University of New Mexico
UNM Digital Repository

Computer Science ETDs | Engineering ETDs

Spring 5-13-2017

Search in T cell and Robot Swarms: Balancing Extent and Intensity

George M. Fricke
University of New Mexico

Follow this and additional works at: https://digitalrepository.unm.edu/cs_etds

 Part of the Artificial Intelligence and Robotics Commons

Recommended Citation
Fricke, George M.. "Search in T cell and Robot Swarms: Balancing Extent and Intensity." (2017). https://digitalrepository.unm.edu/cs_etds/83

This Dissertation is brought to you for free and open access by the Engineering ETDs at UNM Digital Repository. It has been accepted for inclusion in Computer Science ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.

George Matthew Fricke

Candidate

Computer Science

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Melanie E. Moses, Chairperson

Judy L. Cannon

David H. Ackley

Helen J. Wearing

Search in T cell and Robot Swarms: Balancing Extent and Intensity

by

George Matthew Fricke

B.A., Anthropology, Appalachian State University, 1996
M.S., Computer Science, University of New Mexico, 2003
B.S., Mathematics, University of New Mexico, 2012

DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Computer Science

The University of New Mexico

Albuquerque, New Mexico

February 8th, 2017

©2017, George Matthew Fricke

Permission has been granted by the respective copyright holders for the reproduction of previously published material in this dissertation. Please see Appendix C for documentation.

Dedication

To my wife, Suzanne, for her support, encouragement, and patience; to my children Henry, Leo, Owen, and Tristan for reminding me that robots are cool; and to my parents for teaching me that there is so much to discover about the world.

'There are some four million different kinds of animals and plants in the world. Four million different solutions to the problems of staying alive.'
— David Attenborough

Acknowledgments

Thank you to my advisor, Melanie Moses, for her support and the valuable lessons she has taught me. Melanie has allowed me the flexibility to explore all aspects of my research without compartmentalisation. I have recorded the behaviour of foraging ants in the New Mexico desert, seen T cells searching the lymph nodes of mice via laser light, written software to control NASA robots designed to be a first step in swarm exploration of Mars, and learned some of the maths underlying the physics of atomic motion that govern particle interaction rates. I believe that this global view of search and motion provides a unique perspective on the problem of designing spacial search processes and understanding natural search systems. Thank you Melanie for allowing me this breadth of investigation.

Thanks to Kenneth Letendre who took the very simple, and slow, ant foraging simulation I wrote following our field observations, incorporated a genetic algorithm to optimise the search parameters and produced an ant inspired search algorithm. His work demonstrated to me that biologically inspired algorithms can be effective, and how they improve our understanding of the biological system from which we borrow.

Thanks also to Joshua Hecker for creating robot ants. Without his work, along with Karl Stollis, Jake Nichol, and Kurt Leught, developing the iAnt and the NASA Swarmie, all my work would be confined to simulation. Seeing algorithms enacted in simulated worlds is rewarding, but seeing the result in the physical world, and making the physical/virtual barrier ever more permeable, is thrilling.

Thank you to the undergraduates I have had the pleasure of working with: Justyna Tafoya, Antonio Griego, and Linh Tran. They have consistently exceeded my expectations. Justyna has been invaluable in helping me think through the analysis of T cell-DC interactions. Antonio was the first to learn the ARGoS simulation environment and has become the go-to person for ARGoS implementation questions. Antonio also helped implement the gripper control plugin for Gazebo. Linh has worked with me almost from the beginning on the DDSA algorithm and it has been a pleasure to watch her give talks and posters on the subject.

Thanks to Janie Byrum and François Asperti-Bousin for imaging of T cells and DCs, and Sarah Black for helping to write the hotspot analysis code. Their hard work and technical expertise has been essential in gathering the data we need to analyse T cell search strategies.

Finally, thanks to my dissertation committee, Melanie Moses, Judy Cannon, Dave Ackley and Helen Wearing for volunteering their time to evaluate this work and provide suggestions for improvement.

Search in T cell and Robot Swarms: Balancing Extent and Intensity

by

George Matthew Fricke

B.A., Anthropology, Appalachian State University, 1996
M.S., Computer Science, University of New Mexico, 2003
B.S., Mathematics, University of New Mexico, 2012
Ph.D., Computer Science, University of New Mexico, 2017

Abstract

This work investigates effective search and resource collection algorithms for swarms. Deterministic spiral algorithms and Lévy search processes have been shown to be optimal for single searchers (Baeza-Yates et al., 1993; Viswanathan et al., 1996). By generalising these approaches to swarms and measuring the effectiveness of the resulting search patterns in computer models, we find that the intensity-extent trade-off, formalised as the fractal dimension of the search pattern, can be used to adapt search to common challenges in swarm search.

