

# Stochastic Modelling of Lymphocyte Dynamics and Interactions

Mark Stephen Day

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds  
Department of Applied Mathematics

May 2013

The candidate confirms that the work submitted is his own, except where work which has formed part of a jointly-authored publication has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms appropriate credit has been given within this thesis where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

©2013 The University of Leeds and Mark Day

# Joint Publications

Work from the following jointly-authored publications is included in this thesis:

- Susanna Celli, Mark Day, Andreas Müller, Carmen Molina-París, Grant Lythe and Philippe Bousso. *How many dendritic cells are required to initiate a T cell response?* *Blood* **120**:3945. (2012).

All the experimental work, described in section 6.2 of this thesis, is the work of Dr. Susanna Celli, Dr. Andreas Müller and Dr. Philippe Bousso. Mathematical models in the paper were developed with the guidance of Dr. Grant Lythe and Prof. Carmen Molina-París. The main body of the paper was written by Dr. Philippe Bousso. The first draft of the appendix was written by Mark Day, and the final draft was written by Dr. Grant Lythe.

Chapter 6 of this thesis is based on this paper.

- Day M, Lythe G. *Timescales of the adaptive immune response*. In: Molina-París C, Lythe G, editors. *Mathematical Models and Immune Cell Biology*. Springer; 2011.

The work in this book chapter is the joint work of Mark Day and Dr. Grant Lythe. The first draft of the chapter was written by Mark Day, and the final draft was written by Dr. Grant Lythe.

Chapter 3 uses concepts introduced in this book chapter.

# Acknowledgements

First and foremost, I thank my supervisors Grant Lythe and Carmen Molina-París for their continued support and guidance over the last 5 years of my PhD and Master research. It has been a privilege to work with them and other members of the mathematical immunology group.

I thank my collaborators Susanna Celli and Philippe Bousso at the Institut Pasteur, Paris, for the experimental work they carried out to aid our joint work, and answering all my questions regarding it.

Further thanks must go to Ross McQueenie and Pasquale Maffia at the University of Glasgow, for providing data sets from cell imaging experiments for me to analyse, and to Daniel Coombs for pointing me in the right direction with my work on first passage processes.

In addition, I thank my Father and the rest of my family for the support they have provided over the course of my PhD and throughout my life, which has been invaluable.

I thank my friends for providing a fun environment in which to relax, and the other students in my office for providing a constructive environment to work in.

I am also thankful for the opportunities I have been given to attend, and present at, conferences around the world. This would not have been possible without funding provided by EPSRC and a FP7 European grant. I am also thankful for the PhD scholarship provided to me by the school of mathematics, from the EPSRC.

# Abstract

The immune system protects the body against invading pathogens. For an immune response to occur, a T cell must encounter a rare antigen-presenting-cell (APC) presenting its cognate antigen. The time it takes for this encounter depends upon how quickly the T cell is moving, as well as how many APCs carrying the T cells cognate antigen are present.

First passage processes are used to derive an equation for the encounter time of a T cell with one of  $N$  APCs. Using this time, a rate of encounter is established, and used throughout this thesis. The encounter rate is dependent upon the radius of the lymph node, the effective radius of the APC, and the diffusivity of the T cell. However, the diffusivity of T cells has not been clearly established. *In vivo* imaging data is used to develop a systematic method for determining the diffusivity of a population of T cells. Due to *in vivo* imaging experiments having a limited sized imaging volume, a confinement effect is observed. The expected squared displacement of imaged cells is calculated, and the level at which a confinement plateau should be observed is determined.

T cell activation, in lymph nodes, relies upon encounters with APCs, but the number of APCs required to initiate a T cell response is currently unknown. Using mathematical models, in combination with experimental work, the probability of T cell-APC encounters can be quantified. The probability of a T cell, residing in the lymph node for twenty four hours, to interact with APCs is calculated. Extrapolating the developed models to later times and lower cell numbers than can be achieved experimentally, a minimum number of APCs required to initiate a T cell response, for typical human T cell precursor frequencies, is estimated.

It has been proposed that regulatory T cells suppress effector T cells via a three way interaction with APCs, as a method of preventing autoimmunity. A stochastic model of these interactions is developed and explored. The steady state of the system is found to depend upon the rate of encounter of T cells and APCs, as well as the number of APCs. Stochastic effects are observed in the model, which affect the state of the system, and are not observed in a deterministic approach.

Interactions between T cells and APCs, in lymph nodes, are crucial for initiating cell-mediated adaptive immune responses. However, how these interactions cause activation of the T cells is not yet fully understood. Three hypotheses have been proposed for the method of T cell activation. These hypotheses are investigated, and models developed, in an attempt to quantify the observed stages of the activation process. It is found that experimental results can, in part, be explained by a probabilistic approach.



# Contents

Joint Publications . . . . .	iii
Acknowledgements . . . . .	iv
Abstract . . . . .	v
Contents . . . . .	vii
List of figures . . . . .	xiv
List of tables . . . . .	xx
List of abbreviations . . . . .	xxii
<b>1 Immunology introduction</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 The human immune system . . . . .	1
1.3 Dendritic cells . . . . .	2
1.4 T cells . . . . .	4
1.5 Lymph nodes . . . . .	6
1.6 T cell selection . . . . .	8
1.7 Regulatory T cells . . . . .	8
1.8 Experimental techniques . . . . .	9
1.8.1 Two-photon laser scanning microscopy . . . . .	9
1.8.2 Flow cytometry . . . . .	11
1.9 Objectives of the thesis . . . . .	12

<b>2</b>	<b>Introduction to stochastic spatial modelling and numerical methods</b>	<b>13</b>
2.1	Introduction . . . . .	13
2.2	Spatial modelling . . . . .	14
2.2.1	How do T cells move? . . . . .	14
2.2.2	The Wiener process . . . . .	16
2.2.3	Computational algorithm . . . . .	17
2.3	Error reduction algorithms . . . . .	21
2.3.1	Target encounters . . . . .	21
2.3.2	Boundary collisions . . . . .	23
2.3.3	Exponential timestepping . . . . .	24
2.4	Gillespie algorithm . . . . .	26
2.4.1	Physical basis of the stochastic formulation [84] . . . . .	28
2.4.2	The stochastic simulation algorithm . . . . .	29
2.4.3	Computational algorithm . . . . .	29
2.4.4	Simple example . . . . .	30
2.5	Green's functions . . . . .	31
2.6	Other modelling efforts . . . . .	31
2.6.1	Cellular Potts model of the lymph node . . . . .	31
2.6.2	Movement on the reticular network . . . . .	32
2.7	Chapter summary . . . . .	33
2.7.1	Chapter 3 . . . . .	33
2.7.2	Chapter 4 . . . . .	33
2.7.3	Chapter 5 . . . . .	34
2.7.4	Chapter 6 . . . . .	34
2.7.5	Chapter 7 . . . . .	34

2.7.6	Chapter 8 . . . . .	34
<b>3</b>	<b>First passage processes</b>	<b>37</b>
3.1	Introduction . . . . .	37
3.2	Poisson's Equation [184] . . . . .	40
3.3	Central target . . . . .	41
3.3.1	Particle in a random starting position . . . . .	42
3.4	Target 'close' to centre . . . . .	43
3.4.1	Random particle starting position . . . . .	44
3.4.2	Random target within $b$ of centre . . . . .	46
3.5	Randomly placed target . . . . .	47
3.5.1	Solution using Green's functions . . . . .	47
3.5.2	Derivation of SGMFPT equation . . . . .	50
3.5.3	Testing the random target equation . . . . .	54
3.5.4	The case when $R \neq 1$ . . . . .	55
3.6	Two fixed targets . . . . .	56
3.7	Equation for multiple targets . . . . .	58
3.7.1	Two randomly placed targets . . . . .	59
3.7.2	$N$ randomly placed targets . . . . .	60
3.8	Testing the $N$ target equation . . . . .	62
3.9	Summary . . . . .	65
<b>4</b>	<b>The effects of a limited imaging volume on <i>in vivo</i> imaging data</b>	<b>67</b>
4.1	Introduction . . . . .	67
4.2	Expected squared displacement . . . . .	69
4.2.1	One dimension . . . . .	69

4.2.2	Three dimensions . . . . .	73
4.3	Expectation of the normalised displacement squared . . . . .	73
4.3.1	One dimension . . . . .	74
4.3.2	Three dimensions . . . . .	77
4.4	Application to imaging volume . . . . .	79
4.5	Comparison to data . . . . .	81
4.5.1	Underestimation of plateau . . . . .	83
4.6	Summary . . . . .	84
<b>5</b>	<b>Analysis of <i>in vivo</i> imaging data</b>	<b>87</b>
5.1	Introduction . . . . .	87
5.2	About the data . . . . .	89
5.3	Basic analysis . . . . .	90
5.3.1	Displacement squared . . . . .	91
5.3.2	Turning angles . . . . .	93
5.3.3	Persistent motion . . . . .	94
5.3.4	Tissue drift . . . . .	95
5.4	Further analysis . . . . .	96
5.4.1	Sorting data . . . . .	96
5.4.2	Displacement squared against time . . . . .	97
5.4.3	Direction of motion . . . . .	101
5.5	Procedure for calculating diffusion coefficient . . . . .	104
5.5.1	Application to data sets . . . . .	105
5.6	Processing data . . . . .	106
5.7	Summary . . . . .	107

<b>6</b>	<b>How many dendritic cells are required to initiate a T cell response?</b>	<b>109</b>
6.1	Introduction . . . . .	109
6.2	Experimental design/results . . . . .	110
6.3	Mathematical model . . . . .	112
6.4	Computational model . . . . .	113
6.4.1	Over-crowding effect . . . . .	114
6.4.2	Correction of overestimation . . . . .	115
6.4.3	Estimation of parameters . . . . .	116
6.4.4	Comparison to experimental results . . . . .	117
6.4.5	Extrapolation of experimental results . . . . .	118
6.4.6	Minimum number of DCs required to initiate an T cell response . . . . .	120
6.5	Summary . . . . .	121
<b>7</b>	<b>Cross-regulation model</b>	<b>123</b>
7.1	Introduction . . . . .	123
7.1.1	Model dynamics . . . . .	126
7.2	Model 1: competition only . . . . .	128
7.2.1	Reactions . . . . .	128
7.2.2	ODEs . . . . .	128
7.3	Model 2: R suppression only . . . . .	130
7.3.1	Reactions . . . . .	130
7.3.2	ODEs . . . . .	131
7.4	Model 3: R suppression E promotion . . . . .	131
7.4.1	Reactions . . . . .	132
7.4.2	Effector cell dynamics . . . . .	134

7.4.3	Regulatory cell dynamics . . . . .	136
7.4.4	ODEs with effector and regulatory cells . . . . .	138
7.5	Analysis of Model 3 . . . . .	138
7.5.1	Carrying Capacity . . . . .	139
7.5.2	Solution of ODEs and Gillespie algorithm . . . . .	140
7.5.3	Effector and regulatory cell dynamics . . . . .	141
7.5.4	Phase portrait . . . . .	143
7.5.5	Effect of $\kappa_+$ and $\kappa_-$ . . . . .	144
7.5.6	The effect of $\kappa_+$ on cell numbers . . . . .	146
7.5.7	Stochastic effects of initial ratio of cells . . . . .	148
7.6	Summary . . . . .	150
<b>8</b>	<b>Mechanisms of T cell activation</b>	<b>155</b>
8.1	Introduction . . . . .	155
8.1.1	APC changes . . . . .	156
8.1.2	Signal integration . . . . .	157
8.1.3	Probabilistic approach . . . . .	157
8.1.4	T cell swarming . . . . .	157
8.2	Modelling approaches . . . . .	158
8.2.1	APC changes . . . . .	158
8.2.2	Signal integration . . . . .	159
8.2.3	Probabilistic approach . . . . .	160
8.3	APC changes approach . . . . .	160
8.3.1	Model IA . . . . .	161
8.3.2	Model IC . . . . .	162

8.4	Signal integration approach . . . . .	163
8.4.1	Model IIA . . . . .	163
8.4.2	Model IID . . . . .	164
8.5	Probabilistic approach . . . . .	165
8.5.1	Model IIIA . . . . .	166
8.5.2	Model IIIB . . . . .	170
8.6	Summary . . . . .	172
<b>9</b>	<b>Concluding remarks</b>	<b>175</b>
	<b>Appendix</b>	<b>180</b>
A	Spatial modelling . . . . .	181
A.1	Random placement of cells . . . . .	181
A.2	Non-overlapping APCs . . . . .	182
A.3	Unbinding . . . . .	183
A.4	Cell movement . . . . .	184
A.5	Closest APC . . . . .	185
A.6	Encounter during a timestep . . . . .	185
B	Code for encountering one of $N$ targets . . . . .	186
	<b>Bibliography</b>	<b>191</b>



# List of figures

1.1	Role of dendritic cells in the immune response. . . . .	3
1.2	T and B cell activation. . . . .	5
1.3	Outline of function of the different classes of T cells. . . . .	6
1.4	Lymph node diagram. . . . .	7
1.5	Diagram of flow cytometry. . . . .	11
2.1	Illustration of sampling volume of an APC. . . . .	18
2.2	Illustration of overestimation of MFPT on the surface of the target. . . . .	21
2.3	Error comparison with and without target boundary algorithm. . . . .	22
2.4	Illustration of overestimation of MFPT on the boundary of the domain. . . . .	23
2.5	Error comparison with and without surface boundary algorithm. . . . .	24
2.6	Error comparison for Gaussian and exponential timestepping. . . . .	26
3.1	Diagram to illustrate derivation of Poisson's equation. . . . .	40
3.2	Diagram showing a Brownian path to a target located in the centre of a sphere. . . . .	41
3.3	Problem setup with target 'close' to the centre of a sphere. . . . .	44
3.4	Comparison of numerical results and the MFPT equation for target 'close' to centre of a sphere. . . . .	45
3.5	Comparison of numerical results and the SGMFPT for a target 'close' to the centre of a sphere. . . . .	46

3.6	Comparison of (3.4.25) to numerical simulations. . . . .	47
3.7	Diagram showing a Brownian path to a target located in a uniform random position within a sphere. . . . .	48
3.8	Diagram of the problem setup with a uniform randomly placed target. . . . .	48
3.9	Diagram of problem setup when using Green's functions. . . . .	50
3.10	Comparison of numerical results and the SGMFPT for a target in a known non-central position. . . . .	54
3.11	Comparison of numerical results and the TGMFPT for one non-central target. . .	55
3.12	Comparison of numerical results and the TGMFPT for one non-central target with $R \neq 1$ . . . . .	56
3.13	Comparison of numerical results and the SGMFPT for two targets in known positions. . . . .	58
3.14	Comparison of numerical results and the TGMFPT for two, three, and four targets in uniform random positions. . . . .	63
3.15	Comparison of numerical results and the TGMFPT for varying numbers of uniform random targets. . . . .	64
4.1	$\mathbb{E}[x_t^2]$ in one dimension with a fixed starting position. . . . .	72
4.2	$\mathbb{E}[x_t^2]$ in one dimension with a uniform random starting position. . . . .	72
4.3	$\mathbb{E}[\mathbf{X}_t^2]$ in three dimension with a fixed starting position. . . . .	74
4.4	$\mathbb{E}[\mathbf{X}_t^2]$ in three dimension with a uniform random starting position. . . . .	74
4.5	Normalised $\mathbb{E}[(x_t - x_0)^2]$ in one dimension with a fixed starting position. . . . .	75
4.6	Survival probability in one dimension. . . . .	76
4.7	Normalised $\mathbb{E}[(x_t - x_0)^2]$ in one dimension with a uniform random starting position. . . . .	77
4.8	Normalised $\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]$ in three dimensions with a fixed starting position. . .	78
4.9	Survival probability in three dimensions. . . . .	79

4.10	Normalised $\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]$ in three dimensions with a uniform random starting position. . . . .	80
4.11	$\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]$ in three dimensions for a volume smaller in the $z$ dimension than the $x$ and $y$ dimensions. . . . .	80
4.12	Normalised $\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]$ in three dimensions for an imaging volume similar in size to those used in biological experiments. . . . .	81
4.13	Comparison of the predicted expectation and a simulated imaging experiment. . .	82
4.14	Comparison of the predicted expectation and a simulated imaging experiment with a volume the same size as used in biological experiments. . . . .	82
4.15	Illustration of underestimation and correction for the underestimation. . . . .	84
5.1	Snapshot of TPLSM <i>in vivo</i> imaging . . . . .	89
5.2	Basic analysis of the whole data set. . . . .	90
5.3	Mean displacement squared against time plot for whole data set. . . . .	92
5.4	Turning angles of whole data set. . . . .	94
5.5	Illustration of turning angles. . . . .	94
5.6	Mean scalar product for whole data set. . . . .	95
5.7	Tissue drift in data set. . . . .	96
5.8	Plots of cell tracks. . . . .	98
5.9	Displacement squared plots with and without tissue drift. . . . .	98
5.10	Plots of cell tracks for slow and faster cells. . . . .	99
5.11	Displacement squared plots for slow and faster cells. . . . .	100
5.12	Plots of cell tracks for fast and intermediate cells. . . . .	100
5.13	Displacement squared plots for fast and intermediate cells. . . . .	101
5.14	Turning angles of slow and faster cells. . . . .	102
5.15	Turning angles of fast and intermediate cells. . . . .	102

5.16	Dot product of slow and faster cells. . . . .	103
5.17	Dot product of fast and intermediate cells. . . . .	103
5.18	Data set; LN2. Diffusivity plots. . . . .	105
5.19	Data set; LN4. Diffusivity plots. . . . .	105
5.20	Data set; Ear1. Diffusivity plots. . . . .	106
6.1	Experiment recording percentage of activated T cells. . . . .	111
6.2	Flow cytometry output. . . . .	111
6.3	Quantifying number of DCs using GFP labelling. . . . .	112
6.4	Results of two photon imaging experiment. . . . .	112
6.5	Screenshots of spatial model simulations. . . . .	114
6.6	Effect crowding of DCs has on the TGMFPT equation. . . . .	115
6.7	Analysis of displacement squared. . . . .	116
6.8	Recruitment probability for $2 \times 10^5$ DCs. . . . .	117
6.9	Diagram showing the setup of the computational model. . . . .	118
6.10	Recruitment probability per T cell. . . . .	119
6.11	Probability of at least one T cell encountering a DC, in twenty four hours. . . . .	119
6.12	Minimum number of DCs required to achieve a 50% probability of at least one T cell encountering a DC. . . . .	120
6.13	Minimum number of DCs required to achieve a probability $p$ of at least one T cell encountering a DC. . . . .	121
7.1	Plots showing the solution of ODEs and Gillespie algorithm for Model 1. . . . .	129
7.2	Plots showing the solution of ODEs and Gillespie algorithm for Model 2. . . . .	132
7.3	Diagram of interactions that can occur in Model 3. . . . .	133
7.4	Diagram showing the reactions that occur in Model 3 . . . . .	135
7.5	Dynamics of Model 3 with only E cells. . . . .	137

7.6	Comparison of ODE model to Gillespie algorithm for Model 3, with $\kappa_+ = 12$ .	140
7.7	ODEs and two realisation of the Gillespie algorithm with different outcomes for $\kappa_+ = 3$ .	141
7.8	Comparison of number of E and R cells for $\kappa_+ = 12$ .	142
7.9	Comparison of number of E and R cells for $\kappa_+ = 3$ .	142
7.10	Phase portraits of the ODE system.	143
7.11	Plots showing the steady states different sets of initial conditions settle in.	144
7.12	Plot showing the steady states obtained for varying values of $\kappa_+$ and $\kappa_-$ .	145
7.13	Plot showing how the steady state numbers of E and R cells change with varying the value of $\kappa_+$ .	146
7.14	Plot showing the number of E and R cells in the various steady states as the value of $\kappa_+$ changes.	147
7.15	Stochastic effects of varying the starting ratio of E and R cells.	148
7.16	Realisations of the Gillespie algorithm for high initial ratios of R cells.	149
7.17	Realisations of the Gillespie algorithm for low initial ratios of R cells.	150
8.1	APC changes Model IA.	161
8.2	APC changes Model IC.	162
8.3	Signal summing Model IIA.	164
8.4	Signal summing Model IID.	165
8.5	Model IIIA, ODEs, Spatial, and Gillespie.	169
8.6	Model IIIA with switch from short to long bindings just after 8 hours.	170
8.7	Trajectory of Model IIIA.	171
8.8	Model IIIB, ODEs, Spatial, and Gillespie.	173



## List of tables

3.1	Table listing the parameters used to represent the positions of particles and targets. Vector refers to a three-dimensional position and radius refers to the distance from the centre of the domain. . . . .	39
3.2	Table listing notation used to represent first passage processes. Prescribed implies the starting position is known, unprescribed that the starting position is unknown. All random positions are uniform random and exclude the volume occupied by the target in the case of the particles position. . . . .	39
7.1	Tables showing notation and rates used in Chapter 7. . . . .	127

# List of abbreviations

APC	Antigen presenting cell
DC	Dendritic cell
E cell	Effector T cell
FRC	Fibroblastic reticular network
GFP	Green florescent protein
LN	Lymph node
MFTP	Mean first passage time
MHC	Major histocompatibility complex
ODE	Ordinary differential equation
R cell	Regulatory T cell
RN	Reticular network
TCR	T cell receptor
TGMFTP	Target global mean first passage time
TPLSM	Two photon laser scanning microscopy
Treg	Regulatory T cell
SGMFTP	Starting global mean first passage time
WT	Wild type

# Chapter 1

## Immunology introduction

### 1.1 Introduction

This thesis begins with a basic introduction to the human immune system. This is followed by a more in-depth description of the cell types that are model in this thesis, focusing on there function and interactions with other cell types. The main cell types used for modelling purposes in this thesis, and therefore described in this chapter, are: Dendritic cells, T cells, and a specific subset of T cells known as regulatory T cells. The chapter describes other aspects of the immune system, such as lymph nodes, which is the area of the immune system where all the modelling in this thesis takes place. The chapter also describes experimental techniques which are used by my collaborators and one of which is modelled in this thesis. The chapter concludes with an outline of the objectives of this thesis.

The main aim of this chapter is to give the reader an overview of the basic immunological concepts and processes required to understand the work carried out in this thesis. This chapter only begins to scratch the surface of how the immune system operates. For the reader wanting to know further details about immunology I refer them to the many excellent textbooks published on the topic, such as [115].

### 1.2 The human immune system

The human body includes three layers of defence against microbes; the first of these are physical and chemical barriers, preventing microbes entering the body, such as the skin, and the mucosal

epithelial lining of the airways and gut [163]. Only when these barriers are overcome, or evaded, does the immune system come into play.

The next layer of defence is the innate immune system [5,75,115], which is able to respond quickly (within hours) to microbial infection with phagocytic cells, such as macrophages. Macrophages are able to ingest and kill microbes by producing degenerative enzymes and toxic chemicals. They also secrete cytokines, which are molecules that cause inflammation and recruit cells of the adaptive immune system [163].

If an invading microbe cannot be eliminated by the innate immune system, the adaptive immune system becomes activated. The adaptive immune system is more efficient at eliminating infections as it acts in a specific manner, but it takes days rather than hours to get started. The cells of the adaptive immune response have the ability to distinguish a specific pathogen and focus their response against it. Cells involved in the adaptive immune response have specialised antigen receptors on their surface that are able to recognise and respond to individual antigens on the surface of microbes [163]. The body contains billions of lymphocytes that possess a vast repertoire of antigen receptors, allowing the immune system to recognise virtually any antigen a person is exposed to. The adaptive immune response is also responsible for the generation of memory cells, which allow a second infection by the same pathogen to be dealt with more quickly and with a greater cell numbers [163].

The major types of lymphocytes are: B cells, which upon activation differentiate into antibody-secreting plasma cells, T cells, which differentiate into a variety of classes with differing functions, which will be discussed later, and natural killer cells, these lack the antigen specificity and act in an unspecific manner [163].

### **1.3 Dendritic cells**

Dendritic cells (DCs) have long finger-like protrusions used for sampling their surrounding area. Immature DCs migrate through the bloodstream to enter tissues, and carry specialised receptors on their surface that are able to recognise common features of many pathogens, which they are able to phagocytose. Phagocytosis is the process of engulfing and breaking down an invading microbe. DCs also take up large amounts of extracellular fluid and its contents [163]. The pathogens they take up are digested, but their main role is not the clearance of microbes, but rather to activate other cells to clear the infections. This is achieved by presenting antigens from the

invading microbes on their surface (Figure 1.1). DCs then migrate to a nearby secondary lymphoid organ [113, 228, 235, 238], such as a lymph node (LN), where there is a high concentration of lymphatic cells of varying specificities [241]. T cells that recognise the antigen presented on a DC can then be activated by the DC due to the expression of co-stimulatory molecules, which are also required for T cell activation [36, 155, 156, 165, 185, 186]. Recognition of free antigen alone is not sufficient to activate T cells, the co-stimulatory molecules expressed by the DCs are also required. Cells that are able to present antigens to inactive T cells and activate them, are known as antigen-presenting-cells (APCs). Macrophages are also capable of presenting antigen. However, DCs are specialised in presenting antigen and initiating the adaptive immune response, which is why they are called professional APCs.

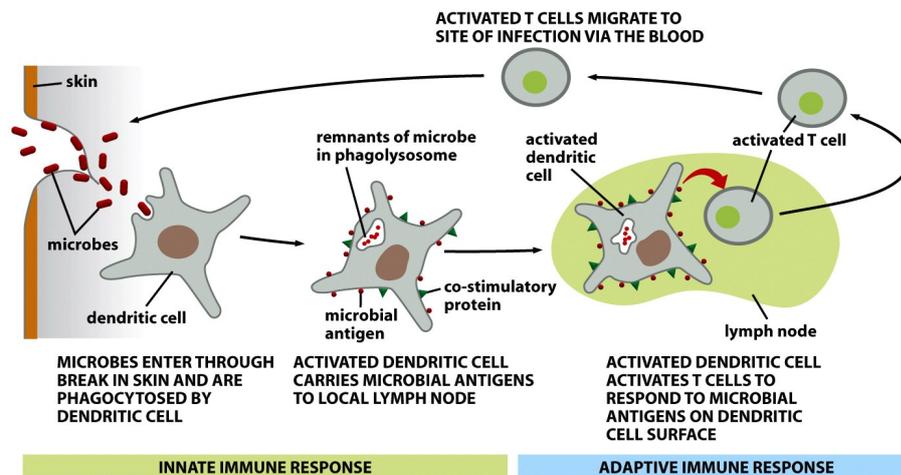


Figure 1.1: Role of dendritic cells in the immune response. Source: Figure 25-5 Molecular Biology of the Cell (©Garland Science 2008 [163]).

A key feature of DCs, which allows efficient sampling within the LN, by T cells, is their extensive probing activity through the rapid motion of their dendrites. This results in the volume surveyed by a DC becoming three-fold larger than its size [153]. The dendrites of DCs will also actively extend towards a T cell [32].

DCs are a subclass of APCs, and a distinction between DCs and APCs is not made in this thesis. The two terms will therefore be used interchangeably.

## 1.4 T cells

T cells are divided into two main subclasses, distinguished by a cell surface protein they express on their surface from the time they mature in the thymus. Cytotoxic T cells express the CD8 protein, and are involved in killing infected cells. The class of T cells involved in activating cells they are cognate to, rather than killing them, express the CD4 protein, and includes helper T cells [163].

In the absence of an infection most lymphocytes circulating in the body are small featureless cells with few cytoplasmic organelles. These inactivated lymphocytes are known as naive. Lymphocytes are activated by binding to an APC presenting antigen that is cognate to their receptor. In the case of T cells, which is the cell type studied in this thesis, this is the T cell receptor (TCR) [77, 78] (Figure 1.2). After a T cell is activated, it undergoes clonal expansion and the daughter cells differentiate into several functional classes [34,54]. There are three major classes of T cells, which fall into the categories of CD4<sup>+</sup> or CD8<sup>+</sup> T cells. Cytotoxic T cells, which express the CD8 protein, are responsible for killing cells that are infected with viruses or other intracellular pathogens. Helper/effector T cells, which express the CD4 protein, provide additional signals that activate antigen stimulated B cells (Figure 1.2) to differentiate and produce antibody [47, 173]. Regulatory T cells, which express the CD4 protein, regulate the activity of other lymphocytes and help control the immune response. During the course of the immune response some of the activated T cells (and B cells) differentiate into memory cells. These are responsible for the long-lasting immunity that results from the exposure to a disease or vaccination. Memory cells allow a rapid response to reinfection by the same pathogen and can rapidly differentiate into another T cell type. The newly activated T cells must then migrate to the site of the infection to assert their function [139, 160].

Unlike macrophages and DCs that express receptors able to recognise general features shared by many pathogens, each T cell expresses a receptor that is highly specific. Each T cell clonotype emerging from the thymus differs from the others in its receptor specificity. The diversity originates from gene rearrangement in the production of the T cells [158], ensuring that collectively T cells are able to recognise millions of different pathogens.

Antibodies produced by B cells are able to target pathogens in the blood and extracellular spaces, however some bacteria, parasites and all viruses replicate inside cells, where antibodies cannot detect them. The responsibility for eliminating these invaders falls to the T cells. Cytotoxic T cells are able to target cells infected with virus, as the infected cells express antigen derived from

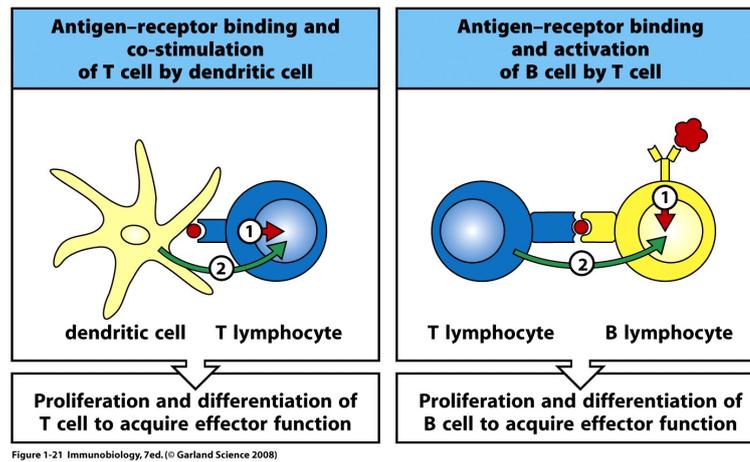


Figure 1.2: T and B cell activation. (©Garland Science 2008 [163])

the viruses multiplication process on their surface [163].

Naive CD4 T cells are able to differentiate into different types of effector T cells [2, 85], the three major types being Th1, Th2, and Th17 (Figure 1.3), although more have been described [102, 185, 220]. All these cell types are involved in combating bacterial infections. Th1 cells are characterised by the secretion of IFN- $\gamma$  and are important activators of macrophages, natural killer cells and CD8<sup>+</sup> T cells [214]. Th1 cells are able to deal with macrophages that have become infected by intracellular bacteria that have evaded the macrophage degradation process. They also function as helper cells, stimulating the production of antibodies by interacting with B cells and producing co-stimulatory signals. Th2 cells secrete the cytokines IL-4, IL-13 and IL-25, which are important for barrier defence at mucosal and epithelial surfaces. Th2 cells activate eosinophils, basophils, mast cells, and alternatively activated macrophages [227]. Th17 cells produce an array of cytokines and are responsible for regulating acute inflammation [130]. They also work with neutrophils in the defence against extracellular bacteria [101, 142].

Naive T cells circulate from the bloodstream into LNs, the spleen and mucosa-associated lymphoid tissues, and then back into the blood. This process allows an individual T cell to sample the peptides presented by thousands of DCs daily [42, 180]. The efficiency at which T cells sample antigen presenting DCs is very high, which is crucial for the initiation of the adaptive immune response, as only one naive T cell in  $10^4$ - $10^6$  is specific for a particular antigen [42].

CD4<sup>+</sup> T cells are dependent on persistent antigen presentation by DCs for full activation and differentiation, and renewed contacts between daughter cells and antigen-bearing DCs has been observed [54, 171].

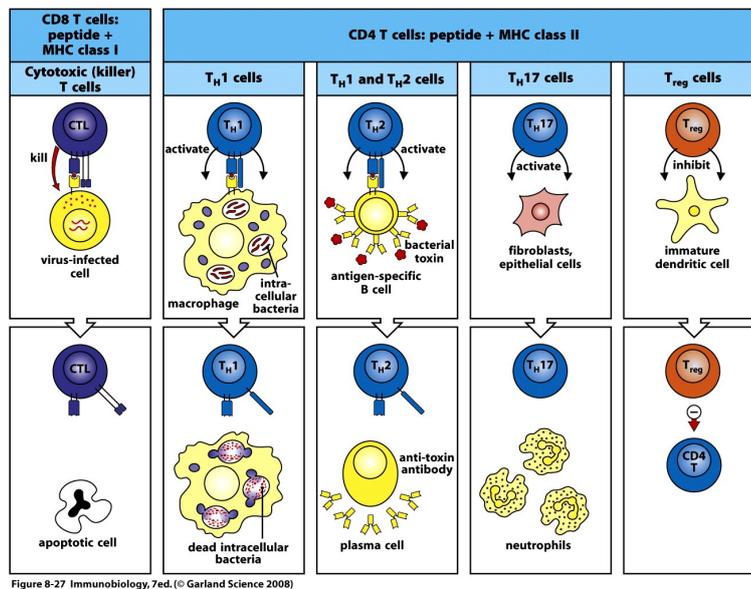


Figure 1.3: Outline of function of the different classes of T cells. (©Garland Science 2008 [163])

## 1.5 Lymph nodes

Lymphocytes circulate in the blood and the lymph and are also found in lymphoid tissues and lymphoid organs. The lymphoid organs are organised into the central/primary and peripheral/secondary lymphoid organs. The primary lymphoid organs are the bone marrow and thymus, and are the sites where lymphocytes are produced and mature. Secondary lymphoid organs include the LNs, spleen and Peyer's patches of the intestine, and are the sites of lymphocyte interactions.

Lymph nodes are located at the points of convergence of the vessels of the lymphatic system, and are highly organised lymphoid organs. The afferent lymphatic vessels drain fluid carrying antigen-bearing cells from infected tissues, as well as pathogens, to the LNs. Free antigens are able to diffuse through the extracellular fluid to the LNs, and DCs actively migrate there following chemokine signals [15]. The same chemokines also attract other lymphocytes, which enter from the blood via high-endothelial-venules [9, 10, 162]. Within the LN, B cells are located in the follicles, which make up the outer cortex of the LN, and T cells are located in the surrounding paracortical areas [97], often referred to as T cell zones.

Within the LN, the majority of naive T cells are highly motile, and display complex trajectories [123, 156] akin to a random walk. Such behaviour allows T cells to thoroughly explore the LN, however, the elements dictating T cell trajectories are unclear at present. It has been proposed

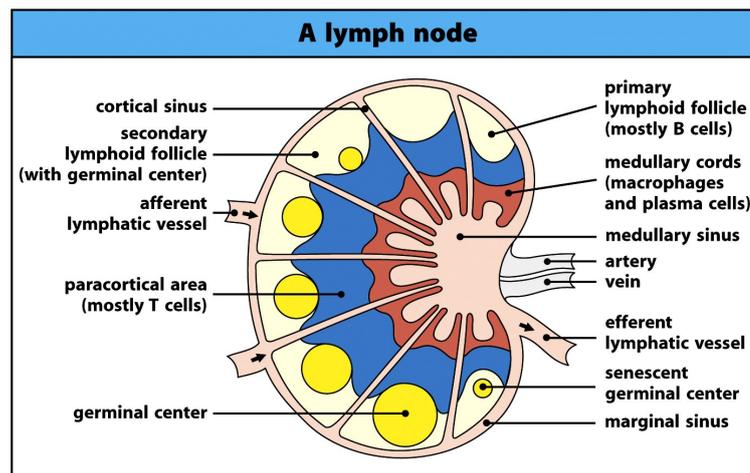


Figure 1.4: Lymph node diagram. (©Garland Science 2008 [163])

that guidance cues dictating T cell direction might involve the fibroblastic reticular network and/or chemokine signals [66]. The fibroblastic reticular network is a network comprising of reticular fibres, fibrous extracellular matrix bundles, and fibroblastic reticular cells, and supports the entire LN [91]. The role of chemokines in influencing T cell migration has been challenged by observations of T cell trajectories using two-photon imaging, in which T cell trajectories were found to be essentially random [155, 156], with a random walk behaviour being found for longer timescales. *In vivo* tracking of T cells within LNs found the mean absolute displacement of T cells increased proportionally with the square root of time, indicating a random walk behaviour. The mean displacement can therefore be characterised by a motility/diffusion coefficient, this was found by Miller *et al.* [156] to be  $67\mu m^2 min^{-1}$  for T cells.

The high motility of T cells, combined with the less motile DCs, which probe their surrounding area with their dendrites, leads to DCs being able to sample a large fraction of the T cell repertoire in a short period of time. It has been estimated that between 500 and 5000 T cells are sampled by each DC every hour [42, 51, 153].

Studies have investigated the dynamics of interactions between antigen-bearing DCs and antigen-specific T cells. Stoll *et al.* [212] reported cell-cell contacts and prolonged interactions between T cells and DCs, followed by activation, dissociation and rapid migration of T cells away from the DC. A similar sequence of events was also observed *in vivo* within the LN, where DC-T cell interactions were observed with  $CD4^+$  [154, 201] and  $CD8^+$  [42, 110, 152] T cells. Further investigation is required to achieve a consensus for the duration of DC-T cell interactions [110, 201].

## 1.6 T cell selection

The immune system is able to combat virtually any pathogen it is presented with. To achieve such a large coverage, it must contain a repertoire of T cells able to recognise as many pathogens as possible. This is achieved through gene rearrangement. Developing T cells undergo a process of gene rearrangement to generate a unique receptor. This process allows the development of a large number of TCR structures, able to recognise billions of different antigens. As this process is random, there is inevitably receptors generated that are not functional, or self reactive. These receptors are screened through a process of positive and negative selection [106, 137, 200, 211] in the thymus.

A functional T cell must have a functional receptor, it must therefore recognise self-peptide:MHC complexes to some extent, to be able to receive a survival signal in the periphery. A self-peptide:MHC complexes is a complex of the body's own antigen and a cell surface molecule. This ensures only thymocytes expressing a useful T cell receptor can survive and mature. This involves receiving a weak interaction with pMHC complexes that include thymic self-peptide, as some foreign antigen peptide is likely to exist that has a much stronger binding with such a receptor in the periphery. T cells without a functional receptor will not be able to undergo weak interactions and will die from neglect.

Only 10% of T cells pass through positive selection. The other 90% of T cells die by neglect, as they do not have a functional receptor. Those that do survive, undergo negative selection [107, 175]. T cells whose receptors recognise self-peptide:MHC complexes too strongly are given a signal to stimulate cell death, by apoptosis [211]. Negative selection deletes potentially self-reactive thymocytes, thereby generating a repertoire of peripheral T cells that is largely self tolerant [209, 211]. Less than 5% of the originally generated T cells survive thymic selection. Inevitably some cells escape negative selection. These self-reactive cells are controlled by peripheral tolerance mechanisms [194].

## 1.7 Regulatory T cells

The immune system is able to discriminate between self and non-self antigens, inhibiting autoimmune responses, but allowing effective immune responses against microbial antigens. The immune system is able to establish and sustain unresponsiveness to self antigens, including

physical elimination of self reactive cells, through a process of thymic selection (Section 1.6). There is also substantial evidence the T cell mediated suppression of self-reactive T cells is another essential mechanism of self tolerance [65, 141, 191, 202]. Tolerance to self-antigen is primarily achieved through thymic selection, but some auto-reactive T cells escape thymic selection or recognise antigens that are not expressed in the thymus. This regulation is carried out by a distinct subset of T cells [81] known as regulatory T cells (Tregs), which are functionally defined as a T cell that inhibits an immune response by influencing the activity of other cell types [81]. There are two groups of regulatory T cells. One subset becomes committed to the Treg lineage during development in the thymus, these are natural Tregs [192]. The other subset develops in the periphery from uncommitted naive T cells [148].

Natural Tregs are CD4<sup>+</sup>, and express the  $\alpha$  chain of the IL-2 receptor (CD25) [11, 193], giving them a high affinity for the IL-2 cytokine. They also express the FOXP3 receptor [76, 109, 125]. Natural Tregs make up about 10-15% of the CD4<sup>+</sup> cells in circulation in the human body. Natural Tregs are potentially self-reactive T cells, that are selected in the thymus. The mode of action of Tregs *in vivo* is still not fully understood, and there is a large array of literature which investigates this [183, 217, 223, 224]. Tregs are believed to act in either a contact-dependent fashion with other cells, or through the secretion of cytokines, such as IL-10 and TGF- $\beta$ .

The function of (natural) Tregs is to suppress the activation of self-reactive T cells. Lack of natural Tregs is known to be the cause of several autoimmune syndromes.

## 1.8 Experimental techniques

### 1.8.1 Two-photon laser scanning microscopy

Two-photon laser scanning microscopy (TPLSM) allows the interactions of T cells and DCs to be viewed within a LN *in vivo* [41, 156]. Most T cells are located near the surface of the LN, in the paracortical area, which is approximately 150-200 $\mu$ m below the organ's surface in mice. A TPLSM experiment is carried out by acquiring sequential images of a three-dimensional volume of a LN that contains fluorescently labelled cells. This is achieved by recording the fluorescent signals emitted by the cells at successive focal planes, and repeating this process every 10-30 seconds for a period of up to a few hours [40]. The data set generated can be viewed as a time-lapse movie, and important information can be extracted from these videos by manual or

automated analysis. TPLSM achieves the same optical sectioning effect as confocal microscopy, but it uses a different optical principle that increases depth penetration, and reduces phototoxicity and photo-bleaching [46], making it ideally suited to imaging T cell activation at a single-cell level [70]. Phototoxicity is the damage caused to cells due to the laser light. Photo-bleaching is the bleaching of the fluorophores that are used to label cells of interest. A fluorophore is a fluorescent chemical compound that is able to re-emit light upon light excitation. In TPLSM two extremely short (femtosecond) pulses of light are emitted, of infrared frequency, which are absorbed almost simultaneously by fluorophores, rather than a single photon of visible light emitted in confocal microscopy. The energy of the photon decreases with increasing wavelength, so the two infrared photons together provide a similar amount of energy to a blue photon, causing a green photon to be emitted by the fluorophores, as it would in regular fluorescence. Using two photon excitation results in fluorescence only at the focal point, due to the fluorescence emission increasing as the square of light intensity, rather than a linear increase, as observed in confocal microscopy. This confines photo-bleaching and phototoxicity effects to the focal plane. Using infrared light allows deeper penetration into the tissue due to its longer wavelength.

There is less phototoxicity using TPLSM compared to other imaging methods. However, the levels of phototoxicity are not insignificant [108], and are the cause of the limited sized imaging volume used in TPLSM. The size of an imaging volume is dependent on the experimental setup, with  $x$  and  $y$  dimensions typically being up to  $250\mu m$ , and the  $z$  dimension typically being up to  $35\mu m$ . The depth that can be imaged is determined by the number of planes in the  $z$ -stack. If the  $z$ -planes are further apart, a greater depth of image can be achieved, however if they are too far apart, cells can reside in the spaces between the planes and will not be observed. The number of  $z$ -planes imaged also affects the time between imaging points. If numerous  $z$ -planes are taken the microscope must scan a larger area before restarting at the first plane for the next timepoint, resulting in the time between successive timepoints increasing. Taking a smaller time increment allows the path of cells to be imaged at a higher resolution, but has the drawback of reducing the volume that can be imaged, and increasing the speed at which phototoxicity occurs, thereby reducing the overall experiment duration. A balance of  $z$ -stack separation, track resolution and experiment duration must be achieved. TPLSM also requires the LN to be exposed for imaging to occur. Experiments take place on living mice [152, 155], but results in the mouse being sacrificed, therefore subsequent imaging on the same mouse cannot be achieved. Recent advances are experimenting with transplanting LNs to a mouse's ear, allowing imaging without harming the mouse, and subsequent imaging on the same mouse to be carried out. This method

also avoids effects that may occur from operating on the mouse, which may affect the experiment's outcome. For an detailed explanation of TPLSM, please refer to these reviews [46,79].

## 1.8.2 Flow cytometry

Flow cytometry is an experimental method that is able to measure and analyse multiple physical characteristics of cells, as they flow in a stream of fluid through a beam of light [163]. The system is also able to sort cells depending on their characteristics, by applying a charge and deflecting them. The characteristics that can be measured include; a cell's relative size, relative granularity or internal complexity, and relative fluorescent intensity. These are measured by recording how a cell scatters incident laser light and emits fluorescence. Cells being examined are labelled with fluorescent markers, which fluoresce depending on the characteristics of each labelled cell. The number of markers that can be detected by flow cytometry depends upon the complexity of the machine used. The scattered and fluorescent light is collected by a system of appropriately positioned lenses, and a combination of beam splitters and filters steer the light to the appropriate detectors. An electrical signal is produced by the detectors proportional to the optical signal striking them. The data is sorted and analysed by a computer, and can provide information on sub-populations within the sample. Cells ranging in size from 0.2 to 150 micrometers can be examined using flow cytometry. This experimental technique is used by my collaborators to obtain some of the results presented in Chapter 6.

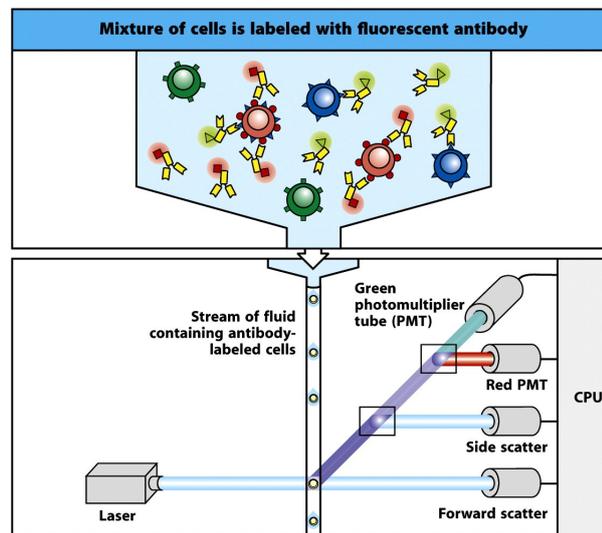


Figure A-26 part 1 of 2 Immunobiology, 7ed. (© Garland Science 2008)

Figure 1.5: Diagram of flow cytometry. (©Garland Science 2008 [163])

## 1.9 Objectives of the thesis

In this section, the initial objectives of the thesis are described. Whether or not these objectives have been met will be evaluated in Chapter 9.

The objectives of this thesis are:

1. Study the theory of first passage processes to apply to encounters between T cells, moving by Brownian motion within a lymph node, and encounters of APCs. Using this theory, develop a mean first passage time equation for a uniform randomly placed particle to locate a uniform randomly placed target.
2. Study the effect of confinement in a limited sized imaging volume, seen in two photon microscopy experiments, on the observed mean squared displacement of the imaged cells. An accurate understanding of artefacts introduced by two-photon imaging technology is crucial to correctly interpreting the data produced.
3. Develop a method to calculate the diffusion coefficient of T cells, using data obtained from two-photon microscopy experiments. The size of the diffusion coefficient is key to the response rate of the immune system.
4. Using concepts developed in objectives one, two and three, along with experimental data, I aim to determine the minimum number of DCs required to initiate a T cell response, given typical T cell precursor frequencies. The number of DCs present influences the probability of activation of the T cell response.
5. Extensive literature [49, 133] is available on deterministic models of interactions between effector and regulatory T cells. Knowing the populations of T cells are usually quite small, I intend to investigate the effects of a stochastic approach on the interactions of effector and regulatory T cells.
6. In references [40,42], three hypotheses are proposed for the activation mechanism of T cells: APC changes, T cell signal integration, and a probabilistic approach . I intend to test these hypotheses using analytical and computational models, with the aim of determining if any of the proposed hypotheses can explain experimental results.

## Chapter 2

# Introduction to stochastic spatial modelling and numerical methods

### 2.1 Introduction

The role of mathematical modelling in immunology was recognised as early as the 1960's [96,145] and 1970's [24, 71, 104, 105, 132, 134, 177, 231], and immunology was one of the first biological areas to adopt mathematical modelling. Mathematical models in immunology have been used in areas ranging from antigen receptor interactions [178], to vaccination [7, 114, 124, 150, 161, 170]. Since the introduction of mathematical immunology, the application of mathematical models has changed from conceptual models of basic reactions to bioinformatic analysis of high throughput data and simulation-based analysis of complex molecular and cellular systems. Over the last couple of decades, mathematical immunology has shifted from deterministic ODE-based models to stochastic models. A deterministic approach predicts an outcome with absolute certainty. Therefore, it is not able to cope with the inherent variability present in biological system. A stochastic model provides a probability of an outcome for a given set of initial conditions. Stochastic models of biological systems are important as they allow random variability to occur, which is often seen in biological processes. In population level models, where the population size is sufficiently large, a deterministic formulation can be used, and will be used in this thesis. However, for small population sizes, extinction of the population may occur, and then it is more realistic to use a stochastic formulation to model the variation in size, allowing fluctuations to occur that might result in extinction.

This chapter introduces the main mathematical techniques used in this thesis. Spatial modelling is used throughout, therefore a large proportion of this chapter focuses on spatial modelling, and methods of reducing error between analytical calculations of first passage times and spatial modelling simulations. The chapter begins by reviewing the various methods proposed for how T cells move, proposed in [92, 121, 123, 153, 225, 232]. In this thesis T cells are assumed to move by Brownian motion. A definition of the Wiener process is given [151] and related to movement by Brownian motion. The chapter goes on to describe the computational algorithm developed in this thesis, which is used throughout the thesis to simulate cell movement. Errors occurring in the computational algorithm due to using a discrete timestep size are then addressed. These errors can occur on the boundary of the target and the boundary of the domain. Algorithms developed in [119] are introduced to reduce these two sources of error and the efficiency of these algorithms are tested against analytical calculations of mean first passage times. A method of reducing the error in computational simulations further by using exponential time-stepping [117–119] is then reviewed. The aim of reducing the amount of error in the spatial simulations carried out is to allow a larger timestep size to be used without significantly reducing the accuracy of the simulation. Using a larger timestep size allows a faster implementation of the simulation, reducing on computational recourses required.

The Gillespie algorithm [83, 84] is then reviewed as a way of computationally simulating a stochastic process. The Gillespie algorithm is used in this thesis to simulate interactions between cells and allows a computationally faster implementation than a spatial algorithm when a spatial aspect is not explicitly required.

Finally, other modeling efforts undertaken by various people as a way of simulating cell movement are reviewed. This include the Cellular Potts model and a model of movement on the reticular network. The chapter concludes with a summary of Chapters 3-8 in this thesis.

## **2.2 Spatial modelling**

### **2.2.1 How do T cells move?**

The question of how T cells move within a lymph node (LN) is a much debated topic. *In vivo* imaging techniques, using two-photon microscopy, have allowed movies of T cell movement and interactions to be generated. This imaging gives the impression that T cells are freely moving

within the LN. However, it must be remembered that only a limited number of cells have been labelled, and in reality the environment is very crowded with other cells, the reticular network, and blood vessels. The movement of cells depends upon their intrinsic motility as well as collisions and interactions with other cells and their surroundings [232].

Various models for T cell movement have already been proposed: (1) Simple Brownian motion, in which a T cell's movement is completely unbiased and follows a random walk [153], (2) persistent motion, in which a T cell moves in one direction for a short period of time then chooses a new random direction to move in, and repeats the process [232], (3) movement on the reticular network, which proposes T cells move along a pre-defined network within the LN [92, 121, 123, 225].

Persistent motion has been observed in *in vivo* experiments [232], where T cells were observed to move in a series of repetitive lunges, repeatedly balling-up and extending, with a period of about two minutes, moving an average distance of about  $20\mu m$  in each cycle, with peak velocities of  $25\mu mmin^{-1}$ . After each lunge, a cell would pause and then take off in another direction. T cells were found to travel in a fairly consistent direction during each lunge, and even to continue to move in a consistent direction over several cycles. However, after each pause, there is a high probability a cell will change direction. Therefore, T cells display a random walk behaviour when visualised over tens of minutes, even though on shorter times scales their motion appears to be more linear.

Some two-photon microscopy experiments have found T cells move along the strands of the fibroblastic reticular cell (FRC) network [16]. Individual T cell trajectories obtained from these studies suggest motion is random and not directed by chemical gradients over large distances. It was found that the mean displacement of a T cell from its starting position increases linearly with the square root of time, indicating a random walk nature of movement. It was concluded that, even though T cells appear to move along the FRC network, due to the structure of the network, and the frequent crossing points, movement along the network appears to be a random walk [16]. The reticular network model has the advantage of reducing the volume cells can move within, thereby potentially increasing the rate of encounters. It does however raise lots of other problems in setting up the mathematical system, such as: How many nodes where cells can change direction should there be? How many connections should there be between nodes? How far apart should nodes be? How do cells pass each other on the network? Methods of modelling the network have been devised in [73]. However, as the number of nodes and connections tend to infinity the model in [73] will approach a Brownian motion model.

In this thesis, T cell movement is modelled as a simple Brownian motion, as this is the simplest of the theories to model, and the general consensus of the experimental results reviewed in this section and models reviewed in Section 2.6 seems to be that cell movement is like Brownian motion on long timescales.

Beauchemin *et al.* [22] developed a model to simulate T cells moving along the FRC network by having cells move for an amount of time at a given velocity in a straight line, then pause for a given time and choose a new direction. They also used the same model to simulate cells moving without a pause time, and found both approaches agreed with experimental data, the latter just needing a slightly higher velocity. They conclude that, the addition of a pause time does not significantly improve the agreement between the experimental data and numerical results. Thus, the spatial modelling in this thesis is implemented without a pause time.

A characteristic of Brownian motion is that a particle may at first move rapidly away from its starting point but then may backtrack, resulting in its absolute displacement increasing at a progressively slower rate as time increases. The mean displacement, of a population of particles undergoing Brownian motion, from the origin grows as function of the square-root of the time. For example: in  $1ms$  a population of uniformly randomly moving particles may have a mean displacement of  $1\mu m$ , in  $100ms$  this increases to  $10\mu m$ , and to move  $100\mu m$  will take a mean time of  $10s$ . This relationship is determined by the diffusion coefficient,  $D$ , of the particle, giving a formula for the mean distance,  $d$ , travelled in time,  $t$ , of

$$d = \sqrt{6Dt}.$$

The factor 6 applies for diffusion in three dimensions, it changes to 4 in two dimensions, and 2 in one dimension.

### 2.2.2 The Wiener process

Let us begin by considering a one-dimensional symmetric random walk such that, in each time unit, a particle is equally likely to take a unit step either to the left or the right. Now suppose we speed up this process by taking smaller and smaller timesteps in smaller and smaller time intervals. If we go to the limit in the correct manner Brownian motion is obtained.

Brownian motion is also known as the Wiener process, or the Wiener-Einstein process [221]. The term Brownian motion acknowledges Robert Brown, who in 1827 noticed pollen grains

suspended in water moved constantly in a zigzag motion under a microscope. In 1905 Albert Einstein stated laws governing Brownian motion using principles from the kinetic energy theory of heat. However, it was Norbert Wiener's research in 1923 that laid the mathematical foundation for Brownian motion. Therefore, the process is often named after Wiener. Melsa and Sage [151] related the Wiener process to the concept of Brownian motion.

Suppose  $W(t)$  is the displacement of a small particle from the origin at time  $t$ . The displacement that occurs in the time  $t_1$  to  $t_2$  is small, and the frequency of direction changes is large. By applying the central limit theorem to the sum of a large number of these small disturbances, it can be assumed that  $W(t_2) - W(t_1)$  has a Gaussian density. The density of the particle's displacement does not depend on the time of the observation, it just depends on the length of the time interval. Therefore the probability density of the displacement from time  $t_1$  to  $t_2$  is the same as from time  $t_1 + t$  to time  $t_2 + t$ .

**Definition:** The stochastic process  $\{W(t)\}, t \in [0, \infty)$ , is a Wiener process if  $W(t)$  depends continuously on  $t$ ,  $W(t) \in (-\infty, \infty)$ , and the following three conditions hold [8]:

- For  $0 \leq t_1 < t_2 < \infty$ ,  $W(t_2) - W(t_1)$  is normally distributed with mean zero and variance  $t_2 - t_1$ , that is,  $N(0, t_2 - t_1)$ .
- For  $0 \leq t_0 < t_1 < t_2 < \infty$ , the increments  $W(t_1) - W(t_0)$  and  $W(t_2) - W(t_1)$  are independent.
- $\text{Prob}\{W(0) = 0\} = 1$ .

The Wiener process can easily be expressed in multiple dimensions ( $B_t$ ), as each of the Cartesian components of the Wiener process in  $n$  dimensions is an independent Wiener process,

$$B_t = (W_t^{(1)}, W_t^{(2)}, \dots, W_t^{(n)}),$$

where  $W_t^{(i)}, i = 1, \dots, n$ , is a one-dimensional Wiener process.

### 2.2.3 Computational algorithm

To model interactions of cells moving by Brownian motion within a LN, it is necessary to devise a spatial model of cell movements and interactions. This spatial model will be used in various forms throughout this thesis. Algorithms used in the spatial model, described below, are given in the

appendix. To create a model in which T cells move, within a LN, in search of antigen-presenting cells (APCs), the following steps need to be implemented.

### Step 0 (Initialisation):

- Firstly, place APCs, of effective radius  $b$ , and T cells, modelled as point particles, uniformly randomly within a sphere of a given radius,  $R$  (A.1). The APC's radius accounts for the APC having dendrites, which are assumed to sample a spherical region of radius equal to the mean dendrite length [17](Figure 2.1). The radius of a T cell is added to the APC's radius to give an effective radius for the APC, allowing T cells to be modelled as point particles.

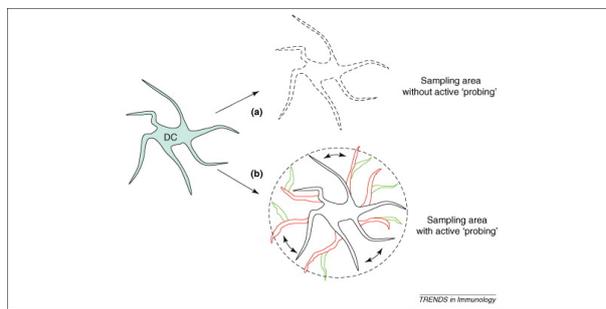


Figure 2.1: Illustration of sampling volume of an APC. Reproduced from [17].

### Step 1 (Movement):

- APCs remain stationary and T cells move by Brownian motion, simulating a random walk.
- A fixed discrete timestep size,  $\delta t$ , is chosen and each T cell picks a distance to move in the  $x$ ,  $y$  and  $z$  dimension from a Gaussian distribution with mean zero and variance  $2D\delta t$ , which is obtained from the definition of the Wiener process stated in section 2.2.2 with the relation  $x_t = \sqrt{2D}W_t^{(1)}$ ,  $D$  is the diffusivity of the T cells (A.4).
- If a T cell moves outside the radius of the LN (Figure 2.2), it is reflected back in by multiplying each of the  $x$ ,  $y$  and  $z$  positions by  $\frac{2R}{r} - 1$ , where  $r$  is the T cell's distance from the centre of the LN.

### Step 2 (Encounters):

- The position of each T cell relative to its closest APC is checked to determine if they have encountered each other.

- If the separation distance is less than  $b$ , which incorporates the size of the APC and the T cell, then an encounter will have occurred.
- If the separation distance is greater than  $b$ , an algorithm is used to calculate the probability of an encounter occurring during a timestep. If this probability is greater than a uniform random variable, between zero and one, an encounter will have occurred (A.6).
- Upon encountering an APC, a T cell can bind to the APC. If this occurs, a time to remain bound is picked from a given distribution, with a set mean binding time. An exponential distribution is used in this thesis.

### Step 3 (Unbinding):

- After the simulation time is incremented, the same time increment must be subtracted from the amount of time bound cells are to remain bound.
- If the time to remain bound is less than  $\delta t$ , the T cell is unbound from the APC it is attached to.
- When a T cell unbinds various scenarios may occur. The T cell can be placed uniformly randomly within the LN, and a new uniform random position chosen if they occupy the same volume as that occupied by the APCs. This is useful for testing first passage time equations that assume an approximately uniform random starting position of a particle. A T cell can be placed at the edge of an APC's zone of attraction. This seems the most likely biological outcome, but as the movement of the T cell is uniformly random, there is a high probability of the T cell rebinding. This can be overcome by having a deactivation time for T cells after binding, in which they cannot rebind. If this time is long enough, the T cell should be in an uniform random position when it is able to bind again.

### Step 4 (Update):

- To improve the efficiency of simulations, instead of each T cell checking its distance from every APC after each timestep, a T cell only needs to check its distance from the APC it is closest to (A.5).
- This is achieved by compiling a list of the closest APC to each T cell. The list of closest APCs does not need to be updated every timestep. The update frequency will depend upon

the diffusivity of the T cells, and the number of APCs. A larger diffusivity and/or a greater number of APCs will require the update interval to be smaller to prevent the closest APC to a T cell significantly changing.

- We must now return to step 1, and continue in this way until some predefined criteria are met, such as a specific time being exceeded or cell numbers reaching a specific value.

In step 0, when placing the APCs, the arrangement in relation to each other must be considered. APCs can be simply placed uniformly randomly, allowing two APCs to occupy the same volume. This outcome is not desirable in our simulations, therefore the APCs are placed so they either occupy completely separate volumes, or they may be allowed to have a slight overlap equivalent to the radius of the T cell. The overlap amount could even be made larger, as in reality the shapes of the APCs would fluctuate (A.2). This is achieved by first placing the cells uniformly randomly within the domain and replacing the cells that do not obey the desired cell overlap rules. Placing APCs with a position that is dependent upon the position of the APCs already placed results in a distribution of APCs that is not uniform random across the whole of the domain, as each subsequent APC has a smaller and smaller volume which it can occupy. This does not pose a problem for the computational modelling, it does however pose a problem for comparison with analytical models. However, when the volume of the domain is large and the volume of the APCs are small, and low in number, the error caused by this not truly uniform placement of APCs is minimal. For larger APCs numbers, as considered in Chapter 6, the error becomes more significant, and a method of compensating for this error is proposed in Chapter 6.

In step 3, the simplest way to reduce the binding times is to reduce the binding times of each cell by the timestep size, every time the simulation time is incremented, however this proves to be computationally inefficient. A more efficient way of reducing the binding times is to record the amount of time since the last binding or unbinding event occurred ( $t_{\text{passed}}$ ), and if this is greater than the minimum binding time of all the bound cells, their binding time will be reduced by  $t_{\text{passed}}$ . This procedure must also be undertaken before binding a new cell (A.3). In this model, APCs can be given various numbers of binding sites, and interactions can occur depending on the type of cells that are bound to an APC. This occurs, for example, in the cross-regulation model (Chapter 7).

## 2.3 Error reduction algorithms

The implementation of discrete timestepping in spatial simulations introduces errors. The error in simulations is calculated by comparing the mean first passage time (MFPT) obtained from the simulations to the analytical solution for the MFPT (3.3.13). Using the spatial modelling algorithm in the previous section, the MFPT is over estimated, due to the strictly positive probability that an excursion is made inside the target region during a timestep that starts and ends inside the region (see Figure 2.2), or outside the domain (see Figure 2.4) in a timestep that starts and ends inside the domain [19, 23, 86, 87, 117, 118, 144, 157]. To reduce these errors, algorithms are used to calculate the probability of encountering an APC during a timestep, and the probability to have hit the outer boundary during a timestep. These algorithms allow a larger timestep size to be used without introducing significant amounts of error, thereby improving the computational speed of the simulation. Error reduction algorithms are described in the next sections.

### 2.3.1 Target encounters

Figure 2.2 illustrates the situation that results in overestimation of the MFPT at the surface of the target. The blue path can be reproduced by decreasing the timestep size, but a smaller timestep size

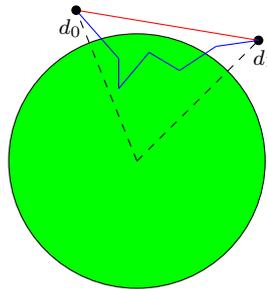


Figure 2.2: Diagram showing the situation that causes overestimation of the MFPT at the target. The green circle represents the APC/target,  $d_0$  and  $d_1$  are the distance from the centre of the target, to the T cell, before and after the timestep. The red line indicates the path the particle has taken in the simulation and the blue line indicates a possible path the particle could have taken if a smaller timestep size was used, resulting in the particle encountering the target.

results in more computational time being needed to execute the program. To determine the size of the error in the spatial code, a simple code that moves one T cell until it finds a stationary APC in the centre of a sphere was implemented. A mean time to encounter the APC over a number

of realisations was calculated for varying timestep sizes and compared to the MFPT calculated analytically (3.3.10), to calculate the error in the simulation. Using this data, a curve of the form  $\beta\sqrt{\delta t}$  could be fitted, as it was suspected the error would change as a function of  $\sqrt{\delta t}$  [86]. This gave an equation for the error of  $13.77\sqrt{\delta t}$  (Figure 2.3). Therefore the error in the simulation is proportional to  $\sqrt{\delta t}$ , where  $\delta t$  is the timestep size.

