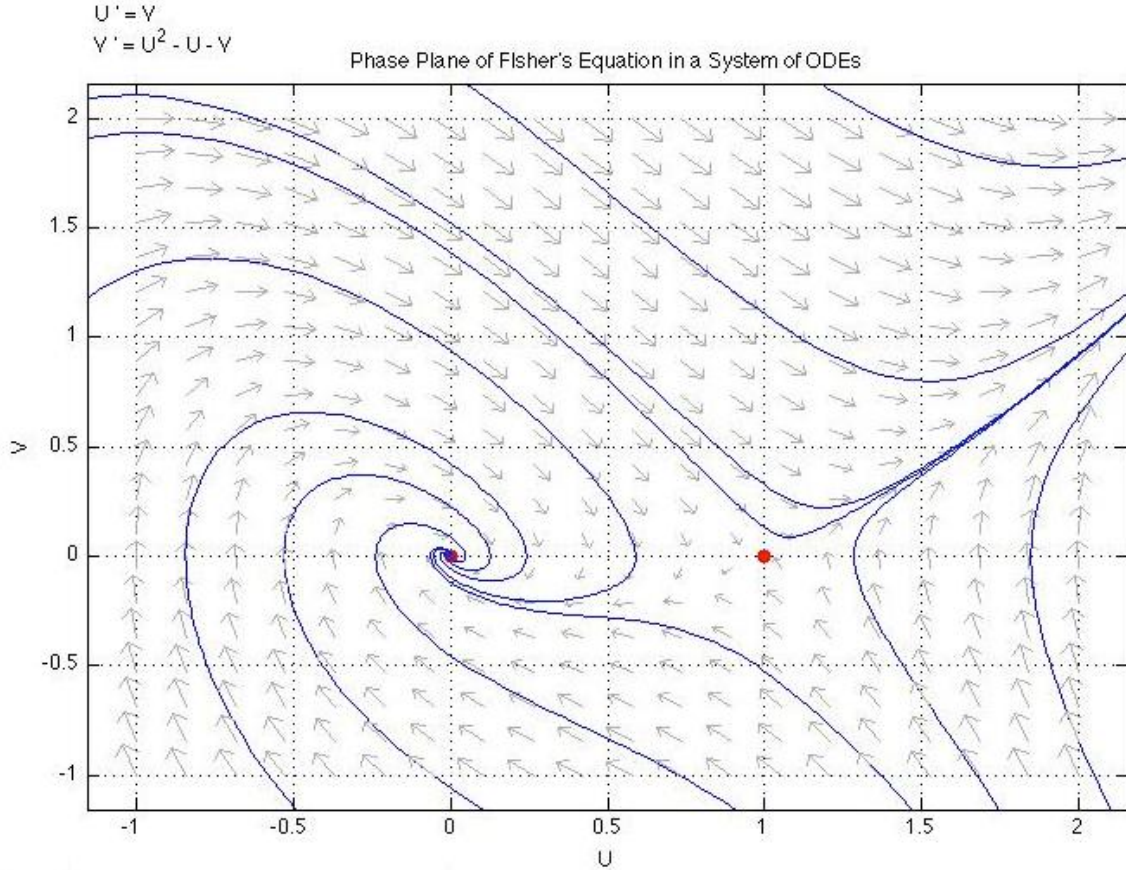
Figure 11: Phase plane of the system of ODEs that describe the travelling wave solutions to Fisher's equation, where $c = 1$. The red dots are the critical points and the blue lines are the phase trajectories. At $(0, 0)$ there is a spiral sink and at $(1, 0)$ there is a saddle point.
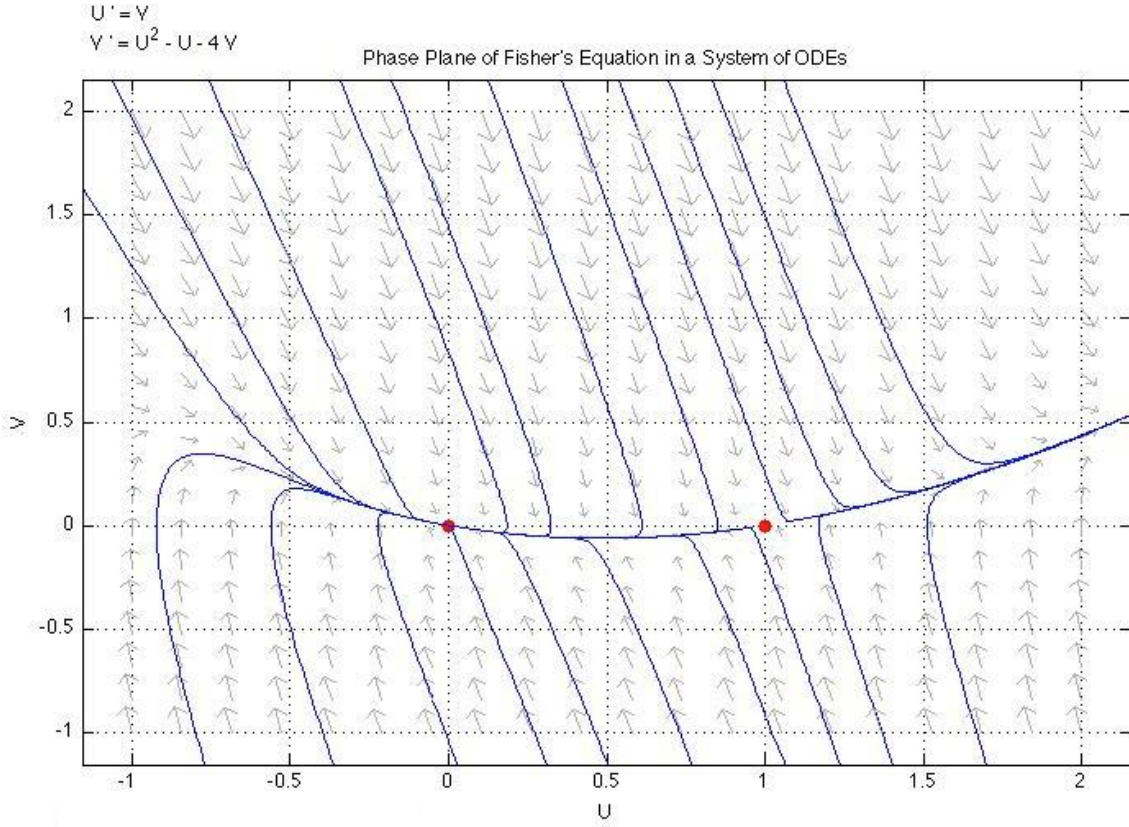
18

Figure 12: Phase plane of the system of ODEs that describe the travelling wave solutions to Fisher's equation, where $c = 4$. The red dots are the critical points and the blue lines are the phase trajectories. At $(0,0)$ there is a stable node and at $(1,0)$ there is a saddle point.

4.2 Solving Fisher's Equation

In order to better understand the travelling wave solutions to Fisher's equation, the following IBVP was solved numerically using MATLAB's pdepe function:

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2} + U(1 - U) \qquad 0 \le x \le 100 \qquad t \ge 0$$
$$U(0, \tau) = 1 \qquad\qquad\qquad \tau \ge 0$$
$$U(100, \tau) = 0 \qquad\qquad\qquad \tau \ge 0$$
$$U(x, 0) = \begin{cases} 1 & x \le 50 \\ 0 & \text{otherwise} \end{cases} \qquad 0 \le x \le 100.$$

The solutions, for varying timesteps, can be seen in Figure 13. From the figure it is clear that a static wavefront is formed which moves throughout space at a constant speed. This wavefront does not hold at the boundaries in order to satisfy the conditions imposed there. As a side note, the solutions did exhibit monotonicity issues. After discussion with Simpson, it was determined that this is due to MATLAB's pdepe function using the method of lines, which can result in physically unrealistic solutions.

19

Solutions to Fisher's equation were investigated for various initial conditions to observe how the logistic source term reacts to overpopulated and underpopulated regions. As expected, provided $U \geq 0$, the solutions continue to form a travelling wavefront. This is due to the logistic source term compensating for regions which do not satisfy the carrying capacity of unity. Likewise, in the event that $U < 0$, the solutions diverge to $-\infty$. It is clear that the logistic source term plays an integral role in forming travelling wave solutions of a specific amplitude.
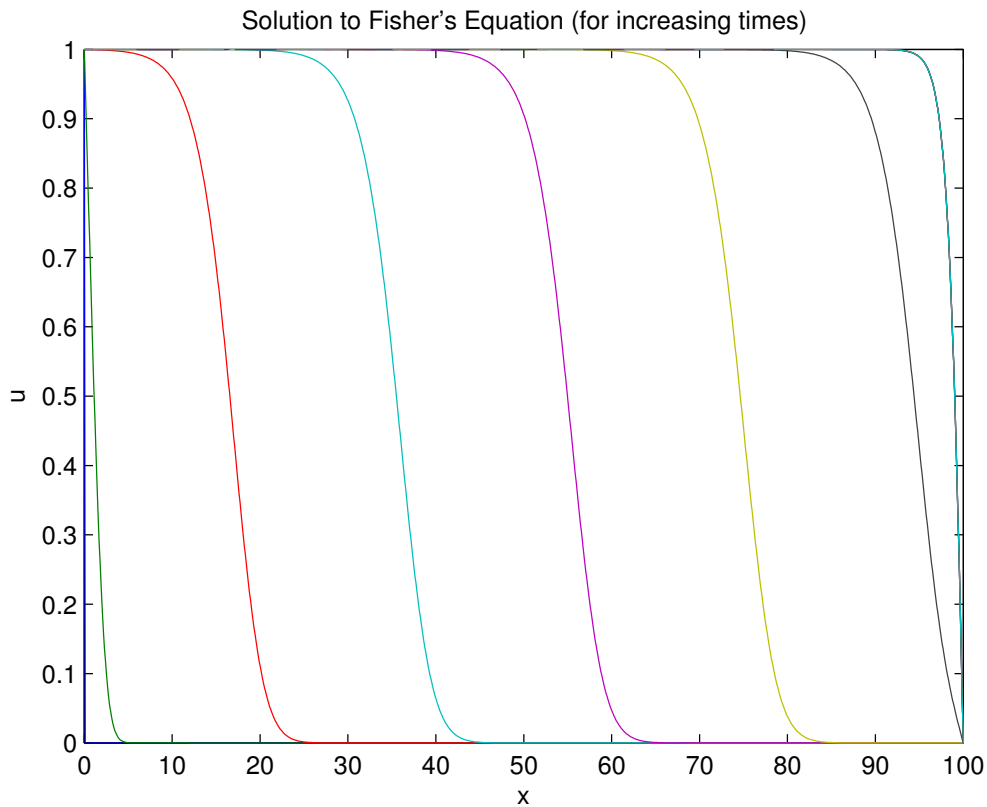
Figure 13: Travelling wave solutions to Fisher's equation for increasing times where $U(0,t) = 1$ and $U(100,t) = 0$. Note carefully that steep wave fronts form at the boundaries to ensure that the conditions imposed there are met.

References

[1] José Canosa. "On a nonlinear diffusion equation describing population growth". In: *IBM Journal of Research and Development* 17.4 (1973), pp. 307–313.

[2] Edward A Codling, Michael J Plank, and Simon Benhamou. "Random walk models in biology". In: *Journal of the Royal Society Interface* 5.25 (2008), pp. 813–834.

[3] Charles M Grinstead and James L Snell. *Introduction to Probability*. American Mathematical Society, 1997.

[4] *Periodic Boundary Conditions*. Virtual Institute of Nano Films. 2012. URL: `http://www.mbnexplorer.com/users-guide/6-numerical-methods-interatomic-interactions/63-boundary-conditions/632-periodic`.

[5] Michael J Plank and Matthew J Simpson. "Lattice-Free Models of Cell Invasion: Discrete Simulations and Travelling Waves". In: *Bulletin of mathematical biology* 75.11 (2013), pp. 2150–2166.

[6] John C Polking. *ODE Software for MATLAB*. Rice University. Apr. 2002. URL: `http://math.rice.edu/~dfield/`.

[7] *Reflective Boundary Conditions*. Virtual Institute of Nano Films. 2012. URL: `http://www.mbnexplorer.com/users-guide/6-numerical-methods-interatomic-interactions/63-boundary-conditions/631-reflective`.

[8] John Renze and Eric W Weisstein. *Law of Large Numbers*. MathWorld–A Wolfram Web Resource. URL: `http://mathworld.wolfram.com/LawofLargeNumbers.html`.

[9] Matthew J Simpson, Kerry A Landman, and Barry D Hughes. "Cell invasion with proliferation mechanisms motivated by time-lapse data". In: *Physica A: Statistical Mechanics and its Applications* 389.18 (2010), pp. 3779–3790.

[10] Matthew Simpson, K Landman, and B Hughes. "Diffusing populations: ghosts or folks". In: *Australasian Journal of Engineering Education* 15.2 (2009), pp. 59–68.