The error occurs because a T cell may encounter an APC during a timestep. To overcome this problem, an algorithm, developed by [86, 119], can be introduced to calculate the probability of the T cell encountering an APC during a timestep. The equation used to calculate the probability of encountering a target during a timestep is

$$\mathcal{P}(\text{encounter}) = \exp\left(-\frac{d_0 d_1}{D\delta t}\right), \quad (2.3.1)$$

where  $d_0$  is the distance from the T cell to the centre of the APC at the start of the timestep,  $d_1$  is the distance at the end of the timestep, and  $D$  is the diffusion coefficient of the T cell. This algorithm can be implemented in the spatial code and a new error calculated by comparison to the MFPT equation (3.3.10). Fitting a curve to the new error gives an equation for the error of  $6.00\delta t$  (Figure 2.3). The error in the code using this error reduction algorithm has now been reduced to be proportional to  $\delta t$ .

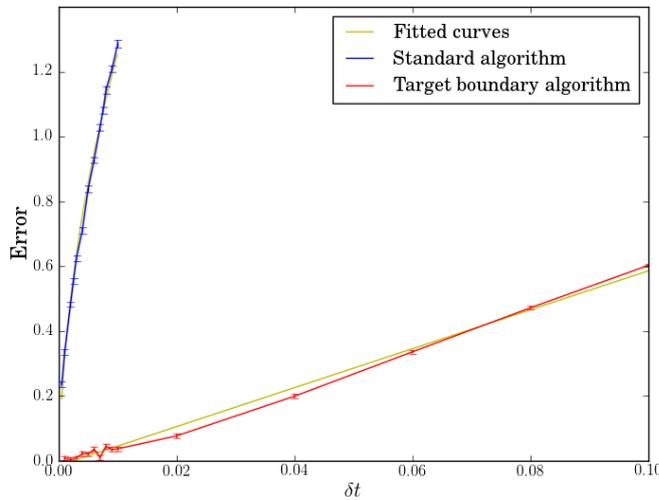


Figure 2.3: Plot showing the error as the timestep size is varied, without the error reduction in blue, with a curve of the form  $\beta\sqrt{\delta t}$  fitted to the data (yellow), and with error reduction in red, with a curve of the form  $\beta\delta t$  fitted to the data (yellow). Parameters used:  $R = 1$ ,  $D = 0.5$ ,  $b = 0.02$ . Number of realisations for each data point =  $10^5$ .

This method greatly reduces the error. However, the algorithm is not exact, and there is still error in the algorithm as it assumes a flat boundary rather than a curved boundary. Error is also introduced from a similar problem occurring on the boundary of the LN that the T cell is moving within, which will be addressed in the next section.

### 2.3.2 Boundary collisions

As with encountering an APC, error is also introduced at the boundary of the LN due to discrete timestepping. That is, during a timestep a T cell could have collided with the boundary, resulting in the MFTP being overestimated. This is illustrated in Figure 2.4.

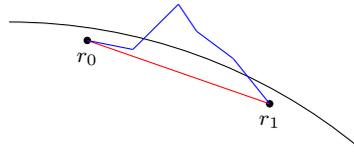


Figure 2.4: Diagram showing the situation that causes overestimation of the MFTP at the boundary of the domain. The black curve represents the boundary,  $r_0$  and  $r_1$  are the distances from the centre of the domain before and after the timestep. The red line shows the path the particle has taken, and the blue line a potential path the particle could have taken if a smaller timestep size was used, resulting in the particle leaving the domain.

This error can be reduced by using an algorithm to calculate the probability of a T cell hitting the boundary during a timestep. To do this we first need to define

$$dw = r_1 - r_0,$$

where  $r_0$  is the distance from the centre of the LN to the T cell at the start of the timestep, and  $r_1$  is the distance at the end of the timestep. The distance the T cell moves away from its starting position in one timestep is a random variable, given by

$$\max = \frac{dw}{2} + \left( \frac{dw^2 - 2dt \log(urv)}{2} \right)^{\frac{1}{2}},$$

[119] where  $urv$  is a uniform random variable, between zero and one [12–14]. To determine if the T cell has hit the boundary we need to check if  $r_0 + \max$  is greater than the radius of the LN ( $R$ ). If this occurs, we calculate a new distance;  $newr = dw - \max + R$ , and multiply each of the  $x$ ,  $y$  and  $z$  components by  $newr/r_1$  to reflect the T cell off the boundary and back into the LN.

We can examine how effective this method is by comparing the error in a spatial simulation, using this algorithm, to the results of Figure 2.3. This comparison is shown in Figure 2.5. Fitting a curve to the error in the data produced with the addition of a surface boundary algorithm gives an equation for the error of  $5.96\delta t^2$ . The error in the code has now been reduced to be proportional to  $\delta t^2$ .

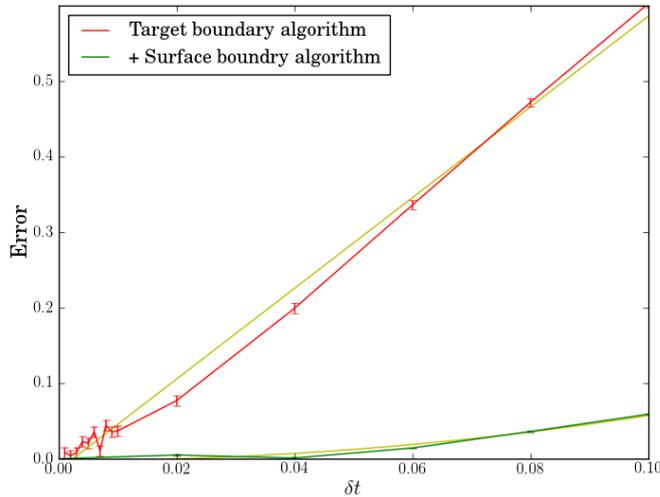


Figure 2.5: Plot showing the results for the error in the time a T cell takes to encounter an APC, when an algorithm is only used to reduce the error on the surface of an APC (red), and when an algorithm is additionally used to reduce the error at the boundary of the LN (green), with curves fitted to the data (yellow). Parameters used:  $R = 1$ ,  $D = 0.5$ ,  $b = 0.02$ . Number of realisations for each data point =  $10^5$ .

### 2.3.3 Exponential timestepping

Most numerical methods that are commonly used [45, 128, 143, 182] for solving stochastic differential equations and in spatial modelling, have a timestep of fixed length. However, it is also possible to have a timestep with a length that is a random variable [90, 168]. For example, a random variable chosen from an exponential distribution. Using exponential timestepping, over fixed timestepping, has the advantage that the probability of encountering a sphere during a timestep can be calculated exactly, using methods from excursion theory [39, 69, 116, 120, 122, 126, 189]. With exponential timestepping [117–119], instead of using timesteps of fixed size, timesteps that are drawn from an exponential distribution with a given mean timestep size are used. To implement

an exponential timestepping algorithm in three-dimensional space, the following steps must be undertaken:

- Define a factor  $mod$  as:  $mod = \sqrt{-2 \log(1 - urv_1)(g_1^2 + g_2^2 + g_3^2) \delta t D}$ , where  $urv_1$  is a uniform random variable between zero and one,  $g_1$ ,  $g_2$  and  $g_3$  are random numbers drawn from a Gaussian distribution with mean zero and variance one,  $\delta t$  is the mean timestep size, and  $D$  is the diffusion coefficient.
- Define the increment in the  $z$  dimension as  $dz_0 = 2urv_2 - 1$ , where  $urv_2$  is another uniform random variable between zero and one.
- Define an angle as  $\theta = 2\pi urv_3$ , where  $urv_3$  is yet another uniform random variable between zero and one.
- The increment in the  $x$  dimension is given by  $dx_0 = \sqrt{1 - dz_0^2} \cos(\theta)$ .
- The increment in the  $y$  dimension is given by  $dy_0 = \sqrt{1 - dz_0^2} \sin(\theta)$ .
- We now need to scale the  $x$ ,  $y$  and  $z$  increments by multiplying by the function  $mod$ . Therefore  $dx = mod \cdot dx_0$ ,  $dy = mod \cdot dy_0$  and  $dz = mod \cdot dz_0$

To determine if a T cell has encountered an APC, or the boundary of the LN, during a timestep, we must first calculate the distance of the T cell from its nearest APC and the distance of the T cell to the centre of the LN. Let us denote these as  $d_0$  and  $r_0$ , respectively. We now move the T cell, as outlined above, and recalculate the distance from the nearest APC, and the distance from the centre of the LN, and define these distances as  $d_1$  and  $r_1$ .

The probability of encountering an APC is calculated by

$$\mathcal{P}(\text{encounter}) = \exp(-\nu(r_e - b)) \frac{\sinh(\nu b)}{\sinh(\nu r_e)},$$

where  $\nu = 1/(\delta t D)$ ,  $b$  is the effective radius of an APC, and  $r_e = \min(d_0, d_1)$  [119]. We then generate a uniform random variable, between zero and one, if it is less than the calculated probability, the T cell is classified as encountering the APC.

Calculating the probability of hitting, and being reflected off, the boundary of the LN is similar. We denote the probability by  $\mathcal{P}(\text{reflect})$  as follows,

$$\mathcal{P}(\text{reflect}) = \exp(-\nu(R - r_r)) \frac{\sinh(\nu r_r)}{\sinh(\nu R)},$$

where  $r_r = \max(r_0, r_1)$ , and  $R$  is the radius of the LN [119]. We then generate a uniform random variable, between zero and one, and if  $\mathcal{P}(\text{reflect})$  is less than the random variable, the T cell is classed as hitting the boundary, and its position must be updated to correct for reflecting off the boundary. An example of a python code that uses exponential timestepping is shown in B.

Figure 2.6 compares exponential timestepping to fixed timestepping with algorithms to reduce the error on the surface of the APC and the boundary of the LN.

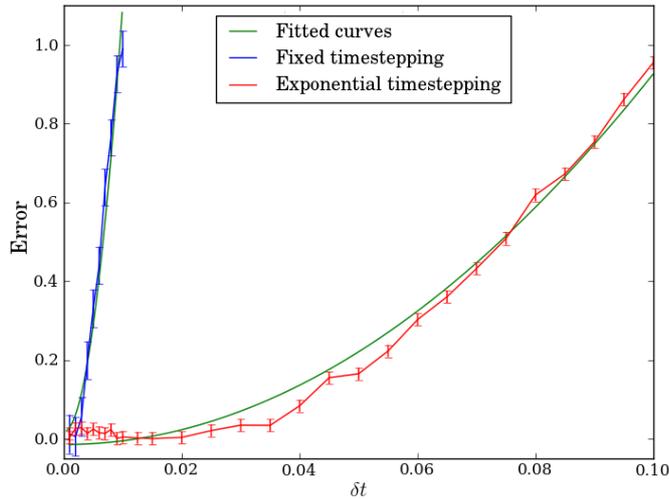


Figure 2.6: Comparison of error using Gaussian timestepping with algorithms to reduce the error on the boundary of the APC and the LN (blue) and exponential timestepping (red). Curves have been fitted to this data (shown in green). Parameters used:  $R = 1$ ,  $D = 0.5$ ,  $b = 0.01$ . Number of realisations for each data point =  $10^5$ .

## 2.4 Gillespie algorithm

Modelling a system using a deterministic approach regards the time evolution of the system as a continuous predictable process that is frequently governed by a set of coupled ODEs. The solution to a system of ODEs, at a given time  $t$ , is fixed for a given set of initial conditions. A stochastic approach regards the time evolution of the system to be continuous, and is governed by a single differential-difference equation, often referred to as the master equation. The Gillespie algorithm [83, 84] allows exact numerical calculations within the framework of the stochastic formalisation without having to deal with the master equation directly. The algorithm uses a

Monte-Carlo procedure to numerically simulate the time evolution of a given system, and it correctly accounts for the inherent fluctuations and corrections that are necessarily ignored in the deterministic formalisation. The algorithm never approximates infinitesimal time increments by finite steps, and is more realistic for modelling biological systems such as numbers of molecules, cells, or individuals, as population numbers can only change in discrete integer steps, unlike in the deterministic approach, which assume population numbers are continuous. Due to the stochastic nature of the Gillespie algorithm it captures behaviour in which populations can switch from one steady state to another, which can include extinction, due to random fluctuations in the system. These fluctuations are especially significant for small cell numbers and are lost in the deterministic approximation.

In this section we set out to solve the following general problem: If a fixed volume contains a spatial uniform mixture of  $n$  cell species which can interact through  $m$  specified reactions/interactions, then given the initial number of cells in these species, at a specific time, what will these cell population levels be at any later time? The most common approach to solving this type of problem is to translate the series of reactions into a set of ordinary differential equations (ODEs). These ODEs are known as the reaction rate equations. Solving them for the cell population numbers, given a specific set of initial conditions, is tantamount to solving the general problem proposed earlier. Analytical solutions to the reaction rate equations can be found only for rather simple cases, and it is usually necessary to solve these system numerically using a computer. The usefulness of this approach for biological systems cannot be denied, however, it does have its drawbacks. The approach assumes the time evolution of a biological process is both continuous and deterministic. However, the time evolution of a biological process is evidently not continuous as cell numbers can only change in discrete integer amounts. Nor is the time evolution a deterministic process as it is impossible, even in principle, to predict the exact population levels at some future time, unless we take into account the exact positions and velocities of all the molecules in the system. In many cases the time evolution of a biological system can, to an acceptable degree, be treated as a continuous deterministic process. However, this should not always be taken for granted as in some case the reaction rate equations inability to describe the fluctuations in the cell population levels can lead to serious shortcomings. Moreover, it is not guaranteed that the reaction rate equations will provide a sufficiently accurate account of the average molecular population levels, as, except for very simple linear systems, the average molecular population levels will not exactly satisfy any closed system of equations [83, 84]. The Gillespie algorithm [83, 84] is a method for computationally solving the time evolution of a biological spatially homogeneous

system that explicitly takes account of the fact that the system is a discrete, stochastic process instead of a continuous, deterministic process. The Gillespie algorithm offers an alternative to the differential reaction rate equations that is free of the difficulties mentioned above.

### 2.4.1 Physical basis of the stochastic formulation [84]

In general a chemical/biological reaction occurs whenever two or more molecules/cells of appropriate kind collide in an appropriate manner. Consider a system comprising of a mixture of two cell types,  $S_1$  and  $S_2$ , which are in equilibrium within a volume,  $V$ . For simplicity it is assumed the cells are spheres of radii  $r_1$  and  $r_2$  respectively. Therefore, a 1-2 collision will occur whenever the distance between the centres of the two cells decreases to  $r_{12} = r_1 + r_2$ . The traditional method to calculate the rate at which these interactions occur is to begin by choosing an arbitrary 1-2 pair of cells, and denoting the speed of cell 1 relative to cell 2 as  $v_{12}$ . We then observe that in the next small time interval  $\delta t$ , cell 1 will sweep out relative to cell 2 a ‘collision volume’  $\delta V_{coll} = \pi r_{12}^2 v_{12} \delta t$  [84] in the sense that if the centre of cell 2 lies within the  $\delta V_{coll}$  at time  $t$ , then the two cells will collide in the time interval  $(t, t + \delta t)$ . We now estimate the number of  $S_2$  cells that have a centre that lies within  $\delta V_{coll}$ , divide that number by  $\delta t$ , and take the limit as  $\delta t \rightarrow 0$  to obtain the rate at which the  $S_1$  cells collide with the  $S_2$  cells. However, as  $\delta V_{coll} \rightarrow 0$ , the number of  $S_2$  cells that have a centre within  $\delta V_{coll}$  will either be 1 or 0, with the zero possibility becoming more and more likely. Therefore, it is physically meaningless to talk about the number of cells with centres laying within  $\delta V_{coll}$  in the limit of vanishingly small  $\delta t$ . However, as the system is in equilibrium, the cells will at all times be distributed randomly and uniformly throughout the volume. Therefore, the probability that the centre of an arbitrary  $S_2$  cell will be located within  $\delta V_{coll}$  at time  $t$  will be given by the ratio  $\delta V_{coll}/V$ , which is true in the limit of vanishingly small  $\delta V_{coll}$ . If we average this ratio over the velocity distribution of the two cell types, it can be concluded that

$$\overline{\delta V_{coll}/V} = V^{-1} \pi r_{12}^2 \overline{v_{12}} \delta t,$$

which equals the average probability that a particular 1-2 cell pair will collide in the next vanishingly small time interval  $\delta t$ . If at time  $t$  there are  $N_1$  of the  $S_1$  cells and  $N_2$  of the  $S_2$  cells in  $V$ , giving a total of  $N_1 N_2$  distinct 1-2 cell pairs, then it follows that the probability that a 1-2 collision will occur within  $V$  in the next infinitesimal time interval  $(t, t + dt)$  is given by

$$N_1 N_2 \pi r_{12}^2 \overline{v_{12}} dt.$$

Although the number of 1-2 collisions occurring in the infinitesimal time interval cannot rigorously be calculated, the probability of a 1-2 collision occurring can be. This means the system must be classified by a ‘collision probability per unit’ time rather than a ‘collision rate’. This is why these collisions constitute a stochastic process instead of a deterministic process.

### 2.4.2 The stochastic simulation algorithm

To be able to simulate a biological system of reactions we require a way of specifying when the next reaction will occur and what kind of reaction it will be. What is required is a method for generating a pair  $(dt, \mu)$ , where  $dt$  is the time increment and  $R_\mu$  is the reaction that will occur, from the set of random pairs whose probability density function is  $P(dt, \mu)$ , given by [83]:

$$P(dt, \mu) = \begin{cases} a_\mu \exp(-a_0 dt) & \text{if } 0 \leq dt < \infty \text{ and } \mu = 1, \dots, M, \\ 0 & \text{Otherwise.} \end{cases} \quad (2.4.2)$$

In which  $a_\mu = h_\mu c_\mu$  and  $a_0 = \sum_{i=1}^M h_i c_i$ , where  $h_\mu$  is the number of distinct  $R_\mu$  cell reactant combinations available in the state  $(N_1, N_2, \dots, N_n)$ , ( $\mu = 1, \dots, m$ ), and  $c_\mu$  is a constant that depends on the physical properties of the reacting cells and the system.

It turns out that there is a simple rigorous way of doing this using uniform random numbers. This achieved as follows [82]: Using two uniform random numbers,  $u_1$  and  $u_2$ , take  $dt = (1/a_0) \ln(1/u_1)$ , and take  $\mu$  to be the integer for which  $\sum_{i=1}^{\mu-1} a_i < u_2 a_0 \leq \sum_{i=1}^{\mu} a_i$ . A rigorous proof can be found in [83]. This procedure generates a random number  $dt$  according to the probability density function  $P_1(dt) = a_0 \exp(-a_0 dt)$  and an integer  $\mu$  according to the probability density function  $P_2(\mu) = a_\mu/a_0$ . The stated result follows, roughly speaking, because  $P_1(dt)P_2(\mu) = P(dt, \mu)$ .

### 2.4.3 Computational algorithm

This section addresses the computational algorithm developed in [83, 84] used to solve the time evolution of a biological system. An example of a generalised Gillespie algorithm for simulating a series of reactions is outlined. For the purposes of this thesis the algorithm is assumed to be simulating interactions between various cell types. The algorithm in this subsection is simulating the interaction between  $n$  populations of cells with population sizes of

$(N_1, N_2, \dots, N_n)$ . These populations interact following the interactions  $(R_1, R_2, \dots, R_m)$  with corresponding rates  $(r_1, r_2, \dots, r_m)$ .

**Step 0 (Initialisation):** Input desired values for the reaction rates  $(r_1, r_2, \dots, r_m)$ , for reactions  $(R_1, R_2, \dots, R_m)$ , and initial population numbers  $(N_1, N_2, \dots, N_n)$ .

**Step 1:** Calculate the ‘reaction parameters’  $(a_1, a_2, \dots, a_m)$  by multiplying the reaction rate by the size of each of the populations required for that reaction to occur, and the number of each of the populations required.

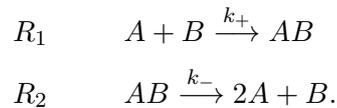
**Step 2:** Calculate two uniform random variables,  $u_1$  and  $u_2$ , in the interval zero to one. Use these numbers to generate: (i) the time elapsed,  $dt$ , for the next reaction to occur by sampling from an exponential distribution,  $dt = -\log(u_1) / \sum_{i=1}^m a_i$ , and (ii) the reaction that has occurred in this time interval. To determine the reaction that has occurred, all possible reactions must be scaled into the interval zero to one. We do this for the first reaction ( $R_1$ ) by assigning the interval between 0 and  $a_1 / \sum_{i=1}^m a_i$ , then for reactions  $R_2$  to  $R_m$  the interval from  $\sum_{i=k}^m a_i / \sum_{i=1}^m a_i$ , where  $2 < k < m$ . The reaction that occurs is then determined by the interval  $u_2$  falls into.

**Step 3:** Adjust the time elapsed by adding  $dt$ , and adjust the population levels to reflect the occurrence of the reaction determined in step 2. Then return to step 1.

This process continues until a desired time or population number is reached.

#### 2.4.4 Simple example

To illustrate the Gillespie algorithm, let us consider a simple example in which we have two cells  $A$  and  $B$ , that bind with rate  $k_+$  and unbind with rate  $k_-$ , producing two cells of type  $A$ . We have the reactions



**Step 0:** Let us start with population size of  $N_A$ ,  $N_B$  and  $N_{AB}$  for populations  $A$ ,  $B$  and  $AB$  respectively, where  $AB$  is the complex of  $A$  and  $B$  bound to each other. We will continue our simulation until the time is greater than some maximum time,  $t_{stop}$ .

**Step 1:** We have  $a_1 = k_+ N_A N_B$  and  $a_2 = k_- N_{AB}$ .

**Step 2:** The  $\sum_i a_i = a_1 + a_2$ , therefore  $dt = -\log(u_1) / (a_1 + a_2)$ . We have the intervals:  $I_1 = \left[0, \frac{a_1}{(a_1 + a_2)}\right]$  for  $R_1$ , and  $I_2 = \left[\frac{a_1}{(a_1 + a_2)}, 1\right]$  for  $R_2$ . The reaction that has occurred is

determined using the uniform random number  $u_2$ .

**Step 3:** We now increment the time by  $dt$ ,  $t = t + dt$ , and adjust the population numbers accordingly. If  $u_2 \in I_1$ ,  $N_A = N_A - 1$ ,  $N_B = N_B - 1$  and  $N_{AB} = N_{AB} + 1$ , and if  $u_2 \in I_2$ ,  $N_A = N_A + 2$ ,  $N_B = N_B + 1$  and  $N_{AB} = N_{AB} - 1$ . We now check if  $t > t_{stop}$ , if it is, we end the simulation, if not, we return to step 1.

## 2.5 Green's functions

There are many great books on Green's functions [21, 233], therefore only a brief description will be given. Green's functions are used to solve inhomogeneous differential equations, with boundary conditions. For an arbitrary differential operator  $L$ , which in this thesis will always be the Laplacian, the Green's function  $G(\mathbf{x}|\mathbf{x}')$  is defined as

$$LG(\mathbf{x}|\mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}'), \quad (2.5.3)$$

where  $\delta$  is the delta function. The solution to  $L\phi(\mathbf{x}) = f(\mathbf{x})$ , is given by

$$\phi(\mathbf{x}) = \int G(\mathbf{x}, \mathbf{x}')f(\mathbf{x}')d\mathbf{x}'. \quad (2.5.4)$$

## 2.6 Other modelling efforts

### 2.6.1 Cellular Potts model of the lymph node

The Cellular Potts model (CPM) is a generalisation of the Ising model, which is used to model magnetic interactions [99]. In the CPM, the domain is divided into a lattice of discrete sites, and cells are described by the lattice sites they occupy [67]. This method allows the cells to have complex shapes that can change. However, it does not model the internal structure of the cell. The motion and interactions of cells is determined by an energy function [57]. The energy function typically depends upon adhesion between cells, volume of a cell, and the concentration of chemical attractants. The method is updated using a modified Metropolis Monte-Carlo algorithm [167]. The Metropolis Monte-Carlo algorithm is a method for obtaining random samples from a probability distribution for which direct sampling is difficult. A modification is applied for use in higher dimensions. The Monte-Carlo algorithm chooses random updates for the lattice positions according to how the energy of the system changes. For example, an optimal cell volume is

normally set, if a cell's volume is below the optimal value, the energy of the system would decrease by increasing the volume of the cell, so the overall volume is closer to the optimal volume. An increase in the contact area between cells will cause an increase in the energy of the system. If a cell moves towards a chemical attractant, the energy of the system will decrease, as the cell is moving in the chemically favoured direction. The algorithm accepts the changes with a probability defined by the change in the overall energy of the system. Lower energy states are accepted with a higher probability than higher energy states.

The Cellular Potts model allows detailed simulation of cell membrane movements. This thesis is concerned with cells interacting, but not with how their membrane structures interact. The cellular Potts model is therefore not an appropriate model for use in this thesis, as such a high level of detail is not required.

### **2.6.2 Movement on the reticular network**

It has been proposed that, rather than moving in three-dimensional space, T cells move along the reticular network (RN), which fills the T cell zone of the LN [92,121,123,225]. It has been reported that the RN acts as a conduit system, transporting soluble antigen to DCs residing in the LN [208], and the cells of the RN secrete the T cell survival factor IL-7 [135]. There is experimental evidence for T cells moving along the RN, and their changes in direction being correlated with branches of the RN [16, 18, 196].

Donovan and Lythe [73] develop a model in which T cells move along a RN, by first constructing a network, and then confining all motion to the network. Using a computational approach, they construct a network by defining sets of vertices and the edges joining them. They assume each vertex in the network is connected to exactly three edges, motivated by the idea of a network growing by branches splitting into two. Various hypotheses are considered for the movement of cells along the edges of the network. They also explore the effect moving on a RN has on the interactions between T cells and APCs, as some APCs reside on the network [129]. It has also been proposed that restricting motion to the RN could be a way of increasing the rate of encounters between cells [56]. It is found that restricting movement to a RN does not necessarily increase the frequency of encounters, compared to movement by Brownian motion in three-dimensional space.

## 2.7 Chapter summary

This thesis focuses on the interactions and dynamics of lymphocytes, within LNs. The stochastic nature of these interactions will be considered by using spatial and Gillespie models. Chapter 3 investigates the first passage processes of the time a T cell takes to locate an APC within a LN. Chapters 4 and 5 focus on *in vivo* imaging experiments, analysing experimental data, and determining the effects of confinement in a limited sized imaging volume when calculating the diffusion coefficient of T cells. Chapter 6 makes use of experiments carried out by the Bousso group, in Paris, to answer the question of, how many DCs are required to initiate a T cell response? Chapter 7 investigates effector and regulatory T cell interactions by developing a stochastic formalisation of the deterministic cross-regulation model [49, 133]. Finally, Chapter 8 will make use of mathematical models to investigate the three hypotheses for T cell activation, proposed in [42].

### 2.7.1 Chapter 3

The work in Chapter 3 aims to determine the time a T cell takes to first encounter an APC within a LN. The time for this encounter to occur is of critical importance to the activation of a T cell response. The chapter builds up on the simple scenario of a particle in a known position attempting to locate a fixed central target, to the case when a particle in a uniform random unknown starting position is attempting to locate one of  $N$  fixed targets located in unknown uniform random positions. This is achieved by solving Poisson's equation, and making use of Green's functions. The work in Chapter 3 greatly benefited from recent theoretical advances in the mathematics describing diffusing objects in a confined region, with one or more targets [58, 61].

### 2.7.2 Chapter 4

The effect of confinement in a limited size imaging volume is investigated in Chapter 4. Since the beginning of the twenty-first century advances have been made in *in vivo* imaging techniques, through the development and improvement of TPLSM. These experiments generate extensive data sets, but are limited by the size of the volume that can currently be imaged. A limited sized imaging volume causes a confinement effect to be observed at later times in the experiments, as cells that take looping paths will remain for longer in the imaging volume. Chapter 4 provides an

estimation of the point at which confinement should be observed, due to a limited volume size.

### 2.7.3 Chapter 5

Chapter 5 makes use of work carried out in Chapter 4 to analyse data obtained from TPLSM experiments. The chapter investigates experimental data to determine the type of motion being observed, and calculate the motility of the cells imaged. A systematic algorithm is developed to determine the type of motion and calculate a diffusion coefficient of cells in a TPLSM data set. The work in this chapter, as well as work in Chapter 4, provides an insight into how the shape, and resolution, of the imaging volume used in TPLSM experiments can be improved to give a more informative data set.

### 2.7.4 Chapter 6

Chapter 6 attempts to answer the question of, how many DCs are required to initiate a T cell response? This work was carried out in collaboration with experimentalists in Paris. Biological experiments allow the probability of activation to be estimated using large numbers of cells, on a short timescale. Mathematical modelling allowed extrapolation of the experimental results to lower cell numbers and longer timescales, enabling a minimum number of DCs required to initiate a T cell response, to be determined.

### 2.7.5 Chapter 7

Chapter 7 extends work done by Léon *et al.* [49, 133], to investigate the effect of stochastic fluctuations on the cross-regulation model. The cross-regulation model is a model of three way interactions between APCs, effector T cells and regulatory T cells. A stochastic model gives a more realistic description of biological processes, allowing variability to be observed in the system. This approach reveals aspects of the system previously not observed the the deterministic model.

### 2.7.6 Chapter 8

In Chapter 8, three hypotheses of T cell activation [40, 42] are investigated. T cells are known to undergo three distinct phases in their activation process: (1) transient interactions, (2) long-lived

interactions, and (3) swarming. Three hypotheses have been proposed to explain this process; APC changes, T cell signal integration, and a probabilistic approach. Chapter 8 models these three hypotheses, with an emphasis on the probabilistic approach.



## Chapter 3

# First passage processes

### 3.1 Introduction

First passage processes [184, 226] have relevance to many practical areas, including chemical [100, 187] and biochemical reactions [30, 31, 63, 138], searching for food by animals [28] and transport processes on complex networks [3, 4, 29, 166, 169, 222]. Mean first passage times (MFPTs) associated with a random walker moving in a bounded domain, with one or more targets are a critical step in transport-limited reactions [187], such as immune cell interactions, for which the time for cells to encounter each other is a limiting factor in the initiation of an immune response. MFPTs are also used in other areas of immunology, such as calculating the arrival time of receptors at localised reactive sites on the surface of immune cells, which is the key step in the signalling process, and therefore activation [62] of the cell upon encounter. This chapter focuses on the larger scale of cell-cell encounters.

The aim of this chapter is to derive an equation for the target global mean first passage time (TGMFPT) of a particle, which represents a T cell, starting in a uniform random position and moving by Brownian motion within a bounded spherical domain that excludes the volume of the target, representing a lymph node (LN), to encounter one or more uniformly randomly placed fixed spherical targets, representing antigen-presenting-cells (APCs). The particles are assumed to move with a constant diffusivity, and the targets are assumed to be stationary. The domain has a reflecting boundary, keeping the particle confined, and the target(s) have absorbing boundaries.

Extensive research has been carried out in the area of MFPTs, and important results have been published regarding the time for a random walker to reach a target site for the first time. This target

could either be within a bounded domain, or, on the surface of the domain. Other first passage related quantities have also been derived, such as splitting probabilities, which is the probability to reach one target before another [3,4,27,31,35,58–62,94,103,149,181,205–207,230]. Particularly useful to the research in this chapter, Condamin *et al.* [60] determine starting global mean first passage times (SGMFPTs), and splitting probabilities, in the presence of one, and two targets, in both discrete and continuous time, on the interior of two and three-dimensional domains. First passage processes for more than two targets [58,59,62] have also been studied. Results of these studies are used in this chapter. Other studies have found expansions for the principal eigenvalues of the Laplacian, MFPT, and splitting probabilities for various domains [59,62], allowing optimal positions of targets to be calculated in order to minimise search time.

This chapter begins by deriving Poisson's equation, which is used to calculate a MFPT for a single target in the centre of a sphere, with a particle undergoing Brownian motion searching for it. A MFPT for a single target being 'close' to the centre of the sphere, where 'close' is defined as being within one target radius, using an approximation proposed by Grigin [93], is investigated. To calculate a solution for a uniformly randomly placed target within the domain, a review of papers by Condamin *et al.* [60] and Chevalier *et al.* [58], in which they derive an equation for a target in a fixed position anywhere within the domain, is undertaken. Using this, an equation is derived for a uniformly randomly placed target, by integrating over all possible starting positions within the domain. The final step is to derive a TGMFPT for multiple targets, in the domain. This is accomplished by making use of equations for  $N$  targets, in fixed positions within the domain, derived in [58,60]. Integrating the equation for multiple fixed targets, over the starting position of each of the targets, results in an equation for multiple uniformly randomly placed targets within a sphere. All the results in this chapter are tested against numerical results, generated using spatial codes (Section 2.2.3).

The results obtained in Sections 3.2 and 3.3 have been derived previously for a multitude of applications. The approximation in Section 3.4 was developed by Girgin [93]. To the best of my knowledge this approximation has not been previously used to calculate the first passage quantities derived in Section 3.4. The work in Section 3.5 to 3.8 to calculate the SGMFPT is a review of work carried out by [58,61]. The calculations to compute the TGMFPT for one and  $N$  targets has not been previously undertaken to the best of my knowledge.

The aim of this chapter is to derive an equation for the time a particle starting in a uniform random position, which excludes the volume occupied by the target, and is undergoing Brownian motion,

takes to encounter a target in a fixed uniform random position. This is analogous to a T cell attempting to encounter an APC within a LN. Obtaining this equation will allow a rate of encounter between T cells and APCs to be calculated for a given set of conditions. This rate can then, and will be, used in this thesis to simulate interactions between T cells and APCs both computationally and analytically.

The SGMFPT refers to the MFPT for a particle starting in a uniform random unknown position, which excludes the volume occupied by the target, to locate a target starting in a fixed known uniform random position. The TGMFPT refers to the MFPT for a particle starting in a uniform random unknown position, which excludes the volume occupied by the target, to locate a target starting in a fixed unknown uniform random position.

<b>Description</b>	<b>Vector</b>	<b>Radius</b>
Initial position of particle	$\mathbf{r}_s$	$r_s$
Target	$\mathbf{r}_T$	$r_T$
Position of particle along path	$\mathbf{r}$	$r$

Table 3.1: Table listing the parameters used to represent the positions of particles and targets. Vector refers to a three-dimensional position and radius refers to the distance from the centre of the domain.

<b>Description of positions</b>	<b>Abbreviation</b>	<b>Notation</b>
Prescribed particle, central target	MFPT	$\tau_0(r_s)$
Prescribed particle, prescribed random target	MFPT	$\tau(r_T, r_s)$
Unprescribed random particle, prescribed random target	SGMFPT	$\bar{\tau}(r_T)$
Unprescribed random particle, unprescribed random target	TGMFPT	$T$

Table 3.2: Table listing notation used to represent first passage processes. Prescribed implies the starting position is known, unprescribed that the starting position is unknown. All random positions are uniform random and exclude the volume occupied by the target in the case of the particles position.

### 3.2 Poisson's Equation [184]

Assuming a symmetric random walk, in one dimension, where there is equal probability of moving in both directions, on the interval  $[x_-, x_+]$ , with time increment  $\delta t$ , the mean time to encounter the boundary,  $t(x)$ , starting from position  $x$ , is the sum of the times for each exit path, multiplied by the probability of that path, averaged over all paths.

$$t(x) = \sum_i P_i(x) t_i(x), \quad (3.2.1)$$

where  $t_i(x)$  is the exit time of path  $i$ , to the boundary, and  $P_i$  is the probability of path  $i$ . Figure 3.1 illustrates (3.2.1).

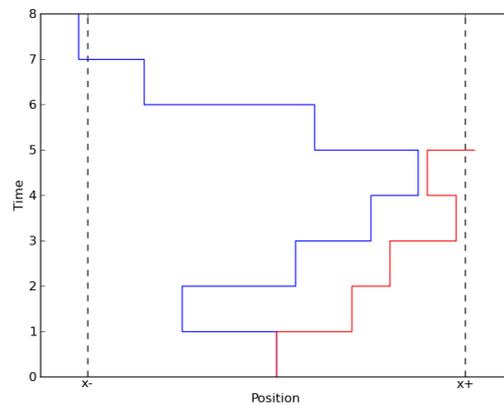


Figure 3.1: Diagram to illustrate (3.2.1). A particle starts at a prescribed position, within the domain, and moves a distance chosen from a uniform distribution, independent of previous movements, in fixed discrete timesteps until it crosses the boundary of the domain.

This mean time, obeys the recursion formula

$$t(x) = \frac{1}{2}[t(x + \delta x) + \delta t] + \frac{1}{2}[t(x - \delta x) + \delta t], \quad (3.2.2)$$

where  $\delta x$  is the increment moved in the  $x$  dimension, and the boundary conditions are  $t(x_-) = t(x_+) = 0$ : the time to encounter the boundary is zero, if starting at the boundary. The relation expresses the mean time starting at  $x$  in terms of the outcome one step in the future. The walk can be viewed as restarting at  $x + \delta x$  or  $x - \delta x$  after time interval  $\delta t$  has passed, each with probability one half. Using the Taylor expansion on  $t(x + \delta x)$  and  $t(x - \delta x)$ , as  $\delta x \rightarrow 0$ , gives

$$t(x + \delta x) = t(x) + \delta x t'(x) + \frac{1}{2} \delta x^2 t''(x) + \dots \quad (3.2.3)$$

$$t(x - \delta x) = t(x) - \delta x t'(x) + \frac{1}{2} \delta x^2 t''(x) + \dots \quad (3.2.4)$$

Substituting equations (3.2.3) and (3.2.4) into equation (3.2.2) gives

$$t(x) = \delta t + t(x) + \frac{1}{2} \delta x^2 t''(x) + \dots \quad (3.2.5)$$

Dividing by  $\delta t$ , and rearranging, gives

$$Dt''(x) = -1, \quad (3.2.6)$$

where  $D$  is the diffusion coefficient. Brownian motion is the limit of a random walk as  $\delta t \rightarrow 0$ , with the limit  $D = \frac{1}{2} \lim_{\delta t \rightarrow 0} \frac{\delta x^2}{\delta t}$ . This can be extended to three dimensions to give Poisson's equation

$$D\nabla^2\tau = -1, \quad (3.2.7)$$

where  $\tau$  is the mean time starting from position  $\mathbf{r}_s$ , which is a three-dimensional vector.

### 3.3 Central target

A MFPT for a particle moving by Brownian motion, within a sphere, to find a target located in the centre of the sphere, can be calculated by making use of Poisson's equation (3.2.7). This problem is illustrated in Figure (3.2).

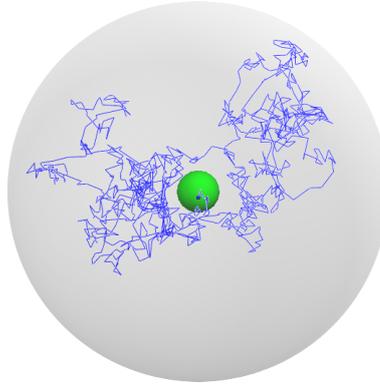


Figure 3.2: Diagram showing a Brownian path (blue) to a target (green) located in the centre of a sphere.

The expected time to encounter a central target,  $\tau_0$ , within a sphere, satisfies Poisson's equation (3.2.7), with a reflecting boundary condition on the surface of the sphere,  $\left. \frac{d\tau_0(r_s)}{dr_s} \right|_{r_s=R} = 0$ , and an absorbing boundary condition on the surface of the target,  $\tau_0(r_s = b) = 0$ , where  $b$  is the size of the target,  $r_s = |\mathbf{r}_s|$  is the distance from the centre to the starting position of the particle, and  $R$  is the radius of the spherical domain.

As the target is in the centre of the sphere there is symmetry in both the azimuthal and polar angles, therefore  $\tau_0(r_s)$  only depends on the distance to the particle's initial position. Using spherical polar coordinates Poisson's equation (3.2.7) becomes

$$\frac{1}{r_s^2} \frac{\partial}{\partial r_s} \left( r_s^2 \frac{\partial}{\partial r_s} \tau_0 \right) = \frac{\partial^2}{\partial r_s^2} \tau_0 + \frac{2}{r_s} \frac{\partial}{\partial r_s} \tau_0 = -\frac{1}{D}. \quad (3.3.8)$$

Equation (3.3.8) has the general solution

$$\tau_0(r_s) = -\frac{1}{6D} r_s^2 - \frac{c_1}{r_s} + c_2, \quad (3.3.9)$$

where  $c_1$  and  $c_2$  are integration constants. Using the boundary conditions to find  $c_1$  and  $c_2$  in (3.3.9), gives

$$c_1 = \frac{R^3}{3D}, \quad c_2 = \frac{1}{6D} \left( b^2 + \frac{2R^3}{b} \right),$$

so that

$$\tau_0(r_s) = \frac{R^3}{3Db} - \frac{R^3}{3Dr_s} - \frac{r_s^2}{6D} + \frac{b^2}{6D}. \quad (3.3.10)$$

### 3.3.1 Particle in a random starting position

The SGMFPT for a particle starting in a uniform random position within the sphere, excluding the volume occupied by the target, to find a central target, can be found by integrating equation (3.3.10) over all the possible starting positions in the sphere, and dividing by the volume of the sphere, minus the volume of the target, as follows:

$$\int_0^{2\pi} \int_0^\pi \sin \theta \int_b^R r_s^2 \tau_0(r_s) dr_s d\theta d\phi = 4\pi \int_b^R \tau_0(r_s) r_s^2 dr_s, \quad (3.3.11)$$

as there is spherical symmetry. Now

$$\frac{4\pi}{6D} \int_b^R \left[ 2R^3 \left( \frac{r_s^2}{b} - r_s \right) + r_s^2 b^2 - r_s^4 \right] dr_s = \frac{2\pi}{45Db} \left[ 2R^5(5R - 9b) + b^5(3b - 5) + 10b^3 R^3 \right]. \quad (3.3.12)$$

Dividing by the volume of the sphere minus the volume of the target;  $(4\pi(R^3 - b^3)/3)$ , gives the equation

$$\bar{\tau}_0 = \frac{R^3}{R^3 - b^3} \left( \frac{R^3}{3Db} - \frac{3R^2}{5D} + \frac{b^2}{3D} - \frac{b^4}{6D} + \frac{b^5}{10D} \right), \quad (3.3.13)$$

where  $\bar{\tau}_0$  is used to represent the SGMFPT. For  $b \ll R$ , this becomes

$$\bar{\tau}_0 \approx \frac{R^3}{3Db} - \frac{3R^3}{5D}. \quad (3.3.14)$$

### 3.4 Target ‘close’ to centre

The aim is to determine an equation for the time a particle moving by random motion, within a sphere, takes to find a uniformly randomly placed target. This proves to be simple when the target is in the centre of the sphere, as there is spherical symmetry. However, if the target is not in the centre there is only symmetry in the azimuthal angle, therefore Poisson’s equation becomes

$$\nabla^2 \tau(r_T, r_s, \theta) = \frac{1}{r_s^2} \frac{\partial}{\partial r_s} \left( r_s^2 \frac{\partial \tau}{\partial r_s} \right) + \frac{1}{r_s^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \tau}{\partial \theta} \right) = -\frac{1}{D}. \quad (3.4.15)$$

Separation of variable leads to the expression

$$\tau(r_T, r_s, \theta) = \frac{A_0}{r_s} + B_0 - \frac{1}{6D} r_s^2 + \sum_{n=0}^{\infty} (A_n r_s^n + B_n r_s^{n+1}) P_n(\cos \theta), \quad (3.4.16)$$

where  $\theta$ , is the polar angle between the particle and target,  $r_T$  is the position of the target, the  $P_n$ ’s are Legendre polynomials, and  $A_0$ ,  $B_0$ ,  $A_n$  and  $B_n$  are constants. As a first approximation (3.4.16) is expanded to order  $\cos \theta$  terms giving

$$\tau(r_T, r_s, \theta) = \frac{A_0}{r_s} + B_0 - \frac{1}{6D} r_s^2 + \left( \frac{A_1}{r_s^2} + B_1 r_s \right) \cos \theta. \quad (3.4.17)$$

At the surface of the sphere, we have a reflecting boundary condition, which implies

$$\frac{\partial \tau(r_T, r_s, \theta)}{\partial r_s} = 0 \quad \text{for} \quad r_s = R, \quad (3.4.18)$$

as in the previous case. The boundary condition on the surface of the target, is not as straight forward as in the previous case. A paper by Grigin [93], proposes that an equation for the position on the surface of a sphere ‘close’ to the centre of another sphere is given by  $r_s = b(1 + \epsilon \cos \theta)$ , where  $\epsilon$  is a small parameter given by  $r_T/b$ , and  $r_T$  is the distance of the target from the centre of the sphere. For this approximation to hold, the condition  $r_T < b$  must be satisfied (see Figure 3.3). Therefore, the approximation is only useful when the target is ‘close’ to the centre of the sphere.

Using this approximation, the boundary condition for the surface of the target is given by

$$\tau(r_T, r_s, \theta) = 0 \quad \text{for} \quad r_s = b(1 + \epsilon \cos \theta). \quad (3.4.19)$$

Using boundary conditions (3.4.18) and (3.4.19), we find

$$\begin{aligned} A_0 &= -\frac{R^3}{3D}, & B_0 &= \frac{1}{6Db} (2R^3 + b^3), \\ A_1 &= \frac{r_T R^3}{3D} \left( \frac{b^3 - R^3}{2b^3 + R^3} \right), & B_1 &= \frac{2r_T}{3D} \left( \frac{b^3 - R^3}{2b^3 + R^3} \right). \end{aligned}$$

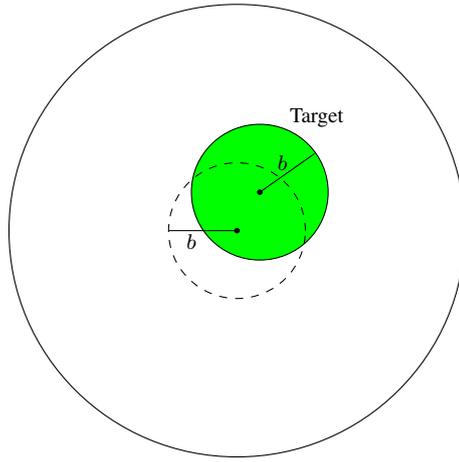


Figure 3.3: Setup of a problem with the centre of a target (green) within one target radius ( $b$ , indicated by dashed line) of the centre of the sphere.

Giving the MFPT equation

$$\tau_c(r_s, r_T, \theta) = \frac{R^3}{3Db} - \frac{R^3}{3Dr_s} - \frac{r_s^2}{6D} + \frac{b^2}{6D} + (b^3 - R^3) \frac{r_T(R^3 + 2r_s^2)}{3Dr_s^2(2b^3 + R^3)} \cos \theta, \quad (3.4.20)$$

where the  $c$  subscript refers to the ‘close’ approximation. When  $r_T \rightarrow 0$ , we approach the case where the target is in the centre,

$$\tau_c(r_s, r_T, \theta) \rightarrow \frac{R^3}{3Db} - \frac{R^3}{3Dr_s} - \frac{r_s^2}{6D} + \frac{b^2}{6D}, \quad (3.4.21)$$

agreeing with (3.3.10). Equation (3.4.20) can be tested by comparing to spatial simulations (Section 2.2.3), for varying angles ( $\theta$ ) around the centre. Figure 3.4 shows (3.4.20) is a better approximation to the MFPT than assuming the target is in the centre, as in (3.3.13).

### 3.4.1 Random particle starting position

The SGMFPT for a particle, starting in a uniform random position within a sphere, excluding the volume occupied by the target, to find a target, that has a centre within radius  $b$  of the centre of the sphere, can be found by integrating (3.4.20) over all possible starting positions of the particle, as follows

$$\int_0^{2\pi} \int_0^\pi \sin \theta \int_b^R r_s^2 \tau_c(r_s, r_T, \theta) dr_s d\theta d\phi = 2\pi \int_0^\pi \sin \theta \int_{(b+r_T)}^R r_s^2 \tau_c(r_s, r_T, \theta) dr_s d\theta. \quad (3.4.22)$$

We have the lower limit of  $(b+r_T)$  on the integral over  $r_s$ , to excluded the particle from starting in a position where the target could be placed. This will lead to a larger volume being excluded than

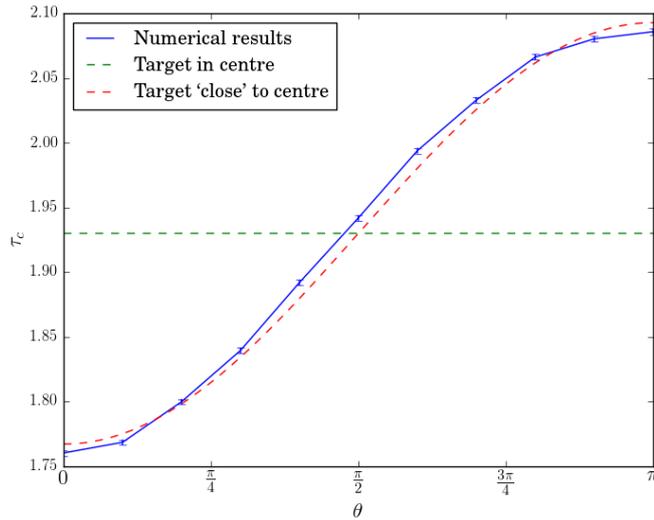


Figure 3.4: Comparison of numerical results from spatial simulations to the MFPT equation for a target in the centre of a sphere (3.3.13) and the MFPT equation for a target ‘close’ to the centre (3.4.20). Parameters used:  $b = 0.2$ ,  $r_s = 0.5$ ,  $r_T = 0.05$ , and  $D = 0.5$ . Number of realisations used to generated each point =  $10^6$ .

the target occupies, giving an upper bound on the SGMFPT. We must also divide by the volume in which the particle could be placed, given by the volume of the sphere minus the volume of  $(b + r_T)$ . The solution of (3.4.22) divided by the volume is given by

$$\begin{aligned} \bar{\tau}_c(r_T) &= \frac{3}{4\pi(R^3 - (b + r_T)^3)} 2 \int_0^\pi \sin \theta \int_{(b+r_T)}^R r_s^2 \tau_c(r_T, r_s, \theta) dr_s d\theta \\ &= \frac{R^3}{R^3 - (b + r_T)^3} \left( \frac{R^3}{3Db} + \frac{r_T^3}{3Db} - \frac{3R^2}{5D} + \frac{r_T^5}{10DR^3} - \frac{r_T^2}{2D} + \frac{r_T^4 b}{2DR^3} + \frac{b^2}{3D} \right. \\ &\quad \left. + \frac{5r_T^3 b^2}{6DR^3} + \frac{r_T^2 b^3}{2DR^3} - \frac{b^5}{15DR^3} \right). \end{aligned} \quad (3.4.23)$$

For  $b \ll R$ , this implies  $r_T \ll R$ , as  $r_T \leq b$ , (3.4.23) becomes

$$\bar{\tau}_c(r_T) \approx \frac{R^3}{3Db} - \frac{3R^2}{5D}, \quad (3.4.24)$$

which agrees with (3.3.14).

Figure 3.5 compares (3.4.23) to numerical simulations (described in Section 2.2.3). Starting positions of the target were chosen to satisfy the condition for this approximation, of being within one target radius of the centre of the sphere. The plots show the equation is a good approximation to the numerical results. There is very little difference between the three cases, implying that the target’s position, within the valid domain, does not have a large effect on the SGMFPT.

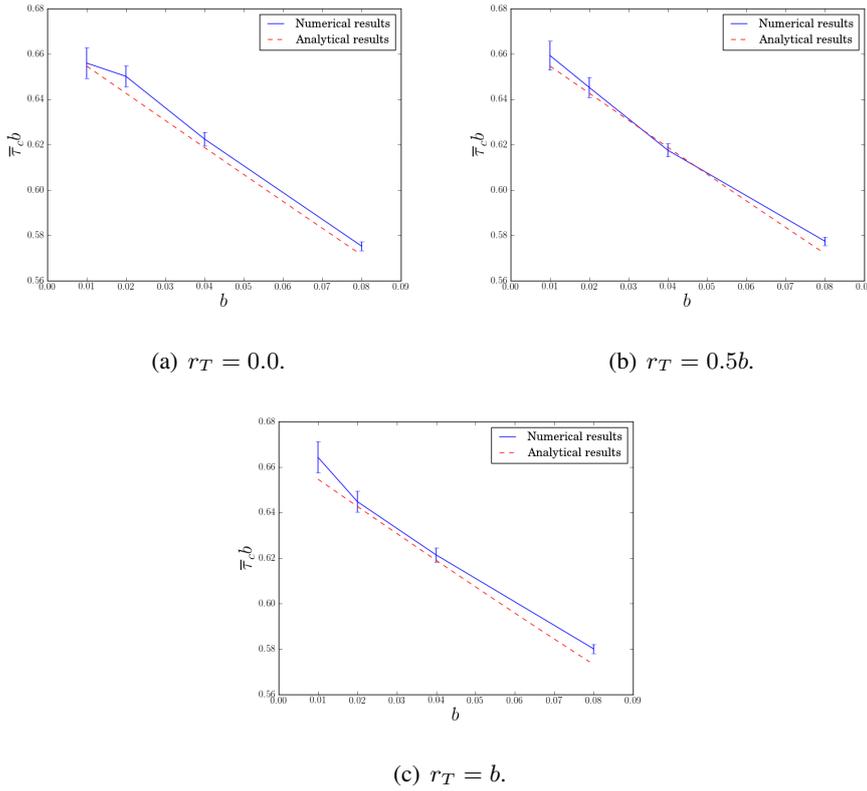


Figure 3.5: Comparison of (3.4.23), dotted red line, against numerical simulations, blue, for varying starting positions as a function of the target size. Parameters used:  $R = 1.0$ , and  $D = 0.5$ . Number of realisations used to generate each point ranges from  $10^4$  to  $10^5$ .

### 3.4.2 Random target within $b$ of centre

The TGMFPT for a target in a uniform random position within the valid region can be found by integrating (3.4.23) over all possible starting positions for the target, between zero and  $b$ . We must then divide by the volume of this region,  $\frac{4}{3}\pi b^3$ , to obtain the TGMFPT,

$$\begin{aligned}
 T_c &= \frac{3}{4\pi b^3} \int_0^{2\pi} \int_0^\pi \sin(\theta) \int_0^b \bar{\tau}_c(r_T) r_T^2 dr_T d\theta d\phi \\
 &= \frac{3}{4\pi b^3} \int_0^b \bar{\tau}_c(r_T) 4\pi r_T^2 dr_T \\
 &= -\frac{1}{10b^3 D} \left[ \frac{43b^5}{30} - \frac{37b^2 R^3}{3} - 6\sqrt{3}R^3 (b^2 + 4bR + R^2) \left( \tan^{-1} \left( \frac{2b+R}{\sqrt{3}R} \right) \right. \right. \\
 &\quad \left. \left. - \tan^{-1} \left( \frac{4b+R}{\sqrt{3}R} \right) \right) 18R^3 (b-R)(b+R) \tanh^{-1} \left( \frac{b(3b+R)}{5b^2 + 3bR + 2R^2} \right) \right]. \quad (3.4.25)
 \end{aligned}$$

For  $b \ll R$ , upon performing a series expansion of the trigonometric terms, (3.4.25) becomes

$$T_c \approx \frac{R^3}{3Db} - \frac{3R^2}{5D}. \quad (3.4.26)$$

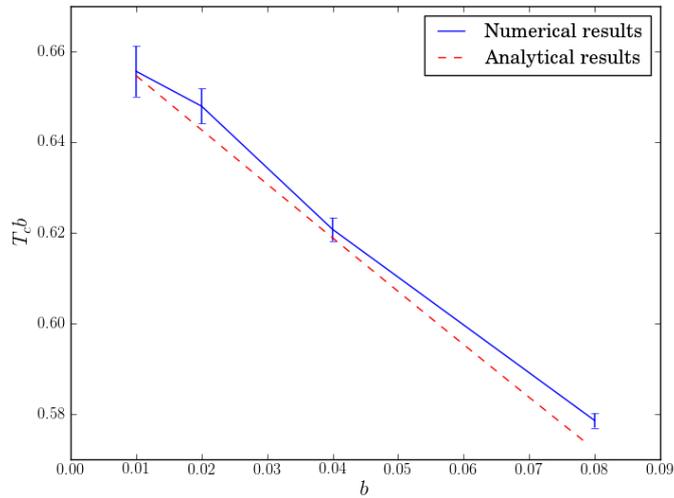


Figure 3.6: Comparison of numerical results and the TGMFPT for a target ‘close’ to the centre of a sphere. Parameters used:  $R = 1.0$  and  $D = 0.5$ . Number of realisations used to generate each point ranges from  $10^4$  to  $10^5$ .

A comparison of (3.4.25) against numerical results can be seen in Figure 3.6. Equation (3.4.25) is found to be a good fit to the numerical results, when the target is placed uniformly randomly with its centre within one target radius of the centre of the sphere and the particle is placed in a uniform random position, which in the equation is not within  $b + r_T$  of the centre, and for the numerical simulations is not within the radius of attraction of the target.

## 3.5 Randomly placed target

Determining an equation for the MFPT when a target is not in the centre of a sphere proves to be more difficult. By using a bispherical coordinate system, the boundary conditions on the target surface can be expressed more easily. However, upon investigation of this method it was found that the system was very complex to solve, and other methods were investigated.

### 3.5.1 Solution using Green’s functions

Chevalier *et al.* [58], derive a formula for the SGMFPT, using Green’s functions, and pseudo Green’s functions, of a point particle, starting in a uniform random position within a spherical

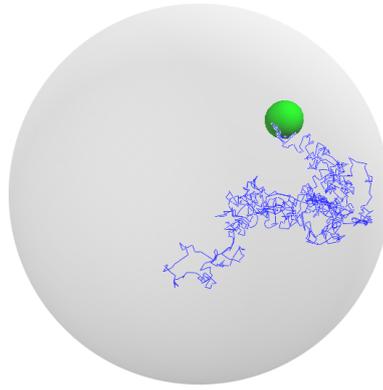


Figure 3.7: Diagram showing a Brownian path (blue), to a target (green) located in a uniform random position within a sphere.

domain of radius one, excluding the volume occupied by the target, moving with a diffusivity  $D$ , to find a target sphere of radius  $b$ , at a radius  $r_T$  from the centre of the domain. This is illustrated in Figure 3.8. This section reviews the work carried out by [58, 61] to derive a SGMFPT. I then make use of this work to derive a TGMFPT given by (3.5.37).

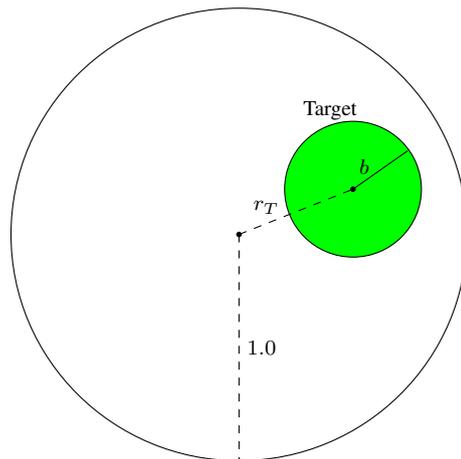


Figure 3.8: Diagram of the problem setup with a uniform randomly placed target.

The equation found for the SGMFPT [58] is

$$\bar{\tau}(\mathbf{r}_T) = \frac{1}{3Db} + \frac{4\pi}{3D} H^*(\mathbf{r}_T|\mathbf{r}_T), \quad (3.5.27)$$

where

$$H^*(\mathbf{r}_T|\mathbf{r}_T) = \frac{1}{4\pi(1-r_T^2)} + \frac{1}{4\pi} \log\left(\frac{1}{1-r_T^2}\right) + \frac{r_T^2}{4\pi} - \frac{7}{10\pi}. \quad (3.5.28)$$

Giving

$$\bar{\tau}(\mathbf{r}_T) = \frac{1}{3Db} + \frac{1}{3D} \left( \frac{1}{1-r_T^2} - \log(1-r_T^2) + r_T^2 - \frac{14}{5} \right). \quad (3.5.29)$$

The derivation of (3.5.29) is given in the next subsection. As  $r_T \rightarrow 0$ , which is the case when the target is in the centre, equation (3.5.29) becomes

$$\bar{\tau}(\mathbf{r}_T) = \frac{1}{3Db} - \frac{3}{5D}. \quad (3.5.30)$$

This is equivalent to the case when the target is in the centre (3.3.13), with  $b \ll R$ , and  $R = 1$ .

To compute the TGMFPT for a particle starting in a uniform random position, which excludes the volume occupied by the target, to find a target, also placed in a uniform random position, the part of (3.5.29) that is dependent on  $r_T$ , must be integrated over all possible locations of the target. Let us introduce  $\gamma$  as follows

$$\gamma = \frac{3}{4\pi(1-b)^3} \int_0^{(1-b)} \left( \frac{1}{1-r_T^2} - \log(1-r_T^2) + r_T^2 - \frac{14}{5} \right) 4\pi r_T^2 dr_T, \quad (3.5.31)$$

so that

$$\gamma = \frac{3}{(1-b)^3} \int_0^{(1-b)} \left( \frac{r_T^2}{1-r_T^2} - r_T^2 \log(1-r_T^2) + r_T^4 - \frac{14r_T^2}{5} \right) dr_T. \quad (3.5.32)$$

Equation (3.5.32) can be split into three integrals

$$\begin{aligned} I_1 &= \int_0^{(1-b)} \left( \frac{r_T^2}{1-r_T^2} \right) dr_T \\ &= \frac{1}{2} \left( -2(1-b) + \log \left( \frac{2-b}{b} \right) \right), \end{aligned} \quad (3.5.33)$$

$$\begin{aligned} I_2 &= - \int_0^{(1-b)} r_T^2 \log(1-r_T^2) dr_T \\ &= -\frac{1}{9} (2b^3 - 6b^2 + 12b - 8 - 3b(b^2 - 3b + 3) \log(b(b-2)) + 6 \log(2-b)) \end{aligned} \quad (3.5.34)$$

$$\begin{aligned} I_3 &= \int_0^{(1-b)} r_T^4 - \frac{14r_T^2}{5} dr_T \\ &= \left( -\frac{b^5}{5} + b^4 - \frac{16b^3}{15} - \frac{4b^2}{5} + \frac{9b}{5} - \frac{11}{15} \right). \end{aligned} \quad (3.5.35)$$

Combining integrals  $I_1$ ,  $I_2$ , and  $I_3$  gives

$$\begin{aligned} \gamma &= \frac{1}{30(1-b)^3} (30b(b^2 - 3b + 3) \log(b(2-b)) - 45 \log(b) - 15 \log(2-b) - 18b^5 + 90b^4 \\ &\quad - 116b^3 - 12b^2 + 132b - 76). \end{aligned} \quad (3.5.36)$$

Therefore, the TGMFPT is given by

$$T = \frac{1}{3Db} - \frac{\gamma}{3D}, \quad (3.5.37)$$

where  $R = 1$ . For  $b \ll 1$ , this becomes

$$T = \frac{1}{3Db} - \frac{1}{90D}(45 \log(b) + 15 \log(2) + 76). \quad (3.5.38)$$

### 3.5.2 Derivation of SGMFPT equation

The following derivation of the SGMFPT equation (3.5.29), is taken from the references by Chevalier [58], and Condamin [60]. In this section  $\mathbf{r}_s$ , is the starting position of the particle,  $\mathbf{r}_T$ , is the position of the target, and  $\mathbf{r}$  is a position along the path of the target, where boldface represents three-dimensional spatial vectors. This is illustrated in Figure 3.9.

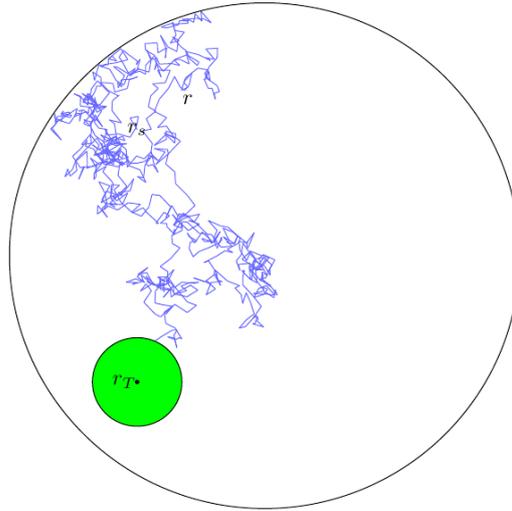


Figure 3.9: Diagram of problem setup. The green circle represents the target, and the blue path is the path of the particle.

The MFPT,  $\tau(\mathbf{r}_s)$ , for a particle, starting in position  $\mathbf{r}_s$ , to find a target in a domain,  $\mathcal{D}$ , satisfies the following equations [184, 188],

$$D\nabla^2\tau(\mathbf{r}_s) = -1 \text{ if } \mathbf{r}_s \in \mathcal{D}^*, \quad (3.5.39)$$

$$\tau(\mathbf{r}_s) = 0 \text{ if } \mathbf{r}_s \in S_T, \quad (3.5.40)$$

$$\partial_n\tau(\mathbf{r}_s) = 0 \text{ if } \mathbf{r}_s \in S_R, \quad (3.5.41)$$

where  $\partial_n$  is the normal derivative to the boundary of the domain. The first equation, (3.5.39), is Poisson's equation, and  $\mathcal{D}^*$  is the domain  $\mathcal{D}$ , minus the area occupied by the target. The second equation, (3.5.40), is an absorbing boundary condition on the surface of the target,  $S_T$ . The third

equation, (3.5.41), is a reflecting boundary condition on the boundary of the domain,  $S_R$ . To solve this problem, the Green's function  $G(\mathbf{r}|\mathbf{r}_s)$  is introduced, which is defined by

$$-\nabla^2 G(\mathbf{r}|\mathbf{r}_s) = \delta(\mathbf{r} - \mathbf{r}_s) \text{ if } \mathbf{r} \in \mathcal{D}^*, \quad (3.5.42)$$

$$G(\mathbf{r}|\mathbf{r}_s) = 0 \text{ if } \mathbf{r} \in S_T, \quad (3.5.43)$$

$$\partial_n G(\mathbf{r}|\mathbf{r}_s) = 0 \text{ if } \mathbf{r} \in S_R, \quad (3.5.44)$$

where  $\delta$  is the Dirac delta function with the property,  $\int_{\mathcal{D}^*} \delta(\mathbf{r} - \mathbf{r}_s) d\mathbf{r} = 1$ , and  $\mathbf{r}$  is the position of the particle along its path.