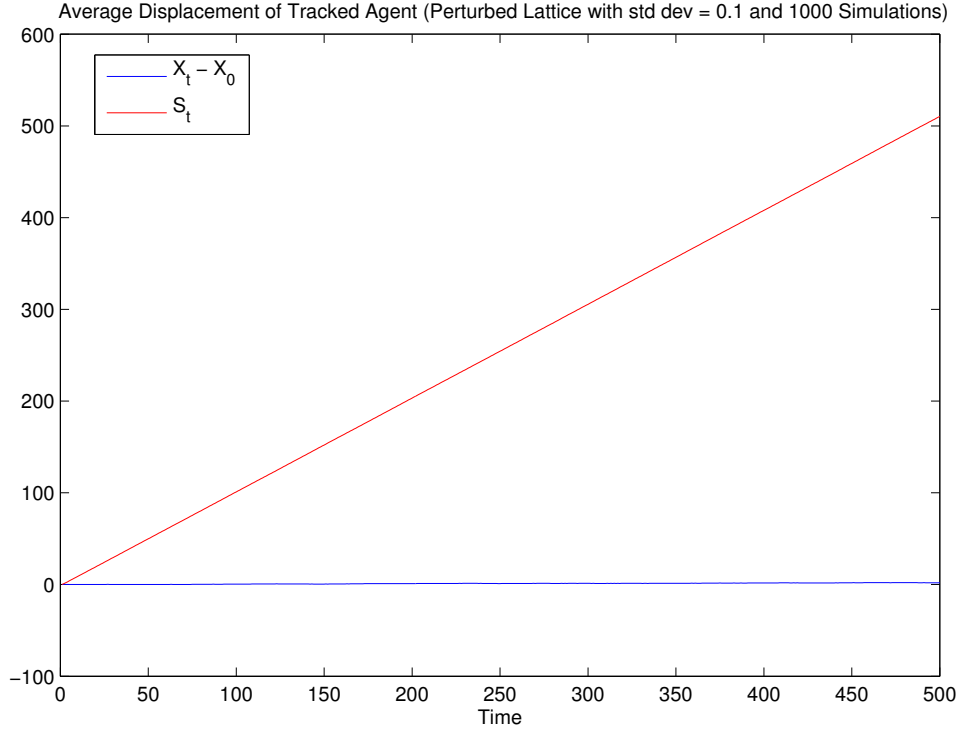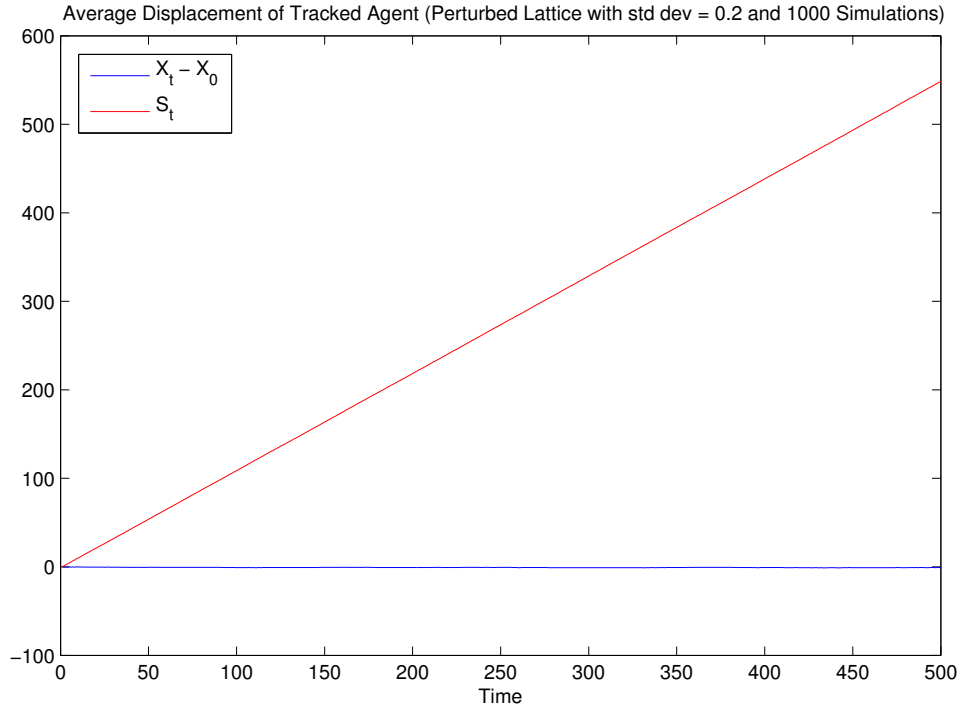# A   Additional Mean Squared Displacement Results

Figure A.1: The progression of S_t and $X_t - X_0$ for a single agent over time on a perturbed lattice averaged over 1000 identically-prepared random walks. The perturbed lattice had $\Delta x = 1$ and perturbations were sourced from a Gaussian distribution with mean zero and a standard deviation of 0.1. Linear least-squares fits were made from this data where the resultant functions were constrained to pass through $(0, 0)$. The fitted gradient for S_t was 1.02 and for $X_t - X_0$ it was 0.0 (rounded to two decimal places).
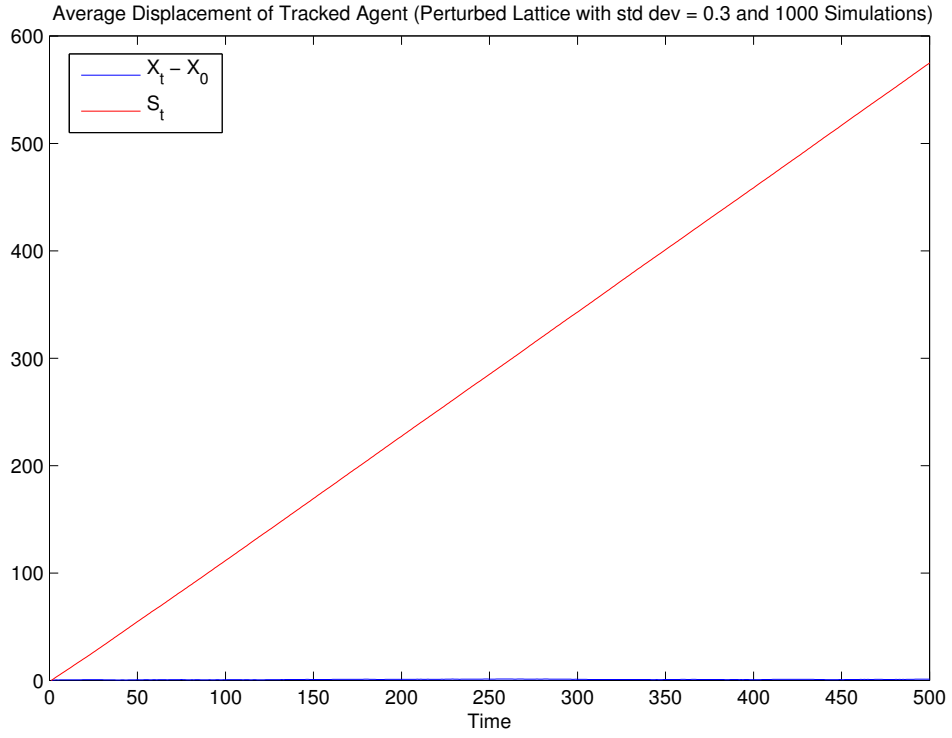
Figure A.2: The progression of S_t and $X_t - X_0$ for a single agent over time on a perturbed lattice averaged over 1000 identically-prepared random walks. The perturbed lattice had $\Delta x = 1$ and perturbations were sourced from a Gaussian distribution with mean zero and a standard deviation of 0.2. Linear least-squares fits were made from this data where the resultant functions were constrained to pass through $(0,0)$. The fitted gradient for S_t was 1.10 and for $X_t - X_0$ it was 0.0 (rounded to two decimal places).
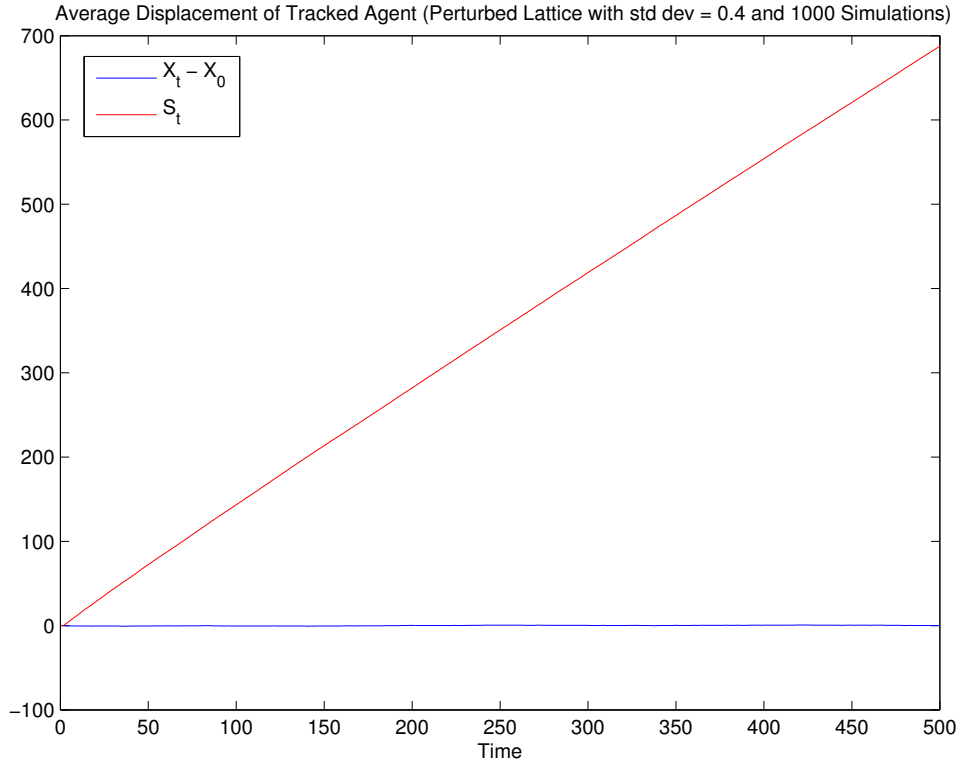
Figure A.3: The progression of S_t and $X_t - X_0$ for a single agent over time on a perturbed lattice averaged over 1000 identically-prepared random walks. The perturbed lattice had $\Delta x = 1$ and perturbations were sourced from a Gaussian distribution with mean zero and a standard deviation of 0.3. Linear least-squares fits were made from this data where the resultant functions were constrained to pass through $(0, 0)$. The fitted gradient for S_t was 1.15 and for $X_t - X_0$ it was 0.0 (rounded to two decimal places).