Green's second identity [213] states

$$\int_V (\phi_1 \nabla^2 \phi_2 - \phi_2 \nabla^2 \phi_1) dV = \oint_S (\phi_1 \partial_n \phi_2 - \phi_2 \partial_n \phi_1) dS, \quad (3.5.45)$$

where the first integral is over the volume of the domain and the second integral is over the surfaces of the target and domain. We must now choose  $\phi_1$  and  $\phi_2$ . Let  $\phi_1 = \tau(\mathbf{r})$  and  $\phi_2 = G(\mathbf{r}|\mathbf{r}_s)$ . This gives

$$\int_{\mathcal{D}^*} (\tau(\mathbf{r}) \nabla^2 G(\mathbf{r}|\mathbf{r}_s) - G(\mathbf{r}|\mathbf{r}_s) \nabla^2 \tau(\mathbf{r})) dV = \int_{S_a \cup S_r} (\tau(\mathbf{r}) \partial_n G(\mathbf{r}|\mathbf{r}_s) - G(\mathbf{r}|\mathbf{r}_s) \partial_n \tau(\mathbf{r})) dS. \quad (3.5.46)$$

From the boundary conditions (3.5.43) and (3.5.44), the right-hand-side becomes zero. Using (3.5.39) and (3.5.42), equation (3.5.46) becomes

$$\int_{\mathcal{D}^*} (-\tau(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_s) + G(\mathbf{r}|\mathbf{r}_s) \frac{1}{D}) dV = 0. \quad (3.5.47)$$

From the definition of the Dirac delta function, we get

$$\int_{\mathcal{D}^*} \tau(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_s) dV = \tau(\mathbf{r}_s), \quad (3.5.48)$$

when  $\mathbf{r} = \mathbf{r}_s$ , as  $\mathbf{r}_s$  is not contained within the domain,  $\mathcal{D}^*$ . Therefore the MFPT is given by

$$\tau(\mathbf{r}_s) = \frac{1}{D} \int_{\mathcal{D}^*} G(\mathbf{r}|\mathbf{r}_s) dV. \quad (3.5.49)$$

The function  $G(\mathbf{r}|\mathbf{r}_s)$  must now be determined.  $G(\mathbf{r}|\mathbf{r}_s)$  is a function constructed from pseudo-Green's functions, that gives a constant for the occupation time across all of the domain, except for a peak in the occupation time at the starting position of the particle, as on average it will spend the most time there, and a dip in the occupation time near the target, as if the particle is near the target there is a high probability it will encounter it.  $G(\mathbf{r}|\mathbf{r}_s)$  can be written as

$$G(\mathbf{r}|\mathbf{r}_s) = H(\mathbf{r}_T|\mathbf{r}_T) - H(\mathbf{r}_T|\mathbf{r}_s) + H(\mathbf{r}|\mathbf{r}_s) - H(\mathbf{r}|\mathbf{r}_T), \quad (3.5.50)$$

where

$$H(\mathbf{r}|\mathbf{r}_T) = \begin{cases} \frac{1}{4\pi|\mathbf{r} - \mathbf{r}_T|} + H^*(\mathbf{r}|\mathbf{r}_T) & \text{if } |\mathbf{r} - \mathbf{r}_T| > b, \\ \frac{1}{4\pi b} + H^*(\mathbf{r}_T|\mathbf{r}_T) & \text{if } |\mathbf{r} - \mathbf{r}_T| < b, \end{cases} \quad (3.5.51)$$

where  $H^*$  will be defined later. The second case is how the  $H(\mathbf{r}_T|\mathbf{r}_T)$  term is expressed. The  $G$  function in (3.5.50) is a function of three variables constructed from the  $H$  functions, which are functions of two variables. In (3.5.50), the  $H(\mathbf{r}|\mathbf{r}_s)$  term accounts for the peak in occupation time near the starting position and the  $H(\mathbf{r}|\mathbf{r}_T)$  term is the dip in occupation time near the target. The  $H$  terms are the pseudo-Green's function [21], which are needed to allow (3.5.50) to satisfy (3.5.42)–(3.5.44), and satisfy

$$-\nabla^2 H(\mathbf{r}|\mathbf{r}_s) = \delta(\mathbf{r} - \mathbf{r}_s) - \frac{1}{V} \text{ if } \mathbf{r} \in \mathcal{D}, \quad (3.5.52)$$

$$\partial_n H(\mathbf{r}|\mathbf{r}_s) = 0 \text{ if } \mathbf{r} \in S_r, \quad (3.5.53)$$

$$H(\mathbf{r}|\mathbf{r}_s) = H(\mathbf{r}_s|\mathbf{r}), \quad (3.5.54)$$

$$\int_{\mathcal{D}} H(\mathbf{r}_s|\mathbf{r}) d\mathbf{r}_s \equiv 0, \quad (3.5.55)$$

where the fourth equation can be derived from the previous three and  $V$  is the volume of the domain. The choice of  $G(\mathbf{r}|\mathbf{r}_s)$  is the simplest one that satisfies equations (3.5.42)–(3.5.44).

Checking if (3.5.50) satisfies (3.5.42), we get

$$\begin{aligned} -\nabla^2 G(\mathbf{r}|\mathbf{r}_s) &= -\nabla^2 H(\mathbf{r}_T|\mathbf{r}_T) + \nabla^2 H(\mathbf{r}_T|\mathbf{r}_s) - \nabla^2 H(\mathbf{r}|\mathbf{r}_s) + \nabla^2 H(\mathbf{r}|\mathbf{r}_T) \\ &= +\delta(\mathbf{r}_T - \mathbf{r}_T) - \frac{1}{V} - \delta(\mathbf{r}_T - \mathbf{r}_s) + \frac{1}{V} + \delta(\mathbf{r} - \mathbf{r}_s) - \frac{1}{V} - \delta(\mathbf{r} - \mathbf{r}_T) + \frac{1}{V} \\ &= \delta(\mathbf{r} - \mathbf{r}_s) \end{aligned}$$

using (3.5.52), and that  $\delta(\mathbf{r}_T - \mathbf{r}_T)$ ,  $\delta(\mathbf{r}_T - \mathbf{r}_s)$  and  $\delta(\mathbf{r} - \mathbf{r}_T)$  are only non zero in the target domain, which is excluded from  $\mathcal{D}^*$ . Checking if (3.5.50) satisfies (3.5.44), we get

$$\begin{aligned} \partial_n G(\mathbf{r}|\mathbf{r}_s) &= \partial_n H(\mathbf{r}_T|\mathbf{r}_T) - \partial_n H(\mathbf{r}_T|\mathbf{r}_s) + \partial_n H(\mathbf{r}|\mathbf{r}_s) - \partial_n H(\mathbf{r}|\mathbf{r}_T) \\ &= 0, \end{aligned}$$

using (3.5.53). Equation (3.5.43) can however only be approximately satisfied [58, 60].

We can now compute (3.5.49),

$$\begin{aligned}
\tau(\mathbf{r}_s) &= \frac{1}{D} \int_{\mathcal{D}^*} G(\mathbf{r}|\mathbf{r}_s) d\mathbf{r} \\
&= \frac{1}{D} \int_{\mathcal{D}^*} [H(\mathbf{r}_T|\mathbf{r}_T) - H(\mathbf{r}_T|\mathbf{r}_s) + H(\mathbf{r}|\mathbf{r}_s) - H(\mathbf{r}|\mathbf{r}_T)] d\mathbf{r} \\
&= \frac{1}{D} \int_{\mathcal{D}^*} [H(\mathbf{r}_T|\mathbf{r}_T) - H(\mathbf{r}_T|\mathbf{r}_s)] d\mathbf{r} \\
&= \frac{V}{D} [H(\mathbf{r}_T|\mathbf{r}_T) - H(\mathbf{r}_T|\mathbf{r}_s)] + \mathcal{O}(b^3 G_0(b)). \tag{3.5.56}
\end{aligned}$$

Using (3.5.54) and (3.5.55), as well as using the fact that the integral over  $\mathcal{D}^*$  is almost equal to the integral over  $\mathcal{D}$ , with the relative order of magnitude of the correction being  $b^3/V$ , in three dimensions, where  $b$  is the radius of the target.

Equation (3.5.56) is the MFPT for a particle starting at a position  $\mathbf{r}_s$ . The SGMFPT, for a particle starting in a uniform random position, which excludes the volume occupied by the target, is obtained by integrating over the volume of the domain, for all starting positions of the particle,  $\mathbf{r}_s$ , and is defined as

$$\begin{aligned}
\bar{\tau}(\mathbf{r}_T) &= \frac{1}{V} \int_{\mathcal{D}} \tau(\mathbf{r}_s) d\mathbf{r}_s \\
&= \frac{1}{V} \frac{V}{D} \int_{\mathcal{D}} [H(\mathbf{r}_T|\mathbf{r}_T) - H(\mathbf{r}_T|\mathbf{r}_s)] + \mathcal{O}(b^3 G_0(b)) d\mathbf{r}_s \\
&= \frac{1}{D} \int_{\mathcal{D}} [H(\mathbf{r}_T|\mathbf{r}_T) + \mathcal{O}(b^3 G_0(b))] d\mathbf{r}_s \\
&= \frac{V}{D} H(\mathbf{r}_T|\mathbf{r}_T) + \mathcal{O}(b^3 G_0(b)). \tag{3.5.57}
\end{aligned}$$

Again using (3.5.54) and (3.5.55).

In our case the bounded domain,  $\mathcal{D}$ , is a sphere, for which the pseudo-Green's function is known exactly [21] and reads

$$H(\mathbf{r}_T|\mathbf{r}_T) = G_0(\mathbf{r}_T - \mathbf{r}_T) + H^*(\mathbf{r}_T|\mathbf{r}_T), \tag{3.5.58}$$

with

$$G_0(\mathbf{r}_T - \mathbf{r}_T) = \frac{1}{4\pi b}, \tag{3.5.59}$$

and

$$H^*(\mathbf{r}_T|\mathbf{r}_T) = \frac{r_T}{4\pi|\mathbf{r}_T - r_T^2\mathbf{r}_T|} + \frac{1}{4\pi} \log A(\mathbf{r}_T) + \frac{1}{8\pi}(r_T^2 + r_T^2) - \frac{7}{10\pi}, \tag{3.5.60}$$

where

$$A(\mathbf{r}_T) = \frac{2r_T}{r_T - r_T(\mathbf{r}_T \cdot \mathbf{r}_T) + |\mathbf{r}_T - r_T^2\mathbf{r}_T|}. \tag{3.5.61}$$

Setting  $r_T = |\mathbf{r}_T|$ , gives

$$H(\mathbf{r}_T|\mathbf{r}_T) = \frac{1}{4\pi b} + \frac{1}{4\pi(1-r_T^2)} + \frac{1}{4\pi} \log\left(\frac{1}{1-r_T^2}\right) + \frac{r_T^2}{4\pi} - \frac{7}{10\pi}. \quad (3.5.62)$$

Substituting (3.5.62) into (3.5.57) gives

$$\bar{\tau}(\mathbf{r}_T) = \frac{1}{3Db} + \frac{1}{3D} \left( \frac{1}{(1-r_T^2)} - \log(1-r_T^2) + r_T^2 - \frac{14}{5} \right),$$

where  $R = 1$ . This is (3.5.29), which we set out to derive.

### 3.5.3 Testing the random target equation

Equation (3.5.29) can be tested against spatial simulations of a particle moving by Brownian motion within a sphere, until it encounters a target in a fixed position. Figure 3.10 shows a plot of the target size,  $b$ , against the SGMFPT,  $\bar{\tau}$ , multiplied by the size of the target, for varying starting positions of the target.

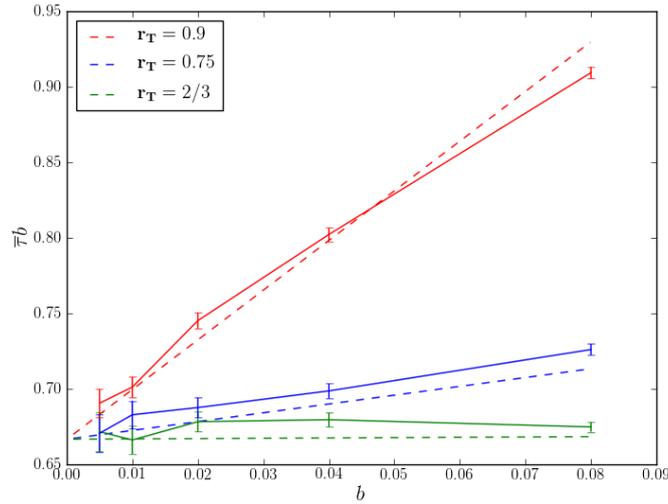


Figure 3.10: Target in a known starting position,  $r_T$ , from the centre of the sphere. Solid lines represent numerical simulations, and dashed lines represent equation (3.5.29). Parameters used:  $D = 0.5$  and  $R = 1$ . Number of realisations used to produce each point ranges from  $3 \times 10^3$  to  $6 \times 10^4$ .

The aim is to obtain a formula for the TGMFPT of a uniform randomly placed particle, excluding the volume occupied by the target, to find a uniform randomly located target. Equation (3.3.13) assumes the target is in the centre of the sphere, and (3.5.37) assumes the target is placed uniform

randomly, as desired. Figure 3.11 shows these two equations, along with numerical simulations, plotted for a varying target size,  $b$ . From Figure 3.11, it can be seen that (3.5.37) is a much better fit to the numerical results than (3.3.13).

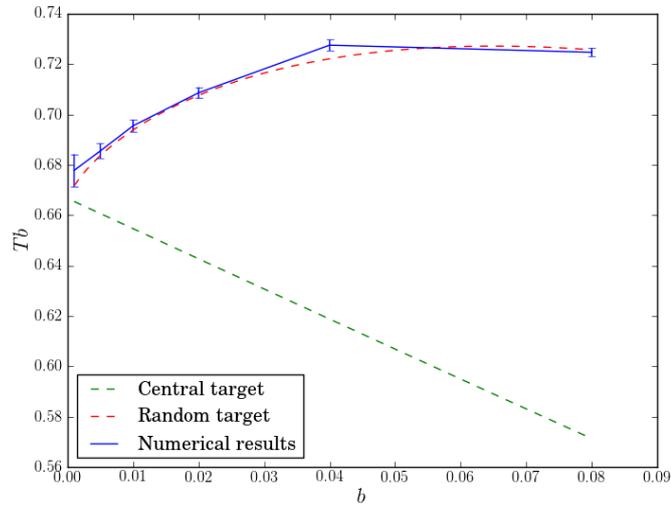


Figure 3.11: Comparison of (3.3.13) and (3.5.37) to numerical results from spatial simulations with target and particle placed in uniform random positions. Parameters used:  $D = 0.5$  and  $R = 1$ . Number of realisations used to produce each point ranges from  $10^4$  to  $2 \times 10^5$ .

### 3.5.4 The case when $R \neq 1$

In (3.5.38) it is assumed that the radius of the domain,  $R$ , is equal to one, for simplicity. To obtain an equation in which  $R \neq 1$ , a simple scaling of the other two parameters must be performed.  $b$  and  $D$  need to be divided by  $R$  and  $R^2$  respectively to obtain

$$T = \frac{R^3}{3Db} - \frac{R^2}{90D} \left( 45 \log \left( \frac{b}{R} \right) + 15 \log(2) + 76 \right). \quad (3.5.63)$$

This scaling can be verified by comparing (3.5.63) to numerical simulations. Figure 3.12 shows the scaling appears to be correct. The numerical results being at higher values than the analytical equation, is due to the timestep size used not being small enough. The timestep size was chosen to reduce the computational time required to execute the simulations.

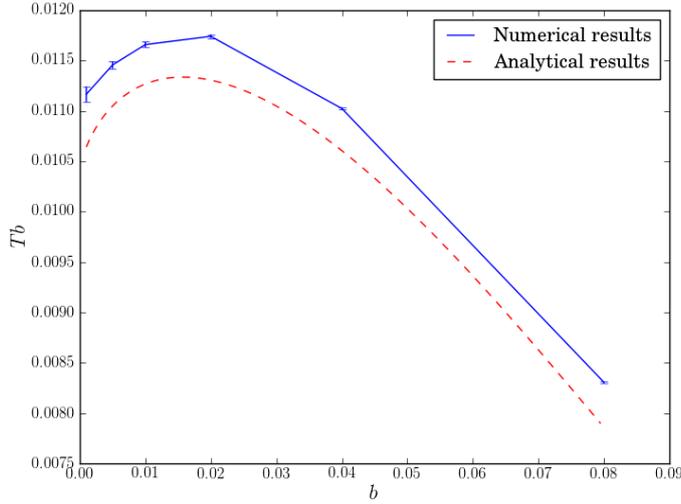


Figure 3.12: Comparison of (3.5.63), when  $R \neq 1$ , to numerical simulations. Parameters used:  $R = 0.25$ ,  $D = 0.5$ . Number of realisations used to produce each point ranges from  $2 \times 10^4$  to  $2 \times 10^6$ .

### 3.6 Two fixed targets

The next question to answer is: what is the SGMFPT for a particle to find one of multiple targets? Let us first consider the case when there are two absorbing spherically symmetric targets, in the domain. We must construct our Green's function  $G(\mathbf{r}|\mathbf{r}_s)$  from pseudo-Green's functions, taking into account that only one of the two targets will be encountered. In order to perform this calculation we need to introduce splitting probabilities for the probability of encountering target one before target two, and vice-versa. Splitting probabilities can be found by solving Laplace's equation with boundary conditions for the surface of the domain and the targets. The splitting probability  $P_i(\mathbf{r})$ ,  $i = 1, 2$  [184], for encountering target  $i$ , satisfies

$$D\nabla^2 P_i = 0, \quad (3.6.64)$$

$$\partial_n P_i(\mathbf{r}) = 0 \text{ if } \mathbf{r} \in \partial\mathcal{D}, \quad (3.6.65)$$

$$P_i(\mathbf{r}) = 1 \text{ if } \mathbf{r} \in \partial T_i, \quad (3.6.66)$$

$$P_i(\mathbf{r}) = 0 \text{ if } \mathbf{r} \in \partial T_j, \ j \neq i, \quad (3.6.67)$$

where  $\partial_n$  is the normal derivative,  $\partial\mathcal{D}$  is the surface of the domain, and  $\partial T_i$  and  $\partial T_j$  are the surfaces of targets  $i$  and  $j$ . The Green's functions for this equation can be written as

$$D\nabla^2 G(\mathbf{r}|\mathbf{r}_s) = \delta(\mathbf{r} - \mathbf{r}_s), \quad (3.6.68)$$

$$\partial_n P_i(\mathbf{r}) = 0 \text{ if } \mathbf{r} \in \partial\mathcal{D}, \quad (3.6.69)$$

$$G(\mathbf{r}|\mathbf{r}_s) = 1 \text{ if } \mathbf{r} \in \partial T_i, \quad (3.6.70)$$

$$G(\mathbf{r}|\mathbf{r}_s) = 0 \text{ if } \mathbf{r} \in \partial T_j, j \neq i. \quad (3.6.71)$$

Using Green's second identity (3.5.45), we find

$$P_i(\mathbf{r}_s, \mathbf{r}_T^{(1)}, \mathbf{r}_T^{(2)}) = - \int_{\partial T_i} \partial_n G(\mathbf{r}|\mathbf{r}_s) d\mathbf{r}, \quad (3.6.72)$$

$G(\mathbf{r}|\mathbf{r}_s)$  can now be written as a function of pseudo-Green's functions [61] as follows:

$$G(\mathbf{r}|\mathbf{r}_s) = \rho_0(\mathbf{r}_s) + H(\mathbf{r}|\mathbf{r}_s) - P_1 H(\mathbf{r}|\mathbf{r}_1) - P_2 H(\mathbf{r}|\mathbf{r}_2), \quad (3.6.73)$$

where  $\rho_0$  is yet to be determined,  $\mathbf{r}_T^{(1)}$ , and  $\mathbf{r}_T^{(2)}$ , are the positions of the two targets. The  $H(\mathbf{r}|\mathbf{r}_s)$  function accounts for a peak in probability near the starting position of the particle, the negative  $H$  terms account for dips in the probability near each target, and the  $P_i$  coefficients are the probability of being at that target. The pseudo-Green's functions are defined as in (3.5.52)–(3.5.55). Expression (3.6.73) satisfies (3.6.64) and (3.6.65). Equations (3.6.66), (3.6.67) and (3.6.73), along with the condition that the splitting probabilities must sum to one, give the following set of equations [61]:

$$\rho_0 + H_{1s} - P_1 H_{11} - P_2 H_{12} = 0, \quad (3.6.74)$$

$$\rho_0 + H_{2s} - P_1 H_{12} - P_2 H_{22} = 0, \quad (3.6.75)$$

$$P_1 + P_2 = 1, \quad (3.6.76)$$

where  $H_{ij} = H(\mathbf{r}_i|\mathbf{r}_j)$ , and the solution of the  $H$  function is the same as in (3.5.58). Expressions for  $P_1$ ,  $P_2$ , and  $\rho_0$  can be obtained from (3.6.74)–(3.6.76), giving

$$P_1 = \frac{H_{1s} + H_{22} - H_{2s} - H_{12}}{H_{11} + H_{22} - 2H_{12}}, \quad (3.6.77)$$

$$P_2 = \frac{H_{2s} + H_{11} - H_{1s} - H_{12}}{H_{11} + H_{22} - 2H_{12}}, \quad (3.6.78)$$

$$\rho_0 = P_1 H_{11} + P_2 H_{12} - H_{1s}, \quad (3.6.79)$$

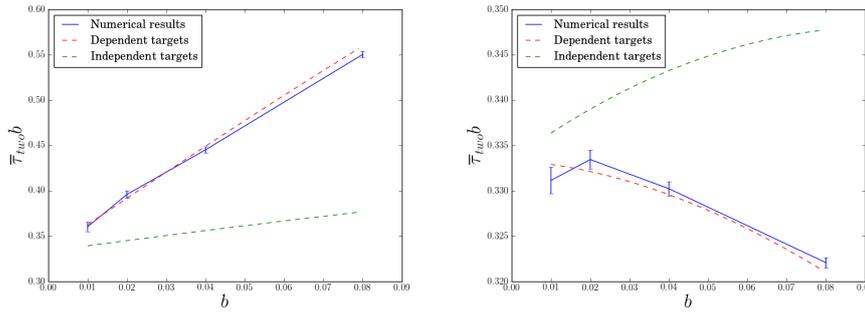
Substituting into (3.5.49) gives

$$\tau_{two} = \frac{V}{D} \frac{(H_{11} - H_{1s})(H_{22} - H_{2s}) - (H_{12} - H_{2s})(H_{12} - H_{1s})}{H_{11} + H_{22} - 2H_{12}}. \quad (3.6.80)$$

Averaging over all starting positions of the particle by integrating (3.6.80) over  $\mathbf{r}_s$ , using the fact that all  $H$  functions with an  $s$  subscript integrate to zero by (3.5.55), gives

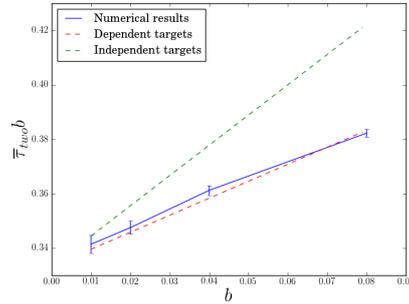
$$\bar{\tau}_{two} = \frac{V}{D} \frac{H_{11}H_{22} - H_{12}^2}{H_{11} + H_{22} - 2H_{12}}. \quad (3.6.81)$$

The results of testing (3.6.81) against numerical simulations are shown in Figure 3.13. In Figure 3.13, dependent target equation, refers to (3.6.81), and independent target equation assumes that the targets have no influence on each other, and is obtained by calculating the SGMFPT for each of the targets using (3.5.37), and combining the solutions using the formula  $1/(1/\bar{\tau}_1 + 1/\bar{\tau}_2)$ , where  $\tau_1$  and  $\tau_2$  are the SGMFPTs for target one and two, respectively.



(a) Targets close together.

(b) Intermediate case.



(c) Targets far apart.

Figure 3.13: Comparison of (3.6.80) and (3.5.37), to numerical simulations. The three cases examine the effect of the target's positions, relative to each other. Parameters used:  $R = 1$  and  $D = 0.5$ . Number of realisations used to produce each point ranges from  $5 \times 10^3$  to  $3 \times 10^5$ .

### 3.7 Equation for multiple targets

A general equation for the time to encounter one of  $N$  targets is given in [58] as

$$\bar{\tau}_N = \bar{\tau}_N^{(-1)} + \bar{\tau}_N^{(0)} + \bar{\tau}_N^{(1)}, \quad (3.7.82)$$

where

$$\bar{\tau}_N^{(-1)} = \frac{1}{3DS}, \quad (3.7.83)$$

$$\bar{\tau}_N^{(0)} = \frac{4\pi}{3DS^2} \left( \sum_{k=1}^N \sum_{j \neq k} b_j b_k H_{jk} + \sum_{k=1}^N b_k^2 H_{kk}^* \right), \quad (3.7.84)$$

and

$$\begin{aligned} \bar{\tau}_N^{(1)} = \frac{16\pi}{3DS^3} & \left[ \left( \sum_{k=1}^N b_k^2 H_{kk}^2 \right)^2 + \sum_{l=1}^N b_l^2 H_{ll}^2 \sum_{j \neq k} b_j b_k H_{jk} \right. \\ & \left. + \frac{S^2}{4\pi} \sum_{k=1}^N \sum_{l \neq k} \bar{P}_j^{(1)} b_k H_{jk} - S \left( \sum_{k=1}^N \sum_{j \neq k} b_j b_k^2 H_{jk} H_{kk}^* + \sum_{k=1}^N b_k^3 H_{kk}^{*2} \right) \right], \quad (3.7.85) \end{aligned}$$

with

$$\bar{P}_j^{(1)} = \frac{4\pi b_j}{S^2} \left[ \sum_{k=1}^N \sum_{l \neq k} b_k b_l H_{lk} - \sum_{k=1}^N \sum_{l \neq j} b_k b_l H_{lj} - \sum_{k=1}^N b_k (b_j H_{jj}^* - b_k H_{kk}^*) \right], \quad (3.7.86)$$

in which  $S = \sum_{i=1}^N b_i$ ,  $b_i$  is the radius of target  $i$ , and  $H_{ij}^* = H_{ij} - G_0(r_i - r_j)$ . In our case we have  $b_i = b$  for all  $i$ , so that  $\bar{P}_j^{(1)} = 0$  for all  $j \in \{1, 2, \dots, N\}$ .

### 3.7.1 Two randomly placed targets

For  $N = 2$  we get

$$\bar{\tau}_{two}^{(-1)} = \frac{1}{6Db}, \quad (3.7.87)$$

$$\bar{\tau}_{two}^{(0)} = \frac{\pi}{3D} \left( H_{11} + H_{22} + 2H_{12} - \frac{1}{2b\pi} \right), \quad (3.7.88)$$

$$\bar{\tau}_{two}^{(1)} = -\frac{2\pi^2 b}{3D} (H_{11}^2 + H_{22}^2 - 2H_{11}H_{22}), \quad (3.7.89)$$

using  $H_{ii}^* = H_{ii} - \frac{1}{4\pi b}$ . Numerically comparing these expressions to (3.6.81), it is found that the two expressions, derived by different methods, give approximately the same result. Expressions (3.7.87)-(3.7.89) will be used in what follows, as they are easier to manipulate and integrate.

To calculate the TGMFPT, averaged over all positions of the two targets within the domain, we need to integrate (3.7.87)-(3.7.89) over the position of the first target, then the second target, and

divide by the volume of the domain for each of the integrals, with  $R = 1$ .

$$\left(\frac{3}{4\pi(1-b)^3}\right)^2 \int_0^{1-b} \int_0^{1-b} (4\pi)^2 \bar{\tau}_{two}^{(-1)} r_1^2 r_2^2 dr_1 dr_2 = \frac{1}{6Db}, \quad (3.7.90)$$

$$\left(\frac{3}{4\pi(1-b)^3}\right)^2 \int_0^{1-b} \int_0^{1-b} (4\pi)^2 \bar{\tau}_{two}^{(0)} r_1^2 r_2^2 dr_1 dr_2 = \frac{1}{2}T - \frac{1}{6Db}, \quad (3.7.91)$$

$$\begin{aligned} \left(\frac{3}{4\pi(1-b)^3}\right)^2 \int_0^{1-b} \int_0^{1-b} (4\pi)^2 \bar{\tau}_{two}^{(1)} r_1^2 r_2^2 dr_1 dr_2 \\ = -b\pi \left( \frac{1}{D(1-b)^3} \beta - \frac{3D}{4\pi} T^2 \right), \end{aligned} \quad (3.7.92)$$

where  $T$  is the TGMFPT to find one uniform randomly placed target given in (3.5.37) and

$$\beta = \int_0^{1-b} 4\pi H_{TT}^2 r_T^2 dr_T. \quad (3.7.93)$$

To calculate (3.7.90)–(3.7.92), we make use of the property that when integrating over all starting positions  $H_{TT} = H_{ii}$ , and  $H_{TT}^2 = H_{ii}^2$ , for all  $i \in \{1, 2, \dots, N\}$ . We also use the following property of the pseudo-Green's functions

$$\int_0^{1-b} H_{ij} dr_i = 0, \quad (3.7.94)$$

given in (3.5.55). Finally we use the result (3.5.37), to write

$$\int_0^{1-b} 4\pi r_T^2 H_{TT} dr_T = D(1-b)^3 T. \quad (3.7.95)$$

Note that in (3.7.90)–(3.7.92), the first two terms sum to give  $T/2$ , which was our first approximation for the TGMFPT of two targets, assuming the targets were independent, implying the TGMFPT scales linearly. The third term is essentially a correction term that takes into account the effects of the other targets. Summing the terms (3.7.90)–(3.7.92) gives the TGMFPT for two targets

$$T_{two} = \frac{1}{2}T - b\pi \left( \frac{1}{D(1-b)^3} \beta - \frac{3D}{4\pi} T^2 \right). \quad (3.7.96)$$

The  $\frac{1}{2}T$  in (3.7.96) gives the TGMFPT assuming the targets are independent, and the second term is a correction term. Using numerical comparisons, it is found that the analytical correction term ( $b\pi(\dots)$ ) is always positive. Therefore,  $T_{two}$  is always less than  $\frac{1}{2}T$ . This implies that, assuming the targets are independent results in an overestimation of the TGMFPT. A plot of (3.7.96) is shown in Figure 3.14(a).

### 3.7.2 $N$ randomly placed targets

To calculate the TGMFPT for  $N$  targets, we again make use of the properties that were used to calculate (3.7.96). Rewriting (3.7.84) and (3.7.85) in a more general form by combining terms

that will integrate to give the same result, they can be made easier to integrate (the general form below is for integration purposes only):

$$\bar{\tau}_N^{(-1)} = \frac{1}{3DNb}, \quad (3.7.97)$$

$$\bar{\tau}_N^{(0)} = \frac{4\pi}{3D(Nb)^2} (NH_{ii}) - \frac{1}{3DNb}, \quad (3.7.98)$$

$$\bar{\tau}_N^{(1)} = -\frac{16\pi^2 b}{3DN^3} \left( N(N-1)H_{ii}^2 - \frac{N(N-1)}{2} 2H_{ii}H_{jj} \right), \quad (3.7.99)$$

where  $i \neq j$ . This general form is obtained by eliminating all the  $H_{ij}$  terms, which will integrate to zero, and examining the expansion of terms upon computing the sums, to combine terms. Integrating (3.7.97)–(3.7.99) over the domain  $N$  times, for the  $N$  targets, and dividing by the volume of the domain  $N$  times, gives

$$\begin{aligned} & \left( \frac{3}{4\pi(1-b)^3} \right)^N \int_0^{1-b} \dots \int_0^{1-b} (4\pi)^N \bar{\tau}_N^{(-1)} r_1^2 \dots r_N^2 dr_1 \dots dr_N \\ & = \frac{1}{3DNb}, \end{aligned} \quad (3.7.100)$$

$$\begin{aligned} & \left( \frac{3}{4\pi(1-b)^3} \right)^N \int_0^{1-b} \dots \int_0^{1-b} (4\pi)^N \bar{\tau}_N^{(0)} r_1^2 \dots r_N^2 dr_1 \dots dr_N \\ & = \frac{1}{N} T - \frac{1}{3DNb}, \end{aligned} \quad (3.7.101)$$

$$\begin{aligned} & \left( \frac{3}{4\pi(1-b)^3} \right)^N \int_0^{1-b} \dots \int_0^{1-b} (4\pi)^N \bar{\tau}_N^{(1)} r_1^2 \dots r_N^2 dr_1 \dots dr_N \\ & = -\frac{4\pi b(N-1)}{N^2} \left( \frac{1}{D(1-b)^3} \beta - \frac{3D}{4\pi} T^2 \right). \end{aligned} \quad (3.7.102)$$

Giving a general formula for  $N$  uniform randomly placed targets of

$$T_N = \frac{T}{N} - \frac{4\pi b(N-1)}{N^2} \left( \frac{1}{D(1-b)^3} \beta - \frac{3D}{4\pi} T^2 \right). \quad (3.7.103)$$

As  $N \rightarrow \infty$ , (3.7.103) becomes linearly dependent on  $N$ , with the second term of (3.7.103) tending to zero. We must now calculate  $\beta$ . This can be done by computing the integral given in

(3.7.93) using Mathematica, giving

$$\beta = \frac{1}{75600\pi(b-2)b^2} \left[ -12600(b-2)b^2 \text{Li}_2\left(\frac{b}{2}\right) + 2(b(b(2b(b(27097 - 27b(b(25(b-9)b + 707) - 679)) - 75845) + 46998) + 525\pi^2 + 58960) - 1050\pi^2 - 107693) + 49245) \right. \\ - 6300) + 105(b-2)b(-30b(2b((b-3)b+3) - 1) \log^2(2-b) + (b(8b(b(b(9(b-5)b+58) + 21) - 111) - 3049) - 120b(b-1)^3 \log(b) + 60) \log(2-b) - 300 \log(b) \\ \left. + b(1824 \log((b-2)^2) - 60 \log\left(1 - \frac{b}{2}\right) \log\left(\frac{b}{2}\right) + \log(b)(8b(b(b(9(b-5)b+58) + 21) - 111) - 30(2b((b-3)b+3) - 3) \log(b) + 729)) - 120(b \log(2) + 2) \tanh^{-1}(1-b)) \right], \quad (3.7.104)$$

where  $\text{Li}_2\left(\frac{b}{2}\right)$  is a polylogarithm, which can be written as a series expansion as follows

$$\text{Li}_s(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^s}, \quad (3.7.105)$$

therefore

$$\text{Li}_2\left(\frac{b}{2}\right) = \sum_{k=1}^{\infty} \frac{\left(\frac{b}{2}\right)^k}{k^2} = \frac{b}{2} + \frac{b^2}{16} + \frac{b^3}{72} + \frac{b^4}{256} + \dots \quad (3.7.106)$$

The SGMFPT equation given in (3.7.82) was found by [58]. To the best of my knowledge, the TGMFPT equation given in (3.7.103) has not been previously derived. Equation (3.7.103) is valid for  $N$  targets. However, because the integration for randomising the starting position of each of the targets has the upper limit of  $1 - b$ , the targets are able to overlap. To obtain an equation for non-overlapping targets, when integrating over all possible starting positions for a target, the volume already occupied by the targets that have been placed must be excluded. This proves to be a more complicated problem to solve, and is not covered in this thesis.

### 3.8 Testing the $N$ target equation

To determine the accuracy of (3.7.103), we will test it against numerical simulations. The numerical simulation requires, (1) placing targets uniform randomly within a sphere, of radius 1, and (2) placing a particle uniform randomly, that moves by Brownian motion, until it encounters one of the  $N$  targets. Plots comparing the numerics for two, three, and four targets, to (3.7.103) are shown in Figure 3.14.

Figure 3.14 shows a comparison of numerical simulations, with standard error bars, (3.7.103), and (3.5.37), assuming target independence. Standard error bars are given by the standard deviation

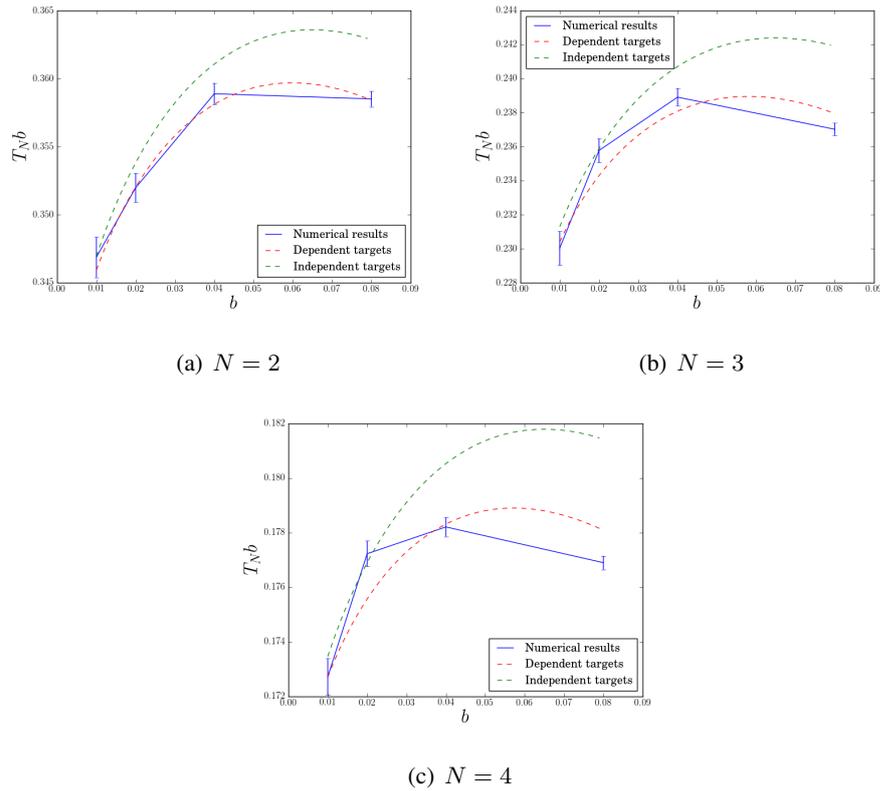


Figure 3.14: Comparison of (3.7.103), assuming target dependence (red), and (3.5.37) assuming target independence (green), to numerical results obtained from spatial simulations. Parameters used:  $R = 1$ ,  $D = 0.5$ . Number of realisations used to produce each point ranges from  $5.5 \times 10^4$  to  $5.5 \times 10^5$ .

of the data, divided by the square root of the number of realisations. In all cases, assuming target dependence (3.7.103) is a better approximation of the numerical results. There is a mismatch between the numerical results and the analytical solution in the  $N = 3$  and  $N = 4$  case due to the placement of the targets. The analytical solution assumes the targets are uniformly randomly distributed, regardless of the position of the other targets. However, the numerical simulations place the targets uniformly randomly in the domain excluding the volume of the other targets. The mismatch is greater for higher  $N$  as the probability of multiple targets occupying the same domain increases as  $N$  is increased.

Figure 3.15 shows a plot of (3.7.103), assuming dependent targets, and (3.5.37), assuming independent targets, compared to numerical results, for varying numbers of targets. The figure shows that the dependent target equation, (3.7.103), is a better approximation to the numerical results. The numerical results in Figure 3.15 are a better fit than those in 3.14 as the simulations

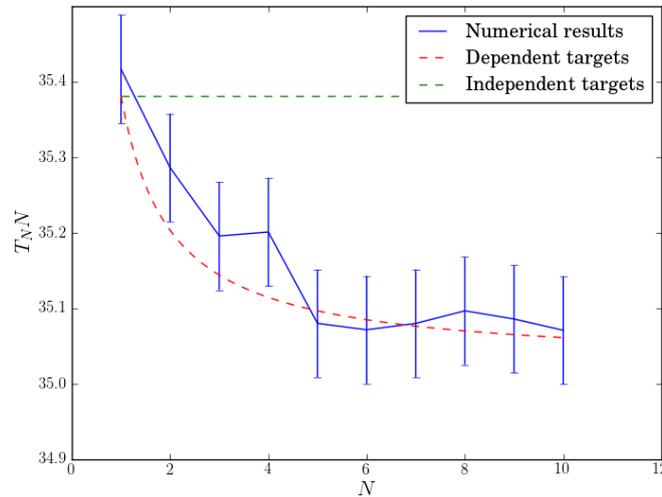


Figure 3.15: Comparison of (3.7.103) and (3.5.37), to numerical simulations obtained from spatial simulations. Parameters used:  $R = 1$ ,  $b = 0.02$  and  $D = 0.5$ . Each numerical point is a mean taken from  $2.5 \times 10^5$  realisations.

place all the targets in uniform random positions, regardless of the positions of the other targets. In Figure 3.15, the error bars appear to be quite large. It must be remembered that the TGMFPT multiplied by the number of targets, is plotted on the  $y$  axis, which is exaggerating the difference between the two equations and the numerical results. If just the TGMFPT is plotted on the  $y$  axis a difference between the numerical results and equation cannot easily be observed and the error bars are tiny. The numerical results in Figure 3.15 were very computationally expensive to produce. The program used to produce the numerical results was coded to run efficiently by using inbuilt python packages, and simulating all ten starting numbers of APCs simultaneously. These methods are described in Appendix B. However, hundreds of days of computing time were still required to achieve the results shown. This was only made possible with use of a high performance computing system (ARC1) made available to me by the University of Leeds. Further computation could be undertaken to reduce the error bars further, but it can already be seen that the numerics are agreeing with (3.7.103), making further computation unnecessary. The program used to produce Figure 3.15 is shown in Appendix B.

### 3.9 Summary

In this chapter, Poisson's equation is used to derive an equation for the mean time it takes a particle, in a prescribed position, to encounter a fixed target in the centre of a spherical domain. This is given by

$$\tau_0(r_s) = \frac{1}{6D} \left( 2R^3 \left( \frac{1}{b} - \frac{1}{r_s} \right) + b^2 - r_s^2 \right).$$

The particle's starting position is then averaged over the domain, to give a SGMFPT equation of

$$\bar{\tau}_0 = \frac{1}{30Db(R^3 - b^3)} (2R^5(5R - 9b) + b^5(3b - 5) + 10b^3R^3).$$

An approximation proposed by Grigin [93], is used to derive an equation for a target in a position 'close' to the centre of the domain, and is found to be

$$\tau_c(r_t, r_s, \theta) = \frac{1}{3Dr} \left( -R^3 + \frac{r_s}{2b}(2R^3 + b^3) - \frac{r_s^3}{2} + \frac{r_T}{r_s} \left( \frac{b^3 - R^3}{2b^3 + R^3} \right) (R^3 + 2r_s^3) \cos(\theta) \right).$$

Averaging over all possible starting positions of the particle gives

$$\bar{\tau}_c(r_T) = \frac{1}{2D(R^3 - (b + r_T)^3)} \left( -R^3(R^2 - (b + r_T)^2) + \frac{1}{3b}(2R^3 + b^3)(R^3 - (b + r_T)^3) - \frac{1}{5}(R^5 - (b + r_T)^5) \right).$$

Averaging over the targets starting position, within the valid region assumed by the approximation, gives

$$\begin{aligned} T_c &= \frac{3}{4\pi b^3} \int_0^b \bar{\tau}(r_T) 4\pi r_T^2 dr_T \\ &= -\frac{1}{10b^3D} \left( \frac{43b^5}{30} - \frac{37b^2R^3}{3} - 6\sqrt{3}R^3 (b^2 + 4bR + R^2) \left( \tan^{-1} \left( \frac{2b + R}{\sqrt{3}R} \right) \right. \right. \\ &\quad \left. \left. - \tan^{-1} \left( \frac{4b + R}{\sqrt{3}R} \right) \right) 18R^3 (b - R)(b + R) \tanh^{-1} \left( \frac{b(3b + R)}{5b^2 + 3bR + 2R^2} \right) \right). \end{aligned}$$

Next, an approach to solve the problem of a target in a uniform random position using Green's functions is used. This approach is outlined in an article, by Chevalier *et al.* [58]. The SGMFPT for a target in a fixed position is given by

$$\bar{\tau}(\mathbf{r}_T) = \frac{1}{3Db} + \frac{1}{3D} \left( \frac{1}{1 - r_T^2} - \log(1 - r_T^2) + r_T^2 - \frac{14}{5} \right),$$

which is the result given in [58]. Averaging the starting position gives the equation

$$T = \frac{R^3}{3Db} - \frac{R^2}{90D} \left( 45 \log \left( \frac{b}{R} \right) + 15 \log(2) + 76 \right).$$

An equation for the time to first encounter one of  $N$  uniform random targets within a spherical domain is found to be

$$T_N = \frac{T}{N} - \frac{4\pi b(N-1)}{N^2} \left( \frac{1}{D(1-b)^3} \beta - \frac{3D}{4\pi} T^2 \right),$$

where  $\beta$  is given in (3.7.104). Throughout the chapter the equations are tested against numerical results generated using a spatial code developed in this thesis 2.2.3. It was found that, for each improvement implemented to the MFTP equations, the equations become better approximations of the numerical results.

## Chapter 4

# The effects of a limited imaging volume on *in vivo* imaging data

### 4.1 Introduction

Imaging experiments using two-photon microscopy (Section 1.8.1) enables detailed visualisation of lymphocyte movements *in vivo* and provide data sets of cell positions over time. Trajectories of the centres of mass of the cells can be plotted to determine the migratory behaviour of cell populations. Plots of cell trajectories are commonly represented in two ways. Firstly, the unshifted coordinates of the cells can be plotted, helping to determine if the cells prefer a particular region. Alternatively, cell positions can be shifted to start from the origin, while maintaining their orientation in space, allowing one to determine if the cells have a preferred direction of migration. If the number of cell tracks observed in all possible directions is approximately equal, it can be inferred that the cells are undergoing random motion. Plotting cell tracks gives qualitative information about the cells, such as the type of motion the cells are undergoing. To determine quantitative information, such as the diffusivity of a population of cells, one must analyse the dynamics of the cells.

A frequently reported migration parameter is cell migration speed. The mean speed between two sequential time frames can easily be estimated by dividing the displacement between the two frames, by the time elapsed in that period. In reality, the trajectory of the cell between two points will not be exactly straight, therefore the estimated mean speed will be an underestimate of the actual mean speed of the cells. Longer time periods between successive points will increase the

error in the estimated mean speed. If the cells are assumed to be undergoing a random walk, their mean displacement in any one dimension will be zero, therefore, a more useful parameter to measure is the diffusivity of a cell, defined as the distance squared a cell moves in a given time.

By plotting the mean displacement squared against time, information about the type of migration, as well as the diffusivity of the cells, can be obtained. A linear increase indicates the cells are undergoing a random walk type motion. A faster than linear increase indicates the cells are undergoing directed migration. On long timescales, directed migration could be due to the cells following a chemokine gradient. A slower than linear mean displacement plot indicates the cells are somehow confined; this could be due to interactions with other cell types keeping them in a specific region. If the cells are undergoing random motion, the diffusivity of the cell population can be determined from the gradient of the mean displacement plot. A mean squared displacement plot is useful for investigating cell motility, but the underlying mechanism of migration cannot reliably be inferred from it, due to multiple underlying microprocesses giving rise to very similar mean squared displacement plots [72, 190, 234]. For example, T cells moving within a lymph node (LN) in randomly orientated steps of fixed duration and speed [22], result in a similar plot to that of persistently moving cells that manoeuvre through a densely packed environment with highly variable speeds [26]. When plotting displacement squared against time, for data obtained from two-photon laser scanning microscopy (TPLSM) experiments, it must be remembered that the imaging area is a finite volume, so that cells that are tracked for longer time periods have remained in the imaging volume for longer. This may be due to the cells having a low diffusivity, or taking a looping path. Cells that leave the imaging volume can no longer be tracked, meaning data is only available on them for a short time periods. This results in more data being available for smaller times, and it also causes a confinement effect, indicated by a plateau forming at later times due to these being the cells that have remained in the imaging volume. This confinement effect occurs because the cells that can be tracked for longer times are the ones that are remaining in the imaging volume.

The aim of this chapter is to calculate the expected squared displacement observed for a population of cells undergoing a random walk in a limited sized imaging volume. This is accomplished by first solving the diffusion equation, to calculate a probability density function of the position of particles at time  $t$ . Using the probability density function, the expected mean squared displacement can be calculated in one and three dimensions. The expectation is then normalised for particles shifted to have an initial position at the origin. Finally, the expectation is compared to numerically

simulated results, and a correction developed to compensate for the large time increments used in biological experiments.

The purpose of this chapter is to be able to estimate the diffusion coefficient of a population of cells by fitting the expected squared displacement calculated analytically to the squared displacement of a population of cells imaged *in vivo*. Comparison of the expected plateau observed due to the limited size of the imaging volume will allow the type of motion occurring to be determined. For example, if a plateau is observed at a lower level than the expected plateau due to a limited size imaging volume, this could be due to the cells being confined by a biological factor, or some other effect occurring. Previous work has been carried out to determine the effect of a limited sized imaging volume on calculating binding time of the observed cells [26]. Work has also been undertaken to increase the accuracy of data collected from TPLSM experiments [25]. However, to the best of my knowledge, the expected squared displacement of a population of cells has not previously been calculated analytically, as is done in this chapter.

## 4.2 Expected squared displacement

### 4.2.1 One dimension

The density function of a particle moving by Brownian motion on the interval  $[-L, L]$ , can be found by solving the diffusion equation

$$\frac{\partial}{\partial t}P(x, t) = D\frac{\partial^2}{\partial x^2}P(x, t), \quad (4.2.1)$$

where  $P(x, t)$  is the density function of being at position  $x$ , at time  $t$ . We have the initial condition  $P(x, 0) = \delta(x - x_0)$ , where  $x_0$  is the starting position of the particle, and  $\delta$  is the Dirac delta function: the probability of being at  $x = x_0$  at time zero is one, and zero elsewhere. The boundary conditions are  $P(L, t) = P(-L, t) = 0$ , as the particle will be absorbed at  $\pm L$ . Thus, the probability of being there is zero.

Equation (4.2.1) can be solved using separation of variables, setting

$$P(x, t) = X(x)T(t), \quad (4.2.2)$$

and substituting into (4.2.1). Rearranging gives

$$\frac{T'}{T} = D\frac{X''}{X}. \quad (4.2.3)$$

By setting both sides equal to a constant,  $a$ , (4.2.3) can be split into two separable ODEs

$$X'' - aX = 0, \quad (4.2.4)$$

$$T' - aDT = 0. \quad (4.2.5)$$

The boundary conditions can only be satisfied, and do not result in a trivial solution, when  $a > 0$ .

Let  $a = k^2$ , then  $X'' - k^2X = 0$ . This has the solution

$$X = c_1 \cos(kx) + c_2 \sin(kx). \quad (4.2.6)$$

Applying the boundary conditions gives

$$X(L) = c_1 \cos(kL) + c_2 \sin(kL) = 0, \quad (4.2.7)$$

$$X(-L) = c_1 \cos(kL) - c_2 \sin(kL) = 0, \quad (4.2.8)$$

where  $c_1$  and  $c_2$  are integration constants. Adding these equations gives  $c_1 \cos(kL) = 0$ . If  $c_1 \neq 0$ , we need  $\cos(kL) = 0$ . Therefore, we have  $kL = \frac{(2n+1)}{2}\pi$ , giving  $k = \frac{(2n+1)}{2}\frac{\pi}{L}$ , where  $n = 0, 1, 2, \dots$ . Those values of  $k$ , give  $c_2 = 0$ . We have the solution

$$X_n(x) = c_1 \cos\left(\frac{(2n+1)\pi}{2} \frac{x}{L}\right). \quad (4.2.9)$$

Solving (4.2.5), gives the solution  $T = c_3 e^{-aDt}$ , where  $c_3$  is an integration constant. For this part of the solution we have  $a = k^2 = \left(\frac{(2n+1)\pi}{2L}\right)^2$ . Therefore

$$T_n = c_3 e^{-D \frac{(2n+1)^2 \pi^2}{4L^2} t}. \quad (4.2.10)$$

Alternatively, (4.2.7) and (4.2.8) can be subtracted from each other to give  $c_2 \sin(kL) = 0$ . If  $c_2 \neq 0$ , we need  $\sin(kL) = 0$ , therefore we have  $kL = (n+1)\pi$ , giving  $k = \frac{(n+1)\pi}{L}$ , where  $n = 0, 1, 2, \dots$ . We use  $(n+1)\pi$ , rather than  $n\pi$ , to avoid division by zero. Those values of  $k$ , give  $c_1 = 0$ , giving a second solution to (4.2.4) of the form

$$X_n(x) = c_2 \sin\left(\frac{(n+1)\pi}{L} x\right). \quad (4.2.11)$$

For this part of the solution, for the  $T$  equation, we have  $a = k^2 = \left(\frac{(n+1)\pi}{L}\right)^2$ . Therefore

$$T_n = c_4 e^{-D \frac{(n+1)^2 \pi^2}{4L^2} t}, \quad (4.2.12)$$

where  $c_4$  is a integration constant. The general solution to (4.2.1) is a linear combination of these two solutions

$$P(x, t) = \sum_{n=0}^{\infty} \left[ a_n(x_0) \cos\left(\frac{(2n+1)\pi}{2} \frac{x}{L}\right) e^{-D \frac{(2n+1)^2 \pi^2}{4L^2} t} + b_n(x_0) \sin\left(\frac{(n+1)\pi}{L} x\right) e^{-D \frac{(n+1)^2 \pi^2}{4L^2} t} \right], \quad (4.2.13)$$

where the  $a_n$ 's and  $b_n$ 's are a combination of the integration constants. The  $a_n$ 's and  $b_n$ 's can be found by applying the initial condition,  $P(x, 0) = \delta(x)$ , and using a Fourier transform:

$$a_n(x_0) = \frac{1}{L} \int_{-L}^L \cos\left(\frac{2n+1}{2} \frac{\pi}{L} x_0\right) \delta(x - x_0) dx_0. \quad (4.2.14)$$

By definition,  $\int_{-L}^L \delta(x - x_0) dx = 1$ , at  $x = x_0$  if  $x_0 \in [-L, L]$ . Therefore, for the case when the particle starts in the centre,  $x = 0$ , we get  $a_n(0) = 1/L$ . For the case when the particle starts at  $x = x_0$ , (4.2.14) gives

$$a_n(x_0) = \frac{1}{L} \cos\left(\frac{2n+1}{2} \frac{\pi}{L} x_0\right). \quad (4.2.15)$$

The  $b_n$ 's are given by

$$b_n(x_0) = \frac{1}{L} \sin\left(\frac{(n+1)\pi}{L} x_0\right). \quad (4.2.16)$$

The mean squared position,  $x_t^2$ , of a particle at time  $t$ , given that the particle is within the domain, is given by

$$\mathbb{E}[x_t^2] = \frac{\int_{-L}^L x^2 P(x, t) dx}{\int_{-L}^L P(x, t) dx}. \quad (4.2.17)$$

We have

$$\int_{-L}^L P(x, t) dx = \sum_{n=0}^{\infty} \frac{4L}{\pi(2n+1)} a_n(x_0) (-1)^n e^{-D \frac{(2n+1)^2}{4} \frac{\pi^2}{L^2} t}, \quad (4.2.18)$$

and

$$\int_{-L}^L x^2 P(x, t) dx = \sum_{n=0}^{\infty} \frac{4L^3}{(2n+1)\pi} a_n(x_0) \left(1 - \frac{8}{(2n+1)^2 \pi^2}\right) (-1)^n e^{-D \frac{(2n+1)^2}{4} \frac{\pi^2}{L^2} t}. \quad (4.2.19)$$

In (4.2.18) and (4.2.19), the terms involving  $b_n(x_0)$  integrate to zero. As time tends to infinity, the  $\mathbb{E}[x_t^2]$  plateaus due to the finite length of the domain. An equation for the plateau can be found by setting  $n = 0$  in (4.2.18) and (4.2.19). This gives

$$\lim_{t \rightarrow +\infty} \mathbb{E}[x_t^2] = \frac{\frac{4L^3}{\pi} a_0(x_0) \left(1 - \frac{8}{\pi^2}\right) e^{-D \frac{1}{4} \frac{\pi^2}{L^2} t}}{\frac{4L}{\pi} a_0(x_0) e^{-D \frac{1}{4} \frac{\pi^2}{L^2} t}} = \left(1 - \frac{8}{\pi^2}\right) L^2. \quad (4.2.20)$$

Equations (4.2.17) and (4.2.20) can be tested by comparison to results simulated numerically using a spatial code. Figure 4.1 shows a plot of  $\mathbb{E}[x_t^2]$  starting at (a)  $x_0 = 0$ , and (b)  $x_0 = L/2$ , along with results of spatial simulations.

### Random starting position

To determine the expectation of the mean squared position, for a particle starting in a uniform random position, we need to average the expectation of  $x_t^2$  over all possible starting positions. In

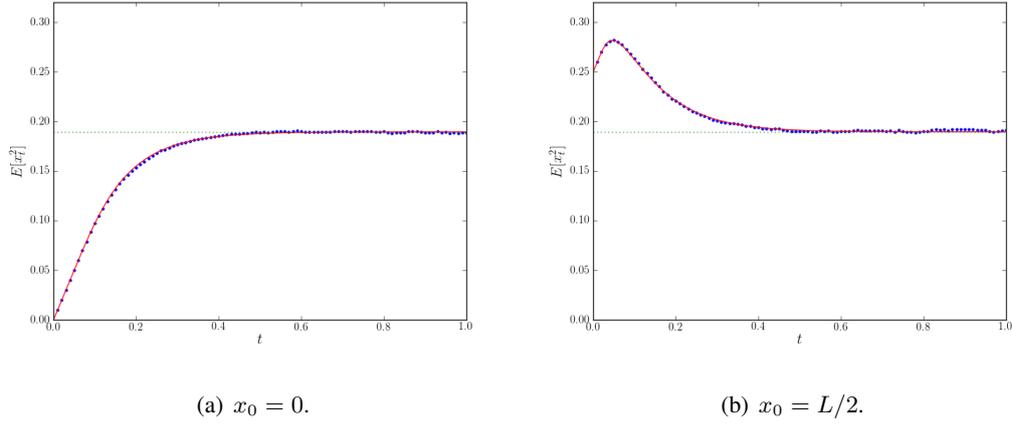


Figure 4.1:  $\mathbb{E}[x_t^2]$ , from (4.2.17), is shown in red, the plateau of  $\mathbb{E}(x_t^2)$  as  $t \rightarrow +\infty$  (4.2.20), is shown in green, and results of numerical simulations are shown in blue. In both simulations  $L = 1.0$  and  $D = 0.5$ .

this case, the  $x_0$  term occurs in the  $a_n$  term (4.2.15), and the  $b_n$  term integrates to zero. Therefore, only the  $a_n$  term needs to be integrated over, giving

$$\bar{a}_n = \frac{1}{2L} \int_{-L}^L a_n(x_0) dx_0 = 2(-1)^n \frac{1}{L} \frac{1}{(2n+1)\pi}. \quad (4.2.21)$$

The expectation of  $x_t^2$  for a particle starting in a uniform random position can be found by using  $\bar{a}_n$  (4.2.21) in (4.2.13), in place of the  $a_n(x_0)$  terms. Figure 4.2 shows  $\mathbb{E}[x_t^2]$  with a particle starting in a uniform random position, with comparison to numerical results.

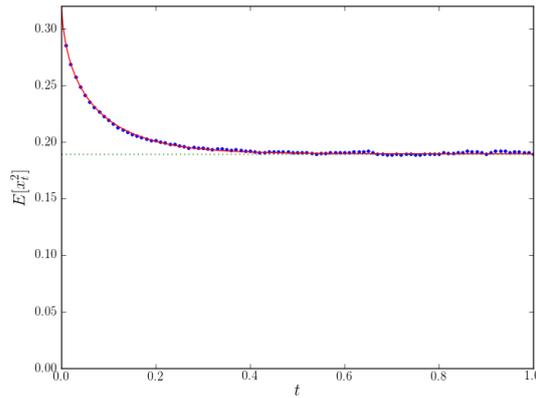


Figure 4.2: Particle starting in a uniform random position.  $\mathbb{E}[x_t^2]$  from (4.2.17), is shown in red, the plateau of  $\mathbb{E}[x_t^2]$  as  $t \rightarrow +\infty$  (4.2.20), is shown in green, and results of numerical simulations are shown in blue. In this simulation  $L = 1.0$  and  $D = 0.5$ .

### 4.2.2 Three dimensions

To calculate the mean squared position of a cell in a three-dimensional volume, with dimensions  $2L_x \times 2L_y \times 2L_z$ , the diffusion equation in three dimensions must be solved. This is given by

$$\frac{\partial}{\partial t} P(x, y, z, t) = D \left( \frac{\partial^2}{\partial x^2} P + \frac{\partial^2}{\partial y^2} P + \frac{\partial^2}{\partial z^2} P \right), \quad (4.2.22)$$

with the initial condition  $P(x, y, z, 0) = \delta(x)\delta(y)\delta(z)$ . However, as the particle is moving by Brownian motion and following a Wiener process, each of the Cartesian components of the Wiener process in  $n$  dimensions is an independent Wiener process (Section 2.2.2). The expectation of the mean squared position of a particle in a three-dimensional volume is the sum of the  $x$ ,  $y$  and  $z$  expectations in one dimension [8]. In the case when the volume is a cube,  $\mathbb{E}[\mathbf{X}_t^2]$  is the sum of three copies of equation (4.2.17). If the volume does not have equal length in the  $x$ ,  $y$  and  $z$  dimensions,  $\mathbb{E}[\mathbf{X}_t^2]$  is the sum of the  $x$ ,  $y$  and  $z$  expectations,

$$\mathbb{E}[\mathbf{X}_t^2] = \mathbb{E}[x_t^2] + \mathbb{E}[y_t^2] + \mathbb{E}[z_t^2], \quad (4.2.23)$$

where  $\mathbf{X}_t$  represents a three-dimensional vector, with  $\mathbf{X}_t = x_t^2 + y_t^2 + z_t^2$ . The plateau is given by the sum of three copies of the one-dimensional plateau (4.2.20), if the size of the volume is equal in all three dimensions. For unequal lengths in three dimensions, the plateau is giving by

$$\lim_{t \rightarrow \infty} \mathbb{E}[\mathbf{X}_t^2] = (L_x^2 + L_y^2 + L_z^2) \left( 1 - \frac{8}{\pi^2} \right), \quad (4.2.24)$$

where  $L_x$ ,  $L_y$  and  $L_z$  are the distances from the origin to the boundary of the volume in the  $x$ ,  $y$  and  $z$  dimensions, respectively. Figure 4.3 shows a plot of  $\mathbb{E}[\mathbf{X}_t^2]$ , starting at (a)  $x_0 = y_0 = z_0 = 0$ , and (b)  $x_0 = L_x/2$ ,  $y_0 = L_y/2$ , and  $z_0 = L_z/2$ , along with results of numerical simulations.

Figure 4.4 shows  $\mathbb{E}[\mathbf{X}_t^2]$  with a particle starting in a uniform random position within a cubic volume, using  $a(x_0)$ ,  $a(y_0)$  and  $a(z_0)$  as given in equation (4.2.21). The  $b_n$  terms integrate to zero, therefore need not be included.

## 4.3 Expectation of the normalised displacement squared

In biological experiments, where data is obtained using TPLSM, plots of displacement squared against time are generated by subtracting a cell's starting position from its position at time  $t$ . To be able to compare our calculations with these plots, an expectation of the normalised particle position must be found by subtracting the initial position of the particle,  $x_0$ .

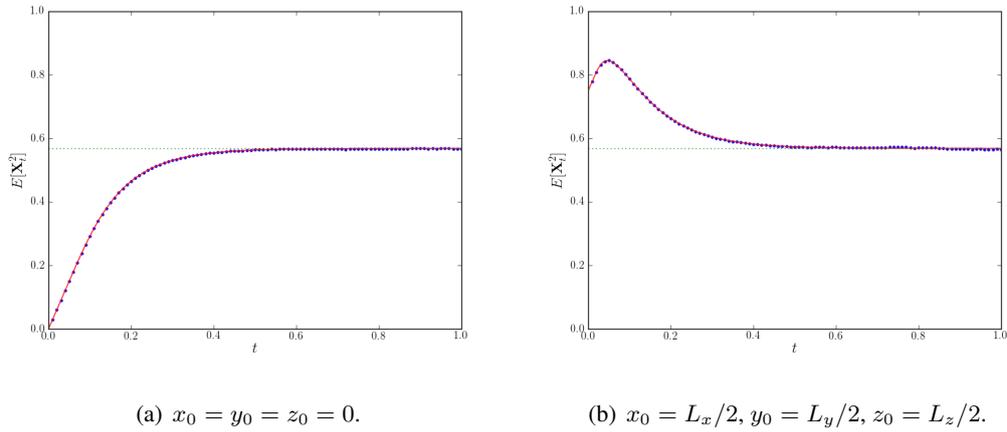


Figure 4.3: Fixed starting positions, in three dimensions.  $\mathbb{E}[\mathbf{X}_t^2]$  from (4.2.23), is shown in red, the plateau of  $\mathbb{E}[\mathbf{X}_t^2]$  as  $t \rightarrow +\infty$  (4.2.24), is shown in green, and results of numerical simulations are shown in blue. In both simulations  $L_x = L_y = L_z = 1.0$  and  $D = 0.5$ .