Figure A.4: The progression of S_t and $X_t - X_0$ for a single agent over time on a perturbed lattice averaged over 1000 identically-prepared random walks. The perturbed lattice had $\Delta x = 1$ and perturbations were sourced from a Gaussian distribution with mean zero and a standard deviation of 0.4. Linear least-squares fits were made from this data where the resultant functions were constrained to pass through $(0,0)$. The fitted gradient for S_t was 1.39 and for $X_t - X_0$ it was 0.0 (rounded to two decimal places).

Figure A.5: The progression of S_t and $X_t - X_0$ for a single agent over time on a perturbed lattice averaged over 1000 identically-prepared random walks. The perturbed lattice had $\Delta x = 1$ and perturbations were sourced from a Gaussian distribution with mean zero and a standard deviation of 0.5. Linear least-squares fits were made from this data where the resultant functions were constrained to pass through $(0,0)$. The fitted gradient for S_t was 1.29 and for $X_t - X_0$ it was 0.0 (rounded to two decimal places).

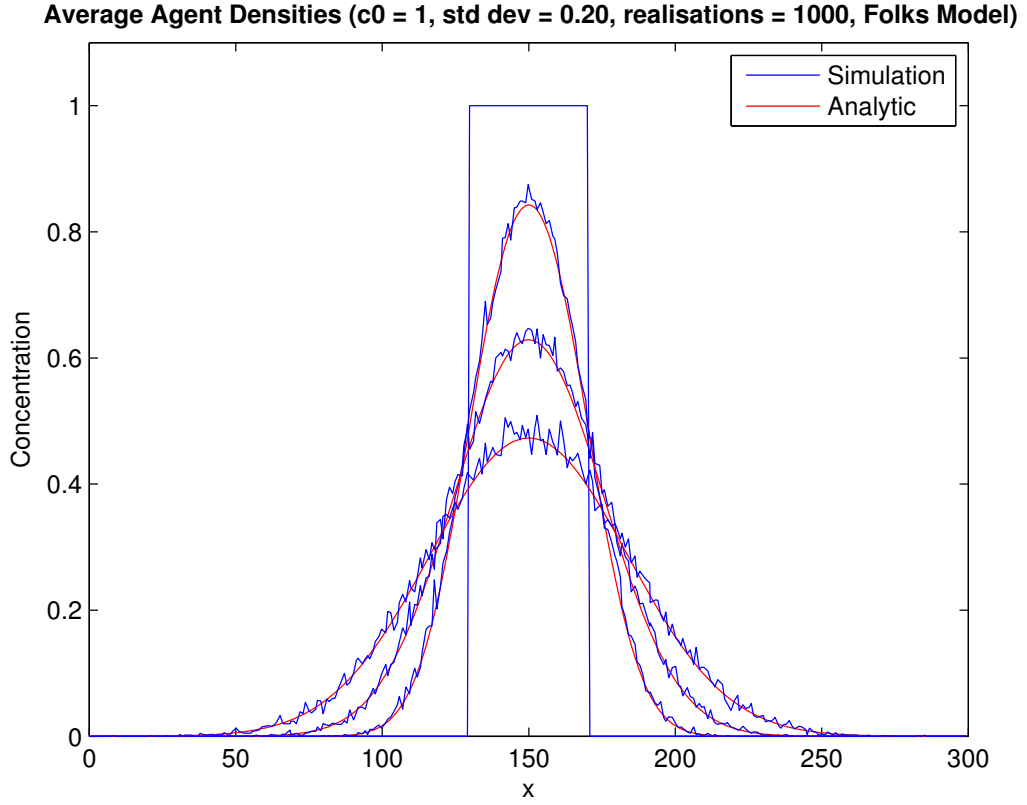# B   Additional Average Population-Densities Results
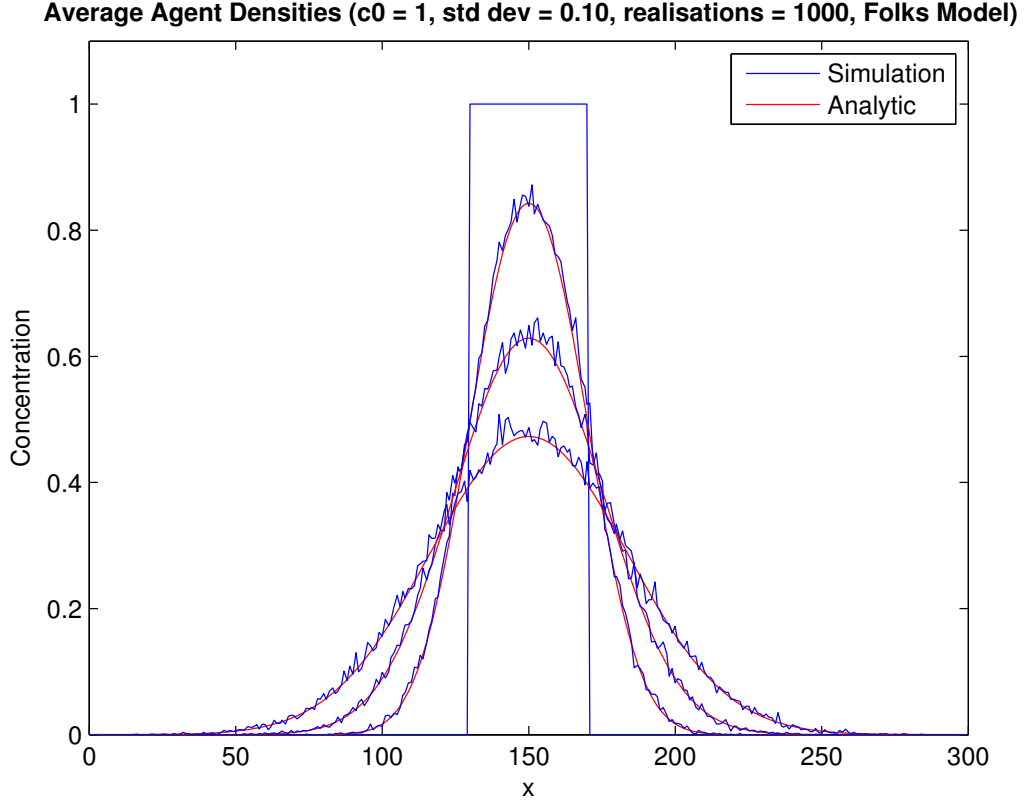
Figure B.1: Comparison of agent density between a random walk model and the diffusion equation. The model had volume exclusion with unbiased motility and no proliferation. The random walks were visualised on a random lattice generated using the perturbation approach where perturbations were sourced from a Gaussian distribution with standard deviation of 0.1. Density data was obtained by averaging the occupancy of each column over 1000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500$, and 1000 and shows agent density diffusing across the lattice.
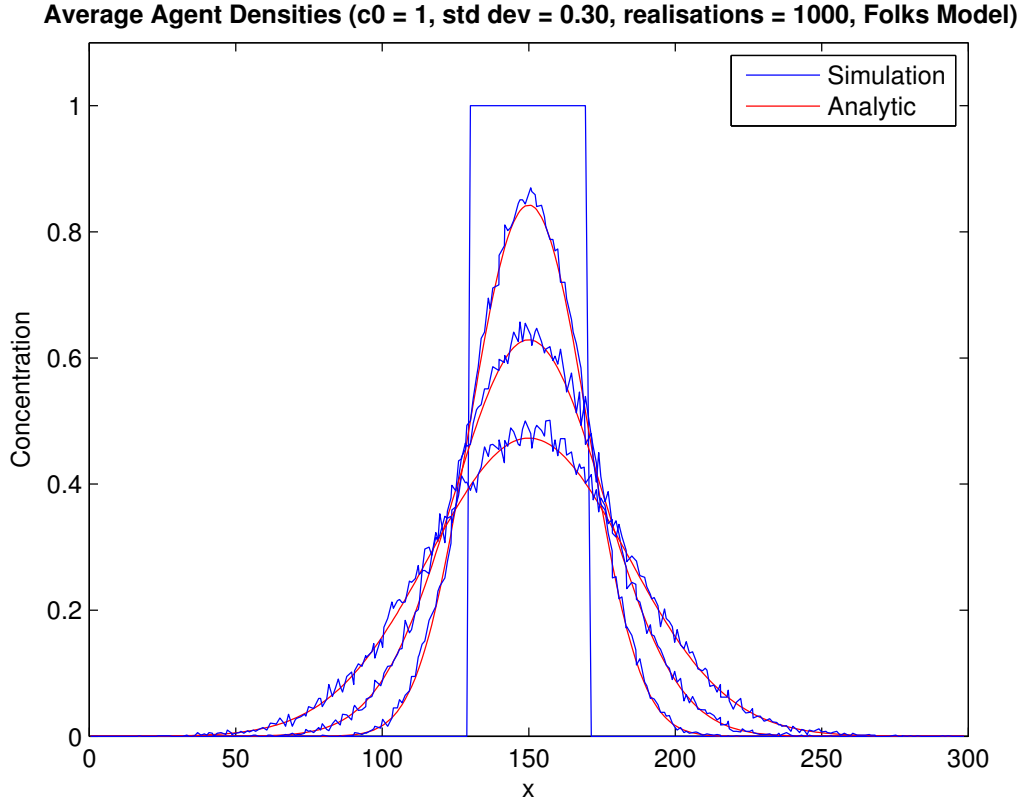
Figure B.2: Comparison of agent density between a random walk model and the diffusion equation. The model had volume exclusion with unbiased motility and no proliferation. The random walks were visualised on a random lattice generated using the perturbation approach where perturbations were sourced from a Gaussian distribution with standard deviation of 0.2. Density data was obtained by averaging the occupancy of each column over 1000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500$, and 1000 and shows agent density diffusing across the lattice.

Average Agent Densities (c0 = 1, std dev = 0.30, realisations = 1000, Folks Model)
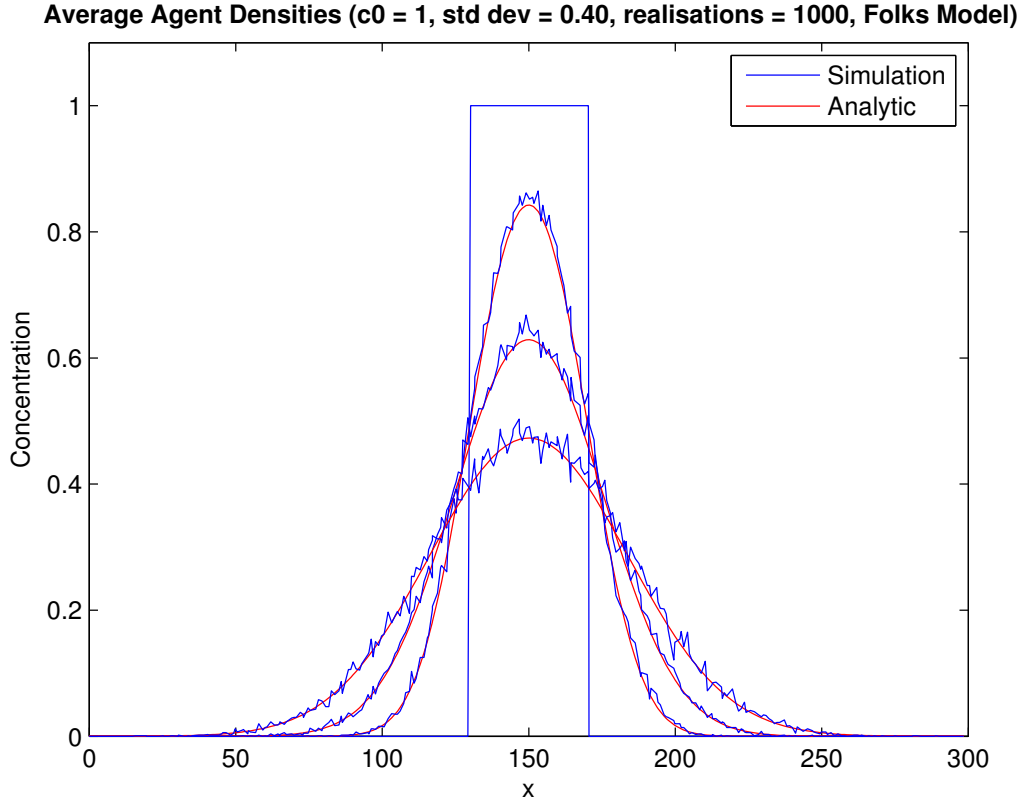
Figure B.3: Comparison of agent density between a random walk model and the diffusion equation. The model had volume exclusion with unbiased motility and no proliferation. The random walks were visualised on a random lattice generated using the perturbation approach where perturbations were sourced from a Gaussian distribution with standard deviation of 0.3. Density data was obtained by averaging the occupancy of each column over 1000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500$, and 1000 and shows agent density diffusing across the lattice.
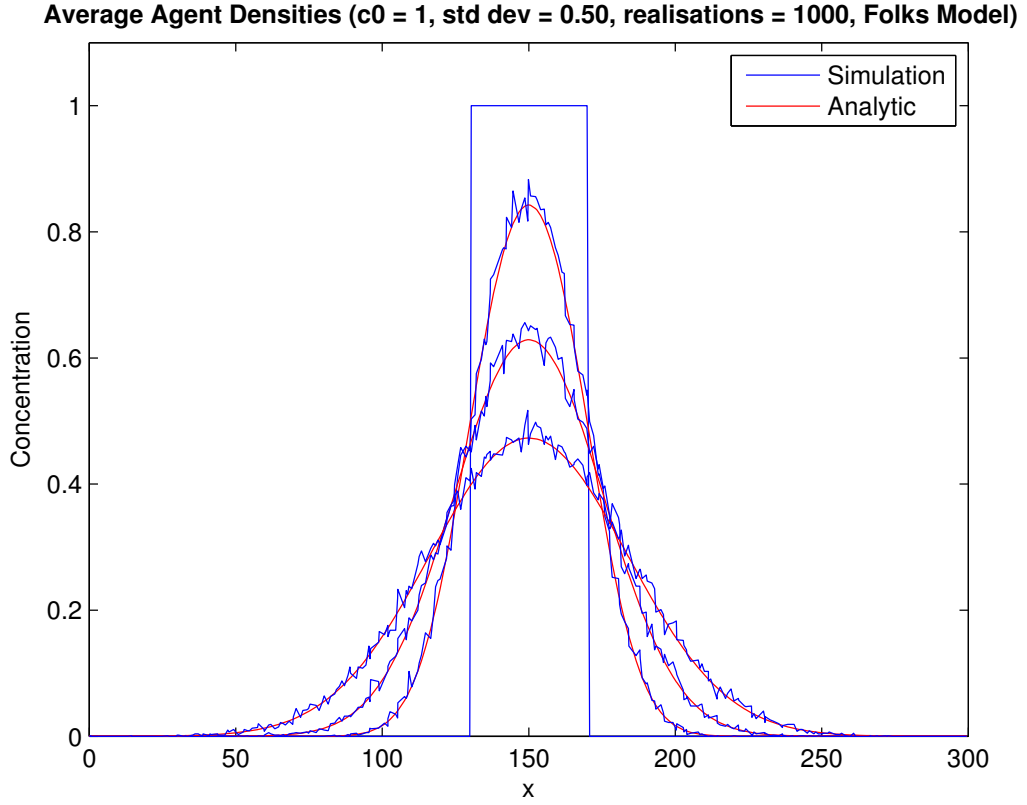
Figure B.4: Comparison of agent density between a random walk model and the diffusion equation. The model had volume exclusion with unbiased motility and no proliferation. The random walks were visualised on a random lattice generated using the perturbation approach where perturbations were sourced from a Gaussian distribution with standard deviation of 0.4. Density data was obtained by averaging the occupancy of each column over 1000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500$, and 1000 and shows agent density diffusing across the lattice.

Figure B.5: Comparison of agent density between a random walk model and the diffusion equation. The model had volume exclusion with unbiased motility and no proliferation. The random walks were visualised on a random lattice generated using the perturbation approach where perturbations were sourced from a Gaussian distribution with standard deviation of 0.5. Density data was obtained by averaging the occupancy of each column over 1000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500$, and 1000 and shows agent density diffusing across the lattice.

C   Additional Average Population-Densities Results (Ghosts Model)