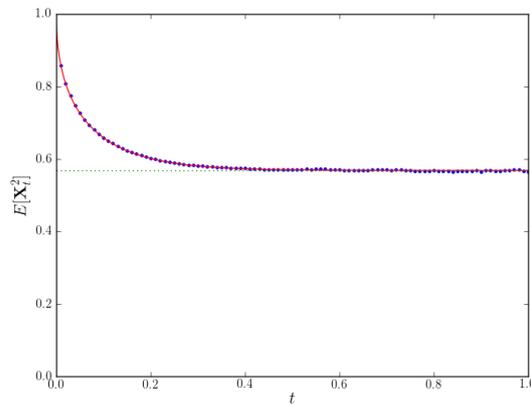


Figure 4.4: uniform random starting position in three dimensions.  $\mathbb{E}[\mathbf{X}_t^2]$  from (4.2.23), is shown in red, the plateau of  $\mathbb{E}[\mathbf{X}_t^2]$  as  $t \rightarrow +\infty$  (4.2.24), is shown in green, and results of numerical simulations are shown in blue. In this simulation  $L_x = L_y = L_z = 1.0$  and  $D = 0.5$ .

### 4.3.1 One dimension

In one dimension, the expectation of  $(x_t - x_0)^2$  can be calculated by splitting the expectation into three parts,

$$\mathbb{E}[(x_t - x_0)^2] = \mathbb{E}[x_t^2] - 2\mathbb{E}[x_t x_0] + \mathbb{E}[x_0^2] = \frac{\int_{-L}^L x^2 P(x, t) dx}{\int_{-L}^L P(x, t) dx} - 2 \frac{\int_{-L}^L x x_0 P(x, t) dx}{\int_{-L}^L P(x, t) dx} + x_0^2. \quad (4.3.25)$$

The first term of (4.3.25) is the same as that given in (4.2.18) and (4.2.19). For the second term we need to compute the integral,

$$\int_{-L}^L x x_0 P(x, t) dx = \sum_{n=0}^{\infty} b_n(x_0) \frac{2L^2}{\pi(n+1)} (-1)^n x_0 e^{-\frac{\pi^2 D(n+1)^2 t}{L^2}}. \quad (4.3.26)$$

The plateau of the expectation function (4.3.25), as  $t \rightarrow +\infty$ , is given by

$$\lim_{t \rightarrow +\infty} \mathbb{E}[(x_t - x_0)^2] = \frac{\int_{-L}^L (x - x_0)^2 P(x, t) dx}{\int_{-L}^L P(x, t) dx} = L^2 \left(1 - \frac{8}{\pi^2}\right) + x_0^2. \quad (4.3.27)$$

Plots of (4.3.25), and the plateau (4.3.27) are shown in Figure 4.5, for specific initial particle positions, along with comparison to numerical simulations.

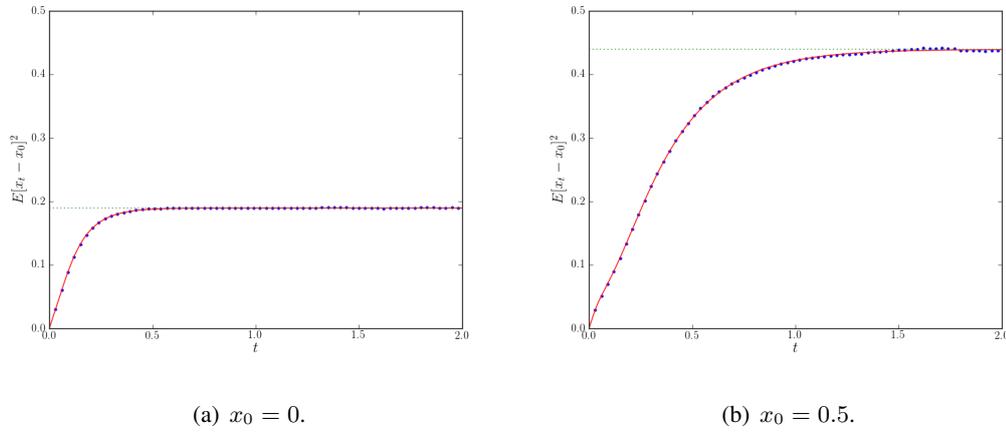


Figure 4.5:  $\mathbb{E}[(x_t - x_0)^2]$  from (4.3.25) is shown in red, the plateau of  $\mathbb{E}[(x_t - x_0)^2]$  as  $t \rightarrow +\infty$  (4.3.27) is shown in green, and results of numerical simulations are shown in blue. In both simulations  $L = 1.0$  and  $D = 0.5$ .

### Random starting position

The  $\mathbb{E}[(x_t - x_0)^2]$  for a particle starting in a uniform random position can be calculated by averaging over all possible starting positions,  $x_0$ , and weighting each starting position by the survival probability for a particle starting at that position remaining in the domain,  $[-L, L]$ , at time  $t$ . The survival probability of a particle in one dimension, starting at  $x_0$  is given by

$$S^{(1)}(x_0, t) = \int_{-L}^L P(x, t) dx = \sum_{n=0}^{\infty} \frac{4L}{\pi(2n+1)} \cos\left(\frac{(2n+1)\pi}{2L} x_0\right) (-1)^n e^{-D \frac{(2n+1)^2}{4} \frac{\pi^2}{L^2} t}. \quad (4.3.28)$$

Figure 4.6 shows comparison of the survival probability (4.3.28) to numerical results, for varying

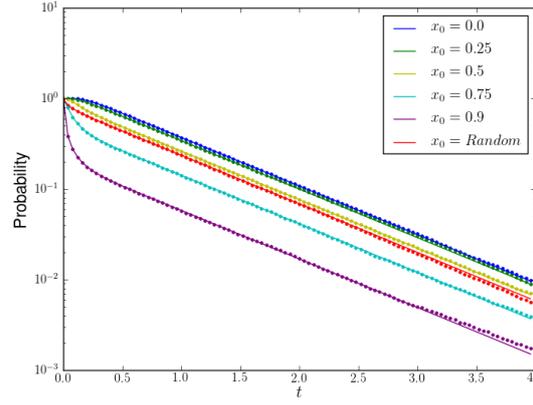


Figure 4.6: Survival probability of particles starting at varying positions,  $x_0$ , indicated in the legend. Solid lines show (4.3.28) and dotted lines show numerical simulations. The results for a uniform random starting position are averaged over multiple realisations.

starting positions in one dimension.

To calculate the mean of  $\mathbb{E}[(x_t - x_0)^2]$  over all starting positions, weighted by the probability of starting in that position, we need to compute

$$\overline{\mathbb{E}}[(x_t - x_0)^2] = \frac{\int_{-L}^L \mathbb{E}[(x_t - x_0)^2] S^{(1)}(x_0, t) dx_0}{\int_{-L}^L S^{(1)}(x_0, t) dx_0}, \quad (4.3.29)$$

where  $\overline{\mathbb{E}}$  represents the expectation averaged over all starting positions. This can be computed directly for the plateau, as  $t \rightarrow +\infty$ , using (4.3.27);

$$\begin{aligned} \lim_{t \rightarrow +\infty} \overline{\mathbb{E}}[(x_t - x_0)^2] &= \frac{\int_{-L}^L (L^2 (1 - \frac{8}{\pi^2}) + x_0^2) S^{(1)}(x_0, t) dx_0}{\int_{-L}^L S(x_0, t) dx_0} \\ &= L^2 \left(1 - \frac{8}{\pi^2}\right) + \frac{\int_{-L}^L x_0^2 S^{(1)}(x_0, t) dx_0}{\int_{-L}^L S^{(1)}(x_0, t) dx_0} \\ &= L^2 \left(1 - \frac{8}{\pi^2}\right) + \frac{\frac{16L^3(\pi^2 - 8)}{\pi^4}}{\frac{16L}{\pi^2}} \\ &= 2L^2 \left(1 - \frac{8}{\pi^2}\right). \end{aligned} \quad (4.3.30)$$

The mean of  $\mathbb{E}[(x_t - x_0)^2]$  over all starting positions, for all  $t$ , cannot be easily calculated analytically. It can be calculated numerically, by choosing a uniform random  $x_0$  within the domain  $-L$ , to  $L$ , calculating  $\mathbb{E}[(x_t - x_0)^2]$ , and multiplying by  $S^{(1)}(x_0, t)$  for that  $x_0$ . We must then sum the expectation, multiplied by the survival probability, for numerous  $x_0$  and divide this sum by the

sum of the survival probabilities for the same  $x_0$ ,

$$\overline{\mathbb{E}}[(x_t - x_0)^2] = \frac{\sum_{x_0} \mathbb{E}[(x_t - x_0)^2] S^{(1)}(x_0, t)}{\sum_{x_0} S^{(1)}(x_0, t)}. \quad (4.3.31)$$

A plot of (4.3.31), the plateau (4.3.30), with comparison to numerical simulations is shown in Figure 4.7.

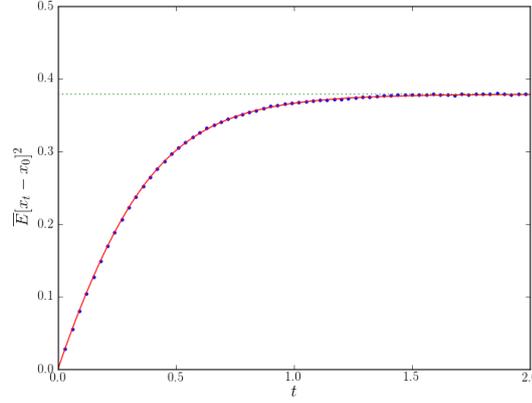


Figure 4.7:  $\overline{\mathbb{E}}[(x_t - x_0)^2]$  (4.3.29), for a particle starting in a uniform random position, is shown in red, the plateau of  $\overline{\mathbb{E}}[(x_t - x_0)^2]$  as  $t \rightarrow +\infty$  (4.3.30), is shown in green, and results of numerical simulations are shown in blue. In this simulations  $L = 1.0$  and  $D = 0.5$ .

### 4.3.2 Three dimensions

To calculate  $\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]$ , the components of the three one-dimensional processes must be summed, giving

$$\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2] = \mathbb{E}[(x_t - x_0)^2] + \mathbb{E}[(y_t - y_0)^2] + \mathbb{E}[(z_t - z_0)^2], \quad (4.3.32)$$

where the one-dimensional expectations are given by (4.3.25). The plateau of the three-dimensional expectation is also given by the sum of three copies of the one-dimensional plateau in  $x$ ,  $y$  and  $z$ , given by (4.3.27),

$$\lim_{t \rightarrow +\infty} \mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2] = (L_y^2 + L_y^2 + L_z^2) \left(1 - \frac{8}{\pi^2}\right) + x_0^2 + y_0^2 + z_0^2. \quad (4.3.33)$$

Figure 4.8 shows a plot of (4.3.32), and the plateau (4.3.33), along with numerical simulations.

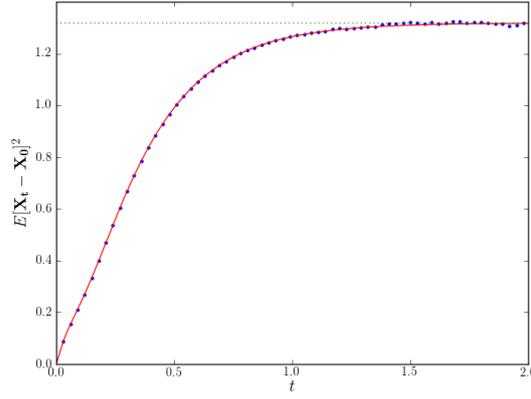


Figure 4.8:  $\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]$  in three dimensions (4.3.32) for a particle starting at  $x_0 = L_x/2$ ,  $y_0 = L_y/2$  and  $z_0 = L_z/2$  is shown in red, the plateau of  $\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]$  as  $t \rightarrow +\infty$  (4.3.33) is shown in green, and results of numerical simulations are shown in blue. In this simulation  $L_x = L_y = L_z = 1.0$  and  $D = 0.5$ .

### Random starting position

For a particle starting in a uniform random position in three dimensions, the expectation must be weighted by the survival probability in three dimensions. The survival probability in three dimension is given by the product of the survival probability in each of the  $x$ ,  $y$  and  $z$  dimensions:

$$S^{(3)}(\mathbf{X}_0) = S^{(1)}(x_0)S^{(1)}(y_0)S^{(1)}(z_0), \quad (4.3.34)$$

where  $S^{(1)}$  is given by equation (4.3.28), and the  $t$  argument has been suppressed. A plot of the survival probability (4.3.34), for various starting positions, is shown in Figure 4.9, with comparison to numerical simulations.

The plateau for the three-dimensional case is calculated in the same way as in the one-dimensional case, by multiplying the plateau for a fixed starting position (4.3.33) by the survival probability (4.3.34), integrating over  $x$ ,  $y$  and  $z$ , and dividing by the integral of  $S^{(3)}(\mathbf{X}_0)$ .

$$\begin{aligned} \lim_{t \rightarrow +\infty} \bar{\mathbb{E}}((\mathbf{X}_t - \mathbf{X}_0)^2) &= \frac{\int_{-L_z}^{L_z} \int_{-L_y}^{L_y} \int_{-L_x}^{L_x} (3L^2 (1 - \frac{8}{\pi^2}) + x_0^2 + y_0^2 + z_0^2) S^{(3)}(\mathbf{X}_0) dx_0 dy_0 dz_0}{\int_{-L_z}^{L_z} \int_{-L_y}^{L_y} \int_{-L_x}^{L_x} S^{(3)}(\mathbf{X}_0) dx_0 dy_0 dz_0} \\ &= 2(L_x^2 + L_y^2 + L_z^2) \left(1 - \frac{8}{\pi^2}\right). \end{aligned} \quad (4.3.35)$$

The mean of  $\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]$ , is given by weighting the expectation for each starting position by the survival probability for that position. The analytical solution cannot easily be calculated, so a

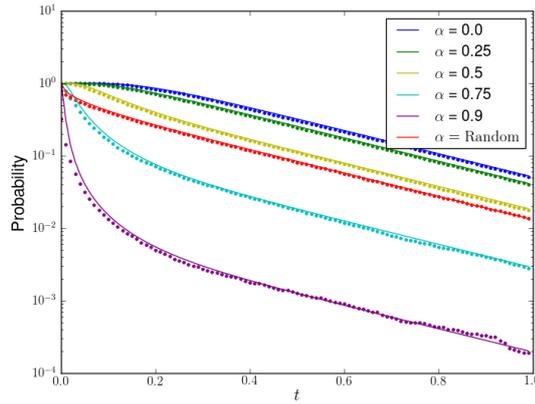


Figure 4.9: Survival probability of a particle in a three-dimensional volume, given by (4.3.34), shown by solid lines, and results of numerical simulations, shown by dotted lines. In these simulations particles have the same starting position in the  $x$ ,  $y$  and  $z$  dimensions, shown in the legend as  $\alpha$ , where  $\mathbf{X}_0 = \alpha(1, 1, 1)$ .  $L_x = L_y = L_z = 1.0$  and  $D = 0.5$ . The uniform random starting position has been averaged over multiple realisations.

numerical calculation is performed in the same way as in the one-dimensional case, by summing over multiple starting positions:

$$\overline{\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]} = \frac{\sum_{\mathbf{X}_0} \mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2] S^{(3)}(\mathbf{X}_0)}{\sum_{\mathbf{X}_0} S^{(3)}(\mathbf{X}_0)}, \quad (4.3.36)$$

where the summation is over various starting positions, for multiple realisations. Figure 4.10 shows a plot of the expectation (4.3.36), and the plateau (4.3.35), for a particle starting in a cube, with comparison to numerical simulations.

## 4.4 Application to imaging volume

The imaging volume used in TPLSM experiments is larger in the  $x$  and  $y$  dimensions than in the  $z$  dimension, due to the way the imaging technique works. It is therefore more likely that a cell will exit via the  $z$  boundary than via one of the  $x$  or  $y$  boundaries. Equation (4.3.36) can be tested using an imaging volume with a  $z$  dimension that is smaller than the  $x$  and  $y$  dimensions, shown in Figure 4.11.

Figure 4.11 shows (4.3.36) is correct for an imaging volume with an unequal size in the  $z$  dimension. We now want to test the accuracy of the calculated expectation value when using

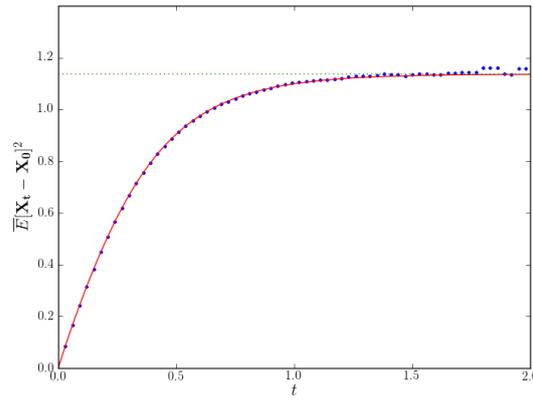


Figure 4.10:  $\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]$  in three dimensions for a particle starting in a uniform random position (4.3.36) is shown in red, the plateau of  $\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2]$  as  $t \rightarrow +\infty$  (4.3.35) is shown in green, and results of numerical simulations are shown in blue. In this simulation  $L_x = L_y = L_z = 1.0$  and  $D = 0.5$ .

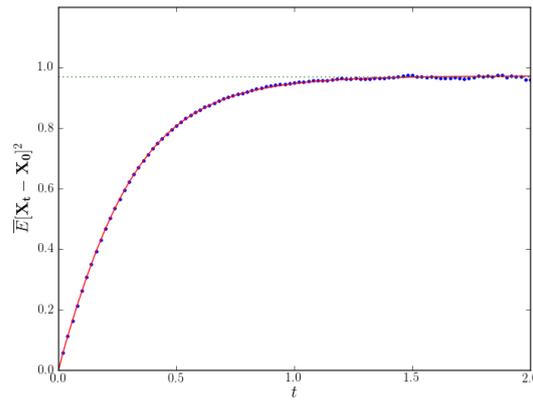


Figure 4.11: Expectation for an imaging volume that is smaller in the  $z$  dimension than the  $x$  and  $y$  dimensions, with particles starting in uniform random positions. Predicted mean expectation (4.3.36), shown in red, the plateau of the mean expectation as  $t \rightarrow +\infty$  (4.3.35), shown in green, and results of numerical simulations are shown in blue. In this simulation  $L_x = L_y = 1.0$  and  $L_z = 0.75$  and  $D = 0.5$ .

parameters that are similar to those found in biological experiments. Using an imaging volume of  $L_x = L_y = 117.275\mu m$ , and  $L_z = 15.4\mu m$ , as used in data provided by a research group in Glasgow, (analysed in Chapter 5), along with a diffusivity of  $20\mu m^2 min^{-1}$ , and an imaging time of 100 minutes, (4.3.36) can be tested.

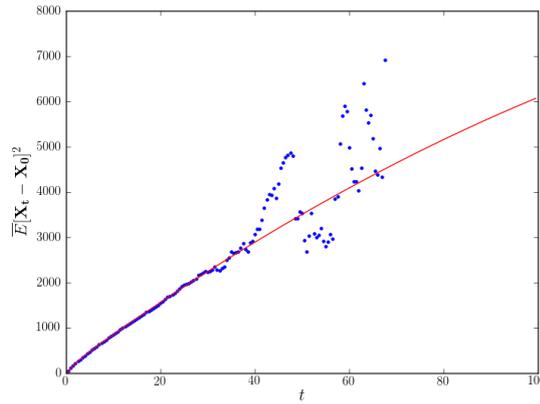


Figure 4.12: Predicted mean expectation (4.3.36), shown in red, with comparison to numerical results generated using a spatial simulation are shown in blue, for an imaging volume of a size similar to those used in biological experiments,  $L_x = L_y = 117.275\mu m$ ,  $L_z = 15.4\mu m$  and  $D = 20\mu m^2 min^{-1}$ . Number of realisations =  $10^8$ .

Figure 4.12 shows the expectation does not reach the confinement plateau by the end of the 100 minutes of imaging, and the numerical results do not extend to the full 100 minutes. There are no numerical results for higher timepoints due to the  $z$  length being much shorter than the  $x$  and  $y$  lengths, resulting in particles exiting via the  $z$  boundary at earlier times. To remain in the imaging volume until a later time, with this imaging volume, a particle must move in the  $x$  and  $y$  dimensions a lot more than the  $z$ . Movement in each dimension is equally likely, therefore this event is uncommon. This results in a plateau, due to confinement in a limited size imaging volume, not being observed. Increasing the diffusivity of the cells would result in the expectation reaching the plateau quicker, however it would also result in cells exiting via the  $z$  boundary quicker and only giving numerical values for even shorter times.

## 4.5 Comparison to data

Using a spatial code written in python, data equivalent to that obtained using TPLSM can be generated. The program, placed cells inside an imaging volume of a specified size and moved the cells following a random walk, recording their position every 30 seconds, as is done in biological experiments. If a cell leaves the imaging volume, it is removed from the simulation. To compare to the equation for the expected displacement squared (4.3.36), we need only consider cells that start in the imaging volume and for only as long as they stay in the volume. Using data generated

in this way, displacement squared against time plots, in which there will be confinement, can be created.

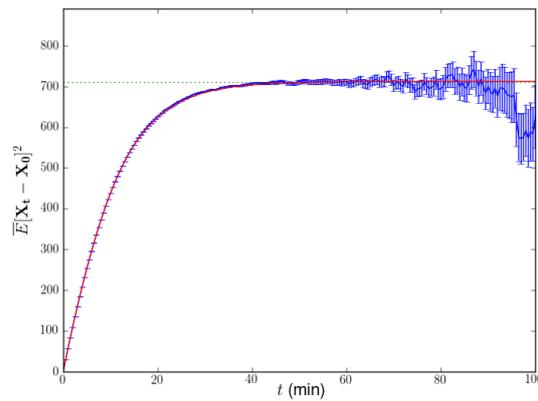


Figure 4.13: Predicted mean expectation (4.3.36) is shown in red, and plateau of the mean expectation as  $t \rightarrow +\infty$  (4.3.35) is shown in green, with comparison to numerical results shown in blue for an imaging volume of  $L_x = L_y = L_z = 25\mu m$ , with  $D = 10\mu m^2 min^{-1}$ .

Figure 4.13 shows the mean displacement squared of cells moving in an imaging volume of  $50\mu m \times 50\mu m \times 50\mu m$ , with a diffusivity of  $10\mu m^2 min^{-1}$ . The simulated data fits very well to the analytical prediction. After a time of 60 minutes, the fit is not quite as good, due to there being less data for later times, as a result of more cells having exited the imaging volume. The data for this plot has been gathered by simulating ten million cells.

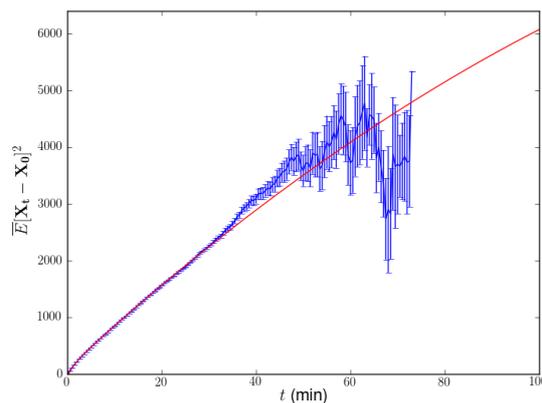


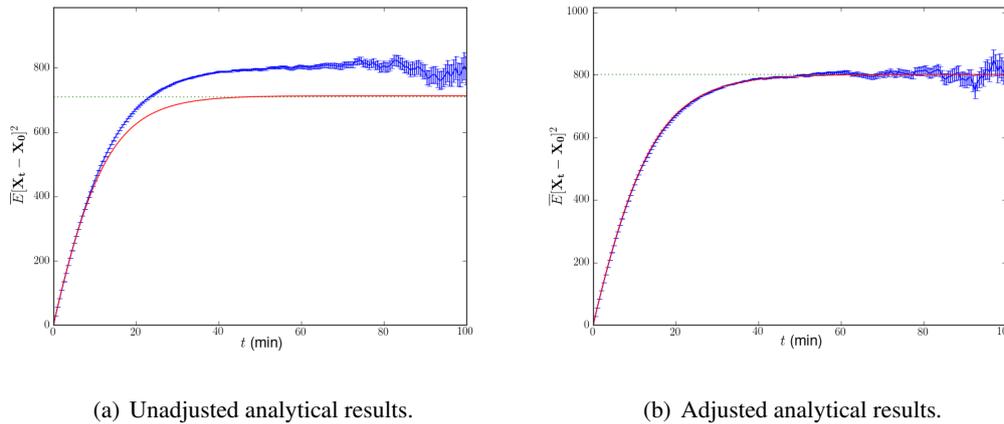
Figure 4.14: Simulated data for an imaging volume the same size as the imaging volume in the data provided by a research group in Glasgow (Chapter 5). The predicted mean expectation (4.3.36) is shown in red, and simulated data is shown in blue.

Figure 4.14 shows simulated data for an imaging volume of  $234.55\mu\text{m} \times 234.55\mu\text{m} \times 30.8\mu\text{m}$ , with a diffusivity of  $20\mu\text{m}^2\text{min}^{-1}$ . For this simulation, the data does not reach the plateau expected for the imaging volume, and there is no data for times beyond 80 minutes. This is due to the  $z$  dimension being a lot smaller than the  $x$  and  $y$  dimensions. If there was data from enough cells, we should, in theory, see some cells that remain in the imaging volume long enough to reach the plateau. However, the data used in Figure 4.14 is gathered by simulating ten million cells, and in biological experiments there are usually around one thousand cells tracked, and this data set is reduced after sorting. Therefore, in a biological experiment a plateau will not be seen due to confinement in the imaging volume, when the  $z$  dimension is a lot smaller than the  $x$  and  $y$  dimensions, which is usually the case.

#### 4.5.1 Underestimation of plateau

When simulating data for Figure 4.13, it was discovered that the plateau the data reached is affected by the update interval of the simulation. In Figure 4.13 data is recorded every 30 seconds of simulated time, but the list of cells that are in the imaging volume is updated every time the cells are moved, this update procedure is required to fit the analytical prediction. However, if the list of cells in the imaging volume is updated every 30 seconds, the plateau of the simulated data is overestimated (Figure 4.15(a)). This is due to cells exiting and then re-entering the imaging volume between the 30 second updates. Therefore, on average, the cells are able to travel further before they are classified as being outside the imaging volume, resulting in the displacement squared being greater. The analytical prediction can be corrected by adding half the mean displacement in one dimension, of a cell in 30 seconds, to the length of each of the boundaries of the imaging volume. Half the mean displacement is used as the cell must exit the imaging volume, then re-enter, for it to continue being tracked. The mean displacement in one dimension, for a cell moving by Brownian motion, is given by  $\sqrt{2D\delta t}$ , where  $\delta t$  is the update interval. The adjustment needs to be added to the  $L_x$ ,  $L_y$  and  $L_z$  values, to correct for the overestimation. The correction results in a better prediction of the plateau (Figure 4.15(b)).

The underestimation must be taken into account when predicting the plateau of data obtained from biological experiments, as cells will be able to exit and re-enter the imaging volume between rescanning of that section.



(a) Unadjusted analytical results.

(b) Adjusted analytical results.

Figure 4.15: Plots showing, (a) the underestimation of the analytical prediction due to updating the list of cells in the imaging volume every 30 seconds, and (b) a simple correction to the underestimation.

## 4.6 Summary

In this chapter mathematical techniques developed have allowed us to determine the expectation of the mean squared displacement, as a function of time, and the plateau as  $t \rightarrow +\infty$ , of a displacement squared against time plot, for particles starting in uniform random positions within an imaging volume of a given size. The aim of this work was to be able to use these predictions to determine the diffusivity of a set of cells by fitting the predictions to the data. However, due to the disparity in size of the  $z$  dimension of the imaging volume used in biological experiments, the displacement squared plots will not lead to a confinement plateau. Therefore, the equation for the plateau cannot be fitted to this type of data. The mean expectation of the displacement squared can be fitted to biological data, to try and determine the diffusivity of the observed cells. Using data provided by a research group in Glasgow (Institute of Infection, Immunity and Inflammation, University of Glasgow), a comparison to biological data could be made. The amount of viable data available in this experiments is quite small, and other factors effect the interpretation of the data, such as: persistent motion and noise in the data. The data also does not reach a plateau due to the shape of the imaging volume. Therefore, fitting the prediction for the mean expected squared displacement does not result in a better prediction of the diffusivity than fitting a straight line. A more in-depth discussion and analysis of *in vivo* imaging data is carried out in Chapter 5.

Although the prediction of the plateau, due to confinement in an imaging volume of limited size, cannot be used as originally intended, it does allow experimental results to be validated.

A displacement squared plot with a plateau below the predicted one indicates the cells are in fact confined, or that there is another factor causing confinement to be observed, such as an underlying population of cells with a lower diffusivity, as found in Section 5.4.2. Comparison of the predictions for the displacement squared to numerically simulated imaging experiments revealed that large intervals in the scanning of the imaging volume cause the predictions to be underestimated. A simple correction can be implemented to compensate for this problem, by adding half the mean distance a cell can move in an update interval to each of the boundaries. The underestimation is most pronounced at the plateau of the displacement squared, but when using imaging volumes similar to those used in biological experiments, the plateau is not reached. There will still be a slight underestimation of the slope to reach the plateau, but this will not be as large as at the plateau. Therefore, the underestimation can most likely be ignored. Numerically simulated imaging experiments illustrated that large amounts of data are required to achieve relatively smooth results. In biological experiments the number of cells tracked is not very large, with usually less than one thousand cells being tracked. To achieve smoother results techniques will need to be developed to extract more data from the tracks already present, this is discussed in Section 5.4.3.



## Chapter 5

# Analysis of *in vivo* imaging data

### 5.1 Introduction

Lymphocytes are constantly moving between, and within, lymph nodes in an attempt to locate antigen presenting cells (APCs) displaying their cognate antigen, to become activated. The time it takes for an activation event to occur depends on the number of APCs presenting cognate antigen. More importantly, it depends on how quickly the lymphocytes attempting to locate the APCs are moving, and if they are being directed towards their target in some way. Therefore, the big question in determining lymphocyte activation time is how ‘fast’ do lymphocytes move within the lymph node.

Advances in imaging technology have allowed lymphocytes to be viewed *in vivo*, in three-dimensional space, deep within intact lymph nodes, using two-photon laser scanning microscopy (TPLSM) [80]. Using these imaging techniques, cell positions over time can be recorded to build up a data set of cell tracks. From the data, one can easily obtain an ‘apparent velocity’ of the cells by simply calculating the distance a cell has moved and dividing by the time to make that movement. However, the ‘apparent velocity’ is not a very useful parameter to report, as time is recorded in discrete intervals, usually of thirty to sixty seconds. Therefore, in the intervening time, the cell may have taken any path between the initial and final position. The cell is unlikely to have taken a straight path, as assumed when calculating a velocity, hence it is often referred to as an ‘apparent velocity’. A more useful parameter to report is the diffusivity of cells, which relates to the displacement squared a cell has moved in a given time. The mean diffusivity of a population of cells can be calculated by plotting a displacement squared as a function of time for

a population of cells, and is determined from the gradient of this plot. However, due to limitations and complexity of the imaging technology, and image processing software, this is more difficult than it first seems. In processing images from TPLSM experiments, it can be difficult to determine what is a cell, and what is a fragment of a cell. Also, cells that are close together can be mistaken for one large cell, and one large cell can be mistaken for multiple smaller cells. A snapshot of an imaging experiment is shown in Figure 5.1, illustrating the ambiguity of what actually is a cell. For a review of the artefacts commonly found in TPLSM data sets, and ways in which data analysis can be used to detect and correct these artefacts, I refer the reader to a paper by Beltman *et al.* [25]. Using displacement squared against time, as well as other analytical techniques, one can also attempt to quantify the type of motion a population of cells are moving by. This motion can be random, directed, persistent or confined [25]. The diffusivity, as well as the type of motion cells are undergoing, is also highly dependent on the conditions that the cells are exposed to. This will depend upon the experimental technique used. By analysing the turning angles between successive movements of a cell, one can quantify if persistent motion is occurring within a population of cells.

Many research groups have attempted to estimate the diffusion coefficient of T cells. These estimates range from a few micrometers squared per minute, to hundreds of micrometers squared per minute. A paper by Miller *et al.* [156], calculates the T cells in their experiment to have a diffusivity of  $67\mu m^2 min^{-1}$ . They find the cells undergo persistent motion on timescales of one to three minutes, and random motion on longer timescales. In another paper by Miller [153], they investigate the diffusion coefficient at various times of the T cell activation process. They found that in the early stage of T cell activation (less than two hours) the cells had a diffusivity of  $9.7\mu m^2 min^{-1}$ , and at a later stage (two to fourteen hours) the diffusivity dropped to  $2.3\mu m^2 min^{-1}$ .

In this chapter, a data set provided by a research group at the University of Glasgow (Institute of Infection, Immunity and Inflammation, University of Glasgow) is analysed in an attempt to calculate the diffusion coefficient of the cells imaged, as well as the type of motion they are undergoing. The chapter begins by analysing basic aspects of the data, such as the total displacement of tracks, and the total time tracks occur for. The data is then broken down, erroneous data discarded, and the remain data are re-analysed to obtain an estimate for the diffusion coefficient of the cells. The analysis enables the development of a systematic method for determining the diffusivity of a population of cells, which is applied to other data sets.

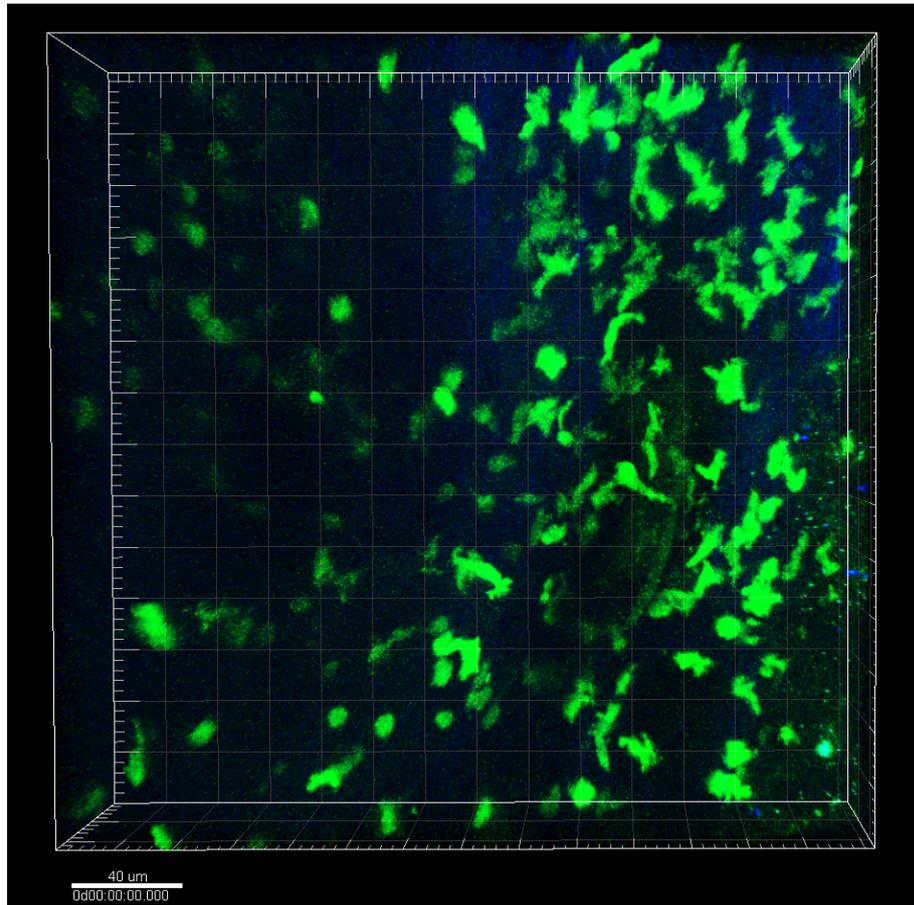


Figure 5.1: Snapshot of TPLSM *in vivo* imaging, at timepoint 0, for the Ear1 data set, described in Section 5.2. T cells are shown in green.

The aim of determining the type of motion T cells are undergoing is to be able to accurately model cell movement *in silico*. If the cells are undergoing random motion, and the diffusivity can be calculated, this can then be used as a parameter in spatial simulations and in MFTP equations derived in Chapter 3.

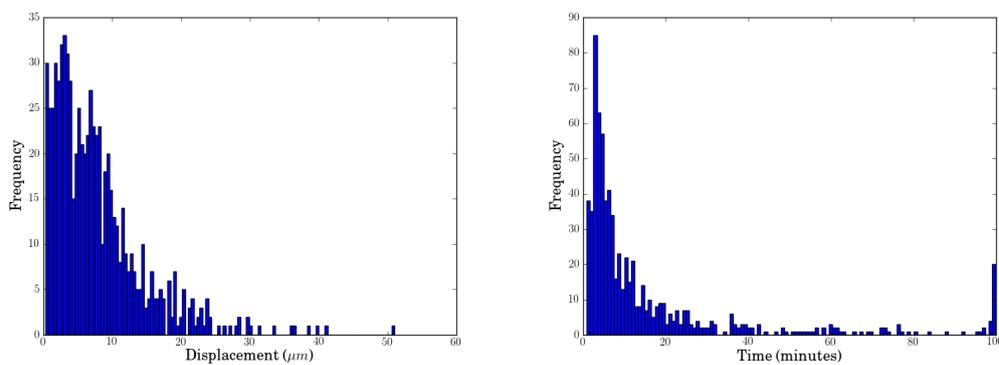
## 5.2 About the data

The main data set analysed in this chapter is for an imaging period of 100 minutes, with timepoints taken every 50.8 seconds. The imaging volume is a box with  $x$ ,  $y$ ,  $z$  dimensions of  $234.55\mu\text{m} \times 234.55\mu\text{m} \times 30.8\mu\text{m}$ . It takes two seconds to scan a frame in the  $x - y$  plane. Therefore, for this data, the  $z$  frames must be about  $1.23\mu\text{m}$  apart. This value is calculated by determining how many planes can be scanned in a timestep (50.8 seconds), and dividing the size of the  $z$  dimension by

this value. The other data sets analysed in Section 5.5.1 only, have the following properties: LN2: has  $x, y, z$  dimensions of  $234.55\mu\text{m} \times 234.55\mu\text{m} \times 32.0\mu\text{m}$ , with a timestep size of 29 seconds, giving a  $z$  plane separation of  $2.29\mu\text{m}$ . The total imaging time is 58 minutes. LN4: has  $x, y, z$  dimensions of  $206.53\mu\text{m} \times 206.53\mu\text{m} \times 22.0\mu\text{m}$ , with a timestep size of 32 seconds, giving a  $z$  plane separation of  $1.38\mu\text{m}$ . The total imaging time is 32 minutes. Ear1: has  $x, y, z$  dimensions of  $282.85\mu\text{m} \times 282.85\mu\text{m} \times 46.0\mu\text{m}$ , with a timestep size of 50 seconds, giving a  $z$  plane separation of  $1.84\mu\text{m}$ . The total imaging time is 33.3 minutes. Figure 5.1 shows a snapshot of a  $z$  slice in an imaging experiment carried out to obtain the Ear1 data set.

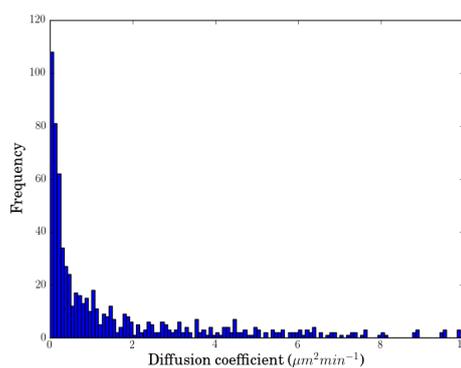
### 5.3 Basic analysis

Basic information about the data can be determined by plotting distributions of the displacement of the cells, the time tracks were imaged for, and the diffusion coefficients of each cell.



(a) Total displacement.

(b) Track duration.



(c) Diffusion coefficients.

Figure 5.2: Basic analysis of the whole data set.

Figure 5.2(a) shows the distribution of displacements of individual cells, from the time they were first tracked, to the time they were last tracked. There appears to be lots of cells with small displacements, and fewer cells with larger displacements. A small displacement does not necessarily mean the cell has not moved far, rather its end position is close to its starting position. A cell could have a long track length, but end up close to where it started. This can be analysed further by investigating the confinement ratio of the cells. The confinement ratio is defined as the ratio of track length and total track displacement, and is sometimes referred to as the meandering index [25]. A small displacement indicates that the cell could be confined to a certain region, resulting in the end position being near to the start position, or the cell could not be moving much, and has not travelled far from where it started. The mean of the total track displacements is  $7.77 \mu m$ , and the standard deviation is  $6.65 \mu m$ .

Figure 5.2(b) shows the distributions of the duration of cell tracks. This is the time from the point when the tracking commenced, to the time when the tracking ended. The figure shows a peak of cells with short durations, and fewer cells with longer durations. A large peak occurs at 100 minutes as this includes all the cells that were tracked for the whole of the imaging period. The mean of the total track durations is 15.99 minutes and the standard deviation is 22.86 minutes.

Figure 5.2(c) shows the distribution of diffusion coefficients for individual cells. The plot has been trimmed at ten minutes on the  $x$ -axis to allow the plot to be viewed more easily. There are a few tracks with diffusion coefficients greater than  $10 \mu m^2 min^{-1}$  that are not displayed on this plot. The diffusion coefficient for each cell has been calculated by

$$\text{Diffusion coefficient} = \frac{\text{Displacement}^2}{6 \times \text{Time interval}}$$

As with the displacement and time interval plots, there are numerous cells with small diffusion coefficients and few with larger diffusion coefficients. The mean diffusion coefficient for all the cells is  $2.37 \mu m^2 min^{-1}$ , and the standard deviation is  $4.25 \mu m^2 min^{-1}$ .

From this basic analysis it is clear that there is a large amount of variation in the data, with all parameters investigated having large standard deviations. The data also has lots of cells that appear to be not moving very far, that are skewing the distributions.

### 5.3.1 Displacement squared

By plotting displacement squared against time, an attempt can be made to quantify the type of motion the cells are undergoing, as well as determining an estimate for the diffusion coefficient

of the population. On a displacement squared against time plot, a straight line indicates Brownian motion, a line that plateaus indicates confinement, and a line that bends upwards indicates a directional bias. However, it is not always clear how to distinguish between these effects, as there are often several competing processes. With our data, confinement is expected, as the cells were only tracked in a small imaging volume. Therefore, cells that remain in the imaging volume for a longer period of time will be those that have not moved as far. A mean displacement squared plot was generated by calculating the displacement squared from each point in a track, to the starting point of the track, and recording the time in which the calculated displacement squared occurred. A mean for each time was then taken. This method allows extra data to be extracted from the cell tracks, rather than just taking the total displacement of the cells.

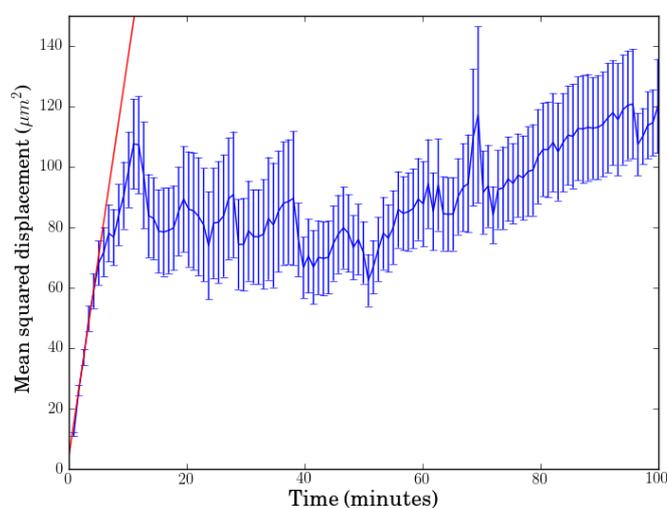


Figure 5.3: Mean displacement squared against time, with standard error bars. The red line is fitted to the first five time points, giving a diffusivity of  $2.33 \pm 0.10 \mu m^2 min^{-1}$ .

Figure 5.3 shows that for small time intervals, approximately less than ten minutes, the cells seem to be undergoing Brownian motion, as there is a linear increase. For larger time intervals the graph flattens out, indicating confinement, as expected. A straight line is seen for small time intervals as the cells have not had enough time to reach the edge of the imaging volume, so confinement has not yet occurred. The red line on the graph is a line of the form  $y = a + bx$ , fitted to the first five timepoints of the graph, where Brownian motion appears to be occurring. A sixth of the gradient of this line gives the diffusion coefficient of the cells. The fitted line has a gradient of 14.0, giving a diffusion coefficient of  $2.33 \mu m^2 min^{-1}$ . Calculating the standard error in the least squares fitting procedure gives an error in the gradient of 0.60, corresponding to an error

in the diffusion coefficient of  $0.10\mu\text{m}^2\text{min}^{-1}$ , making the estimate of the diffusion coefficient  $2.33\pm 0.10\mu\text{m}^2\text{min}^{-1}$ . This value of the diffusion coefficient is quite small, and is at the lower end of the range of diffusion coefficients found in literature [153, 156]. This could be due to tissue drift, which may occur as a result of the experimental setup moving, and the cells are in fact stationary. These ‘stationary’ cells would therefore be skewing the diffusivity of the data set, making it appear smaller than it actually is. This can be investigated further by calculating turning angles, and scalar product, of the cells, as well as the tissue drift in the experiment. Analysing the turning angles, and scalar product, will also help determine if the cells are following a random walk, or moving with persistent motion.

At first glance, the plateau of the displacement squared plot is presumed to be due to the cells being confined in a limited sized imaging volume. However, the level of the plateau is much smaller than the expected plateau, calculated using (4.3.35), of about 10500. Therefore, the plateau observed is not due to the limited sized imaging volume. Upon closer inspection, the ‘plateau’ appears to be made up of drops, then slow increases. This could be the result of multiple populations of cells moving with different diffusivities. Cells with a larger diffusivity will reach the boundary of the imaging volume at smaller times. Once these cells have exited the imaging volume they will no longer be contributing to the mean displacement. Therefore, if there is a population of cells with a lower diffusivity remaining, they will have a larger contribution to the mean. This will cause a drop in the mean squared displacement when the cells with a larger diffusivity exit the imaging volume.

### 5.3.2 Turning angles

To determine if cells move in a persistent direction, or turn back on themselves abruptly, the angle a cell turns between successive time points can be analysed.

In Figure 5.4 a large turning angle (greater than  $90^\circ$ ) indicates a cell has made a shallow turn, and is continuing in a similar direction. A small turning angle (less than  $90^\circ$ ) indicates a cell has made a sharp turn, and is moving in a direction not similar to its previous direction. This is illustrated in Figure 5.5.

Figure 5.4 shows a wide spread of angles, ranging from approximately  $0^\circ$ , meaning the cell is going in the opposite direction, to approximately  $180^\circ$ , meaning the cell is continuing in the same direction. By fitting a sine distribution for the expected angles to the figure, it can be seen that

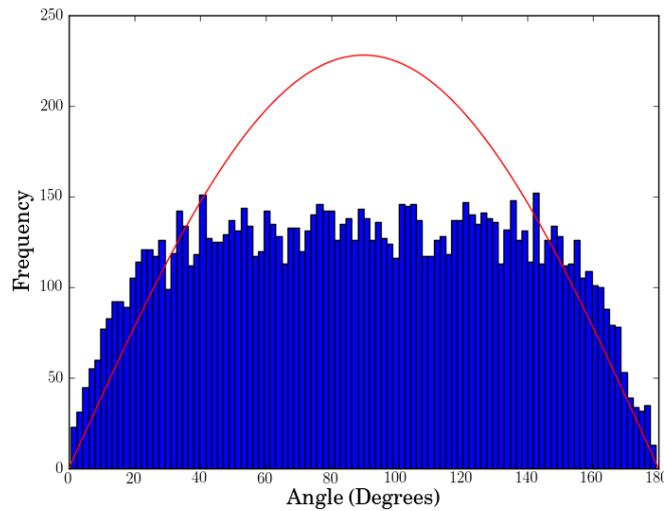


Figure 5.4: Turning angles between successive vectors in tracks for the whole data set. Mean of data is  $89.0^\circ$ . Red curve shows the expected distribution of angles.

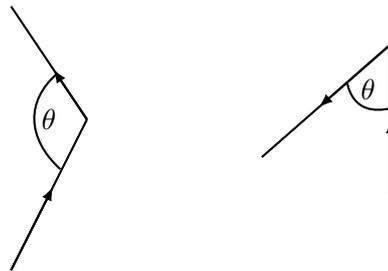


Figure 5.5: Illustration of turning angles. The left diagram shows a shallow turning angle, with  $\theta$  being close to  $180^\circ$ . The right diagram shows a sharp turning angle, with  $\theta$  being close to  $0^\circ$ .

there is more small and large angles than expected, and less angles around  $90^\circ$  than expected. This could be due to cells making small movements back and forth, perhaps due to experimental error.

### 5.3.3 Persistent motion

Persistent motion can be detected by analysing the scalar product of the  $i$ th and  $i + k$ th vectors in a track, where  $k$  is the number of vectors further along the track that is being compared. For example, if  $k = 1$  vectors directly after each other are being compared, and if  $k = 2$  vectors two timesteps apart are being compared. By taking the mean of these scalar products for each  $k$  it can be determined if there is persistent motion and at what point it occurs. A positive scalar

product indicates the two vectors are almost parallel. A scalar product equal to zero indicates the vectors are orthogonal to each other, and a scalar product less than zero indicates the vectors are almost anti-parallel. The scalar product takes into account the magnitude of a vectors, as well as the angle between the vectors. If the cells are moving following a random walk, and do not have any persistent motion, the mean scalar product will be zero. The scalar product is calculated by:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^n a_i b_i,$$

where  $\mathbf{a}$  and  $\mathbf{b}$  are the vectors being compared. In this case we are calculating the scalar product in three dimension, so it is given by the sum of the product of each of the  $x$ ,  $y$  and  $z$  components.

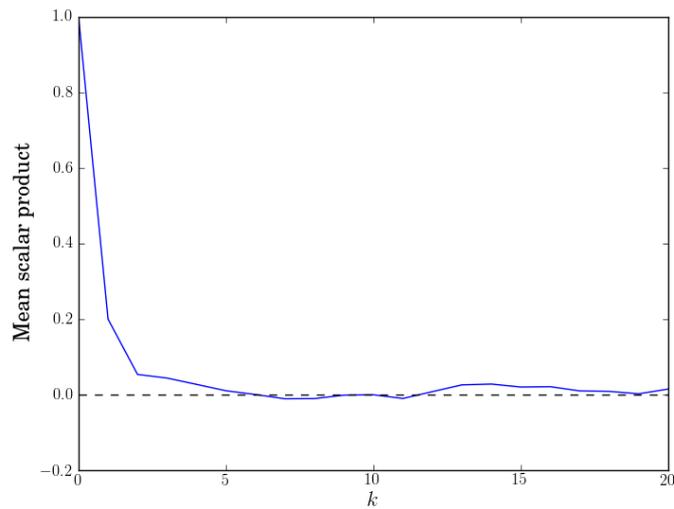


Figure 5.6: Mean scalar product between timepoints of varying size,  $k$ , in the  $x$ ,  $y$  and  $z$  dimensions. Plot has been scaled to start at one for  $k = 0$ .

Figure 5.6 shows for  $k$  less than five, the mean scalar product is greater than zero, and for  $k$  greater than and equal to five, the mean scalar product is close to zero. Indicating that for periods of about four timesteps the cells have persistent motion, but for larger timescales there is no persistent motion. This is consistent with experimental observations [232], where T cells were found to move in a series of repetitive lunges. It can be concluded that the cells movement is well approximated by Brownian motion on timescales of longer than four minutes.

### 5.3.4 Tissue drift

The data analysed have been obtained from an imaging experiment on a live mouse. Even though the mouse is anaesthetised, it may still twitch or make other small movements, resulting in a

shift in the whole imaging volume. Another factor that must be considered is that the whole experimental setup may move or vibrate, resulting in false movements being observed in the data. To determine if there is a directional bias in the experimental procedure, the mean movement in each of the  $x$ ,  $y$  and  $z$  directions can be plotted cumulatively. If there is no drift present, the mean displacement between timepoints should be approximately zero, if the cells are assumed to be undergoing Brownian motion.

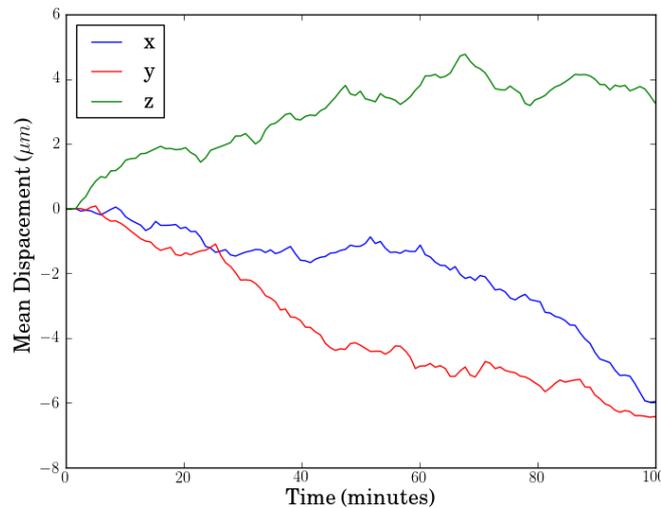


Figure 5.7: Plot of cumulative mean displacement from the mean starting position, as a function of imaging time.

Figure 5.7 shows there is tissue drift in the experiment. This is more pronounced in the  $x$  and  $y$  directions, with the mean cell position being about  $6\mu m$  from where it started, by the end of the experiment, in both the  $x$  and  $y$  dimensions. In the  $z$  direction the mean cell position was about  $3\mu m$  from where it started by the end of the experiment. This artefact can be compensated for by subtracting the amount of drift that has occurred at a timepoint, from the cell's  $x$ ,  $y$  and  $z$  position, at that timepoint.

## 5.4 Further analysis

### 5.4.1 Sorting data

The software that automatically tracks cell movements obtained by TPLSM is a very useful tool, but it does have its drawbacks, and it cannot compete with a human in terms of accuracy. However,

manually tracking cells is very time consuming, so automated software is usually used. One of the main drawbacks of automatic tracking is that the software often cannot distinguish cells that are in close proximity. It can track these clusters as one cell, and upon separation it is possible that the cell identities change. This leads to artefacts occurring in the data, which must be dealt with upon analysis.

Further analysis reveals the data set contains data from a total of 1121 cells. Some of these cells have only one data point, which is of no use for analysing cell tracks. Excluding these leaves 734 tracks (Figure 5.8(a)). These are the set of tracks used for the basic analysis in the previous sections. Of these 734 tracks, some of them only have information up to a timepoint before the end of the imaging time. This is acceptable if the cell being tracked has left the imaging volume, however, there are tracks that have ended part way through imaging and are not near the edge of the imaging volume. Excluding these cells leaves 582 tracks (Figure 5.8(b)). The same situation can happen in reverse, with tracks starting part way through the imaging time, which are not starting near the edge of the imaging volume. Excluding these leaves 498 tracks (Figure 5.8(c)). There is also a subset of tracks that do not have data for all timepoints from the time they are first detected to the time they are last observed. Theoretically, a track could leave and re-enter the imaging volume, which would cause a break in the tracking data. However, the software that tracks the cells is not able to know that a cell entering is one that has previously exited, so would treat the entering cell as a newly observed cell. To overcome this artefact, all tracks that have breaks in their tracking of one or more timepoints were excluded. This leaves 245 tracks (Figure 5.8(d)) to be analysed.

#### 5.4.2 Displacement squared against time

Displacement squared against time plot can now be plotted for the sorted data set.

Figure 5.9(a), in which tissue drift is not accounted for, shows that after 40 minutes the mean displacement of the cells drops, then slowly increases. In Figure 5.9(b), where tissue drift has been accounted for, the increase that occurs after 40 minutes no longer occurs. It can be concluded that this increase is due to tissue drift, and the mean displacement squared of the cells does not change much after 40 minutes, staying around  $15\mu m^2$ . This equates to the cells mean displacement after 100 minutes being about  $4\mu m$ , which is less than the diameter of a T cell,  $7\mu m$ . The way the displacement squared increases up until 40 minutes, and then drops, indicates that there is a population of cells that are barely moving (slow cells), and a population that is moving with

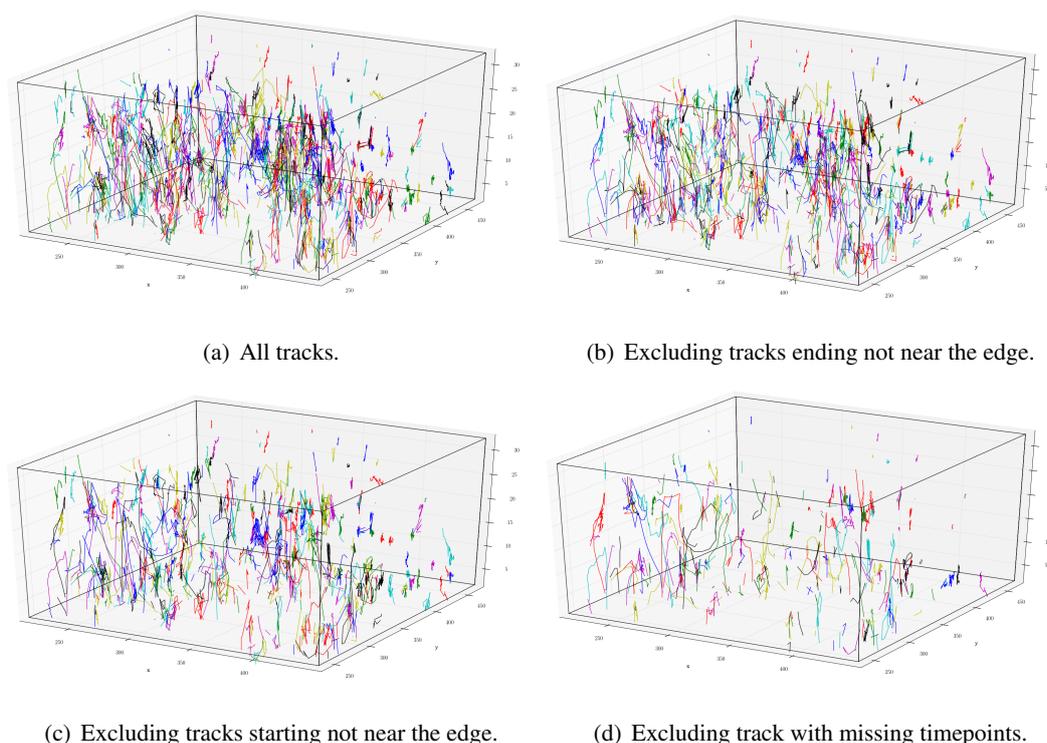


Figure 5.8: Plots of cell tracks in the imaging volume. Box indicates imaging volume at the start of the imaging period.

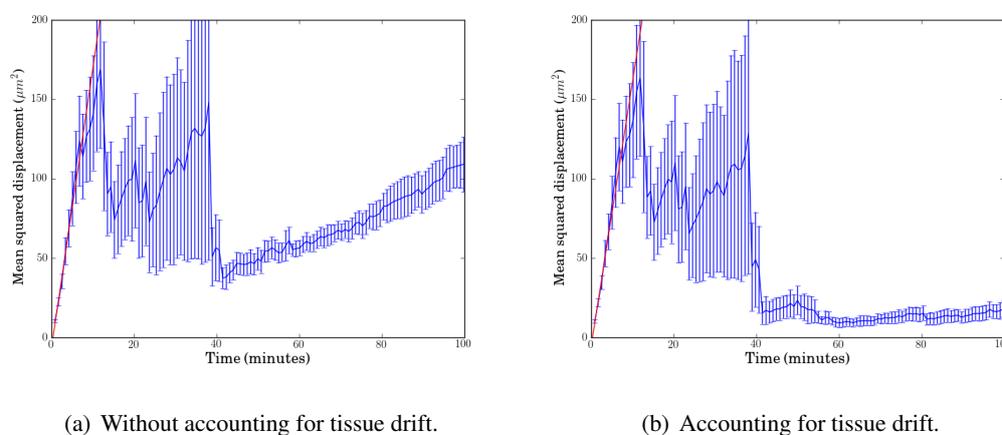


Figure 5.9: Plots of displacement squared against time. The red lines have been fitted to the first few time points, to calculate the diffusivities. In (a),  $D = 2.93 \pm 0.14 \mu m^2 min^{-1}$ , and in (b),  $D = 2.77 \pm 0.21 \mu m^2 min^{-1}$ .

a greater diffusivity. This causes a drop in the mean squared displacement as the faster moving cells are able to leave the imaging volume at earlier times. This results in the slower cells having a larger contribution to the mean after the faster cells have left. It also results in the standard error bars shrinking as there is a smaller amount of variability in the displacement of the cells, as

they are moving slower. The population of slow cells could be bound to stationary structures in the lymph node, such as the cytoskeleton, and their small movement could be due to movement of the cytoskeleton or other factors. In two-photon imaging experiments carried out by Miller *et al.* [156], they found a population of cells attached to structures by long membrane tethers, that were relatively stationary but exhibited probing behaviour. This behaviour suggests the cells were interacting with unseen elements in the environment. Visualisation of the reticular network using CMTMR staining, revealed T cells that adhered to, but did not crawl along, the fibres of the reticular network. In the paper by Miller *et al.* [156] they define stationary cells to be those with a displacement in one timestep of less than two micrometers.

To determine the diffusivity of the faster moving set of cells, the slow set of cells must be excluded, as they will be skewing the mean displacement of the faster-moving set of cells. This was achieved by calculating the mean distance moved in one timestep for each cell and deciding on a cut-off to classify the two populations. Closely examining the data, I found excluding cells with a mean displacement of less than  $1.5\mu m$  in one timestep, eliminated this slow set of cells. Plots of the two sets of tracks are shown in Figure 5.10.

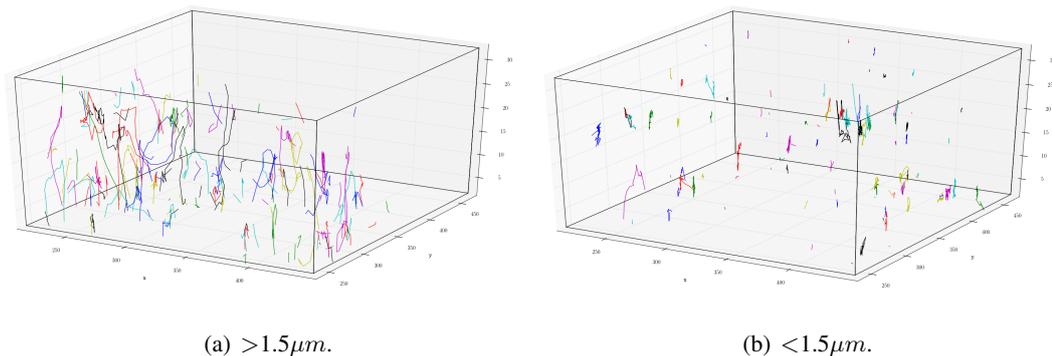


Figure 5.10: Tracks with mean displacement in one timepoint, (a) greater than, and (b) less than  $1.5\mu m$ .

By excluding the slow cells, and re-plotting the displacement squared against time graph, a diffusivity of  $12.22\mu m^2 min^{-1}$  is obtained (Figure 5.11(a)). The slow population of cells has a diffusivity of  $0.32\mu m^2 min^{-1}$  (Figure 5.11(b)). These diffusivities are calculated by fitting straight lines, to the displacement squared plots (Figure 5.11), for points when the cells appear to be undergoing random motion. The displacement squared of the faster cells (Figure 5.11(a)) drops after about 15 minutes, and then slowly increases again, indicating there is yet another population of cells moving slower, that are again going to be skewing the mean of the distribution. Using the

same process, this population of cells moving with an intermediate diffusivity, can be excluded by removing all cells with a mean displacement in one timestep of less than  $2.5\mu m$ . Plotted tracks are shown in Figure 5.12.