Average Agent Densities (c0 = 1, std dev = 0.10, realisations = 1000, Ghosts Model)
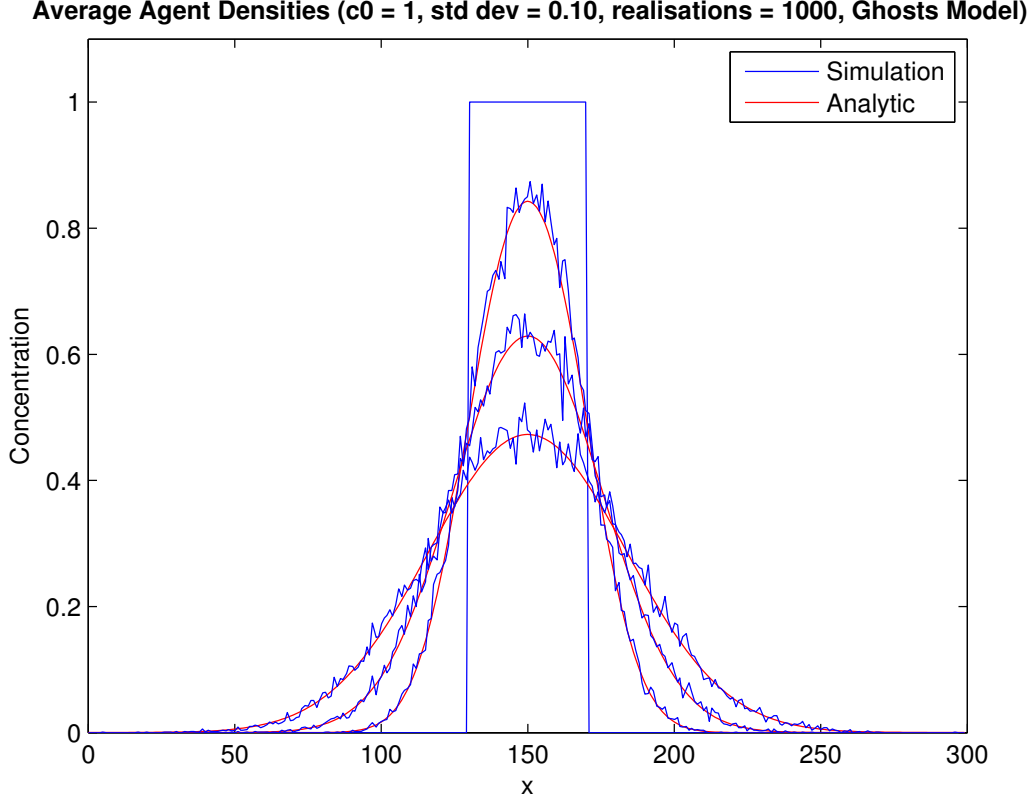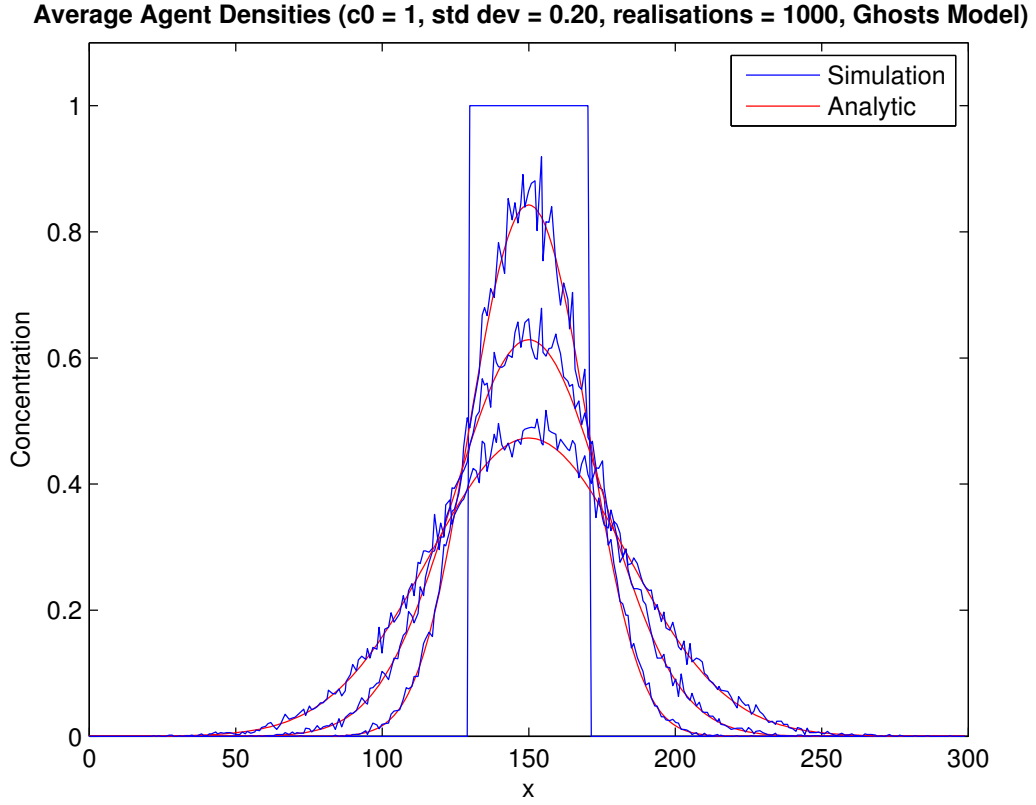

Figure C.1: Comparison of agent density between a random walk model and the diffusion equation. The model had unbiased motility without proliferation and volume exclusion. The random walks were visualised on a random lattice generated using the perturbation approach where perturbations were sourced from a Gaussian distribution with standard deviation of 0.1. Density data was obtained by averaging the occupancy of each column over 1000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500$, and 1000 and shows agent density diffusing across the lattice.

Figure C.2: Comparison of agent density between a random walk model and the diffusion equation. The model had unbiased motility without proliferation and volume exclusion. The random walks were visualised on a random lattice generated using the perturbation approach where perturbations were sourced from a Gaussian distribution with standard deviation of 0.2. Density data was obtained by averaging the occupancy of each column over 1000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500,$ and 1000 and shows agent density diffusing across the lattice.
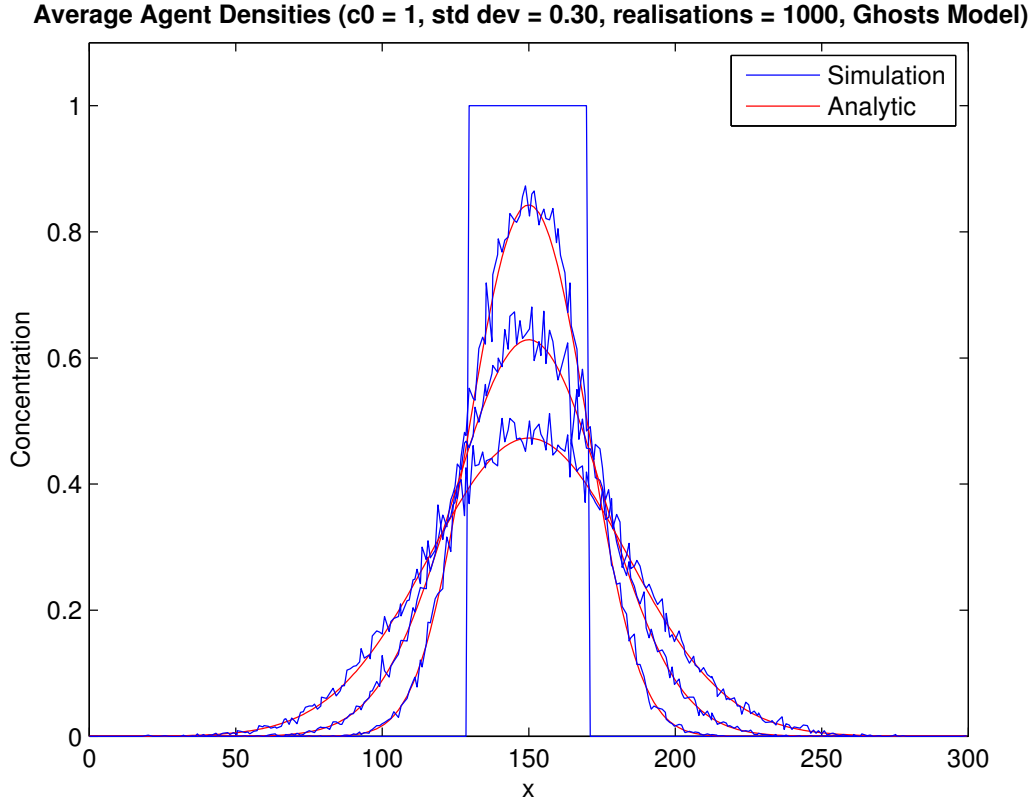
Figure C.3: Comparison of agent density between a random walk model and the diffusion equation. The model had unbiased motility without proliferation and volume exclusion. The random walks were visualised on a random lattice generated using the perturbation approach where perturbations were sourced from a Gaussian distribution with standard deviation of 0.3. Density data was obtained by averaging the occupancy of each column over 1000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500$, and 1000 and shows agent density diffusing across the lattice.
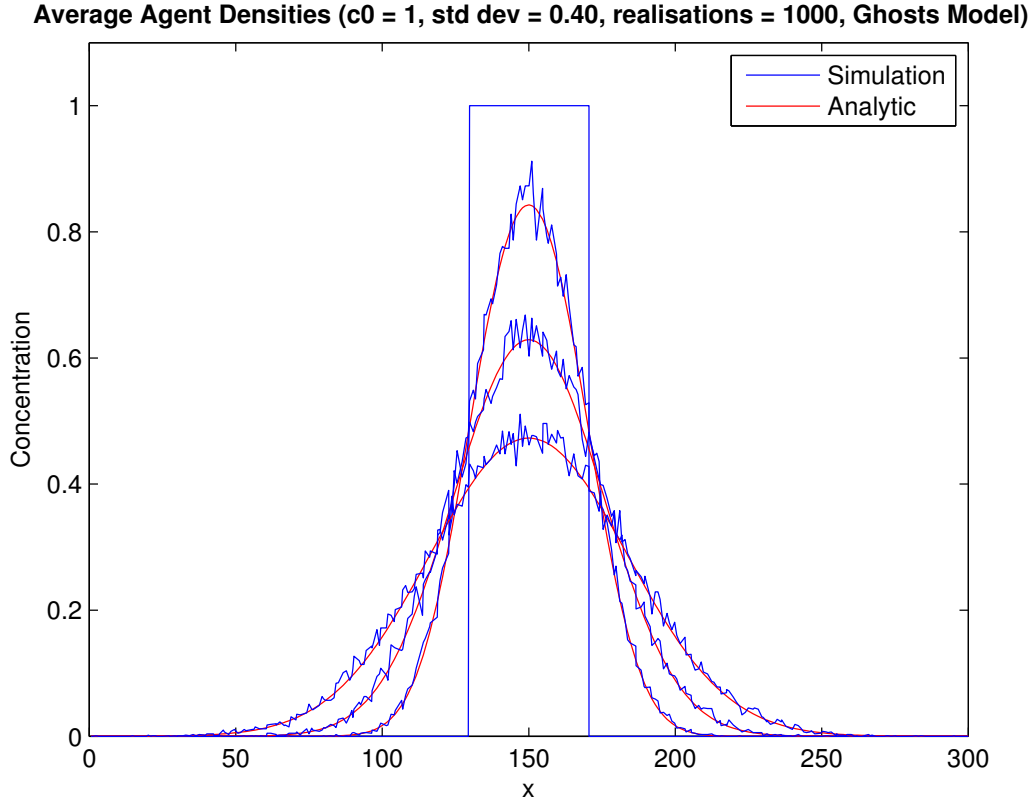
Figure C.4: Comparison of agent density between a random walk model and the diffusion equation. The model had unbiased motility without proliferation and volume exclusion. The random walks were visualised on a random lattice generated using the perturbation approach where perturbations were sourced from a Gaussian distribution with standard deviation of 0.4. Density data was obtained by averaging the occupancy of each column over 1000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500$, and 1000 and shows agent density diffusing across the lattice.
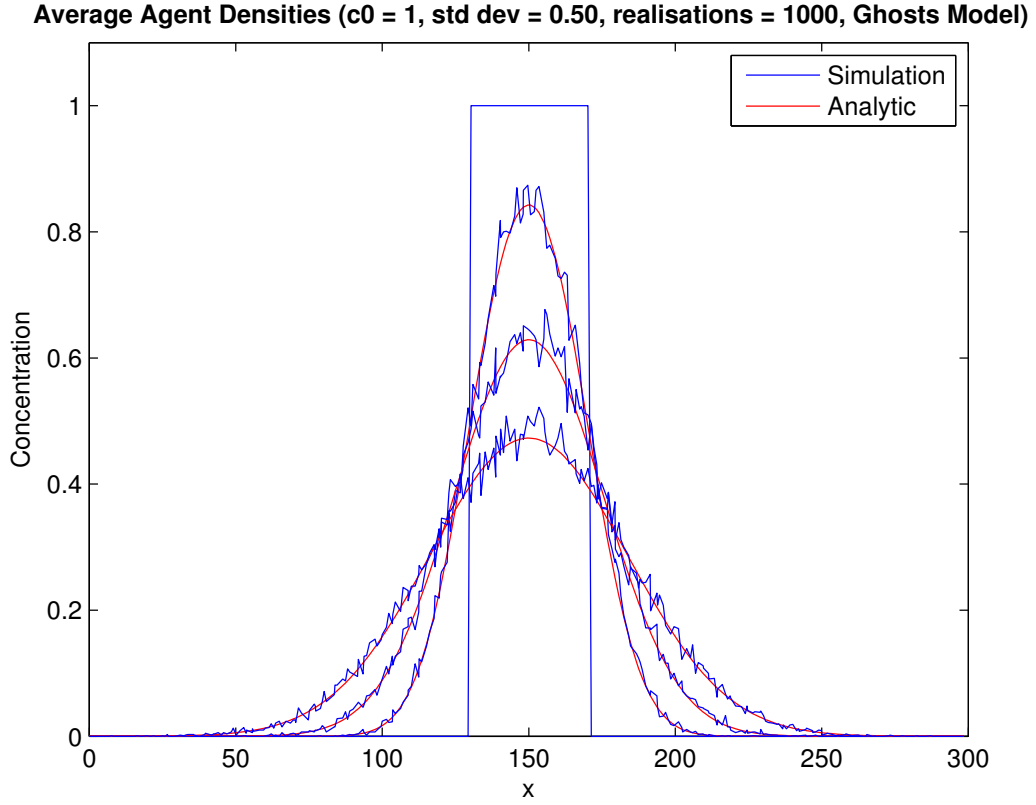
Figure C.5: Comparison of agent density between a random walk model and the diffusion equation. The model had unbiased motility without proliferation and volume exclusion. The random walks were visualised on a random lattice generated using the perturbation approach where perturbations were sourced from a Gaussian distribution with standard deviation of 0.5. Density data was obtained by averaging the occupancy of each column over 1000 identically-prepared realisations satisfying (2) where $C_0 = 1$. Densities are given at $t = 0, 200, 500,$ and 1000 and shows agent density diffusing across the lattice.