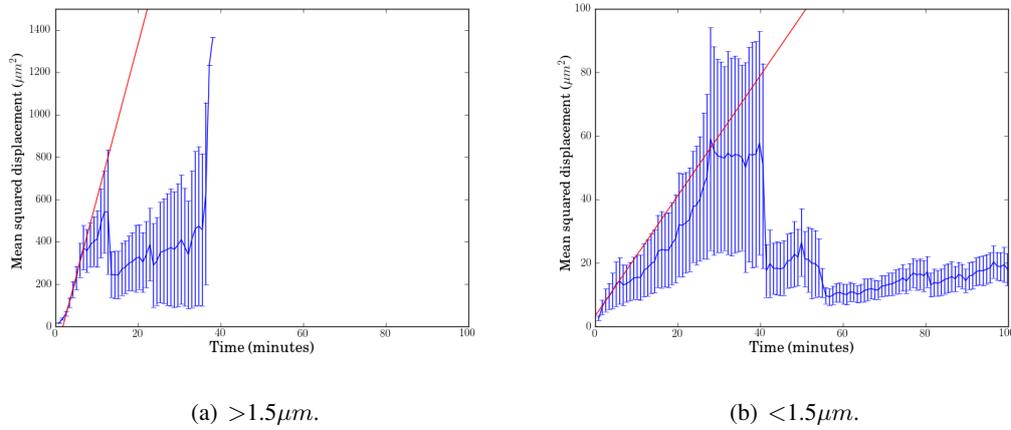


Figure 5.11: Displacement squared against time plots, with lines fitted (red) to calculate the diffusivity. In (a)  $D = 12.22 \pm 0.63\mu m^2 min^{-1}$ , and in (b)  $D = 0.32 \pm 0.02\mu m^2 min^{-1}$ .

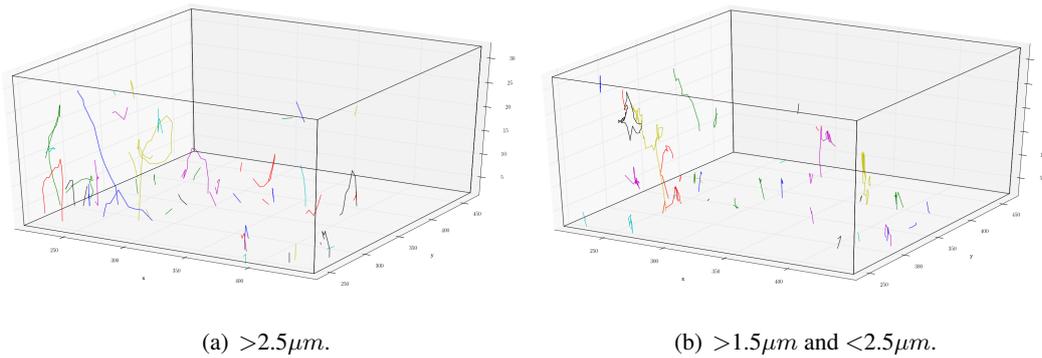


Figure 5.12: Tracks with mean displacement in one timestep, (a) greater than  $2.5\mu m$ , and (b) between  $1.5\mu m$  and  $2.5\mu m$ .

Figure 5.13 shows the cells with larger movement in one timestep (fast cells) have a diffusivity of  $23.84\mu m^2 min^{-1}$ , and the cells with intermediate movement in one timestep have a diffusivity of  $2.90\mu m^2 min^{-1}$ . Figure 5.13(a) shows the displacement squared increases roughly linearly with time, without the drop and increase seen in the other displacement squared graphs, indicating there is not another underlying population of cells in this reduced data set. Due to the limited size of the imaging volume, we would expect the displacement squared to plateau as time tends to infinity. However, due to the  $z$  dimension being a lot smaller than the  $x$  and  $y$  dimensions, most of the cells will escape via the  $z$  boundary, which will happen at shorter times.

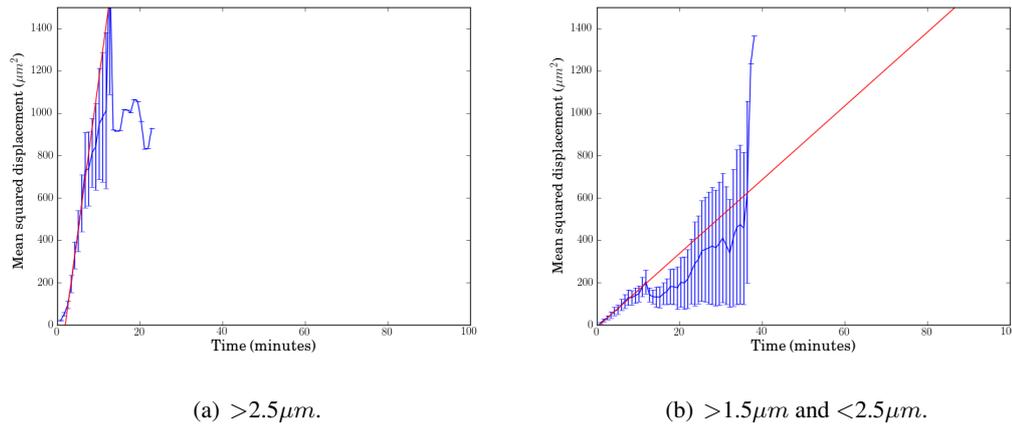


Figure 5.13: Displacement squared against time plots, with lines fitted (red), after persistent motion stops (determined using Figure 5.17), to calculate the diffusivities. In (a)  $D = 23.84 \pm 0.62$ , and in (b)  $D = 2.90 \pm 0.20 \mu m^2 min^{-1}$ .

To observe a confinement effect, cells persisting in the imaging volume until later time points, by moving in the  $x$  and  $y$  dimensions more, must be observed. To observe these tracks, there would need to be a large amount of data points. In this subset there are only 45 tracks remaining, due to sorting the data, which is not enough data to be able to observe these desired tracks. Therefore, we do not observe a plateau. Instead, the displacement squared increases with time, until the cells escape, after which point there is no more data. A slower increase in the displacement squared is observed in Figure 5.13(a) for the first few timesteps, indicating persistent motion in the population of cells. This can be investigated by analysing the mean turning angle between successive movements of the cells, and the scalar product between vectors along a cell's path.

### 5.4.3 Direction of motion

Using only the tracks with a mean displacement in one timestep greater than  $1.5\mu m$ , the turning angles of the cells can be analysed, shown in Figure 5.14(a). This population has a mean of  $111.90^\circ$ , indicating cells generally move in a persistent direction, and are less likely to abruptly turn back on themselves. Figure 5.14(b) shows that, cells with a mean displacement in one timestep of less than  $1.5\mu m$ , have a mean turning angle of  $80.90^\circ$ , indicating these cells make abrupt turns, and are more likely to move back in the direction they came from.

Figure 5.15 shows the analysis of turning angles for tracks with mean displacement in one timestep, (a) greater than  $2.5\mu m$ , and (b) between  $1.5\mu m$ , and  $2.5\mu m$ . Figure 5.15(a) shows

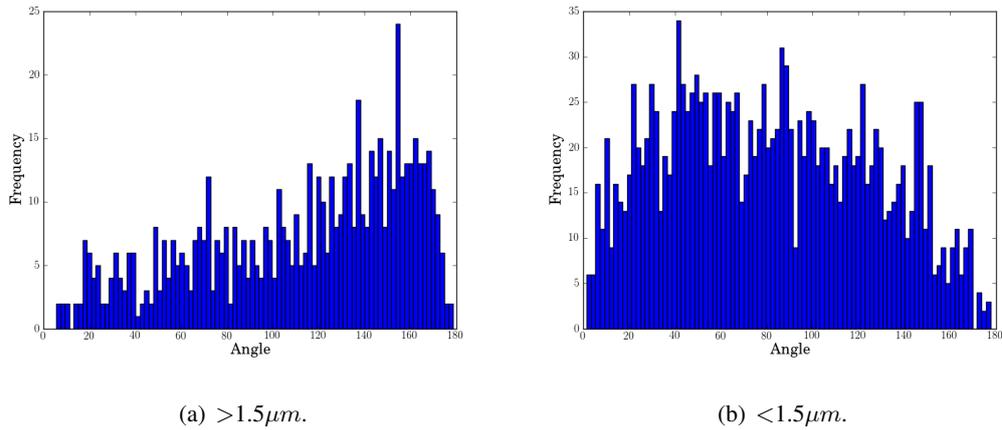


Figure 5.14: Turning angles of tracks with mean displacement in one timepoint, (a) greater than, and (b) less than  $1.5\mu m$ . Plot (a) has a mean of  $111.9^\circ$ , and (b) has a mean of  $80.90^\circ$ .

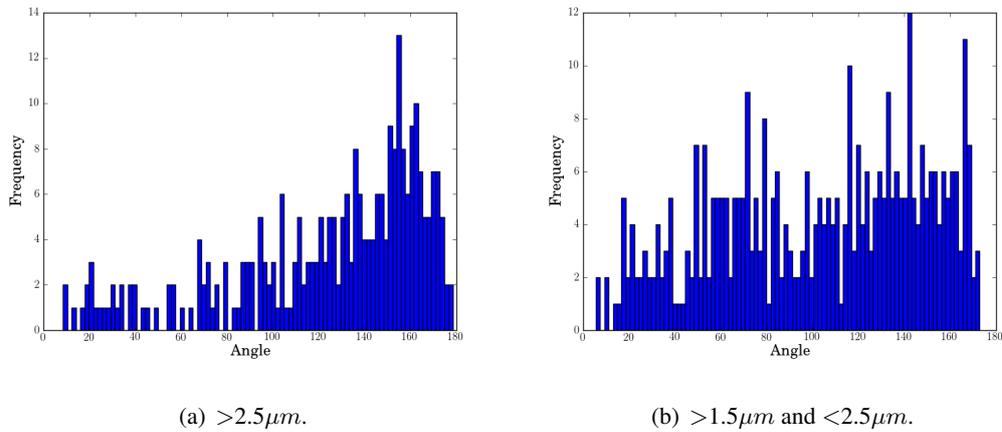


Figure 5.15: Turning angles of tracks with mean displacement in one timepoint, (a) greater than  $2.5\mu m$ , and (b) between  $1.5\mu m$  and  $2.5\mu m$ . Plot (a) has a mean of  $124.35^\circ$ , and plot (b) a mean of  $102.78^\circ$ .

the fast cells to have even less abrupt turning angles, with a mean of  $124.35^\circ$ . The tracks in Figure 5.15(b) have a mean turning angle of  $102.78^\circ$ , implying they also travel in a more persistent direction than the slow cells.

Persistent motion can be detected by calculating the mean scalar product between the  $i$ th and  $i + k$ th vectors along the track. Figure 5.16(a) shows the cells with a mean displacement in one timestep greater than  $1.5\mu m$  are undergoing persistent motion for the first five timesteps, after which there appears to be no persistence. Figure 5.16(b) shows the slow cells are not displaying persistent motion, which is expected for cells that are just moving slightly back and forth. Figure 5.17(a) shows there is again persistent motion for the first five timesteps of the fast cells, after

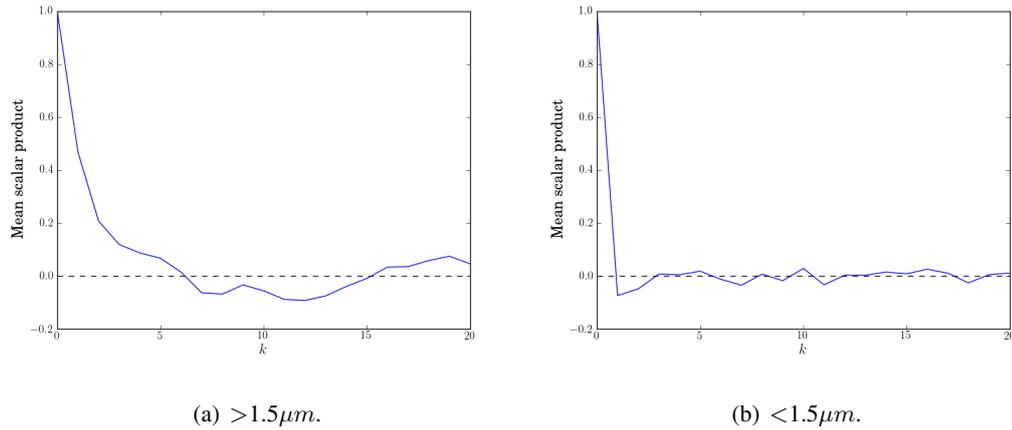


Figure 5.16: Dot product between vectors  $i$  and  $i + k$ , of tracks with mean displacement in one timestep: (a) greater than, and (b) less than  $1.5\mu m$ . Plot has been scaled to start at one for  $k = 0$ .

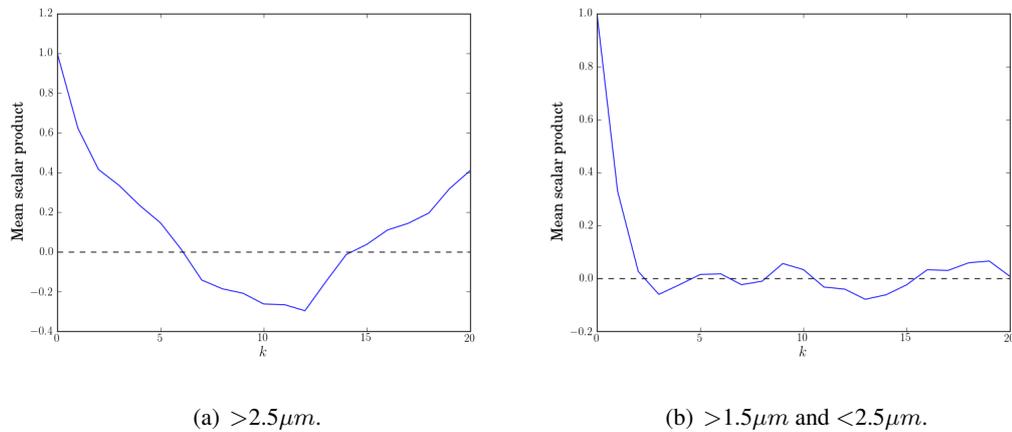


Figure 5.17: Dot product between vectors  $i$  and  $i + k$ , of tracks with mean displacement in one timestep: (a) greater than  $2.5\mu m$ , and (b) between  $1.5\mu m$  and  $2.5\mu m$ . Plot has been scaled to start at one for  $k = 0$ .

which there is no persistence. For the intermediate cells (Figure 5.17(b)), there is persistent motion for the first couple of timesteps, but not for subsequent timesteps.

It can be concluded that, there is persistent motion on short timescales (less than six timesteps), but for longer timescales the cells movement is well approximated by Brownian motion. Ideally, for this data, to extract the diffusion coefficient from a plot of displacement squared against time, the displacement needs to be calculated between points separated by five timesteps, to compensate for the persistent motion in the first five timesteps. In this chapter, displacement squared against time graphs are generated by calculating the displacement from each point to the starting point of the track. Compensating for persistent motion in this method would result in the first five timepoints

being cut off the plot. The line used to calculate the diffusion coefficient is fitted to points after persistent motion is found to stop. Therefore, removing the points where persistent motion is occurring would not change the value of the diffusion coefficient calculated from the fitted line. An alternative method considered for generating displacement squared plots, is to calculate the displacement from every subsequent point in a track to the first point, then the second point, then the third point, and so on. This multiple sampling method generates more data, especially for small time intervals. Using this method would cause a change in the diffusion coefficient if time points being compared are further apart. However, using this method results in long tracks having a larger contribution to the mean of the data, due to them being sampled numerous time. For this reason, this method of multiple sampling was not used in this chapter. This problem could potentially be compensated for by weighting a track's contribution to the mean, for example: all tracks could contribute equal weight to the mean. However, it may be beneficial to take numerous samples from long tracks, as these are the ones with the most data available, so could be the most reliable. Short tracks can be suspicious, due to cells being miss tracked, so it may not be desirable to weight them as heavily as the long tracks.

## 5.5 Procedure for calculating diffusion coefficient

Analysis of the previous data set enabled the development of a systematic procedure to calculate the diffusion coefficient. The procedure developed consists of the following steps:

**Step 1:** Determine the tissue drift in the data, using the method outlined in Section 5.3.4. All following analysis should be carried out compensating for the tissue drift in the experiment.

**Step 2:** Exclude all erroneous data, as outlined in Section 5.4.1.

**Step 3:** Using this sorted data set, plot the displacement squared against time. From this plot, adjust the cut-off for the minimum mean displacement in one timestep until the plot does not appear to plateau. The cut-off is generally between two and four micrometers per timestep.

**Step 4:** Using the trimmed data set, plot the mean scalar product between the  $i$ th and  $i + k$ th vectors in the tracks. Using this plot, determine the point at which persistent motion becomes less apparent (*i.e.* the mean scalar product is approximately zero).

**Step 5:** Fit a straight line to the displacement squared against time plot, starting from the timepoint at which persistent motion is found to stop, until the point at which there is too little data for the values to be conclusive. Using the fitted line, the diffusion coefficient can be calculated by taking one-sixth of the gradient.

### 5.5.1 Application to data sets

The procedure for calculating the diffusion coefficient can be tested by applying it to three data sets that were also provided by the research group in Glasgow. Plots for the mean scalar product and displacement squared, with a fitted line to calculate the diffusion coefficient, are shown in Figures 5.18 to 5.20.

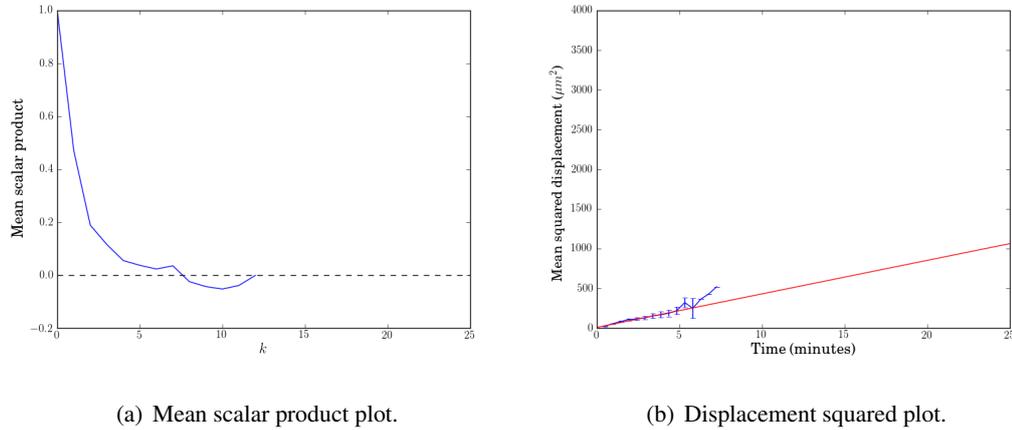


Figure 5.18: Data set LN2. Plots used to determine the diffusivity of the cell population,  $D = 7.03 \pm 0.46 \mu\text{m}^2 \text{min}^{-1}$ .

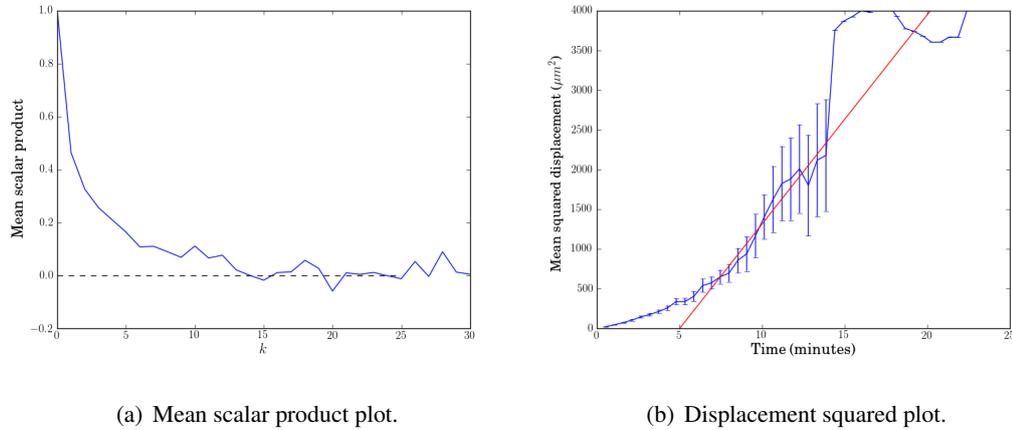
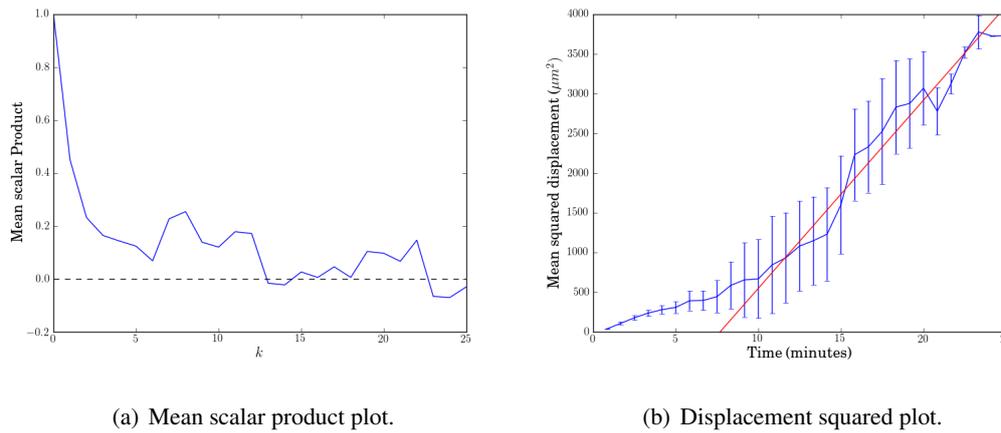


Figure 5.19: Data set LN4. Plots used to determine the diffusivity of the cell population,  $D = 43.94 \pm 3.01 \mu\text{m}^2 \text{min}^{-1}$ .

For these data sets, estimates for the diffusion coefficients are found to be:  $7.03 \pm 0.46 \mu\text{m}^2 \text{min}^{-1}$ ,  $43.94 \pm 3.01 \mu\text{m}^2 \text{min}^{-1}$ , and  $D = 39.49 \pm 2.68 \mu\text{m}^2 \text{min}^{-1}$ , for LN2, LN4 and Ear1, respectively. The amount of tracks that survive the sorting procedure for the data sets is 96, 145, and 35 tracks, respectively.



(a) Mean scalar product plot.

(b) Displacement squared plot.

Figure 5.20: Data set Ear1. Plots used to determine the diffusivity of the cell population,  $D = 39.49 \pm 2.68 \mu\text{m}^2 \text{min}^{-1}$ .

## 5.6 Processing data

To process the data set provided by the Glasgow laboratory, the programming language python was used. The data was imported using the package numpy, and assigned a data structure, allowing different aspects of the data to be easily called. To import the data the following commands were used:

```
import numpy as np
dtype = np.dtype([ ('tc', np.int), ('id', np.int), ('tp', np.int),
('xpos', np.float), ('ypos', np.float), ('zpos', np.float)])
data = np.loadtxt('GlasgowData01.dat', dtype = dtype)
```

In which `tc` is the cell number, `id` is the cell identity, `tp` is the timepoint of the data point, and `xpos`, `ypos` and `zpos` are the  $x$ ,  $y$  and  $z$  position of the cell. Individual cells could then easily be chosen by creating a boolean mask, which reads true for data relating to a specific cell and false for all other data values. Applying this mask to the whole data set, leaves a subset of data for a specific cell. Looping over the whole data set from the minimum to the maximum cell number, and applying a mask as described, allows analysis of each cell track. The following commands were used:

```
for i in range(data['tc'].min(), data['tc'].max()):
    mask1 = data[data['tc']==i]
```

## 5.7 Summary

In all the data sets, the displacement squared plots only have data for part of the total imaging time. This occurs because the cells are primarily escaping from the imaging volume via the  $z$  boundary. To obtain more accurate results, and be able to track cells for a larger fraction of the imaging time, an imaging volume with a larger  $z$  dimension is required. In TPLSM experiments, the  $z$  dimension is scanned in stacks. An increase in the size of the  $z$  dimension can be achieved in several ways. The separation between stacks could be increased, preserving the size of the time increments at which a  $z$  stack is scanned. However, this could lead to cells residing between the  $z$  stacks, and not being scanned. The number of  $z$  stacks could be increased, keeping the size of the separation between stacks. This would lead to the time increment between scanning a stack increasing, as there are more stacks to scan. If the time increment is made too large, less information about cell positions will be available, and cell identities could more easily be mistaken. An alternative solution to achieve a larger  $z$  dimension would be to reduce the size of the  $x$  and  $y$  dimensions. In the data analysed in this chapter, a  $z$  stack takes about two seconds to scan, if the  $x$  and  $y$  dimensions are reduced, this time would be reduced, allowing more  $z$  stacks to be scanned in a given time. This would allow a larger  $z$  dimension to be achieved, without compromising on the  $z$  stack separation, or the size of the time increments. By having a more cube like imaging volume, cells can be tracked for a longer period of time. This seems like a good solution to give more data, but it has its drawbacks. The imaging is carried out in a LN of a live mouse, by exposing the LN. The LN has an outer cortex, which must be penetrated to visualise T cell movement occurring deeper within the LN. If a larger  $z$  dimension is used, light must penetrate deeper within the LN, resulting in more scattering of the light, giving a decreased resolution. This limits the depth at which imaging can be performed. The imaging technology used in these experiments is constantly improving, and with future developments a deeper image will be able to be obtained. If a larger  $z$  dimension cannot be used, the next best improvement to make would be to reduce the  $x$  and  $y$  dimensions, allowing an increase of the resolution in the  $z$  dimension, or the temporal resolution.

To achieve a more accurate analysis of the data, more data points are required. A large amount of data are lost due to the tracks appearing or disappearing part way through the imaging time, whilst not being near the edge of the imaging volume. Data are also lost by the exclusion of tracks that have missing timepoints. To have more data to analyse, these artefacts must be dealt with; it is possible that this can be done on the existing data by further analysis, or it may need investigating experimentally. More data could be obtained by carrying out experiments with exactly the same

setup for the imaging volume and the experimental technique. In this way, data could be merged to give a larger data set.

Analysis of data from TPLSM experiments enabled the development of a method to calculate the mean diffusion coefficient of a data set of cell movements. The diffusion coefficients of the four data sets analysed ranged from 7 to  $44\mu\text{m}^2\text{min}^{-1}$ . This range could be due to inaccuracy in the data, such as how many anomalous tracks needed to be excluded. Variability can also occur due to the experimental setup. This could be due to the size of the imaging volume, the number and separation of  $z$  stacks taken, the time at which timepoints are taken, or the biological setup of the experiment.

## Chapter 6

# How many dendritic cells are required to initiate a T cell response?

### 6.1 Introduction

An adaptive immune response is initiated through encounters between rare naive antigen-specific T cells, and antigen-bearing dendritic cells (DCs), within the lymph node (LN). However, the number of DCs required to initiate a T cell response has never been precisely quantified [42]. Two-photon imaging experiments have revealed that, every hour, 500–5000 T cells come in close proximity with any given DC [42, 51, 153], but only a small proportion of these T cells will be specific for a given antigen. In the naive repertoire, T cells specific for a particular antigen, are presented at low precursor frequencies [6, 159], typically between  $10^{-6} - 10^{-5}$  [37, 42]. Experiments described in this chapter were carried out by Susanna Celli, Andreas Müller and Philippe Bousso, at the Institut Pasteur, Dynamics of Immune Responses Unit, Paris, France. Experiments were performed in a mouse popliteal LN, which contains approximately  $10^6$  T cells. With this number of cells, the given precursor frequencies corresponds to between one and ten T cells. Therefore, in the context of local infection, the number of DCs in the lymph node will influence the probability that a rare antigen-specific T cell becomes activated.

Current experimental techniques are able to measure, *in vivo*, the proportion of T cells activated by DCs in a LN. However, the experiments have their limitations and drawbacks. A roughly known number of DCs can be injected into a mouse, but only a proportion of those will reach the LN of interest. The proportion that reaches the LN can be estimated using *in vivo* imaging, but

methods of counting the DCs are often inaccurate. To accurately measure T cell activation, a large number of cells must be present. Therefore, experiments must start with hundreds of thousands of DCs, to enable tens of thousands to reach the LN, which are required to be able to accurately quantify the proportion of T cells activated, using flow cytometry (Section 1.8.2) or two-photon imaging (Section 1.8.1). These drawbacks result in current experimental techniques not being able to determine how many DCs are required to initiate a T cell response. A T cell response is initiated within twenty four hours. Currently, *in vivo* imaging data cannot be gathered twenty four hours after the experiments begin, to determine the proportion of activated T cells, as desired. Using mathematical modelling, simulations with fewer DCs, and for later times, can be performed.

In this chapter experiments carried out by the Bousso laboratory, in Paris, are reviewed. The experiments use flow cytometry [42, 55], and two-photon imaging [40]. Analytical and computational mathematical models are developed, to calculate the recruitment probability for a T cell. The models are then calibrated using experimental results, allowing extrapolation to lower DC numbers and later times, which cannot be investigated experimentally. In theory, only one T cell is required to initiate a T cell response, as once activated a T cell will produce multiple copies of itself. A model is developed to calculate the probability of at least one T cell becoming activated, allowing the minimum number of DCs required to elicit an immune response to be calculated. Knowing the minimum number of DCs required to initiate an immune response is important for many therapeutic applications, including DC vaccine against cancer.

## 6.2 Experimental design/results

Biological experiments were carried out by the Bousso laboratory (Institut Pasteur, Dynamics of Immune Response Unit, Paris, France). They used an experimental approach achieving synchronisation of antigen presentation [55] in the LN, allowing all the antigen to be presented at a given time. They injected MHC class two double negative (MHC class II<sup>-/-</sup>) mice (mice with no CD4 T cells), with wild type DCs, expressing a mature phenotype, and Marilyn CD4<sup>+</sup> T cells. After twenty four hours, the mouse was injected with Dby peptide, resulting in antigen presentation in the LN by the transferred DCs within minutes of injection (Figure 6.1a). After thirty minutes, the fraction of CD4<sup>+</sup> T cells that had received T cell receptor (TCR) signals was determined, by measuring c-jun phosphorylation, which is a very early marker of antigen recognition in T cells. It was found, that the fraction of T cells that received signal increased with

the number of DCs transferred (Figures 6.1b and 6.2).

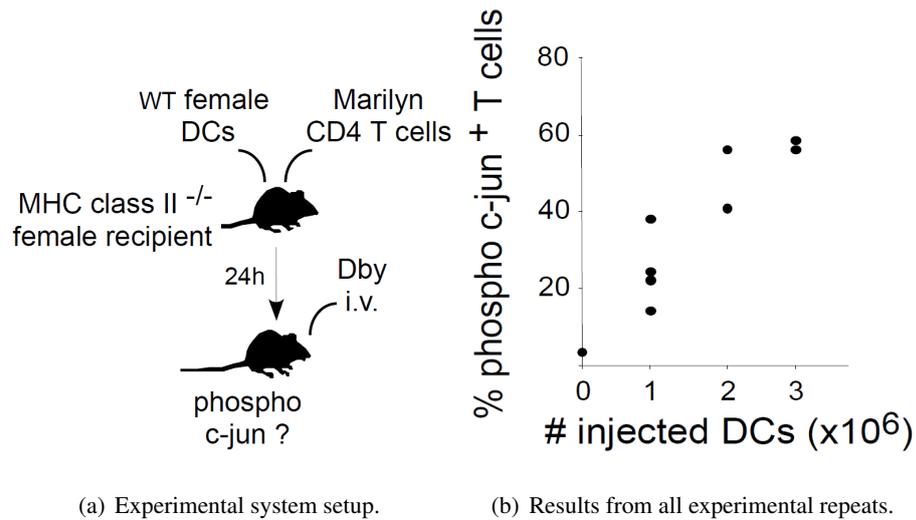


Figure 6.1: Experiment in which varying number of DCs are injected into mice, and the number of activated T cells are recorded using phospho c-jun. Reproduced from [53].

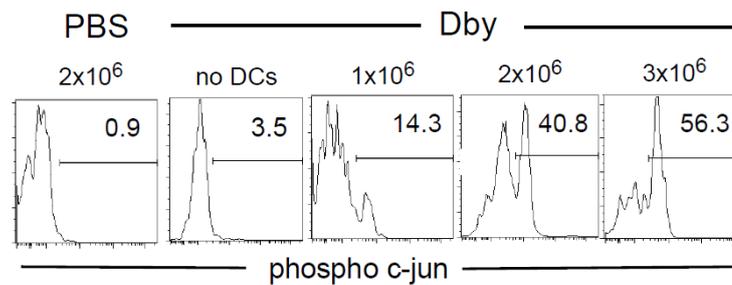


Figure 6.2: Results showing flow cytometry output of the number of T cells expressing phospho c-jun, for varying initial DC numbers. Reproduced from [53].

The number of DCs expressing green fluorescent protein (GFP), at the time of peptide injection was measured to determine how many of the injected DCs reached the LN. For example, they found that  $22286 \pm 797$  DCs were present in the LN one day after injection of  $2 \times 10^6$  DCs, and these DCs induced c-jun phosphorylation in  $48 \pm 7.6\%$  of antigen-specific T cells in thirty minutes (Figure 6.3).

Two photon laser scanning microscopy (TPLSM) was used to measure the time it takes for T cells to encounter, and bind, an antigen-bearing DC, in the same conditions, with 20000 DCs. An accumulation of T cell-DC interactions over time, as more, and more T cells encountered DCs, was observed. In five to ten minutes 50% of T cells formed a stable contact, increasing to 88% by thirty minutes (Figure 6.4).

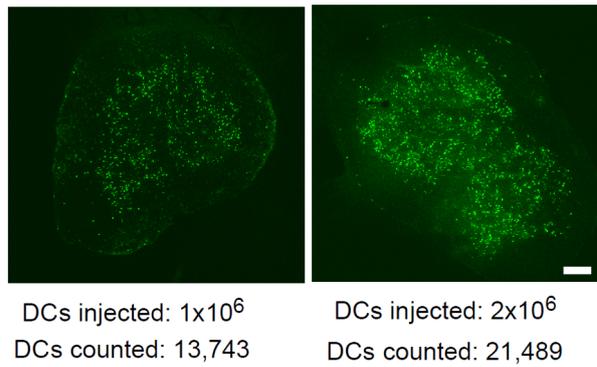
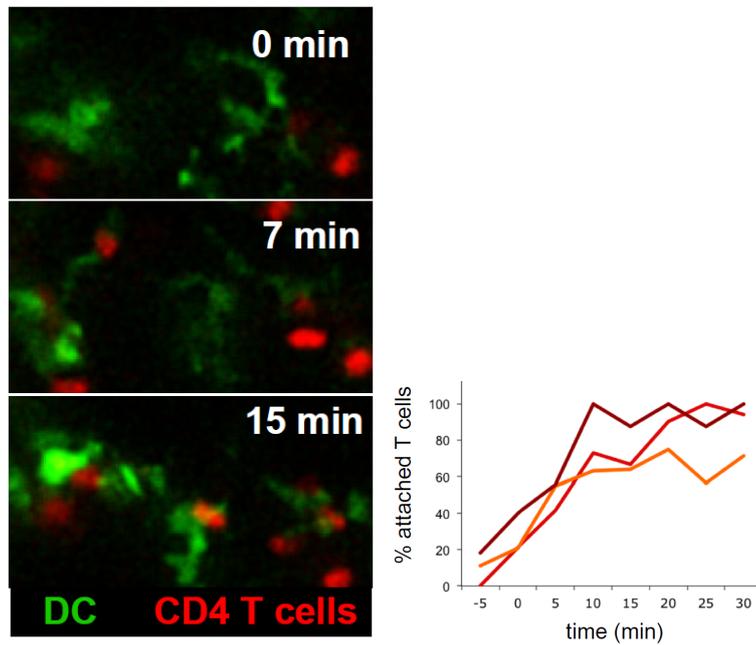


Figure 6.3: Quantifying number of DCs using GFP labelling. Green dots represent DCs. Reproduced from [53].



(a) Image from TPLSM *in vivo* experiment. (b) Plot of results from TPLSM experiments, showing the percentage of attached/activated T cells as a function of time. Each coloured line represents one experiment.

Figure 6.4: TPLSM experiments. Reproduced from [53].

### 6.3 Mathematical model

To create an mathematical model, for comparison to experimental results, the probability of a T cell encountering a DC must be calculated. We assume T cells encounter DCs at a constant rate,

$\alpha$ , given by the inverse of the mean first passage time (3.5.37), which is approximately  $\alpha \approx \frac{3Db}{R^3}$ , where  $D$  is the diffusivity of the T cells,  $b$  is the effective radius of the DCs, and  $R$  is the radius of the LN. The number of T cells that have not encountered a DC will decrease exponentially with time, giving an exponentially decreasing function for the probability of a T cell not encountering a DC. The probability that one T cell, chosen at random, does not encounter a DC, before time  $t$ , is given by

$$\mathcal{P}(\text{A T cell does not encounter a DC before } t) = e^{-\alpha Nt}, \quad (6.3.1)$$

where  $t$  is the imaging time and  $N$  is the number of DCs. The probability that one T cell chosen at random, does encounter a DC, is the complement of (6.3.1), given by

$$\mathcal{P}(\text{recruitment}) = 1 - e^{-\alpha Nt}. \quad (6.3.2)$$

This is a probability per T cell and is independent of the number of T cells.

The probability that at least one of  $M$  T cells encounters a DC, in the imaging period, can be computed by calculating the probability that no T cells encounter an DC and taking the complement of this probability, as was done to calculate (6.3.2). This gives

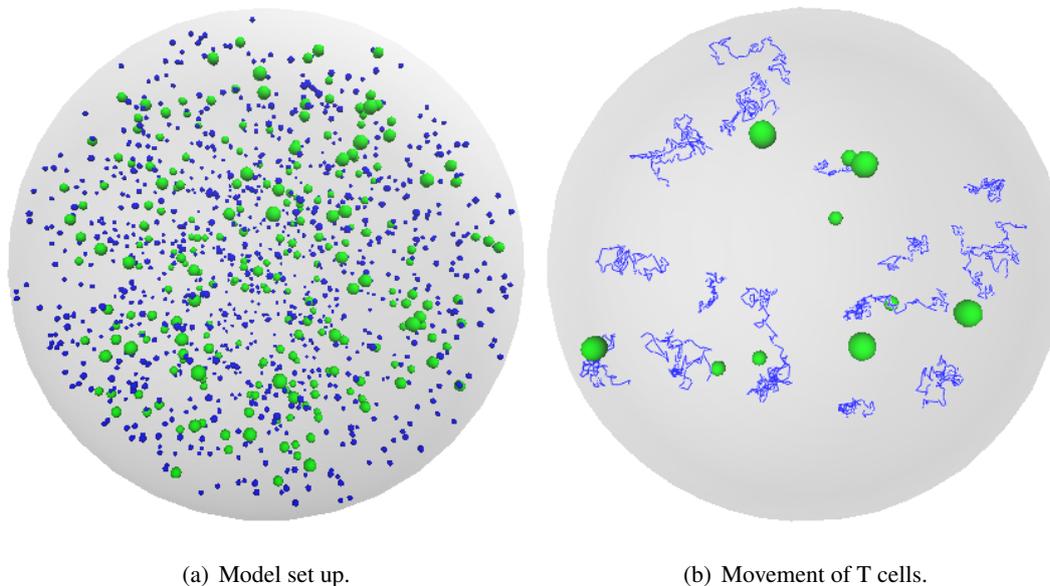
$$\mathcal{P}(\text{At least one T cell encounter an DC before } t) = 1 - e^{-\alpha NMt}, \quad (6.3.3)$$

where  $N$ ,  $\alpha$  and  $t$  are defined as previously, and  $M$  is the number of T cells. This probability depends on the number of T cells present.

Increasing the diffusivity of the T cells, or the the size of the DCs, will increase the encounter rate in a linear manner, but increasing the radius of the LN will decrease the encounter rate in a cubed way. Therefore, changing the size of the LN will have a much greater effect on the probabilities stated in (6.3.2) and (6.3.3), than changing the diffusivity of the T cells or the effective radius of the DCs.

## 6.4 Computational model

Using a variant of the spatial model described in Section 2.2.3, simulations of cells moving in a LN, until they encounter a DC, can be carried out. Figure 6.5 shows the simulation setup. Figure 6.5(a), is a screenshot of a simulation of a crowded environment, with numerous DCs and T cells, in the LN. Figure 6.5(b) shows the tracks of T cells moving by Brownian motion, in an attempt to encounter a DC.



(a) Model set up.

(b) Movement of T cells.

Figure 6.5: Screenshots of spatial model simulations. (a) shows a crowded environment, with randomly placed DCs (green) and T cells (blue). (b) shows tracks of T cells (blue), moving by Brownian motion, in an attempt to encounter DCs (green). The DCs and T cells are placed following a uniform random distribution, but are replaced if they occupy a volume that has already been occupied by a DC.

#### 6.4.1 Over-crowding effect

The target global mean first passage time (TGMFPT) (3.5.37) calculated for one DC uniform randomly placed in a LN, assumes that the remaining volume of the LN is unoccupied. This is the case when there is just one DC, but not when there are multiple DCs. For small numbers of DCs this does not pose a big problem, as the DCs are a lot smaller than the LN, so the space occupied by the DCs is negligible. However, when there are thousands of DCs in the LN, the space they occupy becomes significant, leaving less free space for the T cells to move in, which causes the TGMFPT equation to break down. The equation derived in Section 3.7.2 can be used for multiple targets. However, this equation assumes the DCs can overlap, which is not the case in these simulations. The equation can be used as an approximation, and will be a better approximation than assuming the DCs are independent, meaning they have no influence on each other, but the accuracy will decrease as the number of targets increases, due to the DCs having a higher probability of occupying the same volume, and the amount of free volume for the T cells to move in decreasing. A simple compensation can be implemented in an attempt to deal with the reduced amount of free space for the T cells to move within. This is done by using the

effective free volume, being the free space in the LN, rather than the actual volume of the LN. This is achieved by replacing the  $R^3$  term by  $R^3 - Nb^3$ , where  $N$  is the number of DCs. Figure 6.6 shows the effect on the TGMFPT equation, by comparing the equation with the effective free volume adjustment, against the equation without the adjustment.

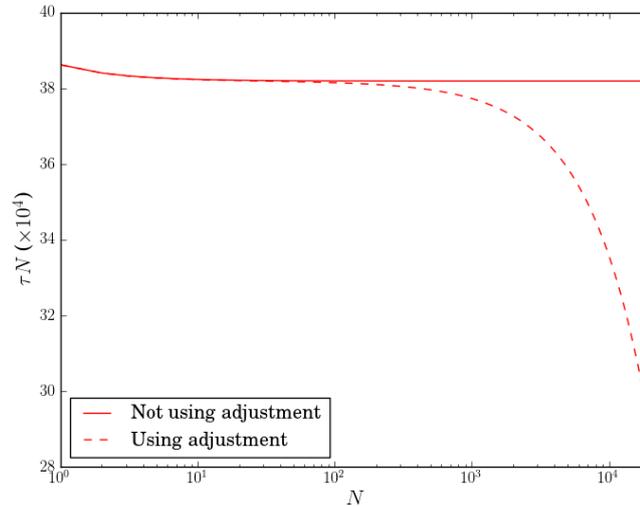


Figure 6.6: Effect crowding of DCs has on the TGMFPT equation.

#### 6.4.2 Correction of overestimation

In the numerical simulations, used to calculate activation probabilities, an over estimation to the analytical prediction was found. This occurs because, when placing the T cells, if they are placed within the zone of attraction of a DC, the position is rejected, and a new position chosen. This results in the T cell not having a uniform random distribution across the lymph node, which the equation for the TGMFPT is assuming. To resolve this problem, T cells were placed uniformly randomly within the lymph node, and replaced if they occupy the effective volume of a DC. Then the T cells were given a period of time (*e.g.*: 20 minutes) to equilibrate, allowing their distribution to become approximately uniform random. Any T cells encountering a DC in this time were removed from the simulation. After the equilibration time, data was recorded assuming the simulation started with the number of T cells present at the end of the equilibration time.

### 6.4.3 Estimation of parameters

Analysis of data sets from imaging experiments carried out by the Bouso laboratory were performed, as in Chapter 5, to determine the diffusivity of the T cells. Figure 6.7 shows displacement squared plots of the data, with fitted lines to calculate the diffusivity.

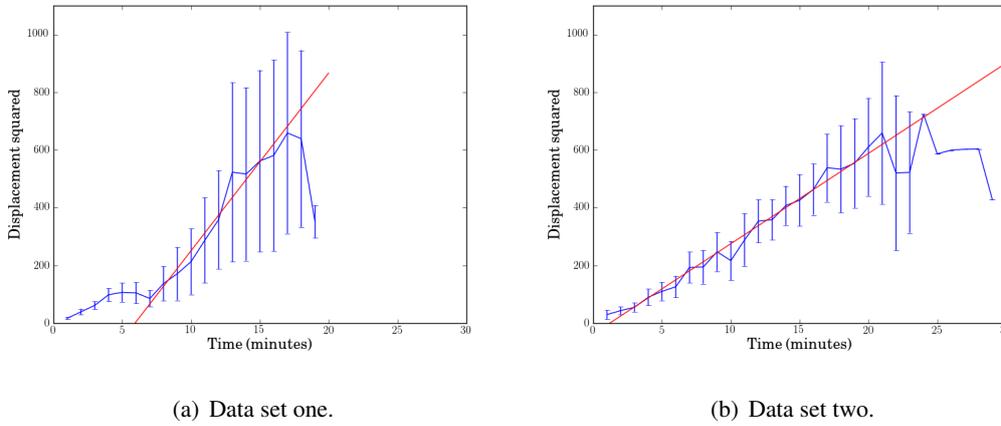


Figure 6.7: Analysis of displacement squared. Diffusivities calculated to be (a)  $10.26 \pm 0.73 \mu m^2 min^{-1}$  and (b)  $5.21 \pm 0.16 \mu m^2 min^{-1}$

The data was not as extensive as the data analysed in Chapter 5, resulting in a large amount of variability in the results. From the analysis, diffusivities of  $10.26 \pm 0.73 \mu m^2 min^{-1}$  and  $5.21 \pm 0.16 \mu m^2 min^{-1}$  were found. During the analysis, the data used to produce Figure 6.7(a), was found to be more reliable. Analysis of the cell tracks for the data used to generate 6.7(b) revealed all the cells to be remaining in the same  $z$  plain throughout the imaging period. For this reason this data set was judge to not be reliable as this should not be happening. Therefore, a diffusivity of  $10 \mu m^2 min^{-1}$  will be used in this chapter. The experiments were performed in a popliteal LN, the radius of which, was measured to be approximately  $400 \mu m$ , although there is uncertainty in this parameter, with some literature claiming the popliteal LN is as large as  $700 \mu m$  in radius [164]. The size of the DCs is a lot more uncertain as, due to their dendrites being able to reach out and contract, their size is constantly changing [17]. T cells are known to have a fairly constant radius of  $3.5 \mu m$  [1]. After consultation with the Bouso laboratory, a radius of  $8 \mu m$  for the DCs was agreed upon. This results in an effective radius for the DCs of  $11.5 \mu m$ , and is the value used for  $b$ . The radius of the LN was chosen to be within the range stated above, with a value of  $500 \mu m$  being used.

#### 6.4.4 Comparison to experimental results

Experiments carried out by the Bousso laboratory found that in 5-10 minutes, 50% of T cells had formed a stable contact, increasing to 88% by 30 minutes. Using (6.3.2) with parameters  $R = 500\mu m$ ,  $b = 11.5\mu m$ ,  $D = 10\mu m^2 min^{-1}$ , and  $N = 2 \times 10^5$ , the probability of recruitment is found to be 50.0% in 10 minutes, increasing to 87.4% in 30 minutes, which is a very good fit to the experimental data. The parameters  $R$ ,  $b$ , and  $D$ , were chosen within realistic ranges. These parameters are used to calculate the parameter  $\alpha$ .

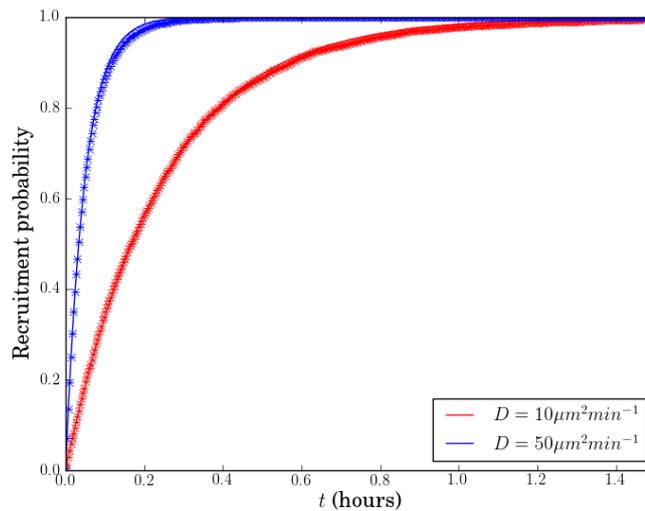


Figure 6.8: Recruitment probability for  $2 \times 10^5$  DCs. Equation (6.3.2), is shown by solid lines, numerical results, with standard error bars, are shown by crosses. Parameters used  $R = 500\mu m$  and  $b = 11.5\mu m$ .

Figure 6.8 shows a comparison of (6.3.2) to numerical simulations, for diffusivities of  $10\mu m^2 min^{-1}$  and  $50\mu m^2 min^{-1}$ . To simulate the numerical results shown in Figure 6.8,  $2 \times 10^5$  DCs were placed in the LN. This proved not to be possible when the DC's effective volumes are not allowed to overlap, as placing them uniformly randomly, within the free volume of the LN, results in a large amount of wasted space. To overcome this problem, the DCs were placed with a minimum separation between their centres of  $2b_1$ , where  $b_1$  is the actual radius of a DC. This results in DC volumes not being able to overlap, but their effective volumes being able to overlap, as the effective volume takes into account the size of the T cells. This setup is shown in Figure 6.9. The analytical equation used is (6.3.2), with an effective free volume used as described in section 6.4.1. The numerical results fit surprisingly well to the theoretical equation.

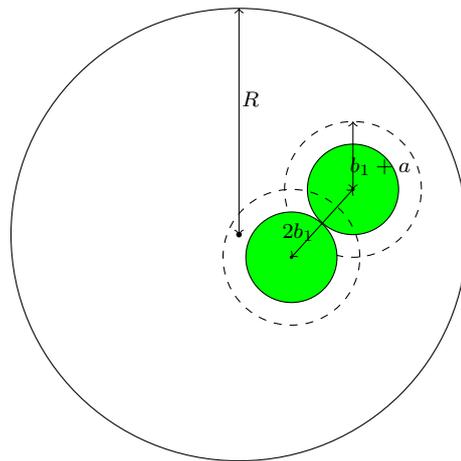


Figure 6.9: Diagram showing the setup of the computational model. Green circles represent DCs, with their effective volume shown by a dashed line. The centres of DCs are separated by twice the DCs actual radius,  $2b_1$ , and  $a$  is the radius of a T cell, incorporated into the DC's effective radius.

#### 6.4.5 Extrapolation of experimental results

Encouraged by the good fit between the model and the activation probabilities obtain from the experimental data, the model was used to calculate recruitment probabilities for lower DC numbers, and longer times, which can not be investigated experimentally. When using  $2 \times 10^5$  DCs, in the biological experiments and mathematical model, the recruitment probability almost reaches 100% in about 30 minutes. These large numbers of DCs are required to carry out biological experiments, but such large numbers are not necessarily required to invoke an immune response. Using the biological data, the model parameters could be calibrated, then, using these parameters, other scenarios can be simulated. The recruitment probability can be calculated for varying numbers of DCs, after six, twelve and twenty four hours. Using (6.3.2) and numerical simulations, as shown in Figure 6.10. Figure 6.10 shows in twenty four hours, about one thousand DCs are required to achieve approximately 100% recruitment probability for a T cell. Twenty four hours is used, as this is the time an immune response takes to be initiated.

In theory, only one T cell needs to be activated for an immune response to occur, as once activated the T cell will undergo clonal expansion. Using (6.3.3), the probability of at least one T cell becoming activated can be calculated, for smaller numbers of T cells, and DCs, than can be achieved experimentally.

Figure 6.11 shows the probability of at least one T cell becoming activated in twenty four hours,

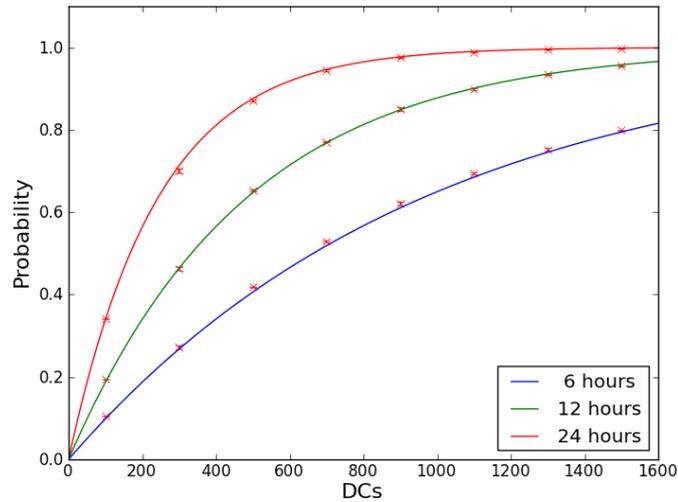


Figure 6.10: Recruitment probability per T cell. Equation (6.3.2) is shown by solid lines, and numerical results, simulated using spatial simulations, by crosses.  $R = 500\mu m$ ,  $D = 10\mu m^2 min^{-1}$  and  $b = 11.5\mu m$ .

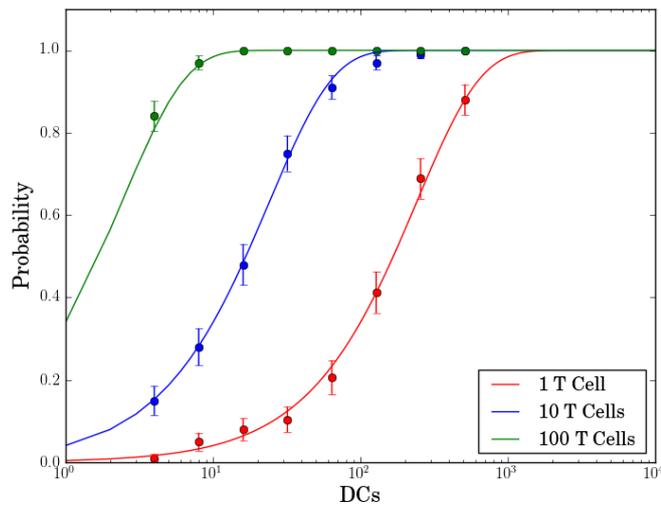


Figure 6.11: Probability of at least one T cell encountering a DC, in twenty four hours. Equation (6.3.3) is shown by solid lines, and numerical results, simulated using spatial simulations, by dots.  $R = 500\mu m$ ,  $D = 10\mu m^2 min^{-1}$  and  $b = 11.5\mu m$ .

compared to numerical simulations, for varying DC numbers, and T cell initial numbers, of one, ten, and one hundred. These initial values of T cells, correspond to precursor frequencies of  $10^{-6}$ – $10^{-4}$ . The precursor frequency expresses the fraction of T cells that are specific to a

particular antigen. A mouse popliteal LN contains approximately  $10^6$  T cells. The figure shows that approximately ten DCs are required to achieve a probability of approximately 100%, that at least one T cell encounters a DC in twenty four hours, with one hundred searching T cells. This number increases to approximately one hundred DCs, for ten T cells, and to approximately one thousand DCs, for just one T cell.

#### 6.4.6 Minimum number of DCs required to initiate an T cell response

Using (6.3.3), the number of DCs required to activate at least one T cell can be determined. Let us assume a T cell response is initiated if the recruitment probability, in a given time period, is greater than 50%. The number of DCs required to achieve a 50% recruitment probability, in time  $T$  can be obtained by rearranging (6.3.3), with the recruitment probability set to  $\frac{1}{2}$ :

$$N^* = \frac{\log 2}{\alpha NT}. \quad (6.4.4)$$

Equation (6.4.4) can be used to calculate the minimum number of DCs required to initiate a T cell response. Typical precursor frequencies of T cells are  $10^{-6}$ – $10^{-5}$ , equate to 1 – 10 T cells. These values give a minimum number of DCs required to be 174 and 17 respectively, in twenty four hours. These numbers increase to 698 and 70 respectively in a six hour period. These results are illustrated in Figure 6.12.

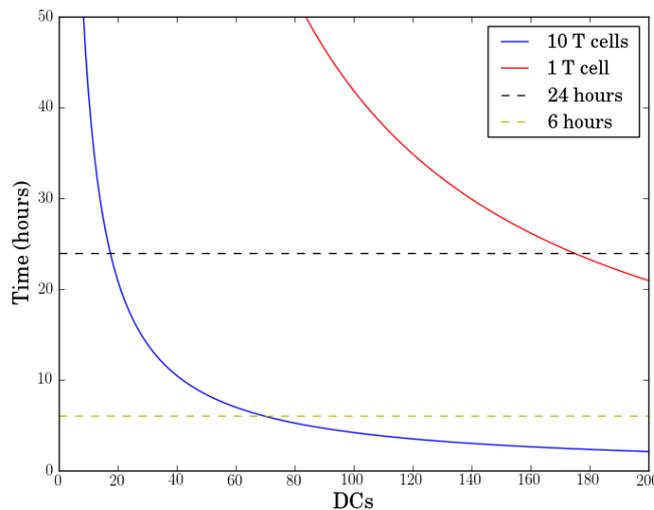


Figure 6.12: Minimum number of DCs required to achieve a 50% probability of at least one T cell encountering a DC.

In Figure 6.12, the minimum number of DCs was calculated to achieve a 50% activation probability. A more general version of (6.4.4) can be derived by including the term  $p$ , the probability of encountering a DC, giving

$$N_p^* = \frac{\log\left(\frac{1}{1-p}\right)}{\alpha NT}. \quad (6.4.5)$$

Using this equation, the minimum number of DCs required to induce a T cell response can be calculated as a fraction of the recruitment probability required. For a T cell precursor frequency of  $10^{-5}$  (10 T cells), the minimum number of DCs required for the probabilities of encounter to be 0.25, 0.5, 0.75 and 0.95, in twenty four hours, are 7, 17, 35, and 75 respectively. This is illustrated in Figure 6.13

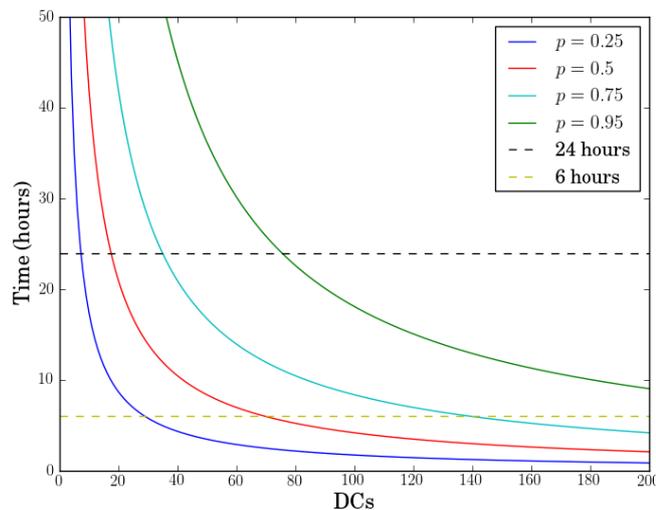


Figure 6.13: Minimum number of DCs required to achieve a probability  $p$ , of at least one T cell encountering a DC.

## 6.5 Summary

In this chapter, it is found, that a minimum of 174 DCs are required to achieve a probability greater than 50%, of initiating and immune response, within twenty four hours, and a precursor frequency of  $10^{-6}$  T cells. This minimum drops to 17, for a precursor frequency of  $10^{-5}$  T cells. Values of one and ten T cells are used, for precursor frequencies of  $10^{-6}$  and  $10^{-5}$ , as the estimated number of T cells in a mouse popliteal LN is  $10^6$ . The dependence on the number of T cells can

be removed from (6.4.4), by letting  $\rho$  equal the density of T cells in the LN, thus  $\rho\phi = N/V$ , where  $V = \frac{4}{3}\pi R^3$ , and the fraction of antigen specific T cells, or precursor frequency, equals  $\phi$ . Doing this, gives the equation

$$N^* = \frac{\log 2}{4\pi} \frac{1}{\rho D b T} \frac{1}{\phi}, \quad (6.5.6)$$

where  $\alpha \approx \frac{3Db}{R^3}$ . This shows that  $N^*$  is inversely proportional to the precursor frequency. In humans, naive antigen specific T cells, are also presented at precursor frequencies of  $10^{-6}$ – $10^{-5}$ . Therefore, the same minimum number of DCs also applies to humans. This result also means that the number of DCs required to initiate a T cell response depends on the precursor frequency, but not the volume of the LN. Therefore, it does not matter if the DCs specific to a particular antigen ( $A^*$ ) are all in one LN or divided between several LN as long as the precursor frequency is the same in all LNs. The human body can only accommodate a limited number of T cells. Therefore, to have a vast coverage by as many antigen specific T cells as possible, only a few T cells specific to any given antigen can be present [53]. The results in this chapter suggest that typical precursor frequencies of  $10^{-6}$ – $10^{-5}$ , allow the diversity of T cells to be maximised, but still allows efficient T cell recruitment by low numbers of DCs. If numbers of DCs fall below this threshold, stimulation would most likely not take place, due to lack of encounters.

This chapter considers an estimate of the minimum required of DCs needed to initiate a T cell response, but higher numbers of DCs could further enhance the magnitude of the response, as well as allowing successive encounter between T cells and DCs [40, 42]. On the other hand, if chemokines are playing a role, to attract T cells to DCs, fewer DCs will be required to initiate a response [50, 111]. DC vaccines can be used to elicit an immune response against cancer cells. This work suggests the efficiency of delivery of DCs may not be of critical importance, as T cell recruitment by DCs is already an efficient process. Optimising the quality and duration of the stimulation may be of greater importance to DC vaccines [53].

## Chapter 7

# Cross-regulation model

### 7.1 Introduction

Millions of different clonotypes of T cells are created by a stochastic recombination process during development [106, 137, 200, 211]. Inevitably some of these clonotypes are self-reactive to the body's own antigens. To alleviate this problem, T cells undergo positive and negative selection, to select out clonotypes that react to the body's own antigens. Even with this process some self-reactive clonotypes slip through the net, which is where regulatory T cells (Tregs) play a role. Although it may have been agreed that Tregs are key players in avoiding autoimmunity, and are currently the topic of intensive experimental research [44, 179, 199], the mechanism of action of Tregs is still not fully known.

Several mechanisms for Treg mediated suppression have been suggested [195, 204]. It has been proposed that suppression by Tregs requires direct cell-cell contact, either between Tregs and effector T cells [203], or antigen-presenting-cells (APCs), Tregs, and effector T cells [183, 217, 223]. Experiments *in vitro* have shown Tregs are unable to suppress the proliferation of T cells when separated by a semi-permeable membrane [217, 223], and culture supernatant of antigen-stimulated Tregs is unable to suppress T cell proliferation, indicating direct cell-cell contact is required. Following direct cell-cell contact, Tregs may kill their target cells by a granzyme-dependant, or perforin-dependent mechanism, or give a negative signal to the target T cells, inhibiting their proliferation [38, 48, 89]. How exactly cell-cell suppression occurs is still unclear. Hypotheses in the literature range from simple competition between Tregs and effector T cells for conjugation to an APC [136, 229], to more complex interactions in which an APC relays

a signal from a Treg to an effector cell [52,215,216].

Tregs do not produce the cytokine IL-2, but do express the high affinity CD25/IL-2 receptor (IL-2R), making Tregs efficient absorbers of IL-2, thereby hindering the activation of T cells that may rely on IL-2 to proliferate [176]. This makes Tregs highly dependent on exogenous IL-2 for their survival in the periphery [109, 146, 174, 236, 240]. A lack of IL-2 leads to fatal autoimmune/inflammatory disease, which is mainly due to lack of Treg survival [140]. Treg suppression via absorption of IL-2 has been shown to be a strong mechanism *in vitro* [68,203]. *In vivo* experiments have shown strong autoimmune responses in mice deficient in IL-2, or the IL-2 receptor, suggesting IL-2 is essential for Treg proliferation, but is not essential for effector T cell proliferation [198]. Experiments have shown *in vitro* and *in vivo* competition for IL-2 does occur [20, 68, 127].

The theory of IL-2 suppression appears to disagree with experiments in which a semi-permeable membrane, separating Tregs from their target cells, prevents suppression [38,48,89], as cytokines diffuse easily in extracellular medium, and can diffuse hundreds of microns in the timescale of tens of minutes. However, measurements of cytokine concentrations in cell culture supernatants are often very low, and very low concentrations are unlikely to trigger cellular responses [33]. Therefore, concentrations of cytokines large enough may only be achieved in micro-environments surrounding secreting cells. Also, cytokines are taken up locally by other cells, which further restricts their spatial range [198].

The APCs that provide the proliferation signal to effector T cells are a potential target of Tregs. Activated Tregs may down-regulate the expression of CD80 and CD86 [197], which provide co-stimulatory signals, on APCs, as well as causing APCs to express the enzyme indoleamine 2,3-dioxygenase [95, 172], which breaks down the amino acid, tryptophan to kynurenine, which is toxic to T cells close to the APC.

The models in this chapter are based on Kalet Léon and Jorge Carnerio's cross-regulation model [49, 133] of interactions between effector (E) cells, regulatory (R) cells, and APCs, and has received experimental support [218,219,239]. Léon *et al.*'s model is a deterministic ODE model. Modelling biological systems using a deterministic approach can lead to unrealistic results, due to the inherent stochastic nature of biological systems. For example, a steady state with only a handful of cells could be achieved in a deterministic model, which in reality is likely to go extinct due to random fluctuations. A more realistic model of a biological system is a stochastic model, taking into account random fluctuations in the population sizes. A stochastic system can

be implemented by listing the reactions that occur in a model, and using a Gillespie algorithm (Section 2.4) to simulate these reactions.

CD4<sup>+</sup> T cells are divided into multiple distinct lineages. However, for this chapter only two lineages are considered: Tregs and conventional T helper cells, also known as effector T cells. Effector T cells control the adaptive immune response by activating, in an antigen specific fashion, other effector cells, such as CD8<sup>+</sup> cytotoxic T cells, B cells and macrophages. Tregs are T cells in charge of suppressing potentially deleterious activity of conventional T helper cells [64].

As with Léon *et al.*'s model, the models in this chapter only consider antigen specific suppression, and all cells in the models are assumed to be of the same clonotype. The main model examined assumes a three-way interaction between an E cell, R cell, and APC, is required for suppression to occur. However, this does not restrict the model to one of just cell-cell contacts, it is also a reasonable approximation for a model in which IL-2 is secreted by E cells, upon binding with an APC, and used to proliferate by R cells if they are in close proximity when the E cell is secreting IL-2. Both cell types would need to be close to the APC, at the same time, for the R cell to absorb the IL-2 being secreted by the E cell, allowing the R cell to proliferate and preventing the E cell from proliferating due to lack of IL-2.

In this chapter, three models are considered: Model 1- competition only, Model 2- R cell suppression only, and Model 3- R cell suppression and E cell promotion. These are the same three model scenarios that are considered by Léon *et al.* For all models ODEs are determined and solved numerically, with comparison to the stochastic system, solved using a Gillespie algorithm. For the third model, which proves to be the most interesting, the system is non-dimensionalised, reducing the number of parameters to two. The effects of varying these parameters are investigated, and two differing sets of parameter values are used for the analysis, allowing comparison of the system in two distinct parameter regimes. Steady states are obtained for the chosen parameter sets, and their stability investigated. The effect of the initial values of cells is investigated by plotting phase planes, to view the trajectories of the system, and considering the steady states achieved for various sets of initial cell numbers, allowing a bifurcation in the system to be observed. The stochastic effects of varying the initial ratio of the two cell types is investigated, and found to be significant for certain parameter regimes.

The aim of this chapter is to develop a mathematical model that can adequately describe the activation process of R cells and suppression of E cells by R cells. A stochastic framework is used to incorporate the inherent stochasticity found in biological system for small cell numbers,

which can be the case in autoimmunity. It is also hoped that a stochastic model will be able to reveal interesting aspects of the system that are not observed in an analytical approach.

### 7.1.1 Model dynamics

In all three models, cells proliferate depending on which other cells have contacted a given APC in the same time period. The rules determining cell proliferation differ for the three model scenarios investigated. Each APC is given two binding sites for T cells to bind to, for simplicity. Both E and R cells bind with APCs at a rate of  $k_+$ , unbind at a rate of  $k_-$ , and die with rate  $\mu$ . In all models, only free cells are able to die, while cells bound to an APC cannot die for the duration of the APC contact. The on rate,  $k_+$ , is given by the inverse of the first passage time for a T cell to encounter an APC (3.5.37), and the off rate,  $k_-$ , is given by the inverse of the mean binding time of a T cell and APC. In the models, T cell proliferation occurs upon unbinding of cells.

$E$  is used to represent effector cells and  $R$  to represent regulatory cells. The state of an APC is expressed as  $A_{ij}$ , where  $i$  is the number of E cells bound, and  $j$  is the number of R cells bound to that APC. Each APC is assumed to have two binding sites. There are six basic populations of APCs:  $A_{00}$ ,  $A_{10}$ ,  $A_{01}$ ,  $A_{11}$ ,  $A_{20}$  and  $A_{02}$ . In model 2, there are extra APC populations in which E cells are suppressed, and in model 3 there are APC populations in which R cells have been promoted. When an E or R cell unbinds from  $A_{11}$ , it leaves the remaining cell suppressed (if it is an E cell, models 2 and 3) or promoted (if it is an R cell, model 3). To distinguish these populations from  $A_{10}$  and  $A_{01}$ , APCs that have an E cell bound, which has been suppressed and will not proliferate on unbinding, are written as  $A_{10}^*$ , and APCs that have an R cell bound, which has been promoted and will proliferate upon unbinding, are written as  $A_{01}^\#$ . The case in which an E or an R cell binds to one of these suppressed/promoted populations of APCs, must now be considered. If an E cell binds to  $A_{10}^*$ , this is represented as  $A_{20}^*$ , indicating one of the E cells has been suppressed and the other has not. The same concept applies to an R cell binding to give  $A_{02}^\#$ . If an E cell binds to  $A_{01}^\#$  it becomes  $A_{11}$ , and if an R cell binds to  $A_{10}^*$  it becomes  $A_{11}$ , as it does not matter if a cell has previously been suppressed or promoted when the  $A_{11}$  complex is formed. For the purpose of writing ODEs, the APC populations can be written in the form  $n_k$ . These populations are expressed in Table 7.1, along with the parameters used in this chapter.

In the systems investigated, the steady states obtained can be split into three categories: extinction, in which no cells remain, autoimmunity, in which only E cells remain, and tolerance, in which both E and R cells remain.

$A_{ij}$	$n_k$	Number of cells bound	Rate	Equals
$A_{00}$	$n_{00}$	No cells	$k_+$	Binding rate
$A_{10}$	$n_{10}$	1 E cell	$k_-$	Unbinding rate
$A_{01}$	$n_{01}$	1 R cell	$\mu$	Death rate
$A_{20}$	$n_{20}$	2 E cells	$N_A$	Total number of APCs
$A_{02}$	$n_{02}$	2 R cells	$t$	Time
$A_{11}$	$n_{11}$	1 E cell, 1 R cell	$\kappa_+$	$k_+N_A/\mu$ (dimensionless binding rate)
$A_{10}^*$	$n_{10}^*$	1 suppressed E cell	$\kappa_-$	$k_-/\mu$ (dimensionless unbinding rate)
$A_{01}^\#$	$n_{01}^\#$	1 promoted R cell	$\tau$	$t\mu$ (dimensionless time)
$A_{20}^*$	$n_{20}^*$	2 E cells, 1 suppressed		
$A_{02}^\#$	$n_{02}^\#$	2 R cells, 1 promoted		

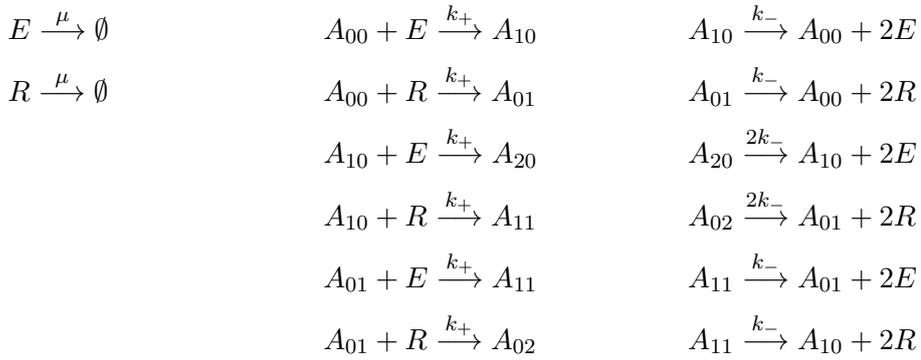
Table 7.1: Tables showing notation and rates used in this chapter.

## 7.2 Model 1: competition only

Firstly, we consider a model in which cells compete for the binding sites of the APCs, with no suppression or promotion. In this model both E cells and R cells proliferate upon unbinding from an APC, regardless of which other cells are bound to the given APC.

### 7.2.1 Reactions

The interactions between cells, in Model 1, can be represented as a system of reactions as follows



### 7.2.2 ODEs

These reactions can be written as the following system of ODEs:

$$\frac{dn_E}{dt} = 2k_-(n_{10} + 2n_{20} + n_{11}) - n_E\mu + k_+n_E(n_{20} + n_{02} + n_{11} - N_A), \quad (7.2.1a)$$

$$\frac{dn_R}{dt} = 2k_-(n_{01} + 2n_{02} + n_{11}) - n_R\mu + k_+n_R(n_{20} + n_{02} + n_{11} - N_A), \quad (7.2.1b)$$

$$\frac{dn_{10}}{dt} = k_-(n_{10} + 2n_{20} + n_{11}) - k_+(n_{10}(2n_E + n_R) + n_E(n_{01} + n_{20} + n_{02} + n_{11} - N_A)), \quad (7.2.1c)$$

$$\frac{dn_{01}}{dt} = k_-(n_{01} + 2n_{02} + n_{11}) - k_+(n_{01}(n_E + 2n_R) + n_R(n_{10} + n_{20} + n_{02} + n_{11} - N_A)), \quad (7.2.1d)$$

$$\frac{dn_{20}}{dt} = -2k_-n_{20} + k_+n_{10}n_E, \quad (7.2.1e)$$

$$\frac{dn_{02}}{dt} = -2k_-n_{02} + k_+n_{01}n_R, \quad (7.2.1f)$$

$$\frac{dn_{11}}{dt} = -2k_-n_{11} + k_+(n_{10}n_R + n_{01}n_E). \quad (7.2.1g)$$

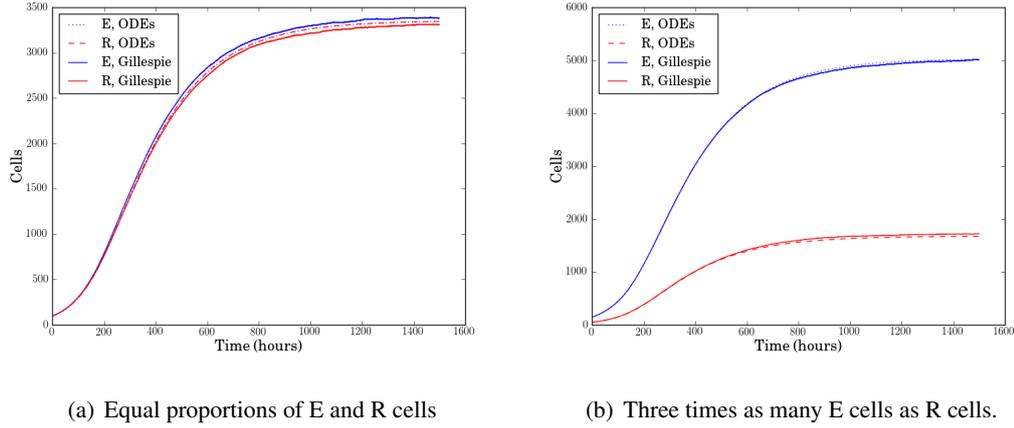


Figure 7.1: Plots showing the solution of ODEs and Gillespie algorithm for Model 1. Parameters used:  $k_+ = 1.2 \times 10^{-5} \text{minutes}^{-1}$ ,  $k_- = 1/60 \text{minutes}^{-1}$ ,  $\mu = 10^{-5} \text{minutes}^{-1}$ , and  $N_A = 25$ . Gillespie results averaged from 100 realisations in both plots.

This model is symmetric, as the populations of E and R cells can be interchanged without effecting the dynamics of the model, therefore neither population has an advantage over the other if the system starts with equal numbers of E and R cells. Figure 7.1 shows for (a) equal initial numbers of E and R cells, the steady state for both cell types is equal, and for (b) unequal initial numbers of E and R cells, the steady state achieved contains cell numbers in equal proportions to the starting proportions of each cell type. In both Gillespie simulations, an average is taken over multiple realisations, and compared to the ODEs. In Figure 7.1(a), the Gillespie results do not agree exactly with the ODEs, this is due to not enough realisations have been carried out. This has been done purposely so the Gillespie and ODEs solutions can be distinguished.

The steady state value of E and R cells can be calculated by first determining the carrying capacity of the system. The carrying capacity is the total number of E and R cells the system can support due to the limited number of APCs and binding sites. The carrying capacity can be found by assuming either the E or R cells go extinct, and solving the system with the other cell type. This gives

$$\text{Carrying capacity} = \frac{k_- \left( k_+ N_A - \mu + \sqrt{k_+^2 N_A^2 - \mu^2} \right)}{k_+ \mu}. \quad (7.2.2)$$

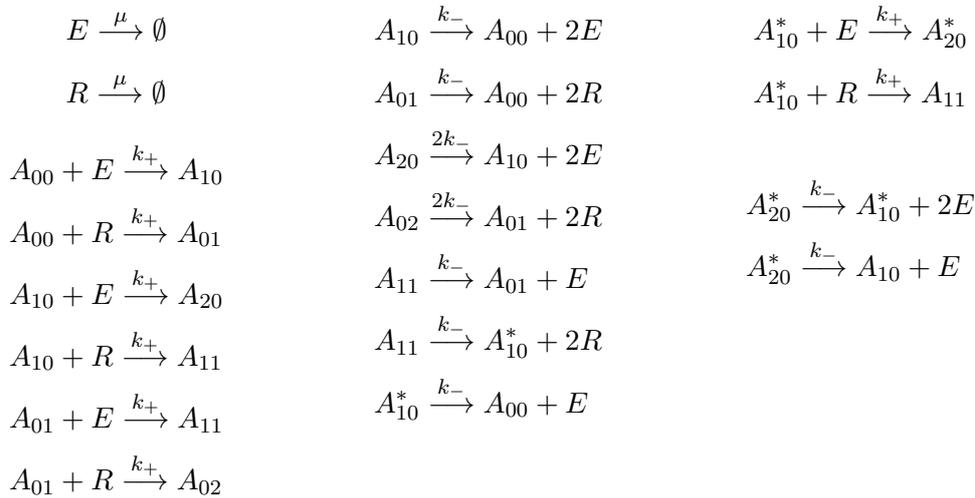
The steady state value of E and R cells is given by the initial fraction of each of the cell types multiplied by the carrying capacity (7.2.2) of the system. This occurs because the system is symmetric, so neither population has an advantage over the other.

### 7.3 Model 2: R suppression only

We now consider a model in which R cells are able to suppress E cells, but do not require the E cells to be able to proliferate. In this model, if an R cell binds to an APC, it receives a signal to proliferate, regardless of what other cells are bound to that APC. If an E cell binds to an APC, and if no R cells are bound to the same APC, it will proliferate. However, if an R cell is bound to the APC at any time during the E cell's binding, the E cell is suppressed and will not proliferate after unbinding.

#### 7.3.1 Reactions

The interactions in this model can be represented by the following system of reactions



### 7.3.2 ODEs

These reactions can be written as the following system of ODEs:

$$\frac{dn_E}{dt} = k_-(2n_{10} + n_{10}^* + 4n_{20} + 3n_{20}^* + n_{11}) - n_E\mu + k_+n_E(n_{20} + n_{20}^* + n_{02} + n_{11} - N_A), \quad (7.3.3a)$$

$$\frac{dn_R}{dt} = 2k_-(n_{01} + 2n_{02} + n_{11}) - n_R\mu + k_+n_R(n_{20} + n_{20}^* + n_{02} + n_{11} - N_A), \quad (7.3.3b)$$

$$\begin{aligned} \frac{dn_{10}}{dt} &= k_-( -n_{10} + 2n_{20} + n_{20}^* ) \\ &\quad - k_+(n_{10}(2n_E + n_R) + n_E(n_{10}^* + n_{01} + n_{20} + n_{20}^* + n_{02} + n_{11} - N_A)), \end{aligned} \quad (7.3.3c)$$

$$\frac{dn_{10}^*}{dt} = k_-( -n_{10}^* + n_{20}^* + n_{11} ) - k_+n_{10}^*(n_E + n_R), \quad (7.3.3d)$$

$$\begin{aligned} \frac{dn_{01}}{dt} &= k_-( -n_{01} + 2n_{02} + n_{11} ) \\ &\quad - k_+(n_{01}(n_E + 2n_R) + n_R(n_{10} + n_{10}^* + n_{20} + n_{20}^* + n_{02} + n_{11} - N_A)), \end{aligned} \quad (7.3.3e)$$

$$\frac{dn_{20}}{dt} = -2k_-n_{20} + k_+n_{10}n_E, \quad (7.3.3f)$$

$$\frac{dn_{20}^*}{dt} = -2k_-n_{20}^* + k_+n_{10}^*n_E, \quad (7.3.3g)$$

$$\frac{dn_{02}}{dt} = -2k_-n_{02} + k_+n_{01}n_R, \quad (7.3.3h)$$

$$\frac{dn_{11}}{dt} = -2k_-n_{11} + k_+((n_{10} + n_{10}^*)n_R + n_{01}n_E). \quad (7.3.3i)$$

Solving these ODEs using an Euler method, as well as the Gillespie algorithm, for this system gives the plots shown in Figure 7.2. Figure 7.2 shows the R cells drive the E cells to extinction, which is expected, as there is no restriction on the number of R cells by suppression from the E cells. The autoimmune steady state, with only E cells, can be achieved by setting the initial number of E cells to be much greater than the number of R cells.

## 7.4 Model 3: R suppression E promotion

Finally, we consider a model in which R cells are able to suppress the proliferation of E cells, and R cells require promotion by E cells to undergo proliferation. In this model, upon an E or an R cell binding to an APC, they proliferate following the rules outlined below, and in Figure 7.3.

For E cells:

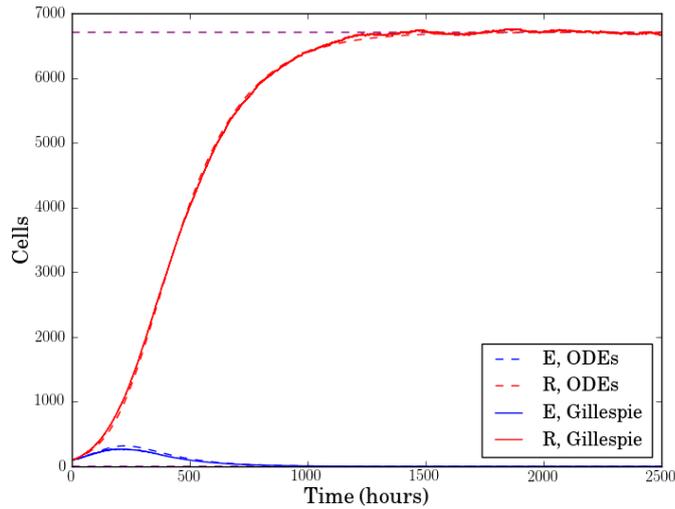


Figure 7.2: Dynamics of Model 2 with  $k_+ = 1.2 \times 10^{-5} \text{minutes}^{-1}$ ,  $k_- = 1/60 \text{minutes}^{-1}$ ,  $\mu = 10^{-5} \text{minutes}^{-1}$ , and  $N_A = 25$ , and initial values of E and R cells set to 100. Purple dotted line shows the steady state of the system. Gillespie results averaged from 100 realisations.

- The E cell proliferates if and only if no R cells have bound to APC in the time the E cell has been bound.

For R cells:

- The R cell proliferates if and only if an E cell has bound to the APC at any point during the time the R cell has been bound. We call this promotion of R cells by E cells.

In this model, there are 10 populations of APCs, plus the free E cells and free R cells. The total number of APCs is constant, therefore the  $A_{00}$  population can be constrained, giving a total of 11 populations, 24 reactions and 3 rates. A system of reactions for this model can be written, following the rules of Figure 7.3, with two binding sites per APC.

### 7.4.1 Reactions

The interactions in Model 3 can be expressed by the following set of reactions.

**Death:** In principle, we can have separate death terms for E and R cells, but for simplicity both death rates are set to be equal. We represent these as



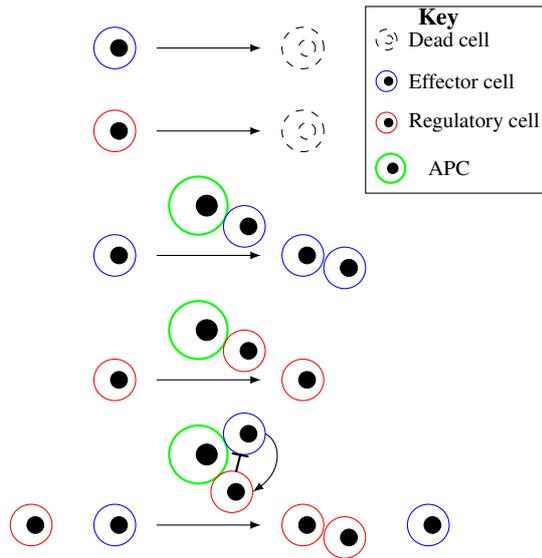
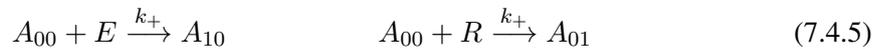
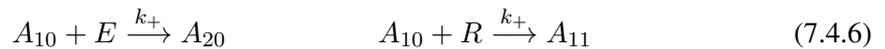


Figure 7.3: Diagram of interactions that can occur in Model 3, blue circles represent E cells, red circles represent R cells, dashed circles represent dead cells, green circles represent APCs, line in shape of ‘T’ indicates suppression of the cell contacting the flat end, and curved arrow indicates promotion.

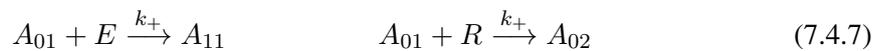
**Binding to an empty APC:** An E, or an R cell, can bind to an empty APC with rate  $k_+$ . Although there are two binding sites available, the rate of a T cell encountering an APC is independent of the number of free binding sites.



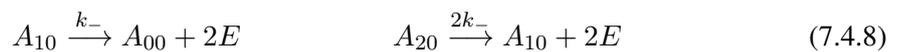
**Binding to  $A_{10}$ :**



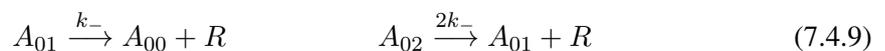
**Binding to  $A_{01}$ :**



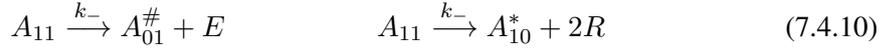
**Unbinding of E cells:** Unbinding and proliferation of E cell to produce two E cells, as suppression has not occurred. Rates are multiplied by the number of cells that can potentially unbind.



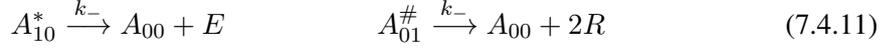
**Unbinding of R cells:** Unbinding of an R cell, without proliferating, to produce one R cell, as promotion has not occurred. Rates are multiplied by the number of cells that can potentially unbind.



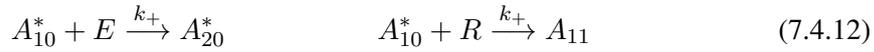
**Unbinding from  $A_{11}$ :** The # indicates the remaining R cell is promoted, and the \* the remaining E cell is suppressed.



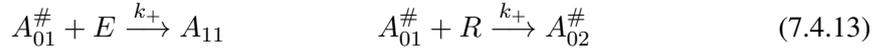
**Unbinding of suppressed/promoted cells:**



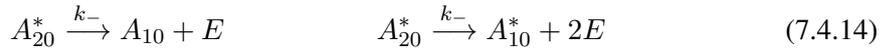
**Binding to  $A_{10}^*$ :** It must be noted that the complex resulting from the first reaction has an E cell that has been suppressed, represented by the \*. For the  $A_{11}$  case, the E cell is always suppressed, therefore the \* is not required.



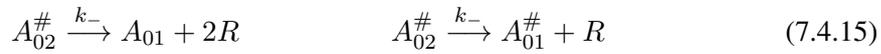
**Binding to  $A_{01}^\#$ :** Again, it must be noted that one of the resulting complexes has an R cell that has been promoted, which is represented by the #.



**Unbinding from  $A_{20}^*$ :** Either the suppressed, or unsuppressed E cell can unbind.



**Unbinding from  $A_{02}^\#$ :** Either the promoted, or un-promoted R cell can unbind.



These reactions can be represented as a diagram, shown in Figure 7.4.

## 7.4.2 Effector cell dynamics

Let us first consider the simple case when there are no R cells, but only E cells and APCs. The variables of the model are the number of free E cells and the numbers of  $A_{00}$ ,  $A_{10}$  and  $A_{20}$ , denoted as  $n_E$ ,  $n_{00}$ ,  $n_{10}$  and  $n_{20}$ , respectively. In this case, we have the following system of ODEs:

$$\frac{dn_E}{dt} = k_-(2n_{10} + 4n_{20}) - k_+(n_{00} + n_{10})n_E - \mu n_E, \qquad (7.4.16a)$$

$$\frac{dn_{00}}{dt} = k_-n_{10} - k_+n_E n_{00}, \qquad (7.4.16b)$$

$$\frac{dn_{10}}{dt} = 2k_-n_{20} + k_+n_{00}n_E - k_-n_{10} - k_+n_E n_{10}, \qquad (7.4.16c)$$

$$\frac{dn_{20}}{dt} = k_+n_E n_{10} - 2k_-n_{20}. \qquad (7.4.16d)$$

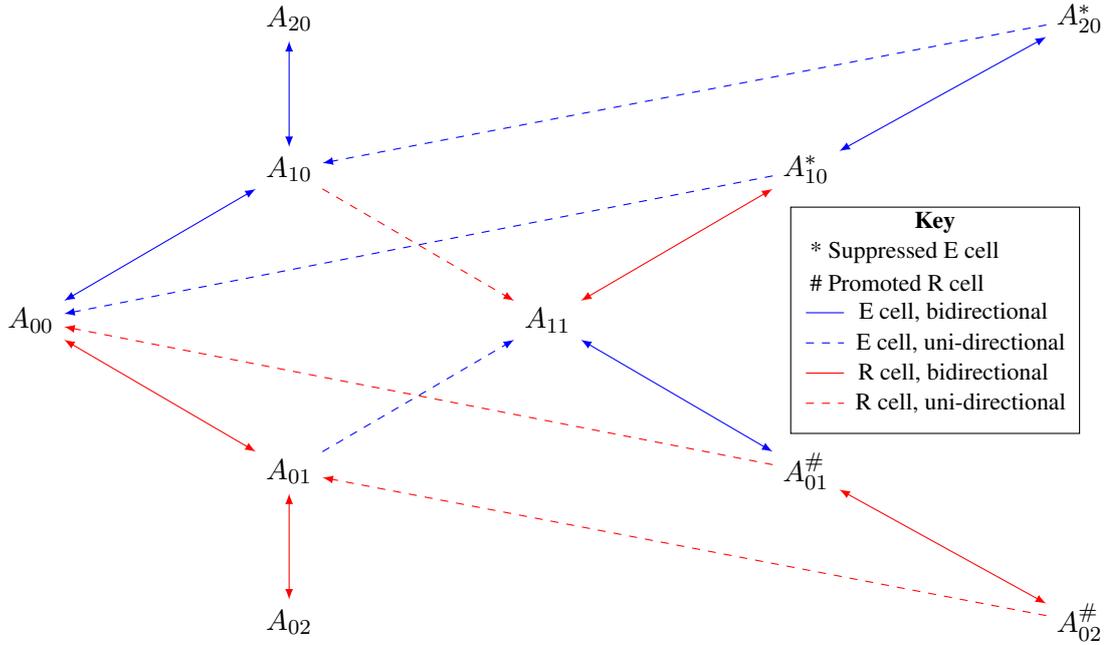


Figure 7.4: Diagram showing the reactions that occur in Model 3. Uni-directional refers to a reaction that occurs in only one direction, and bidirectional to a reaction that is reversible.

The total number of APCs is constant, therefore one of the populations can be eliminated as follows:

$$N_A = n_{00} + n_{10} + n_{20}.$$

Writing  $n_{00} = N_A - n_{10} - n_{20}$ , the equation for  $n_{00}$  (7.4.16a) can be eliminated, leaving a system of three ODEs:

$$\frac{dn_E}{dt} = k_-(2n_{10} + 4n_{20}) + k_+(n_{20} - N_A)n_E - \mu n_E, \quad (7.4.17a)$$

$$\frac{dn_{10}}{dt} = k_-(-n_{10} + 2n_{20}) - k_+(2n_{10} + n_{20} - N_A)n_E, \quad (7.4.17b)$$

$$\frac{dn_{20}}{dt} = -2k_-n_{20} + k_+n_{10}n_E. \quad (7.4.17c)$$

This system can be non-dimensionalised by setting,  $k_+ = \frac{\kappa_+\mu}{N_A}$ ,  $k_- = \kappa_-\mu$ ,  $t = \frac{\tau}{\mu}$  and  $n_i = N_A x_i$ . By non-dimensionalising, the number of parameters in the system is reduced. This yields the following dimensionless equations:

$$\frac{dx_E}{d\tau} = \kappa_-(2x_{10} + 4x_{20}) + \kappa_+(x_{20} - 1)x_E - x_E, \quad (7.4.18a)$$

$$\frac{dx_{10}}{d\tau} = \kappa_-(-x_{10} + 2x_{20}) - \kappa_+(2x_{10} + x_{20} - 1)x_E, \quad (7.4.18b)$$

$$\frac{dx_{20}}{d\tau} = -2\kappa_-x_{20} + \kappa_+x_{10}x_E, \quad (7.4.18c)$$

where  $x$  is used to represent the dimensionless variables. Non-dimensionalising reduces the system to two parameters,  $\kappa_+$  and  $\kappa_-$ . The real, non-negative, steady state solutions of (7.4.18) are

$$\{x_E, x_{10}, x_{20}\} = \{0, 0, 0\}, \quad (7.4.19)$$

and

$$\{x_E, x_{10}, x_{20}\} = \left\{ \frac{\kappa_-}{\kappa_+} \left( \kappa_+ - 1 + \sqrt{\kappa_+^2 - 1} \right), \frac{1}{\kappa_+} \left( -\kappa_+ + 1 + \sqrt{\kappa_+^2 - 1} \right), 1 - \frac{1}{\kappa_+} \right\}. \quad (7.4.20)$$

The stability of the steady states can be investigated by writing the ODE system (7.4.18) in matrix form. The values of  $x_E$ ,  $x_{10}$  and  $x_{20}$  can then be set to the steady state value and eigenvalues of the system found. If these eigenvalues are all negative, the system is stable. Analysing the stability of these steady states reveals the zero steady state to have eigenvalues:

$$\begin{aligned} & -2\kappa_-, \\ & -\frac{1}{2} \left( \kappa_- + \kappa_+ + 1 + \sqrt{\kappa_-^2 + \kappa_- (6\kappa_+ - 2) + (\kappa_+ + 1)^2} \right), \\ & -\frac{1}{2} \left( \kappa_- + \kappa_+ + 1 - \sqrt{\kappa_-^2 + \kappa_- (6\kappa_+ - 2) + (\kappa_+ + 1)^2} \right), \end{aligned}$$

which are all negative. Therefore the steady state is stable when  $\kappa_+ < 1$ . The eigenvalues of the non-zero steady state are rather long, and are not expressed here, but are only real, negative and therefore stable for  $\kappa_+ > 1$ .

This simplified system of ODEs can be plotted, and are compared to stochastic simulations generated using the Gillespie algorithm, shown in Figure 7.5

### 7.4.3 Regulatory cell dynamics

A non-zero steady state with no  $E$  cells and only  $R$  cells, is not possible, as the  $R$  cells rely upon the  $E$  cells to proliferate. This can be shown by forming a system of ODEs with only  $R$  cells, giving the following system of three ODEs:

$$\frac{dx_R}{d\tau} = -x_R + \kappa_-(x_{01} + 2x_{02}) + \kappa_+x_R(x_{02} - 1), \quad (7.4.21a)$$

$$\frac{dx_{01}}{d\tau} = \kappa_-(-x_{01} + 2x_{02}) - \kappa_+x_R(2x_{01} + x_{02} - 1), \quad (7.4.21b)$$

$$\frac{dx_{02}}{d\tau} = -2\kappa_-x_{02} + \kappa_+x_{01}x_R. \quad (7.4.21c)$$

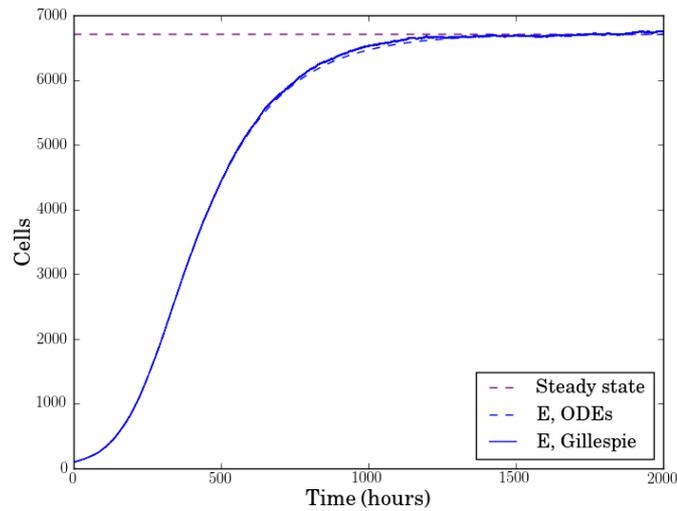


Figure 7.5: Figure showing the ODEs (7.4.17), compared to a Gillespie simulation, for a system with only  $E$  cells.  $\kappa_+ = 3$  and  $\kappa_- = 500/3$ . Gillespie results averaged from 10 realisations.

Setting the three equations equal to zero, and solving for  $x_R$ ,  $x_{01}$  and  $x_{02}$ , the only steady state found is the zero state, therefore  $R$  cells cannot exist on their own. Analysing the stability of this steady state, gives the eigenvalues:

$$\begin{aligned}
 & -2\kappa_-, \\
 & -\frac{1}{2} \left( \kappa_- + \kappa_+ + 1 + \sqrt{(\kappa_- + \kappa_+ + 1)^2 - 4\kappa_-} \right), \\
 & -\frac{1}{2} \left( \kappa_- + \kappa_+ + 1 - \sqrt{(\kappa_- + \kappa_+ + 1)^2 - 4\kappa_-} \right),
 \end{aligned}$$

which are all negative, and therefore the zero steady state is stable for all parameter values.

### 7.4.4 ODEs with effector and regulatory cells

Dimensionless differential equations can be written for the full system, with both E and R cells, for the mean number of cells in each populations, in which  $x$  represents the dimensionless variables:

$$\begin{aligned} \frac{dx_E}{d\tau} &= -x_E + \kappa_-(2x_{10} + x_{10}^* + 4x_{20} + 3x_{20}^* + x_{11}) \\ &\quad + \kappa_+x_E(x_{20} + x_{20}^* + x_{02} + x_{02}^\# + x_{11} - 1), \end{aligned} \quad (7.4.22a)$$

$$\begin{aligned} \frac{dx_R}{d\tau} &= -x_R + \kappa_-(x_{01} + 2x_{01}^\# + 2x_{02} + 3x_{02}^\# + 2x_{11}) \\ &\quad + \kappa_+x_R(x_{20} + x_{20}^* + x_{02} + x_{02}^\# + x_{11} - 1), \end{aligned} \quad (7.4.22b)$$

$$\begin{aligned} \frac{dx_{10}}{d\tau} &= \kappa_-(x_{10} + 2x_{20} + x_{20}^*) \\ &\quad - \kappa_+(x_{10}(2x_E + x_R) + x_E(x_{10}^* + x_{01} + x_{01}^\# + x_{20} + x_{20}^* + x_{02} + x_{02}^\# + x_{11} - 1)), \end{aligned} \quad (7.4.22c)$$

$$\begin{aligned} \frac{dx_{10}^*}{d\tau} &= \kappa_-(x_{10}^* + x_{20}^* + x_{11}) - \kappa_+x_{10}^*(x_E + x_R), \\ \frac{dx_{01}}{d\tau} &= \kappa_-(x_{01} + 2x_{02} + x_{02}^\#) \\ &\quad - \kappa_+(x_{01}(x_E + 2x_R) + x_R(x_{10} + x_{10}^* + x_{01}^\# + x_{20} + x_{20}^* + x_{02} + x_{02}^\# + x_{11} - 1)), \end{aligned} \quad (7.4.22d)$$

$$\frac{dx_{01}^\#}{d\tau} = \kappa_-(x_{01}^\# + x_{02}^\# + x_{11}) - \kappa_+x_{01}^\#(x_E + x_R), \quad (7.4.22f)$$

$$\frac{dx_{20}}{d\tau} = -2\kappa_-x_{20} + \kappa_+x_{10}x_E, \quad (7.4.22g)$$

$$\frac{dx_{20}^*}{d\tau} = -2\kappa_-x_{20}^* + \kappa_+x_{10}^*x_E, \quad (7.4.22h)$$

$$\frac{dx_{02}}{d\tau} = -2\kappa_-x_{02} + \kappa_+x_{01}x_R, \quad (7.4.22i)$$

$$\frac{dx_{02}^\#}{d\tau} = -2\kappa_-x_{02}^\# + \kappa_+x_{01}^\#x_R, \quad (7.4.22j)$$

$$\frac{dx_{11}}{d\tau} = -2\kappa_-x_{11} + \kappa_+(x_{10} + x_{10}^2)x_R + \kappa_+(x_{01}x_{01}^*)x_E. \quad (7.4.22k)$$

## 7.5 Analysis of Model 3

This system can be analysed using Mathematica, to find steady states. A solution cannot be found for the general case, but steady states can be calculated for specific sets of parameter values. To choose parameters to use in the model, the TGMFPT equation (3.5.37) was used, and a rate

obtained from the inverse of the TGMFPT. Using realistic values for  $R$ ,  $D$  and  $b$ , of  $500\mu m$ ,  $40\mu m^2 min^{-1}$  and  $11.5\mu m$ , gives a value for  $k_+$  of approximately  $1.2 \times 10^{-5} minutes^{-1}$ . A value for  $k_-$  of  $1/60 minutes^{-1}$  was chosen, corresponding to a mean binding time of 60 minutes. A death rate,  $\mu$ , of  $10^{-4} minutes^{-1}$  was chosen, corresponding to a mean life span of  $10^4$  minutes, which is approximately one week. These parameters give the dimensionless parameters  $\kappa_- = 500/3$ , and  $\kappa_+ = 0.12N_A$ . We will investigate two values for the number of APCs,  $N_A$ , of 25 and 100, corresponding to  $\kappa_+ = 3$  and  $\kappa_+ = 12$ . For  $\kappa_+ = 3$  the following steady states are found:

- Extinction:  $E = R = 0$ .
- Autoimmunity:  $E = 6706$ ,  $R = 0$ .
- Tolerant:  $E = 6060$ ,  $R = 202$  (unstable).
- Tolerant:  $E = 1064$ ,  $R = 1106$  (stable).

For  $\kappa_+ = 12$  the following steady states are found:

- Extinction:  $E = R = 0$ .
- Autoimmunity:  $E = 31887$ ,  $R = 0$ .
- Tolerant:  $E = 31766$ ,  $R = 40$  (unstable).
- Tolerant:  $E = 398$ ,  $R = 3008$  (stable).

The steady state numbers stated are actual cell numbers, rather than the non-dimensionalised values. Analysis of these steady states reveals for both values of  $\kappa_+$ , the extinction steady state is unstable, the autoimmune steady state is a stable node, and one of the tolerant steady states is a stable spiral and the other is unstable, indicated in parentheses.

### 7.5.1 Carrying Capacity

To determine the total number of E and R cells that can be supported by the system the carrying capacity can be calculated. The carrying capacity of the system can be calculated by setting  $\frac{dx_E}{dt} = \frac{dx_{10}}{dt} = \frac{dx_{20}}{dt} = 0$  and  $x_R = x_{01} = x_{02} = x_{11} = x_{10}^* = x_{01}^\# = x_{20}^* = x_{02}^\# = x_{11} = 0$ , as the carrying capacity of the system is reached when there are no R cells, and the E cell numbers are able to grow to the carrying capacity of the system. The only populations that will be non zero

are  $x_E$ ,  $x_{10}$  and  $x_{20}$ , and when the number of E cells reaches the carrying capacity, this will be a steady state, implying the rate of change of these populations will be zero. Using these conditions the carrying capacity of the system is found to be

$$x_E = \frac{\kappa_-}{\kappa_+} \left( \kappa_+ - 1 + \sqrt{\kappa_+^2 - 1} \right), \quad (7.5.23)$$

for  $\kappa_+ > 1$ .

### 7.5.2 Solution of ODEs and Gillespie algorithm

The ODEs can be solved numerically, using an Euler method. The system can also be solved using the Gillespie algorithm, allowing stochastic effects to be studied.

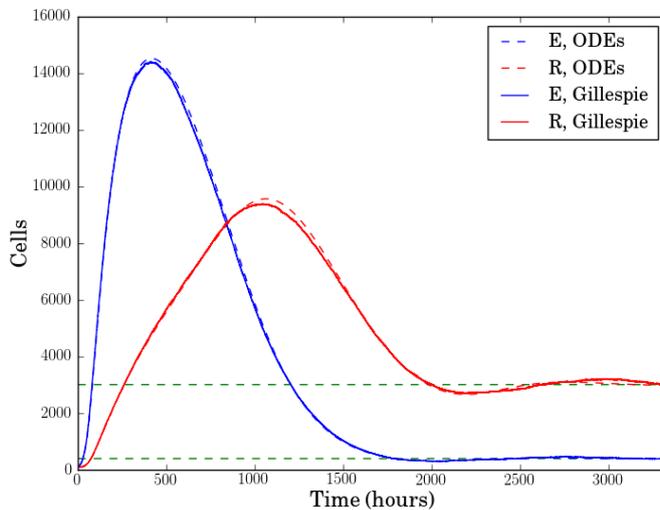
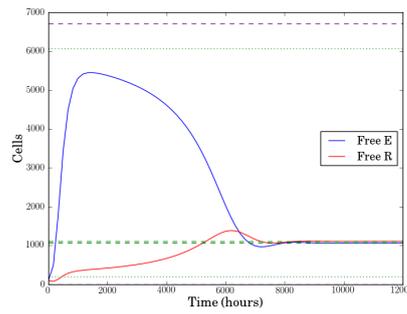


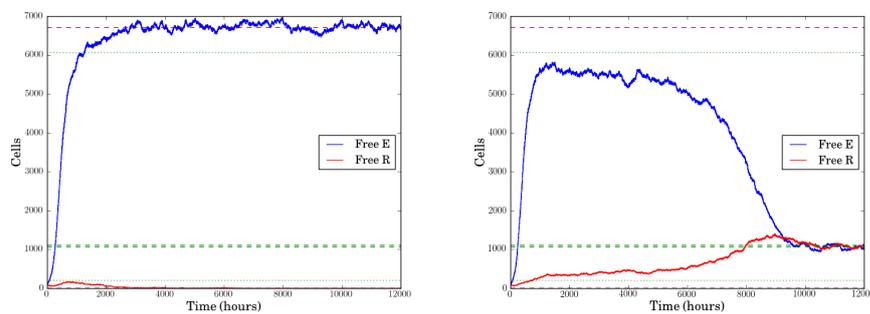
Figure 7.6: Comparison of ODE model to Gillespie algorithm. Horizontal green dashed line represents the stable tolerant state. Parameters used  $\kappa_+ = 12$  and  $\kappa_- = 500/3$ . Initial cells numbers,  $E = R = 100$ . Gillespie results averaged from 100 realisations.

Figure 7.6 shows for  $\kappa_+ = 12$  the ODEs, and the average of numerous Gillespie simulations give the same results, and settle in the stable tolerant steady state. The results for the Gillespie model are averaged over multiple realisations to form a smooth curve. In Figure 7.6, the Gillespie algorithm agrees exactly with the ODEs for this value of  $\kappa_+$ , indicating stochastic effects do not change the steady state achieved in this case, and the ODEs can be used as a mean field approximation to the stochastic system. In the case with  $\kappa_+ = 3$  stochastic effects have a greater contribution, resulting in the Gillespie algorithm settling in different steady states for different realisations, due to the

smaller number of APCs required to give  $\kappa_+ = 3$ . Therefore, an average of multiple realisations could not be compared with the ODEs. Figure 7.7(a) shows the ODEs tend to the unstable tolerant



(a) ODEs.



(b) Gillespie giving the autoimmune steady state (c) Gillespie giving the stable tolerant steady state

Figure 7.7: Figures showing results of (a) the ODE model and (b),(c) two different realisations of the Gillespie algorithm in which different steady states are reached. Horizontal lines represent the steady states. Stable tolerant state: green dashed, unstable tolerant state: green dotted, stable autoimmune state: purple dashed. Parameters used  $\kappa_+ = 3$  and  $\kappa_- = 500/3$ . Initial cells numbers,  $E = R = 100$ .

steady state, but then move away from this steady state and approach stable tolerant steady state. Figures 7.7 (a) and (b) show two different realisations of the Gillespie algorithm for  $\kappa_+ = 3$ , in which different steady states are reached.

### 7.5.3 Effector and regulatory cell dynamics

Plotting the number of E cells against the number of R cells allows the dynamics of the system to be explored. In the case with  $\kappa_+ = 12$ , the dynamics are similar for both the ODEs and one realisation of the Gillespie algorithm (Figure 7.8), showing stochastic effects are small for large

$\kappa_+$ . Figure 7.8 shows the system is attracted towards the stable tolerant steady state.

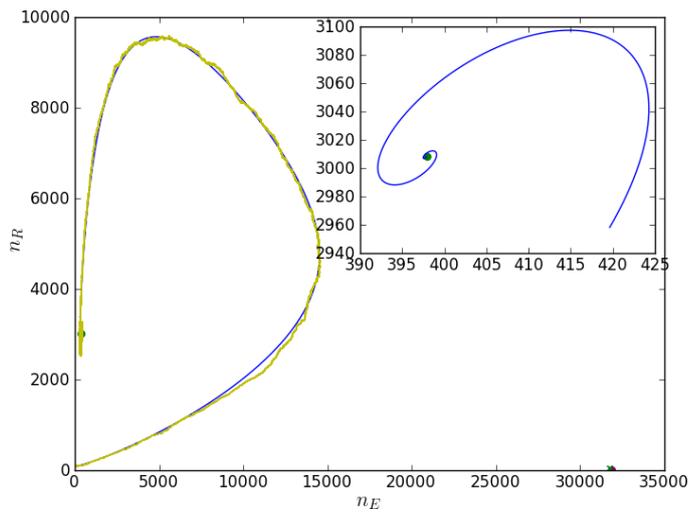


Figure 7.8: Comparison of number of E cells and R cells for ODEs and Gillespie algorithm. Blue line indicates the trajectory of the ODEs and the yellow line the trajectory of one realisation of the Gillespie algorithm. Stable steady states are shown by dots and the unstable steady state by a cross. Parameters used  $\kappa_+ = 3$  and  $\kappa_- = 500/3$ . Initial cells numbers,  $E = R = 100$ .

For  $\kappa_+ = 3$ , the steady state achieved depends upon stochastic fluctuations of the system, with two steady states being observed. Figure 7.9 shows in all cases the system approaches the unstable

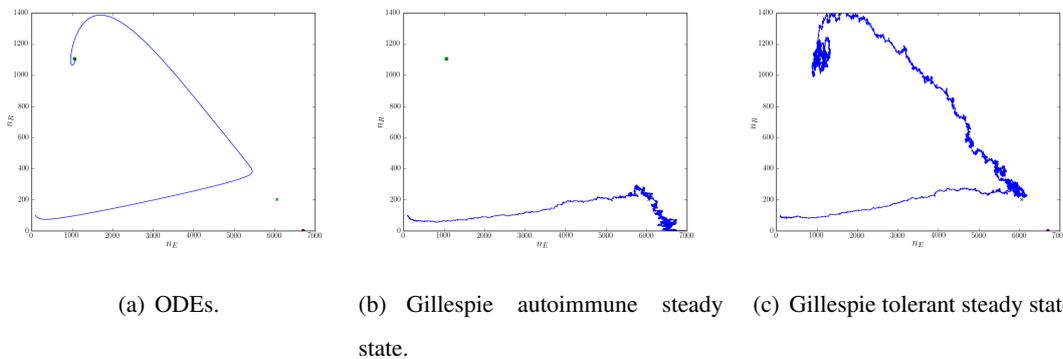


Figure 7.9: Comparison of number of E cells and R cells for (a) the ODE model and (b),(c) two different realisations of the Gillespie model in which different steady states are reached. Stable steady states are shown by dots, and the unstable steady state by a cross. Parameters used  $\kappa_+ = 3$  and  $\kappa_- = 500/3$ . Initial cells numbers,  $E = R = 100$ .

tolerant steady state, but then settles in either the stable tolerant or the autoimmune steady state.

Figures 7.7 and 7.9 illustrate stochastic fluctuations affect the steady state reached, for small values of  $\kappa_+$ . ODEs can be used as a mean field approximation, and will be used in this chapter. However, it must be remembered that stochastic fluctuations play a role, and the steady state achieved in the mean field approximation will only occur in a proportion of the realisations of the stochastic model.

### 7.5.4 Phase portrait

A phase portrait can be used to view the trajectories of the ODE solutions for various initial conditions. Phase portraits were generated using the same code used to solve the ODEs, as used to generate Figures 7.8 and 7.9(a), but for a range of sets of initial conditions. The ODEs were solved for a set number of time steps, and the number of E cells versus R cells for each set of initial conditions was plotted, allowing information about the speed at which the system approaches the steady states, dependent on where in the phase space you are, to be visualised.

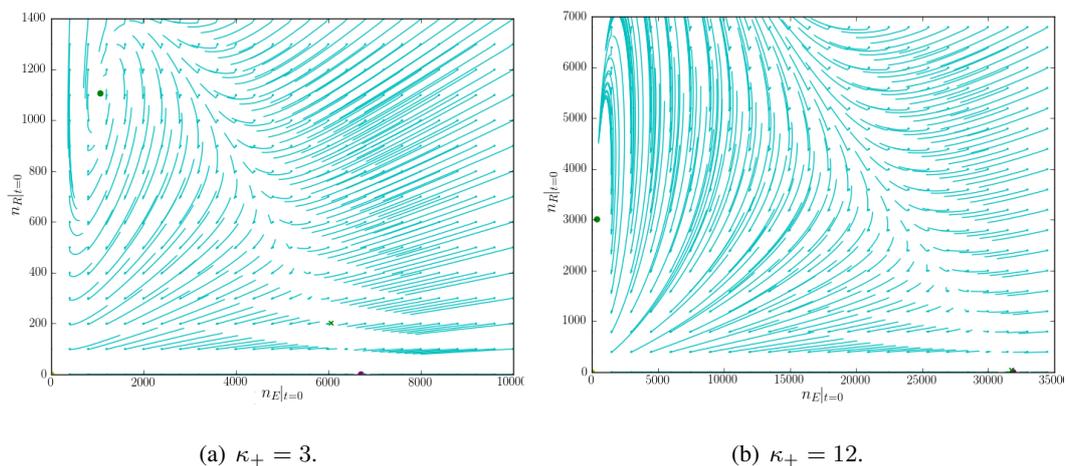


Figure 7.10: Phase portraits of the ODE system. Cyan dots indicate different initial conditions of the trajectories. Steady states are indicated by coloured dots/crosses.

Figure 7.10 shows the system is attracted towards the unstable tolerant steady state, in both cases, and then moves, more slowly, towards either the autoimmune steady state or the stable tolerant steady state.

The steady state various sets of initial conditions settle in can be examined in more detail. This is achieved by calculating the steady state a set of initial conditions settle in, using an Euler method, and plotting the initial condition as dot, with a colour to indicate the steady state it has settled in.

This is then done for multiple sets of initial conditions. Figure 7.11 shows this for (a)  $\kappa_+ = 3$  and (b)  $\kappa_+ = 12$ .

Figure 7.11(a) shows, no set of initial conditions examined settle in the unstable tolerant steady state, and there is a bifurcation point at which the system switches from the autoimmune steady state to the stable tolerant steady state.

Figure 7.11(b) shows, only sets of initial conditions with very low initial numbers of R cells settle in the autoimmune steady state, all other sets of initial conditions examined settle in the stable tolerant steady state, excluding those starting with zero E cells, which go extinct.

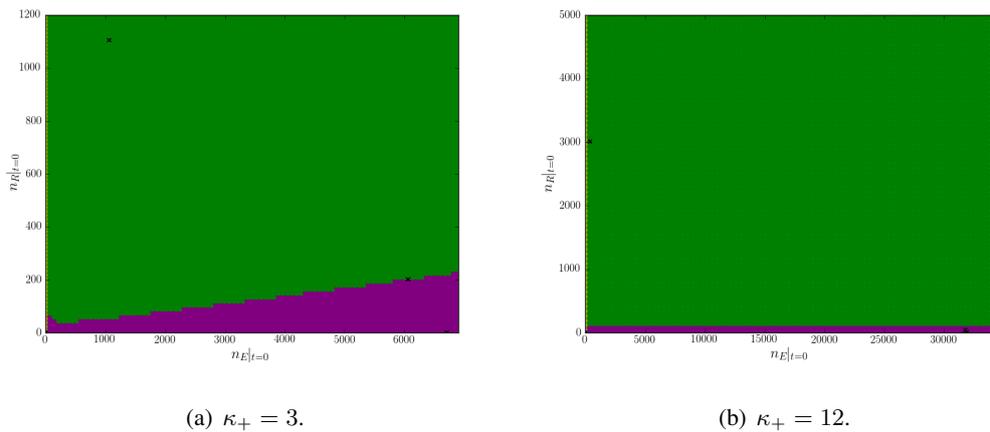


Figure 7.11: Plots showing the steady states different sets of initial conditions settle in (dots). Steady states are indicated by black crosses. Extinction: yellow, autoimmunity: purple, tolerance: green.

### 7.5.5 Effect of $\kappa_+$ and $\kappa_-$

The importance of the various parameters regimes can be viewed by examining the steady states obtained for the various regimes. By non-dimensionalising the system the number of parameters has been reduced to two,  $\kappa_+$  and  $\kappa_-$ . These parameters can be used to investigate the steady states obtained for various set of parameters. Figure 7.12 was generated by solving the ODEs, using Mathematica, for various pairs of parameters. The steady state achieved for a specific pair of parameters is represented as a coloured dot, where the colour indicates the type of steady state, being either extinct, autoimmune, or stable tolerant. In the regimes where the autoimmune steady state is observed, the extinction state is also observed, and likewise, in the regimes where the tolerant steady state occurs, the other two steady states also occur. Therefore, in Figure 7.12

a yellow dot indicates the extinction steady state only, a purple dot indicates the extinction and autoimmune steady states, and a green dot indicates all three steady states are observed.

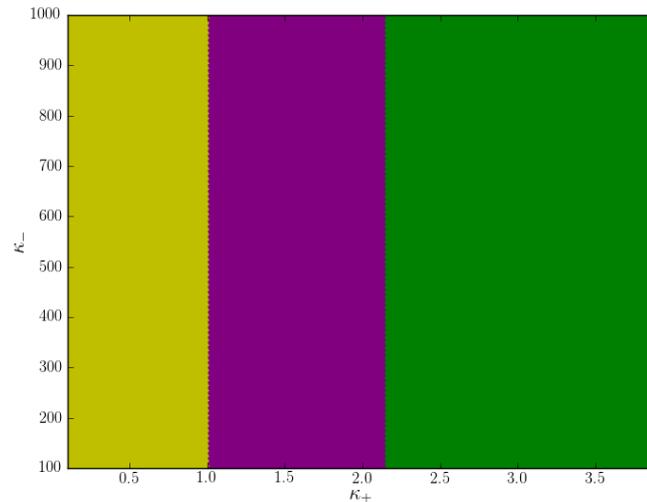


Figure 7.12: Plot showing the steady states obtained for varying values of  $\kappa_+$  and  $\kappa_-$ . Yellow dots indicate the extinction steady state, purple indicate the extinction and autoimmune steady state, and green indicates the extinction, autoimmune and stable tolerant steady states.

Figure 7.12 shows the unbinding rate,  $\kappa_-$ , has no effect on the steady state obtained, and the binding rate,  $\kappa_+$ , is the only parameter that effects the steady state. The value of  $\kappa_+$  at which the system switches from the extinction steady state, to the autoimmune steady state can be calculated by setting the carrying capacity given in (7.5.23) to zero, and is given by

$$\kappa_+ = 1.$$

The number of E cells in the autoimmune steady state is given by the carrying capacity of the system (7.5.23).

From Figure 7.12, the number of APCs required to achieve the autoimmune, and the tolerant steady state can be determined. Using the parameters defined in Section 7.5, we have  $\kappa_+ = 0.12N_A$ , where  $N_A$  is the number of APCs. The autoimmune steady state exists for  $\kappa_+ \geq 1$ . This corresponds to a minimum of 9 APCs being required to achieve the autoimmune steady state. The stable tolerant steady state occurs for  $\kappa_+ \gtrsim 2.15$ . This correspond to a minimum of 18 APCs being required to achieve the stable tolerant steady state.

### 7.5.6 The effect of $\kappa_+$ on cell numbers

The effect of  $\kappa_+$  on the number of E and R cells can be investigated by solving the ODEs, using the Euler method.

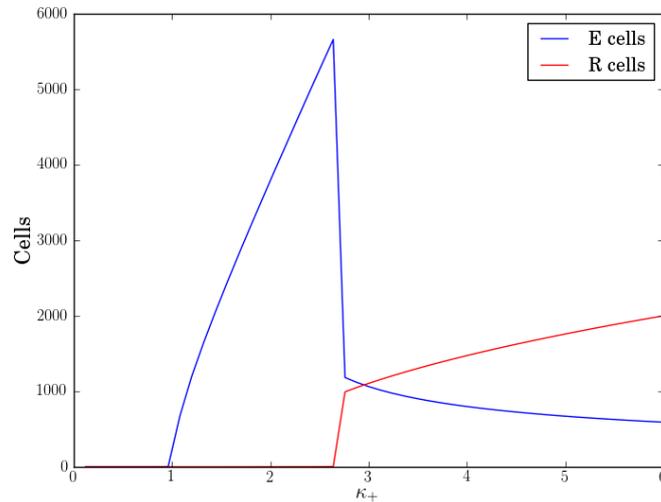


Figure 7.13: Plot showing how the steady state numbers of E and R cells change with varying the value of  $\kappa_+$ . A value of  $500/3$  was used for  $\kappa_-$ . Initial values of  $E = R = 100$  were used.

Figure 7.13 shows the number of E and R cells present in the system at the steady state, for various parameter regimes. In the region corresponding to the extinction steady state in Figure 7.12, there are no E or R cells. As the system moves into the autoimmune steady state observed in Figure 7.12, there are still no R cells, but now the number of E cells at steady state increases as  $\kappa_+$  increases. This increase continues until the region corresponding to the tolerant steady state in Figure 7.12 is reached, at which point a bifurcation occurs, resulting in a sudden drop in E cells and increase in R cells, and co-existence of both E and R cells. In this region, the autoimmune steady state still occurs, but is not the stable steady state for the set of initial conditions used.

Figure 7.13 shows the steady state achieved for initial numbers of E and R cells set to be 100. If different initial cell numbers are used, different steady states can be achieved. Using a Mathematica code, all the steady states that occur for a specific parameter regime, regardless of the initial conditions, can be generated, and their stability determined. Results are shown in Figure 7.14.

In Figure 7.14 the black crosses in both plots correspond to the same unstable steady state. The

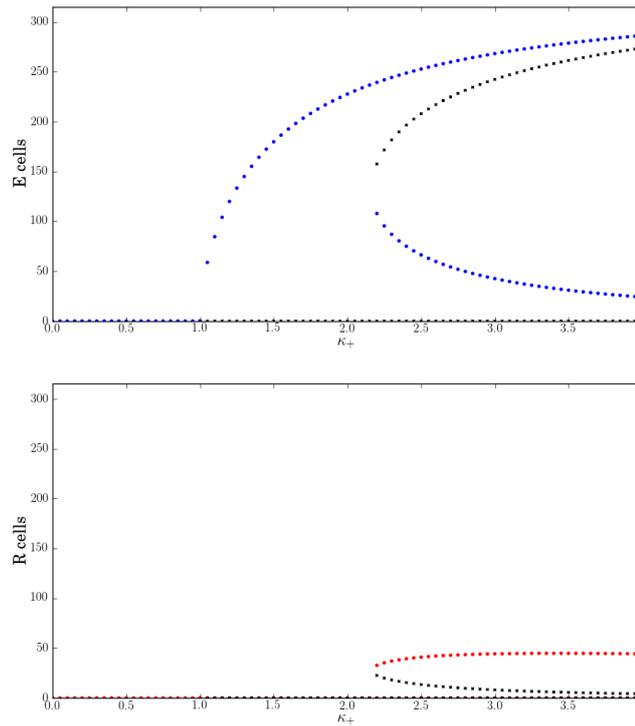


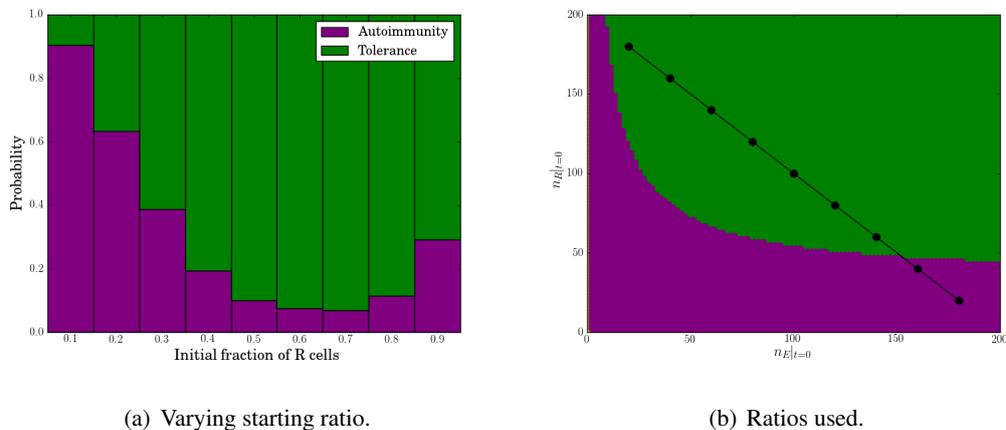
Figure 7.14: Plot showing the number of E and R cells in the various steady states as the value of  $\kappa_+$  changes. Blue/red dots indicate stable steady states, black crosses indicate unstable steady states. A value of  $500/3$  was used for  $\kappa_-$ .

red dots in the bottom plot, that are not on the  $x$ -axis, correspond to the lower blue dots in the top plot. The upper blue dots in the top plot, equate to the autoimmune steady state, and their corresponding points on the bottom plot lie on the  $x$ -axis.

Figure 7.14 shows the extinction steady state is stable only for  $\kappa_+ < 1$ , after which point it becomes unstable. The autoimmune steady state exists for all values of  $\kappa_+ \geq 1$ , and is always stable. The tolerant steady state is stable, and is the same tolerant steady state seen in Figure 7.13, and a second unstable tolerant steady state is observed, in which there is a smaller number of R cells and larger number of E cells. As  $\kappa_+$  increases, the unstable tolerant steady state approach the autoimmune steady state.

### 7.5.7 Stochastic effects of initial ratio of cells

To test the effect of stochastic fluctuations on the system, multiple realisations of the Gillespie algorithm can be computed, to determine the probability of being in a specific steady state, for varying initial ratios of E and R cells.



(a) Varying starting ratio.

(b) Ratios used.

Figure 7.15: Plot (a) shows the effect of varying the starting ratio of E and R cells for  $\kappa_+ = 3$ . Total number of cells (E+R) is 200. Plot (b) shows the steady state the ODEs settle in for various pairs of initial cell numbers, with black dots indicating the initial ratios used in (a). Extinction: yellow, autoimmunity: purple, tolerance: green.

Figure 7.15(a) shows stochastic effects are present in the system, which are not observed in the deterministic model. The initial ratio of E and R cells determine the most dominant steady state seen, but for all initial ratios both the autoimmune and tolerant steady states were observed. The extinction steady state was not observed in any of the simulations. Figure 7.15(a) was obtained using the Gillespie algorithm, running simulations until either the E or R cells went extinct, or the number of E and R cells were within 5% of the stable tolerant steady state. At this point it was assumed the system would settle in that state. Solving the system using the ODEs, for this parameter range and the same initial conditions, gives the autoimmune steady state when the fraction of R cells is less than or equal to 0.2, and the tolerant steady state for higher fractions of R cells, as shown in Figure 7.15(b).

Figure 7.15(a) shows increasing the fraction of R cell reduces the probability of autoimmunity occurring, up until a fraction of R cells of 0.8, at which point the probability of autoimmunity increases again. This can be investigated by examining individual realisations of the Gillespie algorithm, for specific R cell starting ratios, shown in Figure 7.16.

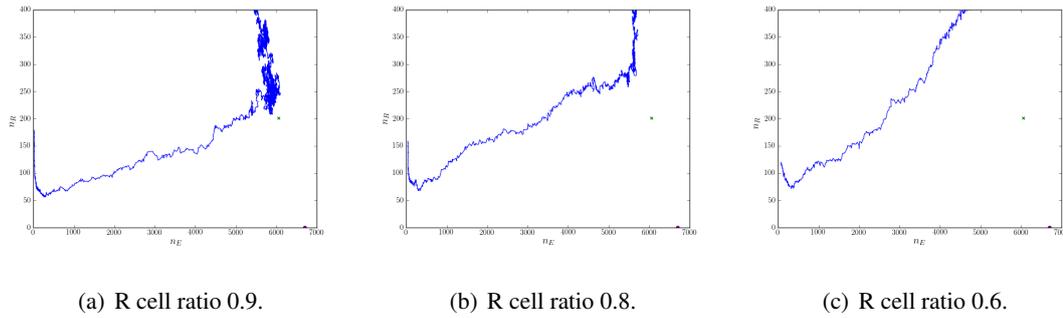


Figure 7.16: Realisations of the Gillespie algorithm for high initial ratios of R cells. The green cross indicates the unstable tolerant steady state.

The increase in the probability of autoimmunity appears to be due to the large number of R cells present, resulting in them occupying the majority of the APC binding sites. This results in the R cells not being able to proliferate, as they are not receiving signal from the E cells, and the E cells dying off quicker as only free cells can die. Thus, the R cells die, as there are fewer E cells to stimulate their proliferation. As the cell numbers drop, the E cells are able to recover quicker as they do not depend upon the R cells to proliferate. As the E cells recover, the R cells can then also recover, but the population levels have dropped to be near the bifurcation point of the system. This results in more realisations settling in the autoimmune steady state, as the system is closer to the bifurcation seen in Figure 7.11(a). Thus, only a small stochastic fluctuation is required to push the system towards the autoimmune steady state. Figures 7.16(a)-(c) show that for larger ratios of R cells, the cell numbers drop to lower levels, and spend longer near the bifurcation point, and for medium ratios of R cells, the system quickly rises above the bifurcation point. The anomalous points in 7.11(a) are caused by this effect.

In the simulations, with lower fractions of R cells (0.1 and 0.2), the system starts below the bifurcation point, hence the ODEs settle in the autoimmune steady state. This explains the higher probability of autoimmunity in Figure 7.15. This is illustrated in Figure 7.17

Performing a similar analysis of the steady states in the case with  $\kappa_+ = 12$ , reveals that for all initial R cell ratios, only the tolerant steady state is observed. This occurs because the bifurcation of the system is at a very low level, as shown in Figure 7.11(b). Therefore, very few initial R cell ratios are near the bifurcation point.

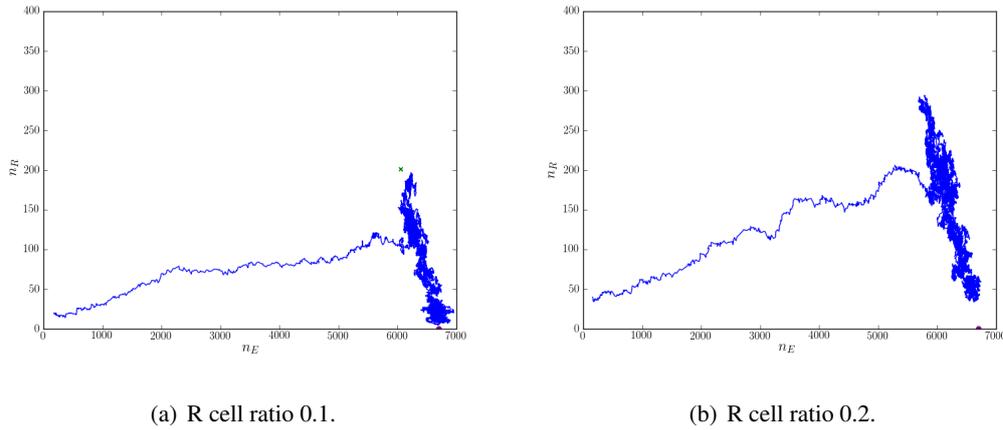


Figure 7.17: Realisations of the Gillespie algorithm for low initial ratios of R cells. The green cross indicates the unstable tolerant steady state.

## 7.6 Summary

In this chapter the cross-regulation model proposed by Léon *et al.* [49, 133] has been investigated, and a deterministic and stochastic version of the model formulated and studied. The chapter investigates three model variants, which were also investigated by Léon *et al.*: competition only, R cell suppression only, and R cell suppression and E cell promotion. The first two models do not exhibit any interesting dynamics, with the competition model settling in a steady state dependent on the initial ratio of cells. In the R suppression model, E cells are driven to extinction in all cases, except when the initial number of E cells is much greater than that of R cells, in which case the R cells go extinct. The third model, in which R cells regulate the number of E cells, and E cells provided a division stimulus to the R cells, proves to be the most interesting from an immunological and mathematical perspective. The model developed by Léon *et al.* is a purely deterministic model, and does not consider stochastic effects. The deterministic models developed in this chapter were formulated in a different way to Léon *et al.*'s models. They were formulated by using the reactions of the system used for the Gillespie model. This allowed direct comparison of the deterministic and stochastic models in this chapter.

In model three, the number of steady states obtained, that are real and positive, is dependent on the value of  $\kappa_+$  only, and has no dependence on  $\kappa_-$ . For  $\kappa_+ \leq 1$ , only the extinction steady state exists. For  $1 < \kappa_+ \lesssim 2.15$ , where the upper boundary is approximate, as it can only be determined numerically, the extinction steady state and one autoimmune steady state exist. For  $2.15 \lesssim \kappa_+$  four steady states exist: the extinction state, the autoimmune state, and two tolerant

steady states, one of which is stable, and the other unstable. These states agree with the findings of Léon *et al.* [49, 133], in which they also find the same four types of steady states. They find the unstable tolerant state to have a large number of E cells, and a small number of R cells, agreeing with the findings of this chapter. They find the stable tolerant steady state to have a small number of E cells and a large number of R cells, which is what is found in this chapter, for larger values of  $\kappa_+$ . For the parameters chosen in this chapter, to achieve the tolerant steady state eighteen APCs are required, and to achieve the autoimmune state nine APCs are required. Below nine APCs, extinction occurs. It must be noted that the parameters used were chosen to be biologically significant, but there are other parameter combinations that could be chosen, which will also be biologically significant, giving different numbers of APCs required to achieve tolerance.

The parameter  $\kappa_+$  equals  $k_+ N_A / \mu$ , therefore if the interaction rate,  $k_+$ , of the cells is too low, autoimmunity or extinction will occur. This makes sense biologically, as if cell interactions occur very infrequently, they will not receive a signal to proliferate, and will die out. If the number of APCs,  $N_A$ , is too low, autoimmunity or extinction will occur. The number of APCs being low reduces the global encounter rate between T cells and APCs, resulting in less T cells proliferating, and T cells being more likely to die out. Finally, if the death rate,  $\mu$ , is too high, this will also result in autoimmunity or extinction, as the T cells will die before they can encounter an APC to receive a signal to proliferate. A combination of these three parameters, that gives a large enough  $\kappa_+$ , are required to prevent autoimmunity or extinction.

Our investigation of stochastic effects in the cross-regulation model has revealed that for large values of  $\kappa_+$ , stochastic effects are not significant. However, for smaller values of  $\kappa_+$ , that are still within the tolerant regime, stochastic effects play a significant role, and can cause the steady state achieved to differ for different realisations. For  $\kappa_+ = 3$ , the system was found to settle in either the autoimmune or tolerant steady state. This occurs because there is a bifurcation at which the system switches between the states. For large  $\kappa_+$  the bifurcation occurs at very low cell numbers, as observed in Figure 7.11(b). But, for smaller values of  $\kappa_+$ , the bifurcation occurs at larger cell numbers, as observed in Figure 7.11(a). Stochastic fluctuations can cause the system to move across the bifurcation point, which would not occur in a deterministic model. The probability of a specific steady state occurring in the stochastic system, was found to depend on the initial ratio of E and R cells. For roughly equal initial proportions of E and R cells, the probability of autoimmunity is low. Decreasing the proportion of R cells, increases the probability of autoimmunity, to the point where the tolerant steady state is rarely seen. For low initial R cell ratios, the ODEs also predict

autoimmunity. For high initial R cell ratios, the ODEs predict tolerance. Yet the stochastic model reveals there is still a significant probability of autoimmunity. Léon *et al.* [49, 133] also found for low R cell ratios, autoimmunity occurs. For large R cell ratios, the ODEs give the tolerant steady state, however, the stochastic model reveals a slight increase in the probability of autoimmunity, occurring due to stochastic fluctuations. This is not observed by Léon *et al.*, due to the purely deterministic nature of their model.

In patients with conditions which reduce their T cell numbers, such as HIV, autoimmunity is unlikely to occur if the ratio of E and R cells remains constant. However, if the number of APCs is depleted, or the death rate of T cells becomes too large, autoimmunity or even extinction could occur. If only a patient's R cells are depleted, this is likely to result in autoimmunity. In IPEX syndrome, which causes a mutation in the FOXP3 gene [237], resulting in R cells not functioning, autoimmunity is observed, which is also found in the model in this chapter for the same situation, with no, or few, R cells.

Upon first devising the models in this chapter, a probability of productive binding for E cells, and R cell was used. If a cell has a weak/unproductive binding, it cannot suppress or promote other cells, and if a cell has a strong/productive binding, it can. Having these parameters allowed an advantage to be given to one of the cell types if desired. This resulted in two extra populations of cells, as in the  $A_{11}$  population it must be remembered if a cell has been promoted or suppressed, depending on the strength of its binding. After initial analysis of the system, it was noticed that there was a way for cells to have multiple chances of having a strong/weak binding. Allowing a cell to change its binding strength during a binding is undesirable. Comparing the system to an agent based model revealed this problem was significant. To overcome the problem the binding state and suppression/promotion state of each cell in a complex must be recorded, resulting in over twenty populations. This added complexity did not result in a better model, and it was decided to keep a simpler model, in which all bindings are productive. Alternative models in which the binding rate of one cell type was multiplied by a factor were investigated. Doing this allowed an advantage to be given to one of the cell types, but again this did not prove to add anything extra to the model.

Model 3, devised in this chapter, has shown interesting results, with only two parameters. Effects not observed by Léon *et al.* [49, 133] are also found, by investigating the effects of stochastic fluctuations on the system. The increased probability of autoimmunity for small numbers of APCs, observed in Figure 7.15(a), due to stochastic effects, was observed in Model 3 in this chapter and

was not found by Léon *et al.*'s deterministic approach.



## Chapter 8

# Mechanisms of T cell activation

### 8.1 Introduction

It is believed that T cells need to undergo sustained signalling to become activated [88]. However, it is still unclear whether sustained signalling requires a prolonged interaction, or can be generated by multiple shorter encounters with different antigen-presenting-cells (APCs) [46]. New insights into cell-cell interactions have been gained using two-photon laser scanning microscopy, to view these interactions *in vivo* [41, 43, 156]. Several studies have tracked the dynamics of T cell-APC contacts upon antigen recognition [42, 110, 152, 154, 212]. It has been proposed that the interactions of T cells and APCs can be split into three phases: transient interactions, long-lived interactions, and swarming [42, 152, 154, 212]. A study has found that the minimum contact time between T cell receptors and APCs for activation of naive T cells is in the range of 20 hours [112]. In another study, the duration of TCR signalling required by naive cells was found to be lower in the presence of mature APCs, being in the range of 6 hours [131].

During phase 1: transient interactions, T cells have productive, but brief (less than ten minutes in duration) serial encounters with APCs displaying foreign peptide [42, 110, 152, 156, 201, 212]. T cells do not stop completely, but instead crawl on the surface of the APC. These contacts are almost indistinguishable from those made with non-cognate APCs, but evidence suggests they are not 'null' and result in some level of activation, with up-regulation of CD69 activation marker being detected [152]. Transient contacts dominated early stages of T cell-APC interactions, occurring in the first 8 hours after antigen presentation [110, 152, 154].

In phase 2: long lived interactions, T cells arrest on APCs, forming contacts lasting several hours.

During these contacts, conjugates appear either completely static or more dynamic, presumably owing to the rapid motion of APCs and/or low level of T cell crawling. This stage is thought to be the key event in T cell activation, and has been observed in all experimental conditions associated with priming. Long lived contacts are mostly observed between 6 and 24 hours after initial antigen presentation. The signalling occurring during sustained interactions is thought to be of functional importance [42].

The final stage: swarming, occurs approximately 25 to 30 hours after initial antigen presentation. T cells undergo rapid motion and brief interactions [152]. In this stage, T cells display looping patterns around APCs [154,156,219], and remain in a confined area of the lymph node. Swarming is thought to occur as a result of APCs releasing the antigen they have been presenting to T cells.

It must be noted that the immunological evidence presented in this section has been gathered from experiments carried out by different laboratories, using different experimental techniques. Hence, various and sometimes contradicting results are observed, with switches from short to long bindings at differing times and differing durations of binding times are reported.

Three types of mechanism have been proposed to explain the three phase interaction process: APC changes, signal integration, and a probabilistic approach [40,42]. The next three subsections review these mechanisms. This chapter investigates the three mechanisms proposed to explain the three stage activation process. Various mathematical models are proposed and investigated for each of the mechanisms.

The aim of this chapter is to develop a model that can adequately describe the experimental observations outlined above. The hope is that one of the mechanisms investigated will be able concisely explain the experimental observations significantly better than the other mechanisms. This would allow a mechanism for T cell activation to be proposed. Knowing the mode of T cell activation will allow immunological models involving activation of T cells to accurately mimic reality. It would also allow more specific targeting of therapeutic treatments.

### **8.1.1 APC changes**

The APC changes model assumes different phases of interactions are due to the state of maturation of the APCs. These changes could be due to the amount of antigen presentation, or co-stimulatory molecule presentation, on the surface of the APCs. It has been suggested that APCs need to mature before they can form long lived interactions. Experiments by Hughes and colleagues [110] found T

cells showed a phase of transient interactions with APCs for the first 15 hours after immunisation with antibodies, followed by a phase of stable interactions. However, when APCs were given extra time to mature *in vivo* T cells were able to rapidly establish long lived contacts with APCs, suggesting APCs and not T cells control the timing of arrest [110]. The increased expression of adhesion molecules and co-stimulatory molecules accompanying APC maturation is thought to contribute to the gradual ability of APCs to induce T cell arrest.

### 8.1.2 Signal integration

The signal integration approach proposes that T cells integrate signals they receive during the initial phase of brief interactions with APCs, and a certain cumulative threshold of signals must be achieved before T cells can gain the ability to arrest on APCs [54, 74, 98, 147, 152, 210]. Studies have found T cells re-encountering APCs in the lymph node were more likely to express the high affinity IL-2 receptor and produce IFN- $\gamma$ , which is a sign of T cell activation.

### 8.1.3 Probabilistic approach

The probabilistic approach suggests T cells have a defined probability of arresting on an APC, following an encounter with an antigen, determined by the strength of stimulation and the affinity of the TCR for the antigen. If the probability is low, T cells will engage in numerous unsuccessful binding attempts before arresting on an APC. T cells that establish long lived contacts will reside on the APC for hours. Therefore, long lived contacts are expected to progressively accumulate until they become the dominant type of interaction.

### 8.1.4 T cell swarming

T cell swarming behaviour after a long binding event is not easily modelled. A potential method of achieving swarming is to lower the diffusion coefficient of a T cell upon unbinding, causing it to remain near an APC for longer, but this would not necessarily result in T cells swarming around an APC, as a T cell may move in a straight line away from the APC. T cells could be confined to an area around an APC, but this would be forcing the T cells to stay near the APC and would not be very natural. Attraction towards an APC could be implemented by adding a chemokine gradient, which T cells respond to after undergoing a long binding, causing them to continually move back towards the APC. However, this would prevent them exiting the lymph

node (LN). A better approach is to give a T cell a probability of moving towards an APC rather than making a random movement, determined by the distance between the T cell and APC. In the models proposed in this chapter swarming will not be considered and only the first two stages of activation will be investigated.

## 8.2 Modelling approaches

In this section, methods for mathematically modelling the three activation mechanisms outlined in the introduction, are proposed.

### 8.2.1 APC changes

The binding time of a T cell on an APC can be modelled to depend on APC changes in various ways. The maturation of APCs could depend on the time an APC resides in a LN. Either the APC must be residing in the LN for a specific amount of time before long lived interactions can occur, or the length of the binding time is determined by the time an APC has been residing in the LN. A model can be implemented in which APCs must achieve a certain number of short interactions before they are able to engage in long lived interactions. Similarly, an APC may need to achieve a certain cumulative binding time before a long lived interaction can occur, or a combination of cumulative interaction time and number of bindings may be required. Below six models are proposed using the ideas mentioned.

**Model IA:** APCs are required to reside for a specific amount of time in the LN before they mature and long bindings occur. Before the threshold time T cells interacting with the specific APC will have short bindings, lasting  $\tau_s$ , and after the threshold has been reached a T cell will have long bindings, lasting  $\tau_l$ . Each APC will act independently, becoming mature after a set amount of time residing in the LN regardless of the state of the other APCs.

**Model IB:** The time a T cell spends bound to an APC is set to be a function of the time the APC has resided in the LN. The binding time will be set to follow a specific distribution, such as an exponential distribution. A maximum binding time can be specified and all binding times given as a fraction of this.

**Model IC:** The cumulative time an APC has spent bound to a T cell is recorded. When this time is greater than a given threshold value, the binding switches from short lived interactions to long

lived interactions.

**Model ID:** This model operates in the same way as Model IC, but instead of a threshold of cumulative binding time, there is a threshold that must be reached on the number of bindings an APC has undergone.

**Model IE:** This model is a combination of Models IC and ID. In this model both a threshold value on the cumulative binding time and the number of bindings must be reached before a long binding can occur.

**Model IF:** The time a T cell spends bound to an APC is a function of either the number of bindings an APC has undergone, or the cumulative time an APC has spent bound to a T cell. An upper limit on the binding time can be set, and all binding times given as a fraction of the upper limit.

In Models IA, IC, ID and IE, the dwell time can be either a fixed value for the short bindings, and another value for the long bindings, or a value drawn from a distribution, such as an exponential or log-normal distribution, with a given mean for each type of binding.

After consultation with experimentalists (Dr. P. Bousso, Paris), it has been advised that APC maturation occurs due to time events rather than binding events. It has been suggested that Model IA, followed by Model IC, are the most realistic and will therefore be investigated.

### 8.2.2 Signal integration

Signal integration assumes T cells sum the signals they receive from interactions with APCs. Signal integration by T cells from encounters with APCs can be modelled in various ways. The amount of time a T cell has been bound for can be summed, and a long binding occurs if the summed time is greater than a given threshold. The number of contacts a T cell has undergone can be summed, and again if this sum is greater than a set threshold a long binding will occur. A combination of these methods can be implemented, or a binding time that depends on the time a T cell has already spent bound, or the number of interactions that have already occurred, could be used. Below four potential models are proposed.

**Model IIA:** T cells start by undergoing short bindings, with a given binding time,  $\tau_s$ . The binding time of a T cell is added to a cumulative total for that specific T cell. Once the cumulative binding time exceeds a given threshold, the specific T cell will undergo long bindings on subsequent encounters, with binding time  $\tau_l$ .

**Model IIB:** This model operates the same as Model IIA, but instead of summing the binding times

the number of bindings are summed. It may also be desirable to add the requirement of a binding being greater than a set duration for it to be classified as productive and contribute to the number of bindings undergone.

**Model IIC:** This model is a combination of Models IIA and IIB. In this model a T cell must have a certain number of bindings and a certain cumulative binding time before a long binding can occur.

**Model IID:** In this model, instead of having a long and short binding time, the binding time of a T cell is continuous, determined by either: the number of bindings it has already undergone, or the cumulative binding time of the T cell. The binding time is given by a function of the chosen parameter, with an upper limit on the length of a binding.

In Models IIA, IIB and IIC the length of the short and long bindings can either be a fixed value, or more realistically, a value chosen from a distribution, such as an exponential or log-normal, with a given mean for short and long bindings.

After consultation with experimentalists (Dr. P. Bousso, Paris), it has been agreed that this approach should be based on a cumulative binding time rather than the number of bindings that have occurred, with Models IIA and IID being chosen.

### 8.2.3 Probabilistic approach

The simplest approach of creating a probabilistic model is to give each cell a fixed probability of forming a long binding upon encountering an APC, regardless of the properties of the APC. By setting long bindings to last a few hours an accumulation of long bindings will occur as time progresses. Further complexity can be added by giving each T cell and/or APC a 'strength', which determines the probability of a long binding. The strength could be made to depend on the number of short bindings a T cell has undergone. T cells could be given specificities and their binding probability made dependent on the similarity of the T cell and APCs specificities. Adding extra complexity to the model is only beneficial if a simple model cannot reproduce the desired outcome. We will therefore begin with a simple model and add complexity if required.

## 8.3 APC changes approach

In this section, APC changes Models IA and IC are examined. These models were chosen as they are thought to be the most realistic. The number of T cells undergoing short bindings with APCs

will be referred to as ‘short bound cells’, and the number of T cells undergoing long bindings with APCs will be referred to as ‘long bound cells’.

### 8.3.1 Model IA

In Model IA, APCs must spend a specific amount of time in the LN before they mature and can participate in a long binding. Before this threshold time, short bindings will occur. This time represents the time it takes for an APC to up-regulate co-stimulatory, and adhesion molecules on its surface. Upon encountering an APC, binding times are chosen from an exponential distribution with mean  $\tau_s$  for short bindings and  $\tau_l$  for long bindings. Upon unbinding from an APC a T cell is replaced uniformly randomly within the valid domain.

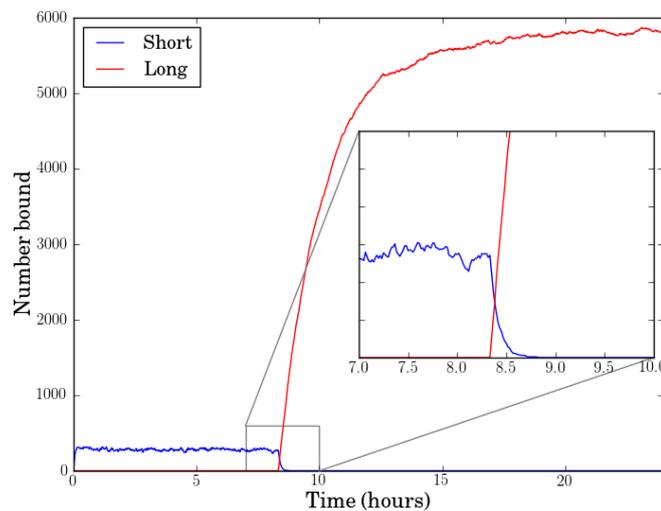


Figure 8.1: Spatial simulation of Model IA with an APC maturation time of 500 minutes. Parameters used:  $R = 400\mu m$ ,  $b = 11.5\mu m$ ,  $D = 50\mu m^2 min^{-1}$ ,  $\tau_s = 5$  minutes,  $\tau_l = 240$  minutes, APCs = 200, and T cells = 9857 after equilibration.

Figure 8.1 shows results of a spatial simulation of Model IA, in which the number of cells undergoing a short binding, at a given time, is shown in blue and those undergoing a long binding in red. In this simulation the time for an APC to mature is set to be 500 minutes, chosen to give a switch from short to long bindings at a point after 8 hours, as found in the literature [110, 152, 154]. The simulation shows a switch, at 500 minutes, in the dominant type of bindings from short to long. Before the switch point there are only short bindings. After the switch point the number of short bindings decays exponentially to zero, and the number of long bindings increases until it

reaches a steady state. In this model the point at which the system switches from the dominant type of binding being short to long is only dependent on the maturation time of APCs.

### 8.3.2 Model IC

In Model IC, APCs must achieve a specific cumulative binding time with T cells before they mature and can undergo long bindings. Before this threshold of cumulative binding times, an APC will undergo short bindings. Each APC acts independently of the other APCs in the LN, therefore APCs will mature at different times rather than simultaneously, as in Model IA. Upon encountering an APC, binding times are chosen from an exponential distribution with mean  $\tau_s$  for short bindings and  $\tau_l$  for long bindings.

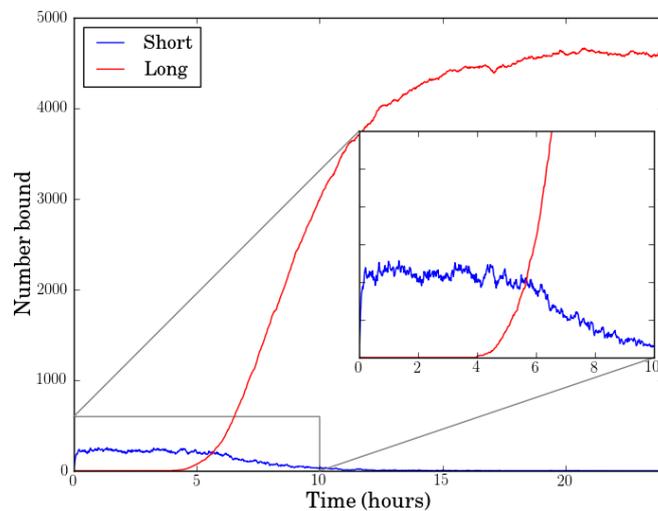


Figure 8.2: Spatial simulation of Model IC with a cumulative binding time needed to undergo long bindings of 500 minutes. Parameters used:  $R = 400\mu m$ ,  $b = 11.5\mu m$ ,  $D = 50\mu m^2 min^{-1}$ ,  $\tau_s = 5$  minutes,  $\tau_l = 240$  minutes, APCs = 200, and T cells = 9857 after equilibration.

Figure 8.2 shows results of a spatial simulation of Model IC. In the simulation APCs must have a cumulative binding time, with T cells, of 500 minutes before they can undergo a long binding. Short binding times are exponentially distributed with a mean of 5 minutes, therefore a mean of 100 bindings will be required to achieve a cumulative binding time of 500 minutes. However, as multiple T cells are able to bind to an APC simultaneously, a cumulative binding time of 500 minutes can be achieved in less than 500 minutes. The figure shows an initial increase in the number of short bound cells with no long bound cells initially. The number of short bound cells

then begins to decay, until they reach zero, and the number of long bound cells increases to a steady state. The switch in the dominant type of binding is seen around 6 hours. In this model APC mature by being in contact with T cells. Therefore, if more T cells are bound to an APC, it will mature faster. This makes the model highly dependent on the number of T cells available to provide maturation signal and the number of APCs competing for the signal from the T cells, as well as the threshold of binding times required. An additive effect of multiple T cells binding to an APC may not be the best approach to take. It may be better to have an APC receive maturation signals at a set rate, regardless of the number of T cells bound to it.

## 8.4 Signal integration approach

In this section, signal integration Models IIA and IID are examined. These models were chosen as they are thought to be the most realistic from an immunological perspective.

### 8.4.1 Model IIA

In Model IIA, each T cell acts independently, moving by Brownian motion, in an attempt to locate an APC. Each T cell keeps a record of the cumulative time it has been bound for. Upon encountering an APC, a T cell will undergo a short binding if the cumulative time is below a given threshold, and a long binding if it is above the threshold. Binding times are chosen from an exponential distribution with mean  $\tau_s$  for short bindings and  $\tau_l$  for long bindings. Once a T cell's cumulative binding time is above the switching threshold, all its bindings will be long, and there is no way in this model for the bindings to revert to being short again.

Figure 8.3 shows the results of a spatial simulation of Model IIA. The figure shows for a cumulative T cell binding time needed for long bindings to occur of 35 minutes, there is a switch in the dominant type of binding from short to long at around 8 hours. The switch point in this model depends on the interaction rate between T cells and APCs, the number of APCs, the length of short bindings, and the cumulative binding time required for a long binding to occur. The cumulative binding time is chosen to be 35 minutes to achieve a switch, from short to long bindings, at around 8 hours, as observed experimentally [110, 152, 154].

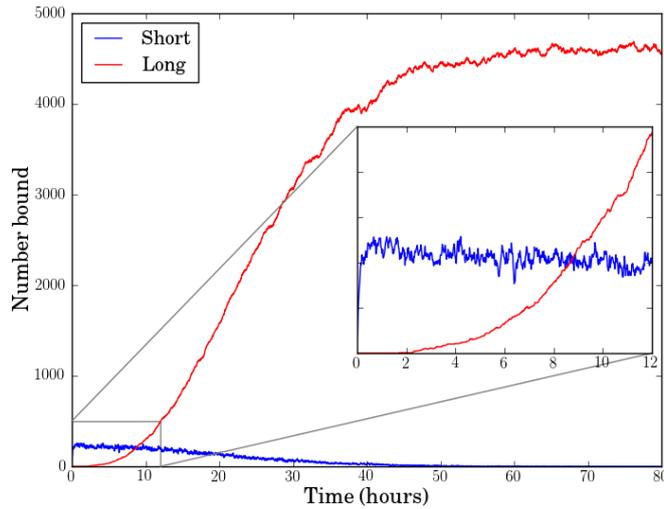


Figure 8.3: Spatial simulation of Model IIA with a cumulative binding time required for long bindings to occur of 35 minutes. Parameters used:  $R = 400\mu m$ ,  $b = 11.5\mu m$ ,  $D = 50\mu m^2 min^{-1}$ ,  $\tau_s = 5$  minutes,  $\tau_l = 240$  minutes, APCs = 200, and T cells = 8025 after equilibration.

#### 8.4.2 Model IID

In Model IID, cells do not have long and short binding as in Model IIA, instead they have a binding time that is a function of the cumulative time a T cell has spent bound to an APC. Upon encountering an APC, a T cell binds for a time given by the function

$$t_b = t_{max} \exp\left(-\frac{1}{\xi(t_c + 1)}\right), \quad (8.4.1)$$

where  $t_{max}$  is the maximum binding time that can be assigned,  $t_c$  is the cumulative time a T cell has spent bound, and  $\xi$  is a scaling parameter. All binding times are given as a fraction of the maximum binding time to prevent binding times becoming too large.  $\xi$  is used to adjust the binding times, with a smaller value giving a slower growth rate. The plus one in the denominator is required for an initial binding to occur.

Figure 8.4 shows an initial increase in the number of short bound cells, followed by a decrease in short bound cells and an increase in the number of long bound cells. In this model a binding time of less than ten minutes is defined as a short binding, and a binding time of greater than ten minutes as a long binding. Ten minutes was chosen for the threshold, as literature states short bindings last less than ten minutes [42, 110, 152, 156, 201, 212]. However, the literature also states

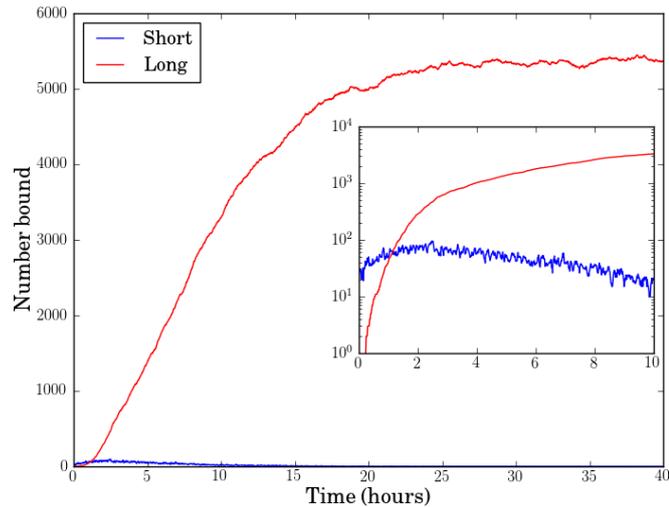


Figure 8.4: Spatial simulation of Model IID with a maximum binding time of 360 minutes, and  $\xi = 0.15 \text{ min}^{-1}$ . Parameters used:  $R = 400 \mu\text{m}$ ,  $b = 11.5 \mu\text{m}$ ,  $D = 50 \mu\text{m}^2 \text{ min}^{-1}$ , APCs = 200, and T cells = 8062 after equilibration.

long bindings last hours [42]. It is therefore not clear how to classify bindings in the intermediate range, and it may be that bindings of intermediate length should not be occurring. The value of  $\xi$  used in this simulation was chosen to give a reasonable plot, but it is not clear how to choose  $\xi$ , or the maximum binding time, or even the function itself. This model has lots of uncertainties and it seems unlikely the binding time will be determined by a continuous process. It is more likely a discrete process caused by up-regulation of receptors on the surface of the T cells.

## 8.5 Probabilistic approach

In this section, two probabilistic models are considered. In the first model (IIIA), upon encounter with an APC, a T cell will undergo a long or short binding with set probabilities. In the second model (IIIB), T cells have the same binding dynamics. However, upon unbinding they enter an inactive state in which they are able to move freely but not bind with APCs until they leave the inactive state and return to being free cells. This inactive state is introduced to make the model more biologically realistic.

### 8.5.1 Model IIIA

In this model, T cells encounter APCs and have a long binding with probability  $p$ , or a short binding with probability  $1 - p$ . The rate of T cells encountering APCs,  $k$ , is given by the inverse of the first passage time for a T cell to encounter an APC in a spherical LN,  $\alpha$  (3.5.37), multiplied by the number of APCs in the LN.  $\alpha$  is approximately equal to  $R^3/3Db$ , in which  $R$  is the radius of the LN,  $D$  the diffusion coefficient of the T cells, and  $b$  the effective radius of the APCs. T cells are not able to die in this model as only short timescales are considered. T cells have binding times of  $\tau_s$  for short bindings and  $\tau_l$  for long bindings. The unbinding rate is given by the inverse of the binding times.

This model has three populations:  $n_f$ , mean number of free cells,  $n_s$ , mean number of short bound cells, and  $n_l$ , mean number of long bound cells. There is no death in this model, so the total number of T cells is fixed at  $M$ , allowing one of the populations to be constrained by setting  $n_f = M - n_s - n_l$ .

We have the ODEs:

$$\frac{dn_f}{dt} = -kn_f + \frac{1}{\tau_s}n_s + \frac{1}{\tau_l}n_l, \quad (8.5.2a)$$

$$\frac{dn_s}{dt} = k(1-p)n_f - \frac{1}{\tau_s}n_s, \quad (8.5.2b)$$

$$\frac{dn_l}{dt} = kp n_f - \frac{1}{\tau_l}n_l. \quad (8.5.2c)$$

Eliminating  $n_f$  results in two ODEs:

$$\frac{dn_s}{dt} = k(1-p)(M - n_l) - n_s \left( \frac{1}{\tau_s} + k(1-p) \right), \quad (8.5.3a)$$

$$\frac{dn_l}{dt} = kp(M - n_s) - n_l \left( \frac{1}{\tau_l} + kp \right). \quad (8.5.3b)$$

A steady state can be obtained by setting  $\frac{dn_s}{dt} = \frac{dn_l}{dt} = 0$  and is found to be

$$n_s^* = \frac{k(1-p)M\tau_s}{1 + k(p(\tau_l - \tau_s) + \tau_s)}, \quad n_l^* = \frac{kpM\tau_l}{1 + k(p(\tau_l - \tau_s) + \tau_s)}. \quad (8.5.4)$$

Analysing the stability of the steady state reveals it is stable for  $\tau_l > \tau_s$ , which always holds. Thus, the steady state is always stable. The eigenvalues are real given that  $\tau_l > \tau_s$ , implying the steady state is a node. From (8.5.4), the ratio of the mean number of short and long bound cells is given by

$$\frac{n_l^*}{n_s^*} = \frac{p}{1-p} \frac{\tau_l}{\tau_s}. \quad (8.5.5)$$

**Solutions of the ODEs**

The ODEs (8.5.3) can be written in matrix form as

$$\frac{d}{dt} \begin{pmatrix} n_l \\ n_s \end{pmatrix} = \begin{pmatrix} -kp - \frac{1}{\tau_l} & -kp \\ -k(1-p) & -k(1-p) - \frac{1}{\tau_s} \end{pmatrix} \begin{pmatrix} n_l \\ n_s \end{pmatrix} + \begin{pmatrix} Mkp \\ Mk(1-p) \end{pmatrix}. \quad (8.5.6)$$

The system can be solved by first finding the solution of the homogeneous part, then solving the inhomogeneous part. The homogeneous part of the system is solved by finding the eigenvalues and eigenvectors of the  $2 \times 2$  matrix in (8.5.6), given by

$$\lambda_1 = -\left(\frac{\gamma + \beta}{2\eta}\right), \quad \lambda_2 = -\left(\frac{\gamma - \beta}{2\eta}\right), \quad (8.5.7)$$

and

$$\begin{pmatrix} \frac{\alpha + \beta}{\omega} \\ 1 \end{pmatrix}, \quad \begin{pmatrix} \frac{\alpha - \beta}{\omega} \\ 1 \end{pmatrix}, \quad (8.5.8)$$

where

$$\gamma = \tau_l + \tau_s + k\tau_l\tau_s, \quad (8.5.9a)$$

$$\zeta = 1 + k(p(\tau_l - \tau_s) + \tau_s), \quad (8.5.9b)$$

$$\eta = \tau_l\tau_s, \quad (8.5.9c)$$

$$\beta = \sqrt{\gamma^2 - 4\zeta\eta}, \quad (8.5.9d)$$

$$\alpha = -k(1-2p)\eta - \tau_l + \tau_s, \quad (8.5.9e)$$

$$\omega = 2k(1-p)\eta. \quad (8.5.9f)$$

The general solutions of the homogeneous part of (8.5.6) is

$$\begin{pmatrix} n_l^{hom}(t) \\ n_s^{hom}(t) \end{pmatrix} = c_1 \begin{pmatrix} \frac{\alpha + \beta}{\omega} \\ 1 \end{pmatrix} e^{\lambda_1 t} + c_2 \begin{pmatrix} \frac{\alpha - \beta}{\omega} \\ 1 \end{pmatrix} e^{\lambda_2 t}. \quad (8.5.10)$$

A solution must now be found for the inhomogeneous part of the system. We guess a solution of the form

$$\begin{pmatrix} n_l \\ n_s \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}. \quad (8.5.11)$$

The derivative with respect to time of (8.5.11) is the zero vector. We must therefore solve

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -kp - \frac{1}{\tau_l} & -kp \\ -k(1-p) & -k(1-p) - \frac{1}{\tau_s} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} Mkp \\ Mk(1-p) \end{pmatrix}. \quad (8.5.12)$$

This gives

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \frac{kMp\tau_l}{\zeta} \\ \frac{kM(1-p)\tau_s}{\zeta} \end{pmatrix}. \quad (8.5.13)$$

We are able to write the general solution to the system as

$$\begin{pmatrix} n_l \\ n_s \end{pmatrix} = c_1 \begin{pmatrix} \frac{\alpha+\beta}{\omega} \\ 1 \end{pmatrix} e^{\lambda_1 t} + c_2 \begin{pmatrix} \frac{\alpha-\beta}{\omega} \\ 1 \end{pmatrix} e^{\lambda_2 t} + \begin{pmatrix} \frac{kMp\tau_l}{\zeta} \\ \frac{kM(1-p)\tau_s}{\zeta} \end{pmatrix}, \quad (8.5.14)$$

where  $c_1$  and  $c_2$  are constants. The eigenvalues  $\lambda_1$  and  $\lambda_2$  are always negative for  $\tau_s < \tau_l$ . Thus, the exponentials in (8.5.14) tend to zero as time tends to infinity, leaving the part of the solution that is not dependent on time, which is the steady state given in (8.5.4).  $c_1$  and  $c_2$  can be found by applying the initial conditions of the number of short and long bound cells at time zero to be zero,

$$n_l(0) = 0, \quad n_s(0) = 0. \quad (8.5.15)$$

We find

$$c_1 = -\frac{kM}{2\beta\zeta}((-\alpha+\beta)\tau_s + p(\omega\tau_l + (\alpha-\beta)\tau_s)), \quad c_2 = \frac{kM}{2\beta\zeta}(-(\alpha+\beta)\tau_s + p(\omega\tau_l + (\alpha+\beta)\tau_s)). \quad (8.5.16)$$

### Analysis of Model IIIA

The number of bound and unbound cells as a function of time can be viewed by solving the ODEs, Gillespie algorithm (see section 2.4), and a spatial code. Results of these simulations are shown in Figure 8.5.

Figure 8.5 shows an initial increase in the number of short bound cells, caused by a higher probability of short bindings occurring. The number of long bound cells quickly takes over, as although this type of binding is less probable, cells that undergo a long binding remain bound for a greater time, allowing long bindings to accumulate. By twenty hours the system

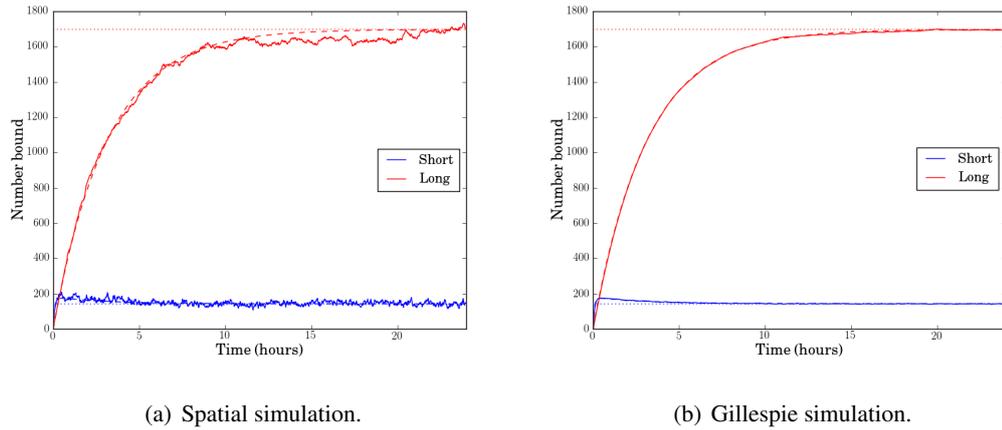


Figure 8.5: Plots showing Model IIIA, simulated using (a) one spatial simulation and (b) the Gillespie algorithm, in both cases with comparison to ODEs, shown by dashed lines. (b) shows the average of one hundred Gillespie simulations. The steady state of the system (8.5.4) is shown by dotted lines. Parameters used:  $R = 400\mu\text{m}$ ,  $b = 11.5\mu\text{m}$ ,  $D = 50\mu\text{m}^2\text{min}^{-1}$ ,  $\tau_s = 5$  minutes,  $\tau_l = 240$  minutes,  $p = 0.2$ , APCs = 200, and T cells = 8060 after equilibration.

approaches a steady state, with the number of long bindings being much greater than the number of short bindings. The results of these simulations show that, at very early times short bindings dominate. However, long bindings dominate at earlier times than observed in biological experiments [110, 152, 154]. The parameters were chosen to be similar to those used in previous chapters, which are thought to be biologically relevant. The short and long binding times were thought to be reasonable according to literature [42, 110, 152, 156, 201, 212]. With these parameters, for short bindings to dominate at times earlier than 8 hours [110, 152, 154], a probability of a long binding occurring must be set around 0.022. With this probability, long bindings dominate at later times but are only slightly more numerous than short bindings. If long bindings are taken to have a length of 20 hours, as found by [112], a probability of a long binding occurring of 0.011 is required to achieve a switch from short to long bindings at a time shortly after 8 hours, shown in Figure 8.6.

In the spatial simulation shown in Figure 8.5(a), binding times are drawn from an exponential distribution, with mean set as  $\tau_s$  or  $\tau_l$ , depending on the type of binding occurring. To initialise the simulation, T cells are placed in uniform random positions within the domain, excluding the volume occupied by the APCs. This results in a distribution of T cells that is not uniformly randomly distributed across the whole domain as the volume of the APCs is excluded. The parameter  $k$  assumes a uniform random distribution across the whole domain. To allow cells to

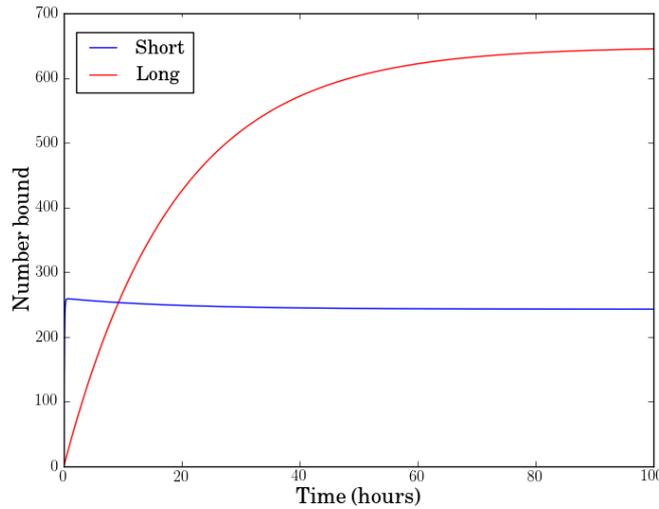


Figure 8.6: Figure showing a case in which there is a shift from short to long binding dominance just after 8 hours. Parameters used:  $R = 400\mu m$ ,  $b = 11.5\mu m$ ,  $D = 50\mu m^2 min^{-1}$ ,  $\tau_s = 5$  minutes,  $\tau_l = 1200$  minutes,  $p = 0.011$ , APCs = 200, and T cells = 10000.

become approximately uniformly randomly distributed the T cells are placed within the domain, excluding the volume occupied by APCs, then an equilibration time is given in which cells move following a random walk, allowing cells to become randomly distributed through the domain. Any T cells encountering an APC during the equilibration time are removed from the simulation. Upon a T cell unbinding from an APC, in the spatial model, it is replaced in a valid ‘random’ position within the domain. This is done to achieve a more accurate comparison to the first passage time equation.

Figure 8.7 shows the trajectory the system follows to reach the steady state. The green and yellow lines represent the nullclines, where the rate of change of short and long bound cells is zero, implying one population is in equilibrium for the corresponding value of the other population. The figure shows the number of short bound cells increases to the nullcline, where the rate of change of short bound cells is zero. The system then approaches the steady state.

### 8.5.2 Model IIIB

In Figure 8.5(a) the spatial model allows an equilibration time to compensate for a not truly uniform random distribution of T cells at the beginning of the simulation. To prevent T cells immediately rebinding after an encounter with an APC, in Model IIIA they are replaced

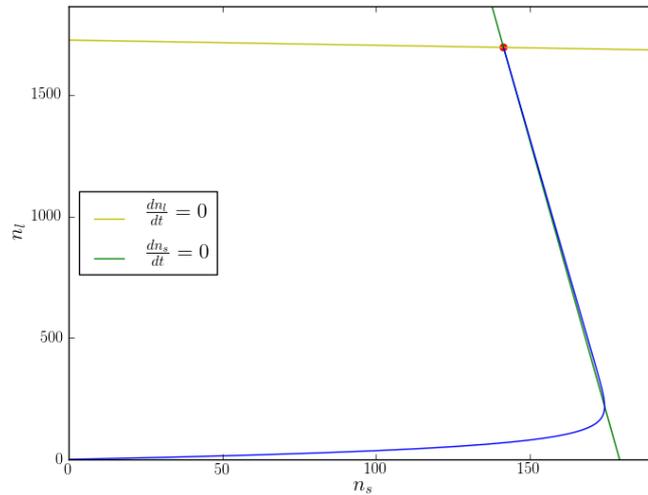


Figure 8.7: Plot showing the trajectory of Model IIIA (blue), for long and short bound cells. Yellow and green lines indicate the nullclines, when the rate of change of long and short bound cells is zero. The intersection of the nullclines is the steady state of the system. Parameters used are the same as in Figure 8.5.

‘randomly’ in a valid position within the domain. This behaviour is unrealistic biologically as a T cell cannot move from one position to another instantaneously. Also, the position the T cell is placed in will not be truly uniform random across the domain as it excludes the volume occupied by the APCs. To alleviate both of these problems, in Model IIIB, cells become inactive after unbinding from an APC for an amount of time drawn from an exponential distribution. During this inactive time a T cell will continue to move, but will not be able to bind to APCs. This allows the T cells to return to a roughly uniform random position within the LN, before they can rebind to an APC. An ODE model can be setup to simulate an inactivation time by adding an extra population to the previous ODE system (8.5.3) to include the inactive cells.

In this model there are four populations:  $n_f$ , mean number of free cells,  $n_s$ , mean number of short bound cells,  $n_l$ , mean number of long bound cells, and  $n_i$ , mean number of inactivated cells. The total number of T cells is given by  $M$ , allowing one of the populations to be eliminated:  $n_f = M - n_s - n_l - n_i$ . Upon unbinding a T cell enters an inactive state for a time  $\tau_i$ , after which

it returns to the population of free T cells. We have the following ODEs:

$$\frac{dn_f}{dt} = -kn_f + \frac{1}{\tau_i}n_i, \quad (8.5.17a)$$

$$\frac{dn_s}{dt} = k(1-p)n_f - \frac{1}{\tau_s}n_s, \quad (8.5.17b)$$

$$\frac{dn_l}{dt} = kpn_f - \frac{1}{\tau_l}n_l, \quad (8.5.17c)$$

$$\frac{dn_i}{dt} = \frac{1}{\tau_s}n_s + \frac{1}{\tau_l}n_l - \frac{1}{\tau_i}n_i. \quad (8.5.17d)$$

Eliminating  $n_f$  gives the three ODEs:

$$\frac{dn_s}{dt} = k(1-p)(M - n_l - n_i) - n_s \left( \frac{1}{\tau_s} + k(1-p) \right), \quad (8.5.18a)$$

$$\frac{dn_f}{dt} = kp(m - n_s - n_i) - n_l \left( \frac{1}{\tau_l} + kp \right), \quad (8.5.18b)$$

$$\frac{dn_i}{dt} = \frac{1}{\tau_s}n_s + \frac{1}{\tau_l}n_l - \frac{1}{\tau_i}n_i, \quad (8.5.18c)$$

in which the parameters are the same as in (8.5.3), with the addition of an inactivation time given by  $\tau_i$ . The steady state for this system is found to be:

$$\begin{aligned} n_s &= \frac{kM(1-p)\tau_s}{1 + k(\tau_i + p\tau_l + (1-p)\tau_s)}, \\ n_l &= \frac{kM\tau_l}{1 + k(\tau_i + p\tau_l + (1-p)\tau_s)}, \\ n_i &= \frac{kM\tau_i}{1 + k(\tau_i + p\tau_l + (1-p)\tau_s)}. \end{aligned} \quad (8.5.19)$$

Analysis of the eigenvalues reveals the steady state to be a stable node.

A plot of the solution of the ODEs can be seen in Figure 8.8, with comparison to a spatial simulation and the mean of one hundred Gillespie simulations of the system.

Figure 8.8 has the same dynamics as Figure 8.5 for the short and long bound cells. The added population of inactive cells does not result in the spatial model being a better fit to the ODEs. However, the behaviour of the two models is qualitatively the same, and Model IIIB is a more realistic version of the biological system. Further investigation is required to determine the cause of the difference between the spatial model and ODEs in Figure 8.8(a).

## 8.6 Summary

In this chapter, three activation mechanisms are investigated: APC changes, signal integration, and probabilistic approach, motivated by hypotheses in literature [40,42]. For APC changes, two

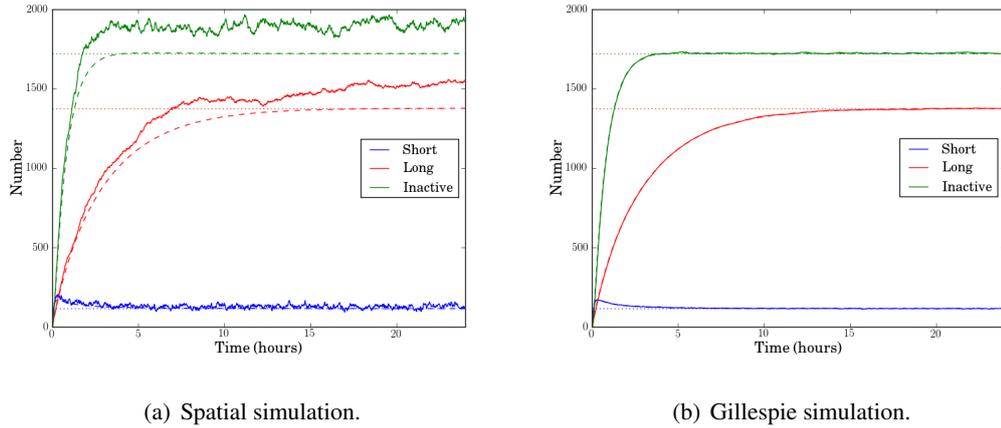


Figure 8.8: Plots showing Model IIIB, simulated using (a) a spatial simulation and (b) the Gillespie algorithm, in both cases with comparison to ODEs, shown by dashed lines. (b) shows the average of one hundred Gillespie simulations. The steady state of the ODE system (8.5.19) is shown by dotted lines. Parameters used:  $R = 400\mu\text{m}$ ,  $b = 11.5\mu\text{m}$ ,  $D = 50\mu\text{m}^2\text{min}^{-1}$ ,  $\tau_s = 5$  minutes,  $\tau_l = 240$  minutes,  $\tau_i = 60$  minutes,  $p = 0.2$ , APCs = 200, and T cells = 8260 after equilibration.

models are developed. The first model (Model IA) assumes APCs must spend a specific amount of time in the LN before they mature and are able to undergo long bindings with T cells. This results in a switch from short to long bindings at a set timepoint, and depends on no other parameters in the model. The second model (Model IC) assumes APCs mature due to contacts with T cells and must have a specific cumulative binding time with T cells before they can undergo long bindings. This model has a more gradual shift from short to long bindings than Model IA, and is dependent on the cumulative binding time required, as well as the number of T cells and APCs in the system. This allows manipulation of the system by adding more T cells or APCs, which could in principle be tested experimentally. This model can be developed further by considering how APCs receive maturation signal from the T cells. For example, APCs may not be able to receive signals from multiple T cells simultaneously, as assumed in Model IC.

Two models are considered for the signal integration approach. The first model (Model IIA) assumes T cells record the cumulative time they have spent bound to an APC and when this value exceeds a given threshold, the specific T cell that has exceeded the threshold can undergo long bindings. This model depends upon the rate of encounter between T cells and APCs, the number of APCs, and the cumulative time required for long bindings to occur. In principle, this system can be investigated experimentally by manipulating the number of APCs. The second model (Model IID) assumes a T cell's binding time is a function of the time a T cell has already spent bound to

APCs. A function was created to give binding times as a proportion of a maximum binding time. This model proved difficult to calibrate and did not produce the results desired. This model could be expanded by considering different functions for the binding times.

A probabilistic approach was considered by developing two models. The first model (Model IIIA) assumes that upon encountering an APC T cells undergo a long binding with probability  $p$  or a short binding with probability  $1 - p$ . In this model, to achieve a switch in the dominant type of binding from short to long, at a time around 8 hours as stated in the literature [110, 152, 154], with a long binding time of 20 hours [112], a probability of a long binding occurring must be set as 0.011. A long binding time of 20 hours seems quite long considering long bindings are observed to occur between 6 and 24 hours after initial antigen presentation [110, 201]. If 20 hour bindings start occurring 6 hours after initial antigen presentation, they will still be occurring 24 hours after the start of antigen presentation. Using a lower long binding time can still achieve a long binding dominance around 8 hours, but the difference between the number of short and long bindings is small. A second probabilistic model (Model IIIB) was developed with the aim of being more biologically realistic. Model IIIA places T cells, after unbinding, approximately uniformly randomly within the domain, which is unrealistic. To alleviate this problem, an inactivation state was added in Model IIIB, which T cells transfer to after unbinding. T cells in the inactive state can move freely but cannot bind, allowing them to return to a roughly uniform random position before being able to bind again. This did not qualitatively change the model from Model IIIA, it did however make the model more realistic. The probabilistic models considered do not seem to be able to satisfactorily explain the experimental observations presented in the introduction, and a more complex model may be required. A model in which T cells return to undergoing shorter bindings should also be investigated, and may be what is required to better explain the experimental results.

This chapter outlines basic models of investigating the programme of activation of T cells. Due to time constraints further investigation could not be undertaken during my PhD. It is my hope that future work, by both mathematicians and experimentalists, will allow the programme of activation of T cells to be determined.

## Chapter 9

# Concluding remarks

This chapter provides a summary of the work in Chapters 3-8, and considers if the initial objectives of the thesis, outlined in Section 1.9, have been fulfilled.

This thesis is concerned with the dynamics and interactions of lymphocytes within lymph nodes (LNs). It focuses on determining how T cells move, and the diffusion coefficient they move with. The diffusion coefficient is a key parameter in mean first passage time (MFPT) equations, used to determine the time a T cell moving following a random walk take to encounter uniformly randomly placed antigen-presenting-cells (APCs). The interaction rate of T cells and APCs determines the time an immune response requires to be initiated. The number of APCs presenting antigen to T cells is a limiting factor to the initiation of an immune response. This thesis also investigates the minimum number of APCs required to initiate a T cell response. The result of interactions between lymphocytes is considered in the cross-regulation model, in Chapter 7, which investigates three-body interactions between APCs, effector T cells, and regulatory T cells, as a mechanism of effector T cell suppression of autoimmunity. The rate of interaction of T cells and APCs is critical to initiating an immune response, but the way in which T cells are activated, upon encountering APCs, is still unknown. Finally, Chapter 8 investigates various hypotheses of T cell activation.

Chapter 3 of this thesis considers first passage processes. Firstly, a MFPT is derived for a particle in a prescribed initial position to locate a fixed target in the centre of a spherical domain, using Poisson's equation. This is generalised to a uniformly randomly placed target to obtain a starting global mean first passage time (SGMFPT). To derive a target global mean first passage time (TGMFPT), for a particle and target located in uniform random initial positions, which excludes the volume occupied by the target, using Poisson's equation, an approximation for the surface of

the target, and a coordinate transformation to bispherical coordinates are considered. The first approach proves to be of limited use, and the second approach proved to be very complex to solve. Making use of recent mathematical advances in the area of first passage processes [58], a method to derive a SGMFPT, involving Green's and pseudo Green's functions was implemented. This resulted in a solution that could be integrated to obtain a TGMFPT. The method was then extended to include multiple uniformly randomly placed targets, to give an equation for the mean time a particle takes to encounter one of  $N$  uniformly randomly placed targets.

Imaging experiments carried out *in vivo* are limited to a finite volume. A limited sized volume results in cells that remain in the volume for a long period of time, being those that are either not moving very fast, or are taking looping paths, making them appear confined. Confinement is observed in displacement squared against time plots as a plateau at later times. Chapter 4 calculates the expectation of the squared displacement of cells, within a cubic domain, to predict the displacement squared as a function of time, for a given set of parameters. It is found that, in imaging volumes smaller in the  $z$  axis than the  $x$  and  $y$  axes, as used in *in vivo* imaging experiments, a plateau is not observed due to particles exiting via the  $z$  boundary. It can be concluded that, in imaging experiments where a confinement plateau is observed, for typical sized imaging volumes, it is not due to a limited sized imaging volume but rather some other form of confinement. This other form of confinement would most likely be immunological confinement, due to chemical signals or other biological factors confining the cells to a particular region.

Chapter 5 analyses data obtained by *in vivo* imaging experiments. Mathematical techniques are applied to determine the type of motion the imaged cells are undergoing, and calculate the diffusion coefficient of the cells. A systematic method is developed to analyse a given data set and determine a diffusion coefficient of the population of cells imaged. It is found that, due to the small size of the  $z$  axis used, viable data is not observed at later imaging times. We conclude that more useful information could be obtained by using a larger  $z$  dimension, or an increased  $z$  or time resolution.

Chapter 6 investigates the minimum number of dendritic cells (DCs) required to initiate a T cell response. A simple model is devised to calculate the probability of T cell activation in a given time, and experimental results are used to calibrate the model. Mathematical modelling allows longer times and lower T cell numbers to be investigated than can be done experimentally. An extension to the model is developed to include the number of T cell searching for DCs in an attempt to become activated. This model is used to calculate a minimum number of DCs required to initiate

a T cell response, in twenty four hours, for typical human T cell precursor frequencies.

Chapter 7 develops a stochastic version of the cross-regulation model, first developed deterministically in [133]. The cross-regulation model postulates that regulatory T cells suppress effector T cells via a three way interaction with APCs, as a mechanism of preventing autoimmunity. A stochastic model is developed showing stochastic effects are present, and significant, for initial ratios of regulatory/effector T cell with a high fraction of regulatory T cells. It is found that, when the initial regulatory T cell ratio is low, there is a high probability of autoimmunity. This probability decreases as the initial ratio of regulatory T cells increases. When the initial ratio of regulatory T cells is very large the probability of autoimmunity is found to slightly increase due to saturation by regulatory T cells. This effect is not observed in a deterministic approach. The binding rate of T cells is found to be the key parameter in determining the steady state of the system, with the other parameters only affecting the timescale.

Chapter 8 investigates three mechanisms proposed to explain the three stage activation process of T cells [40,42]. These mechanisms are: APC changes, T cell signal integration, and a probabilistic approach. For each of the mechanisms two models are considered. We have been able to show that all models can explain the first two stages of the activation process. However, the third stage of the process is not covered in the scope of these models. The simple probabilistic model devised is able to explain experimental results observed, but not as convincingly as could have been hoped for. It is concluded that a more complex probabilistic model is required to explain the experimental results, and the third stage of activation must also be investigated.

In Section 1.9, the initial objectives of this thesis were introduced. The success of this thesis at completing these initial objective will now be analysed. The first objective was to derive an equation for the time a uniformly randomly located particle, within the domain that excludes the volume occupied by the target, moving by Brownian motion, takes to encounter a uniformly randomly placed spherical target, within a spherical domain. Chapter 3 addresses this problem and derives the desired equation, which had not been previously found. An equation is also derived for the time it takes a particle to encounter one of  $N$  uniformly randomly placed targets, which again has not been previously developed. Therefore, the first objective is met and exceeded, having derived an equation beyond what was initially intended.

The second objective was to determine the effects of confinement in a limited sized imaging region. Chapter 4 determines an equation for the expected squared displacement of cells as a function of time, and the expected confinement plateau. This allows results of a given experimental setup to

be predicted and quantified by determining the mean squared displacement plot and plateau that should be observed and comparing to the mean squared displacement plot that is observed. From the work in Chapter 4 and 5, ways of optimising the size and resolution of an imaging volume can be proposed. It can be concluded that objective two has been achieved through this work.

Objective three was to develop a method of calculating the diffusion coefficient of data sets obtained through *in vivo* imaging experiments. This was investigated in Chapter 6, and a method was developed. The method developed is able to calculate the diffusion coefficient of a population of cells in a given data set. I believe there is still scope for further development of the method in the hope of being able to extract more information from the limited sized data sets. Objective three has therefore been partially met, but there is still room for improvement.

The fourth object was to determine the minimum number of DCs required to initiate a T cell response. Through collaboration with experimentalist (Dr. P. Bousso, Paris), and the development of mathematical models, this objective was fulfilled and resulted in a publication in the Blood journal [53]. This work is contained in Chapter 6.

Objective five was to investigate the effects of stochastic fluctuations on the cross-regulation model. In Chapter 7, a stochastic formulation of the cross-regulation model is developed. Stochastic effects were observed that were not seen in the deterministic model. For example, very high initial ratios of regulatory T cells were found to increase the probability of autoimmunity. This fulfils objective five and shows stochastic effects cannot be dismissed in biological models when modelling small numbers of cells.

Finally, the sixth objective of determining the method of T cell activation was addressed in Chapter 7. Various models of T cell activation were investigated. These models were able to explain the first two stages of the T cell activation process: transient interactions and long lived interactions. However, the third stage of the activation process, swarming, was not investigated. As all the investigated models were able to explain the experimental results, to some extent, for the first two stages of the activation process, the activation method that is occurring biologically cannot be determined. To be able to distinguish one of the activation methods as being able to better explain the experimental data, the third stage of the activation process must be considered. Thus, the sixth object was not met. Further work is required in this area to add the third stage of the activation process to the models.

This thesis covers several interlinked projects, which have been investigated to varying degrees.

All areas in this thesis have the potential to be expanded upon, and a few extensions will be outlined here. In Chapter 3, an equation for the time a particle takes to locate one of  $N$  targets is derived. However, the equation assumes the targets can overlap, which is unrealistic. For a small number of targets, in a large domain, the error resulting from this assumption will be very small. But, as the number of targets is increased, the error will also increase. This work can be extended by developing an equation for non-overlapping targets. A further extension would be to add an exit to the domain and determine the time to locate the exit and the probability of encountering a target before exiting the domain. The work in Chapter 5 develops a method of calculating the diffusion coefficient of cells from *in vivo* imaging experiments. This work can be developed by extracting more information from the data available. Instead of just using measurements from a point along a cell track to the start of the track, new sample tracks can be created by segmenting existing tracks into smaller tracks. Segmentation of tracks is achieved by taking a time point and creating a sample track from that timepoint to timepoint 2, 3, 4, etc. Then the next time point along the track can be taken as the starting point and more sample tracks created. However, this results in the contribution of a track to the mean squared displacement being dependent on the length of the track. Therefore if there is an anomalous track it would be exaggerated. A method would need to be developed to prevent this scenario occurring. Chapter 8 introduces basic methods of investigating the hypotheses of T cell activation, but due to time constraints these methods could not be investigated to their full potential. This work could be extended by adding further complexity to the models developed in Chapter 8 and investigating the models proposed but not developed, as they may result in interesting features not previously considered. To fully answer the question of how T cells are activated, the third stage of activation must also be modelled. Also, in the models considered in chapter 8, APCs are assumed to have an infinite number of binding sites. This work could be extended by investigating the effect of limiting the number of binding sites.

The work in Chapters 5, 6 and 8 was carried out using data from or in collaboration with immunologists. Further development of work in this thesis, and the area of mathematical immunology, is only possible by working closely with immunologists. Mathematical models can be motivated and developed by work carried out by immunologists. In turn, the modelling efforts of mathematicians can be used to feedback to immunologists aspects of systems that are found to be interesting from a mathematical point of view, which could then be investigated experimentally. Work carried out by mathematicians can also be used to suggest ways of improving current experimental procedures, as was achieved by the work of Chapters 4 and 5.



# Appendices

## A Spatial modelling

To implement a spatial model of cell movements and interactions within a LN, the programming language python has been used. Sections A.1 to A.6 show the algorithms used to generate various aspects of the spatial model.

### A.1 Random placement of cells

To generate uniform random coordinates for a cell within a sphere of radius  $R$ , we must do the following. A uniform random position in a cube is generated of radius  $R$ . The position is then checked to determine if it is contained within a sphere of equal radius to the cube, if it is, the position is accepted, if not a new position is generated. This process continues until a valid position is chosen.

```
import numpy as np

def randomstart(R):
    rsnow = 2*R*R
    while rsnow > R*R:
        x = np.random.uniform(-R, R)
        y = np.random.uniform(-R, R)
        z = np.random.uniform(-R, R)
        rsnow = x*x + y*y + z*z
    return x, y, z
```

## A.2 Non-overlapping APCs

To create  $N$  non-overlapping APCs, we can make use of the python package ‘numpy’. This enables linear algebra operations to be implemented. In this example, numpy allows the distance to every APC to be calculated simultaneously. Using numpy instead of a loop to calculate the distance to each APC allows a quicker implementation of the algorithm. To generate non-overlapping positions we first select a uniform random position for the APC within the domain. If this is the first APC to be placed the position need not be checked, otherwise we must check the distance from the other APCs that have already been placed. If the minimum of the array of distances to the nearest APCs is greater than double the effective radius of the APC,  $2b$ , the position is accepted, if not a new position is generated. This process continues until valid positions have been chosen for all the APCs needing to be placed. This generates a distribution of APC positions which is not truly uniformly randomly distributed, as each additional APC is placed following a uniform random distribution with the domain that is not already occupied by APCs, rather than the whole domain.

```
import numpy as np

Ax = np.array([])
Ay = np.array([])
Az = np.array([])

for i in range(N):
    nearestdist = 0.0
    while nearestdist < 2*b:
        x, y, z = randomstart(R-b)
        nearestdist = R
        if len(Ax) !=0:
            distance = np.sqrt( (x-Ax)*(x-Ax) + (y-Ay)*(y-Ay)
+ (z-Az)*(z-Az) )
            nearestdist = min( nearestdist, distance.min() )
        else:
            nearestdist = 2*b
    Ax = np.append(Ax, x)
```

```
Ay = np.append(Ay, y)
Az = np.append(Az, z)
```

### A.3 Unbinding

To unbind cells, we must first determine if the time passed since the last unbinding or binding event is greater than the minimum binding time of all the cells bound. If this condition is met, we firstly deal with the cells that are going to unbind. To determine which cells these are we create a mask to exclude the cells that are not bound and another mask to determine which cells have a binding time less than  $t_{passed}$ . Multiplying these two masks gives the cells that need unbinding. We now reset those cells binding times to zero, and replace them in ‘random’, valid positions. To subtract  $t_{passed}$  from the remaining bound cells we need to create a mask of the cells that have a binding time greater than  $t_{passed}$ , and subtract this mask multiplied by  $t_{passed}$  from the array of binding times. We now determine the time until the next unbinding event occurs by finding the minimum of the binding times. This algorithm must also be implemented if a new cell binds.

```
import numpy as np

if tpassed > unbind:
    mask1 = bindingtime > 0
    mask2 = bindingtime < tpassed
    bindingtime = bindingtime - bindingtime*mask1*mask2

    replace = mask1*mask2
    for i in range(N):
        if replace[i] == 1:
            Tx[i], Ty[i], Tz[i] = randomstart(R-a)
            Tclosest[i], Tdist[i] = closest(Tx[i], Ty[i],
            Tz[i], Ax, Ay, Az, R)

    mask3 = bindingtime > tpassed
    bindingtime = bindingtime - tpassed*mask3

    nz = bindingtime*(bindingtime > 0)
```

```
unbind = nz.min()
tpassed = 0.0
```

#### A.4 Cell movement

Moving cells can be done individually for each cell or using the python module numpy, by vector operations. To do this we generate an array of random movements, from a Gaussian distribution, in each of the  $x$ ,  $y$  and  $z$  dimensions, and add these to the cells positions. We now move any cells that have crossed the boundary of the sphere back into the sphere. This is done by calculating the distance of all the cells from the centre, and creating a mask of those cells that have a distance greater than the radius of the sphere. The adjustment factor needed to reflect these cells back into the sphere is calculated, and an array created to multiply the cell positions by, making use of our mask.

```
import numpy as np

def move(x, y, z, D, dt, R):
    size = len(x)
    x += np.random.normal(0, np.sqrt(2*D*dt), size)
    y += np.random.normal(0, np.sqrt(2*D*dt), size)
    z += np.random.normal(0, np.sqrt(2*D*dt), size)
    rsnow = x*x + y*y + z*z
    mask = rsnow > R*R
    mod = ((2*R)/sqrt(rsnow))-1
    mult = np.ones(size)
    mult[mask] = mod[mask]
    x = mult*x
    y = mult*y
    z = mult*z
    return x, y, z
```

## A.5 Closest APC

To find which APC is closest to a given T cell, the distance from the T cell to each APC is calculated and the minimum of this array of distances is found. The index of the closest T cell is determined using the `where` command. The index and the minimum distance to an APC is returned from this algorithm. The minimum distance can then be used to determine if a T cell has encountered an APC or not.

```
import numpy as np

def closest(x, y, z, Ax, Ay, Az, R):
    distance = np.sqrt((x-Ax)*(x-Ax) + (y-Ay)*(y-Ay)
        + (z-Az)*(z-Az))
    dist = distance.min()
    jj = np.where(distance == dist)
    return jj[0][0], dist
```

## A.6 Encounter during a timestep

To calculate the probability of a T cell encountering an APC during a timestep, the position of each T cell to its nearest APC must be calculated before and after the timestep. We firstly form an array of the distance before the timestep by copying the array of distances from the previous loop of the program. Arrays for the  $x$ ,  $y$ , and  $z$  coordinates of the nearest APC to each T cell are created by indexing the arrays of APC positions by the array of the indexes of the closest APC to each T cell. Using these arrays of the coordinates of the closest APCs, the distance for each T cell to its closest APC can be calculated after the timestep. Using the distances before and after the timestep the probability of encountering an APC during a timestep, for each T cell, is calculated, using the equation given in Section 2.3.1. These probabilities are tested to determine if they are greater than a uniform random variable, between zero and one, a positive result is given for encountering an APC.

```
import numpy as np

prevdist = np.copy(Tdist)
```

```

CloseAPCx = Ax[Tclosest]
CloseAPCy = Ay[Tclosest]
CloseAPCz = Az[Tclosest]
Tdist = np.sqrt((Tx - CloseAPCx)*(Tx - CloseAPCx) +
(Ty - CloseAPCy)*(Ty - CloseAPCy) + (Tz - CloseAPCz)*(Tz-CloseAPCz))
prohibit = np.exp(-((prevdist-bb)*(Tdist-bb))/(D*dt))
Random = np.random.random(N)
foundAPC = Random < prohibit

```

## B Code for encountering one of $N$ targets

In this appendix, the code used to calculate the time for a particle to encounter one of  $N$  targets is shown. This code was used to produce Figure 3.15. The code is able to calculate a time to encounter one of one, two, three, up to  $N$  targets, simultaneously. In this code  $N$  is set to be ten. This is achieved by placing the targets and one particle ‘randomly’, as described in A.1. The particle is not placed within a target, and the targets are able to overlap, as is assumed in (3.7.103). Each of the targets is assigned a number from one to  $N$ . The particle moves using exponential timestepping, and the distance to each target is calculated every timestep. An algorithm is used to determine if the particle has encountered a target. If an encounter has occurred, the number of the target is checked, and a encounter time recorded for that number of targets, and all the targets with a greater number. For example, with ten targets, if target seven is encountered the time the encounter occurs at is recorded as the encounter time for seven, eight, nine and ten targets. These targets are then removed and the simulation continues with the remaining targets, in the example given this would be targets one to six. This procedure continues until a encounter time has been found for all targets. If target one is the first target to be encountered, the same encounter time would be given for all the targets and the simulation would end. By using this procedure of numbering the targets, it allows encounter times to be found for one of  $N$  targets simultaneously, rather than running individual simulations for each number of targets. This greatly reduces the computational time required to execute the program. The program also becomes quicker to execute a timestep as targets are encountered and removed, as there are less targets for the particle to check if it has encountered. To achieve an accurate TGMFPT numerous realisations of the program are required. Making the program used as efficient as I could, allowed more realisations to be computed in a given time. Even with making the program efficient, Figure 3.15

required vast amounts of computing power. I was fortunate to have access to the Leeds University high performance computing system, Arc1. Utilising this system, over a years worth of computing time was required to produce Figure 3.15.

```
import numpy as np
import math as m

R = 1.0      # Radius of sphere
D = 0.5      # Diffusion coefficient for particle
NR = 10000   # Number of realisations
b = 0.02     # Radius of targets
dt = 0.0001  # Timestep size
N = 10       # Number of targets

filename = 'MultRandom_sim'+str(sim)+'_N'+str(N)+'b'+str(b)
          +'_dt'+str(dt)+' .dat'

## Random starting position algorithm ##
def randomstart(R):
    rsnow = 2*R*R
    while rsnow > R*R:
        x = np.random.uniform(-R, R)
        y = np.random.uniform(-R, R)
        z = np.random.uniform(-R, R)
        rsnow = x*x + y*y + z*z
    return x, y, z

## Exponential timestepping algorithm ##
def ExpStep(nu):
    urv = np.random.random()
    sgs = m.sqrt(-2*m.log(1-urv))
    g1 = np.random.normal(0, 1)
    g2 = np.random.normal(0, 1)
```

```
g3 = np.random.normal(0, 1)
mod = sgs*m.sqrt(g1*g1+g2*g2+g3*g3)/nu
urv = np.random.random()
dz = (2*urv-1)
angle = 2*m.pi*np.random.random()
dx = m.sqrt(1-dz*dz)*m.cos(angle)
dy = m.sqrt(1-dz*dz)*m.sin(angle)
return mod*dx, mod*dy, mod*dz

## Algorithm to calculate distance to nearest target ##
def distance(tx, ty, tz, ax, ay, az):
    distance = m.sqrt((tx-ax)**2+(ty-ay)**2+(tz-az)**2)
    return distance

nu = 1.0/m.sqrt(dt*D)
N_arg = np.arange(N)

j = 0
while j < NR:
    j += 1

    ## Arrays to store data in ##
    APCHit = np.zeros(N)
    Tdb = np.ones(N)*R
    Tda = np.ones(N)*R

    dist = 0
    ## Set initial conditions inside sphere for the target ##
    wx = np.array([])
    wy = np.array([])
    wz = np.array([])
    for i in range(N):
        x, y, z = randomstart(R-b)
```

```
wx = np.append(wx, x)
wy = np.append(wy, y)
wz = np.append(wz, z)

## Random initial condition inside sphere for the particle ##
while dist < b:
    Tx, Ty, Tz = randomstart(R)
    Dist = np.sqrt((Tx-wx)*(Tx-wx) + (Ty-wy)*(Ty-wy)
                  + (Tz-wz)*(Tz-wz))
    dist = Dist.min()

nstep=-1
## Dynamics of system ##
while len(wx) > 0:
    nstep += 1

## Distances before movement ##
Tdb = np.sqrt((Tx-wx)*(Tx-wx) + (Ty-wy)*(Ty-wy)
              + (Tz-wz)*(Tz-wz))
Rdb = m.sqrt(Tx*Tx + Ty*Ty + Tz*Tz)

## Move particle ##
dx,dy,dz = ExpStep(nu)
Tx += dx
Ty += dy
Tz += dz

## Distances after movement ##
Tda = np.sqrt((Tx-wx)*(Tx-wx) + (Ty-wy)*(Ty-wy)
              + (Tz-wz)*(Tz-wz))
Rda = m.sqrt(Tx*Tx + Ty*Ty + Tz*Tz)

## Check if encounter has occurred ##
```

```

r = np.minimum(Tdb, Tda)
prob = np.exp(-nu*(r-b))*m.sinh(nu*b)/np.sinh(nu*r)
hit = np.random.random(len(wx)) < prob

## Records encounter times ##
for i in range(len(wx)):
    if hit[i] == 1:
        ZeroEntry = (APCHit == 0)
        gti = (n_arg) >= i
        to_add = gti*ZeroEntry*nstep*dt
        APCHit += to_add
        wx = wx[:i]
        wy = wy[:i]
        wz = wz[:i]
        break

## Check if particle has hit the outer boundary ##
dw = Rda-Rdb
urv = np.random.random()
maxx = dw/2.0 + m.sqrt(dw*dw - 2*dt*m.log(urv))/2.0
if Rda + maxx > R:
    newR = dw - maxx + R
    Tx *= newR/Rda
    Ty *= newR/Rda
    Tz *= newR/Rda
    Rda = newR

## Stores results ##
f = open(filename, 'a')
print >> f, ' '.join(str(cell) for cell in APCHit)
f.close()

```

# Bibliography

- [1] A.K. Abbas, A.H.H. Lichtman, and S. Pillai, *Cellular and molecular immunology*, Saunders, 1994.
- [2] A.K. Abbas, K.M. Murphy, A. Sher, et al., *Functional diversity of helper T lymphocytes*, Nature **383** (1996), no. 6603, 787–793.
- [3] E. Agliari, *Exact mean first-passage time on the T-graph*, Physical Review E **77** (2008), no. 1, 011128.
- [4] E. Agliari and R. Burioni, *Random walks on deterministic scale-free networks: Exact results*, Physical Review E **80** (2009), no. 3, 031125.
- [5] R. Alan, B. Ezekowitz, and J. Hoffmann, *Innate immunity: The blossoming of innate immunity*, Current Opinion in Immunology **10** (1998), no. 1, 9–11.
- [6] C. Alanio, F. Lemaitre, H.K.W. Law, M. Hasan, and M.L. Albert, *Enumeration of human antigen-specific naive CD8<sup>+</sup> T cells reveals conserved precursor frequencies*, Blood **115** (2010), no. 18, 3718–3725.
- [7] A.D. Allen, *Modeling AIDS and its treatment with immunostimulation*, Medical Hypotheses **26** (1988), no. 1, 55–58.
- [8] L.J.S. Allen, *An introduction to stochastic processes with applications to biology*, Pearsons Education, Inc., 2003.
- [9] A.O. Anderson and N.D. Anderson, *Studies on the structure and permeability of the microvasculature in normal rat lymph nodes*, The American journal of pathology **80** (1975), no. 3, 387.
- [10] U.H.V. Andrian and C. M'rini, *In situ analysis of lymphocyte migration to lymph nodes*, Cell Communication and Adhesion **6** (1998), no. 2-3, 85–96.

- [11] M. Asano, M. Toda, N. Sakaguchi, and S. Sakaguchi, *Autoimmune disease as a consequence of developmental abnormality of a T cell subpopulation*, *The Journal of experimental medicine* **184** (1996), no. 2, 387–396.
- [12] S. Asmussen and P. Glynn, *Gaussian processes*, *Stochastic Simulation: Algorithms and Analysis* (2007), 306–324.
- [13] S. Asmussen, P. Glynn, and J. Pitman, *Discretization error in simulation of one-dimensional reflecting Brownian motion*, *The Annals of Applied Probability* **5** (1995), no. 4, 875–896.
- [14] S. Asmussen, P.W. Glynn, and H. Thorisson, *Stationarity detection in the initial transient problem*, *ACM Transactions on Modeling and Computer Simulation (TOMACS)* **2** (1992), no. 2, 130–157.
- [15] M. Baggiolini, *Chemokines and leukocyte traffic*, *Nature* **392** (1998), no. 6676, 565–568.
- [16] M. Bajénoff, J.G. Egen, L.Y. Koo, J.P. Laugier, F. Brau, N. Glaichenhaus, and R.N. Germain, *Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes*, *Immunity* **25** (2006), no. 6, 989–1001.
- [17] M. Bajénoff, J.G. Egen, H. Qi, A.Y.C. Huang, F. Castellino, and R.N. Germain, *Highways, byways and breadcrumbs: directing lymphocyte traffic in the lymph node*, *Trends in immunology* **28** (2007), no. 8, 346–352.
- [18] M. Bajénoff, N. Glaichenhaus, and R.N. Germain, *Fibroblastic reticular cells guide T lymphocyte entry into and migration within the splenic T cell zone*, *The Journal of Immunology* **181** (2008), no. 6, 3947–3954.
- [19] P. Baldi, *Exact asymptotics for the probability of exit from a domain and applications to simulation*, *The Annals of Probability* **23** (1995), no. 4, 1644–1670.
- [20] T. Barthlott, H. Moncrieffe, M. Veldhoen, C.J. Atkins, J. Christensen, A. O’Garra, and B. Stockinger, *CD25<sup>+</sup> CD4<sup>+</sup> T cells compete with naive CD4<sup>+</sup> T cells for IL-2 and exploit it for the induction of IL-10 production*, *International immunology* **17** (2005), no. 3, 279–288.
- [21] G. Barton, *Elements of Green’s functions and propagation*, Oxford Science Publications, New York, 1989.

- [22] C. Beauchemin, N.M. Dixit, and A.S. Perelson, *Characterizing T cell movement within lymph nodes in the absence of antigen*, *The Journal of Immunology* **178** (2007), no. 9, 5505–5512.
- [23] M. Beccaria, G. Curci, and A. Viceré, *Numerical solutions of first-exit-time problems*, *Physical review A* **48** (1993), no. 2, 1539–1546.
- [24] G.I. Bell, *Mathematical model of clonal selection and antibody production*, *Journal of theoretical biology* **29** (1970), no. 2, 191–232.
- [25] J.B. Beltman, A.F.M. Marée, and R.J. de Boer, *Analysing immune cell migration*, *Nature Reviews Immunology* **9** (2009), no. 11, 789–798.
- [26] J.B. Beltman, A.F.M. Marée, J.N. Lynch, M.J. Miller, and R.J. De Boer, *Lymph node topology dictates T cell migration behavior*, *The Journal of experimental medicine* **204** (2007), no. 4, 771–780.
- [27] O. Bénichou, C. Chevalier, J. Klafter, B. Meyer, and R. Voituriez, *Geometry-controlled kinetics*, *Nature chemistry* **2** (2010), no. 6, 472–477.
- [28] O. Bénichou, M. Coppey, M. Moreau, PH Suet, and R. Voituriez, *Optimal search strategies for hidden targets*, *Physical review letters* **94** (2005), no. 19, 198101.
- [29] O. Bénichou and J. Desbois, *Exit and occupation times for Brownian motion on graphs with general drift and diffusion constant*, *Journal of Physics A: Mathematical and Theoretical* **42** (2008), no. 1, 015004.
- [30] O. Bénichou, Y. Kafri, M. Sheinman, and R. Voituriez, *Searching fast for a target on DNA without falling to traps*, *Physical review letters* **103** (2009), no. 13, 138102.
- [31] O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez, *Optimizing intermittent reaction paths*, *Phys. Chem. Chem. Phys.* **10** (2008), no. 47, 7059–7072.
- [32] F. Benvenuti, S. Hugues, M. Walmsley, S. Ruf, L. Fetler, M. Popoff, V.L.J. Tybulewicz, and S. Amigorena, *Requirement of Rac1 and Rac2 expression by mature dendritic cells for T cell priming*, *Science Signalling* **305** (2004), no. 5687, 1150.
- [33] H.C. Berg and E.M. Purcell, *Physics of chemoreception*, *Biophysical journal* **20** (1977), no. 2, 193–219.

- [34] J.J. Bird, D.R. Brown, A.C. Mullen, N.H. Moskowitz, M.A. Mahowald, J.R. Sider, T.F. Gajewski, C.R. Wang, and S.L. Reiner, *Helper T cell differentiation is controlled by the cell cycle*, *Immunity* **9** (1998), no. 2, 229–237.
- [35] S. Blanco and R. Fournier, *An invariance property of diffusive random walks*, *EPL (Europhysics Letters)* **61** (2007), no. 2, 168.
- [36] J.M. Blander and R. Medzhitov, *Toll-dependent selection of microbial antigens for presentation by dendritic cells*, *Nature* **440** (2006), no. 7085, 808–812.
- [37] J.N. Blattman, R.m Antia, D.J.D. Sourdive, X. Wang, S.M. Kaech, K. Murali-Krishna, J.D. Altman, and R. Ahmed, *Estimating the precursor frequency of naive antigen-specific CD8 T cells*, *The Journal of experimental medicine* **195** (2002), no. 5, 657–664.
- [38] T. Bopp, C. Becker, M. Klein, S. Klein-Heßling, A. Palmetshofer, E. Serfling, V. Heib, M. Becker, J. Kubach, S. Schmitt, et al., *Cyclic adenosine monophosphate is a key component of regulatory T cell-mediated suppression*, *The Journal of experimental medicine* **204** (2007), no. 6, 1303–1310.
- [39] A.N. Borodin and P. Salminen, *Handbook of Brownian motion-facts and formulae*, Birkhäuser Basel, 2002.
- [40] P. Bousso, *T-cell activation by dendritic cells in the lymph node: lessons from the movies*, *Nat Rev Immunol.* **8** (2008), 675–684.
- [41] P. Bousso, N.R. Bhakta, R.S. Lewis, and E. Robey, *Dynamics of thymocyte-stromal cell interactions visualized by two-photon microscopy*, *Science Signalling* **296** (2002), no. 5574, 1876.
- [42] P. Bousso and E. Robey, *Dynamics of CD8<sup>+</sup> T cell priming by dendritic cells in intact lymph nodes*, *Nature immunology* **4** (2003), no. 6, 579–585.
- [43] B. Breart and P. Bousso, *Cellular orchestration of T cell priming in lymph nodes*, *Current opinion in immunology* **18** (2006), no. 4, 483–490.
- [44] T.M. Brusko, A.L. Putnam, and J.A. Bluestone, *Human regulatory T cells: role in autoimmune disease and therapeutic opportunities*, *Immunological reviews* **223** (2008), no. 1, 371–390.

- [45] K. Burrage and P.M. Burrage, *High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations*, Applied Numerical Mathematics **22** (1996), no. 1, 81–101.
- [46] M.D. Cahalan, I. Parker, S.H. Wei, and M.J. Miller, *Two-photon tissue imaging: seeing the immune system in a fresh light*, Nature reviews. Immunology **2** (2002), no. 11, 872.
- [47] D.J. Campbell, C.H. Kim, E.C. Butcher, et al., *Separable effector T cell populations specialized for B cell help or tissue inflammation*, Nature immunology **2** (2001), no. 9, 876–881.
- [48] X. Cao, S.F. Cai, T.A. Fehniger, J. Song, L.I. Collins, D.R. Piwnica-Worms, and T.J. Ley, *Granzyme B and perforin are important for regulatory T cell-mediated suppression of tumor clearance*, Immunity **27** (2007), no. 4, 635–646.
- [49] J. Carneiro, K. Leon, I. Caramalho, C. Van Den Dool, R. Gardner, V. Oliveira, M. Bergman, N. Sepúlveda, T. Paixao, J. Faro, et al., *When three is not a crowd: a crossregulation model of the dynamics and repertoire selection of regulatory CD4<sup>+</sup> T cells*, Immunological reviews **216** (2007), no. 1, 48–68.
- [50] F. Castellino, A.Y. Huang, G. Altan-Bonnet, S. Stoll, C. Scheinecker, and R.N. Germain, *Chemokines enhance immunity by guiding naive CD8<sup>+</sup>; T cells to sites of CD4<sup>+</sup>; T cell–dendritic cell interaction*, Nature **440** (2006), no. 7086, 890–895.
- [51] D.M. Catron, A.A. Itano, K.A. Pape, D.L. Mueller, and M.K. Jenkins, *Visualizing the first 50 hr of the primary immune response to a soluble antigen*, Immunity **21** (2004), no. 3, 341–347.
- [52] L. Cederbom, H. Hall, and F. Ivars, *CD4<sup>+</sup> CD25<sup>+</sup> regulatory T cells down-regulate co-stimulatory molecules on antigen-presenting cells*, European journal of immunology **30** (2000), no. 6, 1538–1543.
- [53] S. Celli, M. Day, A.J. Müller, C. Molina-Paris, G. Lythe, and P. Bousso, *How many dendritic cells are required to initiate a T-cell response?*, Blood **120** (2012), no. 19, 3945–3948.
- [54] S. Celli, Z. Garcia, and P. Bousso, *CD4 T cells integrate signals delivered during successive DC encounters in vivo*, The Journal of experimental medicine **202** (2005), no. 9, 1271–1278.

- [55] S. Celli, F. Lemaître, and P. Bousso, *Real-time manipulation of T cell-dendritic cell interactions in vivo reveals the importance of prolonged contacts for CD4<sup>+</sup> T cell activation*, *Immunity* **27** (2007), no. 4, 625–634.
- [56] S.B. Charnick and D.A. Lauffenburger, *Mathematical analysis of cell-target encounter rates in three dimensions. effect of chemotaxis*, *Biophysical journal* **57** (1990), no. 5, 1009–1023.
- [57] N. Chen, J.A. Glazier, J.A. Izaguirre, and M.S. Alber, *A parallel implementation of the cellular potts model for simulation of cell-based morphogenesis*, *Computer physics communications* **176** (2007), no. 11, 670–681.
- [58] C. Chevalier, O. Bénichou, B. Meyer, and R. Voituriez, *First-passage quantities of Brownian motion in a bounded domain with multiple targets: a unified approach*, *Journal of Physics A: Mathematical and Theoretical* **44** (2011), no. 2, 025002.
- [59] A.F. Cheviakov, M.J. Ward, and R. Straube, *An asymptotic analysis of the mean first passage time for narrow escape problems: Part II: The sphere*, *Multiscale Modeling & Simulation* **8** (2010), no. 3, 836–870.
- [60] S. Condamin, O. Bénichou, and M. Moreau, *Random walks and Brownian motion: A method of computation for first-passage times and related quantities in confined geometries*, *Physical Review E* **75** (2007), no. 2, 021111.
- [61] S. Condamin, V. Tejedor, R. Voituriez, O. Bénichou, and J. Klafter, *Probing microscopic origins of confined subdiffusion by first-passage observables*, *Proceedings of the National Academy of Sciences* **105** (2008), no. 15, 5675–5680.
- [62] D. Coombs, R. Straube, and M. Ward, *Diffusion on a sphere with localized traps: Mean first passage time, eigenvalue asymptotics, and Fekete points*, *SIAM Journal on Applied Mathematics* **70** (2009), no. 1, 302–332.
- [63] M. Coppey, O. Bénichou, R. Voituriez, and M. Moreau, *Kinetics of target site localization of a protein on DNA: a stochastic approach*, *Biophysical journal* **87** (2004), no. 3, 1640–1649.
- [64] A. Corthay, *How do regulatory T cells work?*, *Scandinavian Journal of Immunology* **70** (2009), 326–336.

- [65] A. Coutinho, S. Hori, T. Carvalho, I. Caramalho, and J. Demengeot, *Regulatory T cells: the physiology of autoreactivity in dominant tolerance and ?quality control? of immune responses*, Immunological reviews **182** (2001), no. 1, 89–98.
- [66] J.G. Cyster, *Chemokines, sphingosine-1-phosphate, and cell migration in secondary lymphoid organs*, Annu. Rev. Immunol. **23** (2005), 127–159.
- [67] J.C. Dallon, *Multiscale modeling of cellular systems in biology*, Current Opinion in Colloid & Interface Science **15** (2010), no. 1, 24–31.
- [68] M. de la Rosa, S. Rutz, H. Dorninger, and A. Scheffold, *Interleukin-2 is essential for CD4<sup>+</sup> CD25<sup>+</sup> regulatory T cell function*, European journal of immunology **34** (2004), no. 9, 2480–2488.
- [69] D.S. Dean and K.M. Jansons, *Brownian excursions on combs*, Journal of statistical physics **70** (1993), no. 5, 1313–1332.
- [70] W. Denk, J.H. Strickler, W.W. Webb, et al., *Two-photon laser scanning fluorescence microscopy*, Science (New York, NY) **248** (1990), no. 4951, 73.
- [71] K. Dietz, *Epidemiologic interference of virus populations*, Journal of mathematical biology **8** (1979), no. 3, 291–300.
- [72] M. Doi and S.F. Edwards, *The theory of polymer dynamics*, vol. 73, Oxford University Press, USA, 1988.
- [73] G.M. Donovan and G. Lythe, *T-cell movement on the reticular network*, Journal of Theoretical Biology (2011).
- [74] M. Faroudi, R. Zaru, P. Paulet, S. Müller, and S. Valitutti, *Cutting edge: T lymphocyte activation by repeated immunological synapse formation and intermittent signaling*, The Journal of Immunology **171** (2003), no. 3, 1128–1132.
- [75] D.T. Fearon, R.M. Locksley, et al., *The instructive role of innate immunity in the acquired immune response*, Science (New York, NY) **272** (1996), no. 5258, 50.
- [76] J.D. Fontenot, M.A. Gavin, and A.Y. Rudensky, *Foxp3 programs the development and function of CD4<sup>+</sup> CD25<sup>+</sup> regulatory T cells*, Nature immunology **4** (2003), no. 4, 330–336.

- [77] K.C. Garcia, L. Teyton, and I.A. Wilson, *Structural basis of T cell recognition*, Annual Review of Immunology **17** (1999), no. 1, 369–397.
- [78] R.N. Germain, *MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation*, Cell **76** (1994), no. 2, 287–299.
- [79] R.N. Germain, M.J. Miller, M.L. Dustin, and M.C. Nussenzweig, *Dynamic imaging of the immune system: progress, pitfalls and promise*, Nature Reviews Immunology **6** (2006), no. 7, 497–507.
- [80] R.N. Germain, E.A. Robey, and M.D. Cahalan, *A decade of imaging cellular motility and interaction dynamics in the immune system*, Science **336** (2012), no. 6089, 1676–1681.
- [81] R.K. Gershon and K. Kondo, *Cell interactions in the induction of tolerance: the role of thymic lymphocytes*, Immunology **18** (1970), no. 5, 723.
- [82] Daniel T Gillespie, *The Monte Carlo method of evaluating integrals*, Tech. report, DTIC Document, 1975.
- [83] D.T. Gillespie, *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*, Journal of computational physics **22** (1976), no. 4, 403–434.
- [84] D.T. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, The journal of physical chemistry **81** (1977), no. 25, 2340–2361.
- [85] L.H. Glimcher and K.M. Murphy, *Lineage commitment in the immune system: the T helper lymphocyte grows up*, Genes & development **14** (2000), no. 14, 1693–1711.
- [86] E. Gobet, *Weak approximation of killed diffusion using Euler schemes*, Stochastic processes and their applications **87** (2000), no. 2, 167–197.
- [87] E. Gobet, *Euler schemes and half-space approximation for the simulation of diffusion in a domain*, ESAIM: Probability and Statistics **5** (2001), no. 1, 261–297.
- [88] M.A. Goldsmith, A. Weiss, et al., *Early signal transduction by the antigen receptor without commitment to T cell activation*, Science (New York, NY) **240** (1988), no. 4855, 1029.
- [89] D.C. Gondek, L. Lu, S.A. Quezada, S. Sakaguchi, and R.J. Noelle, *Cutting edge: contact-mediated suppression by CD4<sup>+</sup> CD25<sup>+</sup> regulatory cells involves a granzyme B-dependent, perforin-independent mechanism*, The Journal of Immunology **174** (2005), no. 4, 1783–1786.

- [90] L.G. Gorostiza, *Rate of convergence of an approximate solution of stochastic differential equations?*, *Stochastics: An International Journal of Probability and Stochastic Processes* **3** (1980), no. 1-4, 267–276.
- [91] E.J. Gretz, C.C. Norbury, A.O. Anderson, A.E.I Proudfoot, and Shaw S., *Lymph-borne chemokines and other low molecular weight molecules reach high endothelial venules via specialized conduits while a functional barrier limits access to the lymphocyte microenvironments in lymphnode cortex*, *J Exp Med.* **192** (2000), 1425 – 1440.
- [92] J.E. Gretz, A.O. Anderson, and S. Shaw, *Cords, channels, corridors and conduits: critical architectural elements facilitating cell interactions in the lymph node cortex*, *Immunological reviews* **156** (1997), no. 1, 11–24.
- [93] A.P. Grigin, *Natural convective diffusion between two nonconcentric spheres*, *Fluid Dynamics* **19** (1984), no. 1, 129–132.
- [94] I.V. Grigoriev, Y.A. Makhnovskii, A.M. Berezhkovskii, and V.Y. Zitserman, *Kinetics of escape through a small hole*, *The Journal of chemical physics* **116** (2002), 9574.
- [95] U. Grohmann, C. Orabona, F. Fallarino, C. Vacca, F. Calcinaro, A. Falorni, P. Candeloro, M.L. Belladonna, R. Bianchi, M.C. Fioretti, et al., *CTLA-4-Ig regulates tryptophan catabolism in vivo*, *Nature immunology* **3** (2002), no. 11, 1097–1101.
- [96] D.L. Groves, W.E. Lever, and T. Makinodan, *Stochastic model for the production of antibody-forming cells*, (1969).
- [97] M.D. Gunn, S. Kyuwa, C. Tam, T. Kakiuchi, A. Matsuzawa, L.T. Williams, and H. Nakano, *Mice lacking expression of secondary lymphoid organ chemokine have defects in lymphocyte homing and dendritic cell localization*, *The Journal of experimental medicine* **189** (1999), no. 3, 451–460.
- [98] M. Gunzer, A. Schäfer, S. Borgmann, S. Grabbe, K.S. Zänker, E.B. Bröcker, E. Kämpgen, and P. Friedl, *Antigen presentation in extracellular matrix: interactions of T cells with dendritic cells are dynamic, short lived, and sequential*, *Immunity* **13** (2000), no. 3, 323–332.
- [99] R. Häggkvist, A. Rosengren, P.H. Lundow, K. Markström, D. Andrén, and P. Kundrotas, *On the ising model for the simple cubic lattice*, *Advances in Physics* **56** (2007), no. 5, 653–755.

- [100] P. Hänggi, P. Talkner, and M. Borkovec, *Reaction-rate theory: fifty years after kramers*, *Reviews of Modern Physics* **62** (1990), no. 2, 251.
- [101] K.I. Happel, P.J. Dubin, M. Zheng, N. Ghilardi, C. Lockhart, L.J. Quinton, A.R. Odden, J.E. Shellito, G.J. Bagby, S. Nelson, et al., *Divergent roles of IL-23 and IL-12 in host defense against Klebsiella pneumoniae*, *The Journal of experimental medicine* **202** (2005), no. 6, 761–769.
- [102] L.E. Harrington, P.R. Mangan, and C.T. Weaver, *Expanding the effector CD4 t-cell repertoire: the Th17 lineage*, *Current opinion in immunology* **18** (2006), no. 3, 349–356.
- [103] B.I. Henry and M.T. Batchelor, *Random walks on finite lattice tubes*, *Physical Review E* **68** (2003), no. 1, 016112.
- [104] R.N. Hiramoto, V.K. Ghanta, J.R. McGhee, and N.M. Hamlin, *A mathematical model for the quantitation of tumor and antibody-forming cell populations within an animal's body*, *The Journal of Immunology* **111** (1973), no. 5, 1546–1553.
- [105] G.W. Hoffmann, *A theory of regulation and self-nonsel self discrimination in an immune network*, *European journal of immunology* **5** (1975), no. 9, 638–647.
- [106] K.A. Hogquist, *Signal strength in thymic selection and lineage commitment*, *Current opinion in immunology* **13** (2001), no. 2, 225–231.
- [107] K.A. Hogquist, T.A. Baldwin, and S.C. Jameson, *Central tolerance: learning self-control in the thymus*, *Nature Reviews Immunology* **5** (2005), no. 10, 772–782.
- [108] A. Hopt and E. Neher, *Highly nonlinear photodamage in two-photon fluorescence microscopy*, *Biophysical Journal* **80** (2001), no. 4, 2029–2036.
- [109] S. Hori, T. Nomura, and S. Sakaguchi, *Control of regulatory T cell development by the transcription factor Foxp3*, *Science Signaling* **299** (2003), no. 5609, 1057.
- [110] S. Hugues, L. Fetler, L. Bonifaz, J. Helft, F. Amblard, and S. Amigorena, *Distinct T cell dynamics in lymph nodes during the induction of tolerance and immunity*, *Nature immunology* **5** (2004), no. 12, 1235–1242.
- [111] S. Hugues, A. Scholer, A. B., A. Nussbaum, C. Combadière, S. Amigorena, and L. Fetler, *Dynamic imaging of chemokine-dependent CD8<sup>+</sup> T cell help for CD8<sup>+</sup> T cell responses*, *Nature immunology* **8** (2007), no. 9, 921–930.

- [112] G. Iezzi, K. Karjalainen, A. Lanzavecchia, et al., *The duration of antigenic stimulation determines the fate of naive and effector T cells*, *Immunity* **8** (1998), no. 1, 89.
- [113] A.A. Itano, M.K. Jenkins, et al., *Antigen presentation to naive CD4 T cells in the lymph node*, *Nature immunology* **4** (2003), no. 8, 733–739.
- [114] M.J. Jacobs, A.E. Van Den Hoek, L.B. Van De Putte, and W.B. Van den Berg, *Induction of tolerance by T-cell vaccination is possible beyond the area of autoimmunity: down-regulation of immunity directed to foreign protein antigens*, *Immunology* **80** (1993), no. 4, 511.
- [115] C.A. Janeway Jr and R. Medzhitov, *Innate immune recognition*, *Science Signalling* **20** (2002), no. 1, 197.
- [116] K.M. Jansons and G.D. Lythe, *Stochastic calculus: Application to dynamic bifurcations and threshold crossings*, *Journal of statistical physics* **90** (1998), no. 1, 227–251.
- [117] K.M. Jansons and GD Lythe, *Efficient numerical solution of stochastic differential equations using exponential timestepping*, *Journal of Statistical Physics* **100** (2000), no. 5, 1097–1109.
- [118] K.M. Jansons and G.D. Lythe, *Exponential timestepping with boundary test for stochastic differential equations*, *SIAM Journal on Scientific Computing* **24** (2003), no. 5, 1809–1822.
- [119] K.M. Jansons and G.D. Lythe, *Multidimensional exponential timestepping with boundary test*, *SIAM J. Sci. Comput.* **27** (2005), no. 3, 793–808.
- [120] K.M. Jansons and L.C.G. Rogers, *Probability and dispersion theory*, *IMA journal of applied mathematics* **55** (1995), no. 2, 149–162.
- [121] E.P. Kaldjian, J.E. Gretz, A.O. Anderson, Y. Shi, and S. Shaw, *Spatial and molecular organization of lymph node T cell cortex: a labyrinthine cavity bounded by an epithelium-like monolayer of fibroblastic reticular cells anchored to basement membrane-like extracellular matrix*, *International immunology* **13** (2001), no. 10, 1243–1253.
- [122] I. Karatzas and S.E. Shreve, *Brownian motion and stochastic calculus*, vol. 113, Springer Verlag, 1991.

- [123] T. Katakai, T. Hara, M. Sugai, H. Gonda, and A. Shimizu, *Lymph node fibroblastic reticular cells construct the stromal reticulum via contact with lymphocytes*, The Journal of experimental medicine **200** (2004), no. 6, 783–795.
- [124] Q.J.A. Khan and D. Greenhalgh, *Hopf bifurcation in epidemic models with a time delay in vaccination*, Mathematical Medicine and Biology **16** (1999), no. 2, 113–142.
- [125] R. Khattri, T. Cox, S.A. Yasayko, and F. Ramsdell, *An essential role for scurfin in CD4<sup>+</sup> CD25<sup>+</sup> T regulatory cells*, Nature immunology **4** (2003), no. 4, 337–342.
- [126] Kiyosi and Itô, *Foundations of stochastic differential equations in infinite dimensional spaces*, vol. 47, SIAM, 1984.
- [127] L. Klein, K. Khazaie, and H. von Boehmer, *In vivo dynamics of antigen-specific regulatory T cells not predicted from behavior in vitro*, Proceedings of the National Academy of Sciences **100** (2003), no. 15, 8886–8891.
- [128] P.E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, vol. 23, Springer, 2011.
- [129] T. Lämmermann and M. Sixt, *The microanatomy of T-cell responses*, Immunological reviews **221** (2008), no. 1, 26–43.
- [130] C.L. Langrish, Y. Chen, W.M. Blumenschein, J. Mattson, B. Basham, J.D. Sedgwick, T. McClanahan, R.A. Kastelein, and D.J. Cua, *IL-23 drives a pathogenic T cell population that induces autoimmune inflammation*, The Journal of experimental medicine **201** (2005), no. 2, 233–240.
- [131] A. Lanzavecchia, G. Lezzi, A. Viola, et al., *From TCR engagement to T cell activation: a kinetic view of T cell behavior*, Cell **96** (1999), no. 1, 1.
- [132] E.W. Larson, J.W. Dominik, A.H. Rowberg, and G.A. Higbee, *Influenza virus population dynamics in the respiratory tract of experimentally infected mice*, Infection and immunity **13** (1976), no. 2, 438–447.
- [133] K. Leon, R. Perez, A. Lage, and J. Carneiro, *Modelling T-cell-mediated suppression dependent on interactions in multicellular conjugates*, Journal of Theoretical Biology **207** (2000), no. 2, 231–254.

- [134] M.I. Levi, K.V. Durikhin, N.N. Basova, L.M. Shmuter, Y.G. Suchkov, A.V. Lipnitskii, and L.G. Gerasyuk, *Cyclic kinetics and mathematical expression of the primary immune response to soluble antigen*, *Folia microbiologica* **18** (1973), no. 3, 229–236.
- [135] A. Link, T.K. Vogt, S. Favre, M.R. Britschgi, H. Acha-Orbea, B. Hinz, J.G. Cyster, and S.A. Luther, *Fibroblastic reticular cells in lymph nodes regulate the homeostasis of naive T cells*, *Nature immunology* **8** (2007), no. 11, 1255–1265.
- [136] G. Lombardi, S. Sidhu, R. Batchelor, and R. Lechler, *Anergic T cells as suppressor cells in vitro*, *Science* **264** (1994), no. 5165, 1587–1589.
- [137] P.E. Love and A.C. Chan, *Regulation of thymocyte development: only the meek survive*, *Current opinion in immunology* **15** (2003), no. 2, 199–203.
- [138] C. Loverdo, O. Benichou, R. Voituriez, A. Biebricher, I. Bonnet, and P. Desbiolles, *Quantifying hopping and jumping in facilitated diffusion of DNA-binding proteins*, *Physical review letters* **102** (2009), no. 18, 188101.
- [139] A.D. Luster, R. Alon, and U.H. Von Andrian, *Immune cell migration in inflammation: present and future therapeutic targets*, *Nature immunology* **6** (2005), no. 12, 1182–1190.
- [140] T.R. Malek, A. Yu, L. Zhu, T. Matsutani, D. Adeegbe, and A.L. Bayer, *IL-2 family of cytokines in T regulatory cell development and homeostasis*, *Journal of clinical immunology* **28** (2008), no. 6, 635–639.
- [141] K.J. Maloy and F. Powrie, *Regulatory T cells in the control of immune pathology*, *Nature immunology* **2** (2001), no. 9, 816–822.
- [142] P.R. Mangan, L.E. Harrington, D.B. O’Quinn, W.S. Helms, D.C. Bullard, C.O. Elson, R.D. Hatton, S.M. Wahl, T.R. Schoeb, and C.T. Weaver, *Transforming growth factor- $\beta$  induces development of the TH17 lineage*, *Nature* **441** (2006), no. 7090, 231–234.
- [143] R. Mannella, *Numerical tricks for studying large rare fluctuations*, *Le Journal de Physique IV* **8** (1998), no. PR6.
- [144] R. Mannella, *Absorbing boundaries and optimal stopping in a stochastic differential equation*, *Physics Letters A* **254** (1999), no. 5, 257–262.
- [145] J.J. Marchalonis and V.X. Gledhill, *Elementary stochastic model for the induction of immunity and tolerance*, (1968).

- [146] A. Marson, K. Kretschmer, G.M. Frampton, E.S. Jacobsen, J.K. Polansky, K.D. MacIsaac, S.S. Levine, E. Fraenkel, H. von Boehmer, and R.A. Young, *Foxp3 occupancy and regulation of key target genes during T-cell stimulation*, *Nature* **445** (2007), no. 7130, 931–935.
- [147] M. Matloubian, C.G. Lo, G. Cinamon, M.J. Lesneski, Y. Xu, V. Brinkmann, M.L. Allende, R.L. Proia, and J.G. Cyster, *Lymphocyte egress from thymus and peripheral lymphoid organs is dependent on SIP receptor 1*, *Nature* **427** (2004), no. 6972, 355–360.
- [148] C.L. Maynard, L.E. Harrington, K.M. Janowski, J.R. Oliver, C.L. Zindl, A.Y. Rudensky, and C.T. Weaver, *Regulatory T cells expressing interleukin 10 develop from Foxp3+ and Foxp3- precursor cells in the absence of interleukin 10*, *Nature immunology* **8** (2007), no. 9, 931–941.
- [149] A. Mazzolo, *Properties of diffusive random walks in bounded domains*, *EPL (Europhysics Letters)* **68** (2007), no. 3, 350.
- [150] A.R. McLean and S.M. Blower, *Modelling HIV vaccination*, *Trends in microbiology* **3** (1995), no. 12, 458–463.
- [151] J.L. Melsa and A.P. Sage, *An introduction to probability and stochastic processes*, Prentice-Hall Englewood Cliffs, 1973.
- [152] T.R. Mempel, S.E. Henrickson, and U.H. von Andrian, *T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases*, *Nature* **427** (2004), no. 6970, 154–159.
- [153] M.J. Miller, A.S. Hejazi, S.H. Wei, M.D. Cahalan, and I. Parker, *T cell repertoire scanning is promoted by dynamic dendritic cell behavior and random T cell motility in the lymph node*, *Proceedings of the National Academy of Sciences of the United States of America* **101** (2004), no. 4, 998–1003.
- [154] M.J. Miller, O. Safrina, I. Parker, and M.D. Cahalan, *Imaging the single cell dynamics of CD4<sup>+</sup> T cell activation by dendritic cells in lymph nodes*, *The Journal of experimental medicine* **200** (2004), no. 7, 847–856.
- [155] M.J. Miller, S.H. Wei, M.D. Cahalan, and I. Parker, *Autonomous T cell trafficking examined in vivo with intravital two-photon microscopy*, *Proceedings of the National Academy of Sciences* **100** (2003), no. 5, 2604–2609.

- [156] M.J. Miller, S.H. Wei, I. Parker, and M.D. Cahalan, *Two-photon imaging of lymphocyte motility and antigen response in intact lymph node*, *Science Signaling* **296** (2002), no. 5574, 1869.
- [157] G.N. Milstein and M.V. Tretyakov, *Simulation of a space-time bounded diffusion*, *Annals of Applied Probability* (1999), 732–779.
- [158] G. Moller, *Origin of major histocompatibility complex diversity*, *Immunology Review* **143** (1995), 5 – 292.
- [159] J.J. Moon, H.H. Chu, M. Pepper, S.J. McSorley, S.C. Jameson, R.M. Kedl, and M.K. Jenkins, *Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude*, *Immunity* **27** (2007), no. 2, 203–213.
- [160] J.R. Mora and U.H. von Andrian, *T-cell homing specificity and plasticity: new concepts and future challenges*, *Trends in immunology* **27** (2006), no. 5, 235–243.
- [161] J. Mossong and C.P. Muller, *Modelling measles re-emergence as a result of waning of immunity in vaccinated populations*, *Vaccine* **21** (2003), no. 31, 4597–4603.
- [162] C. M’Rini, G. Cheng, C. Schweitzer, L.L. Cavanagh, R.T. Palframan, T.R. Mempel, R.A. Warnock, J.B. Lowe, E.J. Quackenbush, and U.H. Von Andrian, *A novel endothelial L-selectin ligand activity in lymph node medulla that is regulated by  $\alpha$  (1, 3)-fucosyltransferase-IV*, *The Journal of experimental medicine* **198** (2003), no. 9, 1301–1312.
- [163] Kenneth M Murphy, *Janeway’s immunobiology 8th ed.*, Garland Science, 2012.
- [164] G.D. Musters, R.P. Kleipool, S. Bipat, and M. Maas, *Features of the popliteal lymph nodes seen on musculoskeletal MRI in a western population*, *Skeletal radiology* **40** (2011), no. 8, 1041–1045.
- [165] G. Napolitani, A. Rinaldi, F. Bertoni, F. Sallusto, and A. Lanzavecchia, *Selected Toll-like receptor agonist combinations synergistically trigger a T helper type 1–polarizing program in dendritic cells*, *Nature immunology* **6** (2005), no. 8, 769–776.
- [166] S. Nechaev and R. Voituriez, *On the plant leaf’s boundary, jupe à godets’ and conformal embeddings*, *Journal of Physics A: Mathematical and General* **34** (2001), no. 49, 11069.

- [167] M.E.J. Newman and G.T. Barkema, *Monte Carlo methods in statistical physics*, Monte Carlo methods in statistical physics/MEJ Newman and GT Barkema. Oxford: Clarendon Press, 1999. **1** (1999).
- [168] N.J. Newton, *An efficient approximation for stochastic differential equations on the partition of symmetrical first passage times*, *Stochastics: An International Journal of Probability and Stochastic Processes* **29** (1990), no. 2, 227–258.
- [169] J.D. Noh and H. Rieger, *Random walks on complex networks*, *Physical review letters* **92** (2004), no. 11, 118701.
- [170] M.A Nowak and A.R. Mclean, *A mathematical model of vaccination against HIV prevent the development of AIDS*, *Proceedings of the Royal Society of London. Series B: Biological Sciences* **246** (1991), no. 1316, 141–146.
- [171] R. Obst, H.M. Van Santen, D. Mathis, and C. Benoist, *Antigen persistence is required throughout the expansion phase of a CD4<sup>+</sup> T cell response*, *The Journal of experimental medicine* **201** (2005), no. 10, 1555–1565.
- [172] C. Oderup, L. Cederbom, A. Makowska, C.M. Cilio, and F. Ivars, *Cytotoxic T lymphocyte antigen-4-dependent down-modulation of costimulatory molecules on dendritic cells in CD4<sup>+</sup> CD25<sup>+</sup> regulatory T-cell mediated suppression*, *Immunology* **118** (2006), no. 2, 240–249.
- [173] T. Okada, M.J. Miller, I. Parker, M.F. Krummel, M. Neighbors, S.B. Hartley, A. O'Garra, M.D. Cahalan, and J.G. Cyster, *Antigen-engaged B cells undergo chemotaxis toward the T zone and form motile conjugates with helper T cells*, *PLoS biology* **3** (2005), no. 6, e150.
- [174] M. Ono, H. Yaguchi, N. Ohkura, I. Kitabayashi, Y. Nagamura, T. Nomura, Y. Miyachi, T. Tsukada, and S. Sakaguchi, *Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1*, *Nature* **446** (2007), no. 7136, 685–689.
- [175] E. Palmer, *Negative selection? clearing out the bad apples from the T-cell repertoire*, *Nature Reviews Immunology* **3** (2003), no. 5, 383–391.
- [176] P. Pandiyan, L. Zheng, S. Ishihara, J. Reed, and M.J. Lenardo, *CD4<sup>+</sup> CD25<sup>+</sup> Foxp3<sup>+</sup> regulatory T cells induce cytokine deprivation–mediated apoptosis of effector CD4<sup>+</sup> T cells*, *Nature immunology* **8** (2007), no. 12, 1353–1362.

- [177] A.S. Perelson, M. Mirmirani, and G.F. Oster, *Optimal strategies in immunology*, Journal of Mathematical Biology **3** (1976), no. 3-4, 325–367.
- [178] A.S. Perelson and F.W. Wiegel, *Theoretical considerations of the role of antigen structure in B cell activation*, Federation proceedings, vol. 40, 1981, p. 1479.
- [179] C.A. Piccirillo and E.M. Shevach, *Naturally-occurring CD4<sup>+</sup> CD25<sup>+</sup> immunoregulatory T cells: central players in the arena of peripheral tolerance*, Seminars in immunology, vol. 16, Elsevier, 2004, pp. 81–88.
- [180] L.J. Picker and E.C. Butcher, *Physiological and molecular mechanisms of lymphocyte homing*, Annual review of immunology **10** (1992), no. 1, 561–591.
- [181] S. Pillay, M.J. Ward, A. Peirce, and T. Kolokolnikov, *An asymptotic analysis of the mean first passage time for narrow escape problems: Part i: Two-dimensional domains*, Multiscale Modeling & Simulation **8** (2010), no. 3, 803–835.
- [182] E. Platen, *An introduction to numerical methods for stochastic differential equations*, Acta numerica **8** (1999), no. 1, 197–246.
- [183] S. Read, S. Mauze, C. Asseman, A. Bean, R. Coffman, and F. Powrie, *CD38<sup>+</sup> CD45RBlow CD4<sup>+</sup> T cells: a population of T cells with immune regulatory activities in vitro*, European journal of immunology **28** (1998), no. 11, 3435–3447.
- [184] S. Redner, *A guide to first-passage processes.*, Cambridge University Press, 2001.
- [185] R.L. Reinhardt, S. Hong, S.J. Kang, Z. Wang, and R.M. Locksley, *Visualization of IL-12/23p40 in vivo reveals immunostimulatory dendritic cell migrants that promote Th1 differentiation*, The Journal of Immunology **177** (2006), no. 3, 1618–1627.
- [186] C. Reis e Sousa, *Activation of dendritic cells: translating innate into adaptive immunity*, Current opinion in immunology **16** (2004), no. 1, 21–25.
- [187] S.A. Rice, *Diffusion-limited reactions*, vol. 25, Elsevier Science, 1985.
- [188] H. Risken, *The Fokker-Planck equation: Method of solution and applications*, Springer-Verlag, Berlin, 1996.
- [189] L.C.G Rogers and D.W. Diffusions, *Markov processes and martingales: Volume 2, ito calculus, J*, Wiley & sons, 1987.

- [190] M. Rubinstein and R. Colby, *Polymer physics (Chemistry)*, Oxford University Press, USA, 2003.
- [191] S. Sakaguchi, *Naturally arising CD4<sup>+</sup> regulatory T cells for immunologic self-tolerance and negative control of immune responses*, *Annu. Rev. Immunol.* **22** (2004), 531–562.
- [192] S. Sakaguchi, M. Ono, R. Setoguchi, H. Yagi, S. Hori, Z. Fehervari, J. Shimizu, T. Takahashi, and T. Nomura, *Foxp3<sup>+</sup> CD25<sup>+</sup> CD4<sup>+</sup> natural regulatory T cells in dominant self-tolerance and autoimmune disease*, *Immunological reviews* **212** (2006), no. 1, 8–27.
- [193] S. Sakaguchi, N. Sakaguchi, M. Asano, M. Itoh, and M. Toda, *Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). breakdown of a single mechanism of self-tolerance causes various autoimmune diseases*, *The Journal of Immunology* **155** (1995), no. 3, 1151–1164.
- [194] S. Sakaguchi, N. Sakaguchi, J. Shimizu, S. Yamazaki, T. Sakihama, M. Itoh, Y. Kuniyasu, T. Nomura, M. Toda, and T. Takahashi, *Immunologic tolerance maintained by CD25<sup>+</sup> CD4<sup>+</sup> regulatory T cells: their common role in controlling autoimmunity, tumor immunity, and transplantation tolerance*, *Immunological reviews* **182** (2001), no. 1, 18–32.
- [195] S. Sakaguchi, K. Wing, Y. Onishi, P. Prieto-Martin, and T. Yamaguchi, *Regulatory T cells: how do they suppress immune responses?*, *International immunology* **21** (2009), no. 10, 1105–1111.
- [196] S.L. Sanos, J. Nowak, M. Fallet, and M. Bajenoff, *Stromal cell networks regulate thymocyte migration and dendritic cell behavior in the thymus*, *The Journal of Immunology* **186** (2011), no. 5, 2835–2841.
- [197] D.M. Sansom, C.N. Manzotti, and Y. Zheng, *What's the difference between CD80 and CD86?*, *Trends in immunology* **24** (2003), no. 6, 313–318.
- [198] A. Scheffold, K.M. Murphy, and T. Höfer, *Competition for cytokines: Treg cells take all*, *Nature immunology* **8** (2007), no. 12, 1285–1287.
- [199] J. Scheffold, A. and Hühn and T. Höfer, *Regulation of CD4<sup>+</sup> CD25<sup>+</sup> regulatory T cell activity: it takes (IL-) two to tango*, *European journal of immunology* **35** (2005), no. 5, 1336–1341.

- [200] E. Sebzda, S. Mariathasan, T. Ohteki, R. Jones, M.F. Bachmann, and P.S. Ohashi, *Selection of the T cell repertoire*, Annual review of immunology **17** (1999), no. 1, 829–874.
- [201] G. Shakhar, R.L. Lindquist, D. Skokos, D. Dudziak, J.H. Huang, M.C. Nussenzweig, and M.L. Dustin, *Stable T cell–dendritic cell interactions precede the development of both tolerance and immunity in vivo*, Nature immunology **6** (2005), no. 7, 707–714.
- [202] E.M. Shevach, *Regulatory T cells in autoimmunity\**, Annual review of immunology **18** (2000), no. 1, 423–449.
- [203] E.M. Shevach, *CD4<sup>+</sup> CD25<sup>+</sup> suppressor T cells: more questions than answers*, Nature Reviews Immunology **2** (2002), no. 6, 389–400.
- [204] E.M. Shevach, *Mechanisms of Foxp3<sup>+</sup> T regulatory cell-mediated suppression*, Immunity **30** (2009), no. 5, 636–645.
- [205] A. Singer, Z. Schuss, and D. Holcman, *Narrow escape, part II: The circular disk*, Journal of statistical physics **122** (2006), no. 3, 465–489.
- [206] A. Singer, Z. Schuss, and D. Holcman, *Narrow escape, part III: Non-smooth domains and riemann surfaces*, Journal of statistical physics **122** (2006), no. 3, 491–509.
- [207] A. Singer, Z. Schuss, D. Holcman, and R.S. Eisenberg, *narrow escape, part I*, Journal of statistical physics **122** (2006), no. 3, 437–463.
- [208] M. Sixt, N. Kanazawa, M. Selg, T. Samson, G. Roos, D.P. Reinhardt, R. Pabst, M.B. Lutz, and L. Sorokin, *The conduit system transports soluble antigens from the afferent lymph to resident dendritic cells in the T cell area of the lymph node*, Immunity **22** (2005), no. 1, 19–29.
- [209] J. Sprent and H. Kishimoto, *The thymus and central tolerance*, Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **356** (2001), no. 1409, 609–616.
- [210] J. Sprent, J.F.A.P. Miller, and G.F. Mitchell, *Antigen-induced selective recruitment of circulating lymphocytes*, Cellular immunology **2** (1971), no. 2, 171–181.
- [211] T.K. Starr, S.C. Jameson, and K.A. Hogquist, *Positive and negative selection of T cells*, Annual review of immunology **21** (2003), no. 1, 139–176.

- [212] S. Stoll, J. Delon, T.M. Brotz, and R.N. Germain, *Dynamic imaging of T cell-dendritic cell interactions in lymph nodes*, *Science Signalling* **296** (2002), no. 5574, 1873.
- [213] W.A. Strauss, *Partial differential equations: An introduction*, vol. 3, Wiley New York, 1992.
- [214] S.J. Szabo, B.M. Sullivan, S.L. Peng, and L.H. Glimcher, *Molecular mechanisms regulating Th1 immune responses*, *Annual review of immunology* **21** (2003), no. 1, 713–758.
- [215] L.S. Taams, E.P.J. Boot, W. van Eden, and M.H.M. Wauben, *'anergic' T cells modulate the T-cell activating capacity of antigen-presenting cells*, *Journal of autoimmunity* **14** (2000), no. 4, 335–341.
- [216] L.S. Taams, A.J.M.L. van Rensen, M. Poelen, C.A.C.M. van Els, A.C. Besseling, J. Wagenaar, W. van Eden, and M.H.M. Wauben, *Anergic t cells actively suppress T cell responses via the antigen-presenting cell*, *European journal of immunology* **28** (1998), no. 9, 2902–2912.
- [217] T. Takahashi, Y. Kuniyasu, M. Toda, N. Sakaguchi, M. Itoh, M. Iwata, J. Shimizu, and S. Sakaguchi, *Immunologic self-tolerance maintained by CD25<sup>+</sup> CD4<sup>+</sup> naturally anergic and suppressive T cells: induction of autoimmune disease by breaking their anergic/suppressive state*, *International immunology* **10** (1998), no. 12, 1969–1980.
- [218] C. Tanchot, F. Vasseur, C. Pontoux, C. Garcia, and A. Sarukhan, *Immune regulation by self-reactive T cells is antigen specific*, *The Journal of Immunology* **172** (2004), no. 7, 4285–4291.
- [219] Q. Tang, J.Y. Adams, A.J. Tooley, M. Bi, B.T. Fife, P. Serra, P. Santamaria, R.M. Locksley, M.F. Krummel, and J.A. Bluestone, *Visualizing regulatory T cell control of autoimmune responses in nonobese diabetic mice*, *Nature immunology* **7** (2005), no. 1, 83–92.
- [220] C.M. Tato and J.J. O'Shea, *Immunology: what does it mean to be just 17?*, *Nature* **441** (2006), no. 7090, 166–168.
- [221] H.M. Taylor and S. Karlin, *An introduction to stochastic modeling*, Academic Press, 1998.
- [222] V. Tejedor, O. Bénichou, and R. Voituriez, *Global mean first-passage times of random walks on complex networks*, *Physical Review E* **80** (2009), no. 6, 065104.

- [223] A.M. Thornton and E.M. Shevach, *CD4<sup>+</sup> CD25<sup>+</sup> immunoregulatory T cells suppress polyclonal T cell activation in vitro by inhibiting interleukin 2 production*, The Journal of experimental medicine **188** (1998), no. 2, 287–296.
- [224] A.M. Thornton and E.M. Shevach, *Suppressor effector function of CD4<sup>+</sup> CD25<sup>+</sup> immunoregulatory T cells is antigen nonspecific*, The Journal of Immunology **164** (2000), no. 1, 183–190.
- [225] T. Ushiki, O. Ohtani, and K. Abe, *Scanning electron microscopic studies of reticular framework in the rat mesenteric lymph node*, The Anatomical Record **241** (1995), no. 1, 113–122.
- [226] N.G. Van Kampen, *Stochastic processes in physics and chemistry*, vol. 1, North holland, 1992.
- [227] D. Voehringer, T.A. Reese, X. Huang, K. Shinkai, and R.M. Locksley, *Type 2 immunity is controlled by IL-4/IL-13 expression in hematopoietic non-eosinophil cells of the innate immune system*, The Journal of experimental medicine **203** (2006), no. 6, 1435–1446.
- [228] U.H. von Andrian, T.R. Mempel, et al., *Homing and cellular traffic in lymph nodes*, Nature Reviews Immunology **3** (2003), no. 11, 867–878.
- [229] H. Waldmann and S. Cobbold, *How do monoclonal antibodies induce tolerance? a role for infectious tolerance?*, Annual review of immunology **16** (1998), no. 1, 619–644.
- [230] M.J. Ward and J.B. Keller, *Strong localized perturbations of eigenvalue problems*, SIAM Journal on Applied Mathematics **53** (1993), no. 3, 770–798.
- [231] J. Wayne Streilein and C.B. Read, *A mathematical model in immunoregulation where the spleen has a pivotal role*, Journal of Theoretical Biology **61** (1976), no. 2, 363–376.
- [232] S.H. Wei, I. Parker, M.J. Miller, and M.D. Cahalan, *A stochastic view of lymphocyte motility and trafficking within the lymph node*, Immunological reviews **195** (2003), no. 1, 136–159.
- [233] H.F. Weinberger, *A first course in partial differential equations: with complex variables and transform methods*, Courier Dover Publications, 2012.
- [234] S. Wieser, G.J. Schütz, et al., *Tracking single molecules in the live cell plasma membrane Do's and Dont's*, Methods **46** (2008), no. 2, 131–140.

- [235] N.S. Wilson, D. El-Sukkari, G.T. Belz, C.M. Smith, R.J. Steptoe, W.R. Heath, K. Shortman, and J.A. Villadangos, *Most lymphoid organ dendritic cell types are phenotypically and functionally immature*, *Blood* **102** (2003), no. 6, 2187–2194.
- [236] Y. Wu, M. Borde, V. Heissmeyer, M. Feuerer, A.D. Lapan, J.C. Stroud, D.L. Bates, L. Guo, A. Han, S.F. Ziegler, et al., *FOXP3 controls regulatory T cell function through cooperation with NFAT*, *Cell* **126** (2006), no. 2, 375–387.
- [237] P.L. Yong, P. Russo, and K.E. Sullivan, *Use of sirolimus in IPEX and IPEX-like children*, *Journal of clinical immunology* **28** (2008), no. 5, 581–587.
- [238] M. Yoshino, H. Yamazaki, H. Nakano, T. Kakiuchi, K. Ryoke, T. Kunisada, and S.I. Hayashi, *Distinct antigen trafficking from skin in the steady and active states*, *International immunology* **15** (2003), no. 6, 773–779.
- [239] P. Yu, R.K. Gregg, J.J. Bell, J.S. Ellis, R. Divekar, H. Lee, R. Jain, H. Waldner, J.C. Hardaway, M. Collins, et al., *Specific T regulatory cells display broad suppressive functions against experimental allergic encephalomyelitis upon activation with cognate antigen*, *The Journal of Immunology* **174** (2005), no. 11, 6772–6780.
- [240] Y. Zheng, S.Z. Josefowicz, A. Kas, T. Chu, M.A. Gavin, and A.Y. Rudensky, *Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells*, *Nature* **445** (2007), no. 7130, 936–940.
- [241] R.M. Zinkernagel, S. Ehl, P. Aichele, S. Oehen, T. Kündig, and H. Hengartner, *Antigen localisation regulates immune responses in a dose-and time-dependent fashion: a geographical view of immune reactivity*, *Immunological reviews* **156** (1997), no. 1, 199–209.