Search extent and intensity lie on a continuum: more intensive patterns search thoroughly in the local area, while extensive patterns cover more area but may miss targets nearby. We show that the most efficient trade-off between search intensity and extent for swarms depends strongly on the distribution of targets, swarm size

and the rate of collision among searchers (Fricke et al., 2016a). The optimal trade-off is also influenced by the target detection error rate. The search can, therefore, be tuned to match conditions common in real-world robot search tasks.

We also demonstrate that our swarm spiral algorithm is an effective strategy for resource collection (Fricke et al., 2016b). Deterministic spiral search strategies for single searchers have been considered unsuitable in the presence of localisation error (Reynolds et al., 2007), but the swarm algorithm performs well even in the presence of localisation error. Since the spiral strategy is effective and easily analysed it makes an ideal benchmark against which to compare stochastic search processes.

Collective search by T cells is a critical component of the adaptive immune response. We characterise T cell search patterns and find that they balance the need to search extensively for rare antigen while maintaining local contacts with antigen-presenting cells. We perform two analyses that demonstrate that T cells interact with their environment during search (Fricke et al., 2013, 2015, 2016c). We also measure the interaction between T cells and Dendritic cells using mutual information and demonstrate non-random spatial association between T cells and their targets.

Contents

List of Figures **xiii**

List of Tables **xvi**

Glossary **xvii**

1 Introduction **1**
1.1 Lost Key Analogy . . . 2
1.2 Relationship to Breadth-First and Depth-First search . . . 4
1.3 Spatial Search in Two Swarm Systems . . . 5
1.3.1 Search in Immunology . . . 6
1.3.2 Search and Foraging in Swarm Robotics . . . 7
1.4 The Role of Simulation . . . 8
1.5 Organisation and Contributions . . . 9

2 Background **13**
2.1 Deterministic Search . . . 15
2.2 Stochastic Search . . . 16
2.3 Lévy Search . . . 17
2.4 Hausdorff Fractal Dimension . . . 19
2.5 Mean Squared Displacement . . . 20
2.6 Brownian Motion . . . 21

3 Analysis of T cell Search in Lymph Nodes **23**
3.1 Author Contribution Statement . . . 23
3.2 Publication Notes . . . 24
3.3 Author Summary . . . 25
3.4 Abstract . . . 26
3.5 Introduction . . . 27

3.6 Results . . . 31
3.6.1 Movement of naïve T cells in lymph nodes is superdiffusive, not Brownian . . . 31
3.6.2 Naïve T cell movement in LNs is not consistent with a Lévy walk . . . 38
3.6.3 T cells balance search for unique individual targets and interactions with multiple targets . . . 43
3.6.4 Naïve T cells show heterogeneity in movement patterns . . . 46
3.6.5 "Hotspots" in the LN environment show differing patterns of T cell motion . . . 46
3.7 Discussion . . . 49
3.8 Materials and Methods . . . 54
3.8.1 Ethics statement . . . 54
3.8.2 Mice . . . 54
3.8.3 T cell observations using two-photon microscopy . . . 55
3.8.4 Distribution fitting . . . 56
3.8.5 Autocorrelation and cross-correlations . . . 57
3.8.6 Mean squared displacement . . . 58
3.8.7 Heterogeneity . . . 59
3.8.8 Search efficiency simulation . . . 59
3.8.9 Identifying hotspots and hot tracks . . . 61
3.9 Acknowledgments . . . 61
3.10 Software . . . 62

4 Spatial Association of T cells and Dendritic Cells 63
4.1 Author Contribution Statement . . . 63
4.2 Overview . . . 64
4.3 Significance . . . 65
4.4 Methods . . . 67
4.5 Results . . . 73
4.6 Conclusions . . . 75
4.7 Software . . . 76

5 Adaptation of Lévy Exponents to Target Configuration 77
5.1 Author Contribution Statement . . . 77
5.2 Publication Notes . . . 77
5.3 Abstract . . . 78
5.4 Summary . . . 78
5.5 Introduction . . . 79
5.5.1 Search in Immunology . . . 81
5.5.2 Lévy Search . . . 85

5.6 Related Work . . . 87
5.6.1 Robotic Lévy Search . . . 87
5.6.2 Lévy Search in Immunology . . . 89
5.6.3 Lévy Search with Heterogeneous Target Distributions . . . 90
5.7 Stochastic Fractal Search . . . 90
5.8 Methods . . . 94
5.8.1 iAnt Robot Platform . . . 94
5.8.2 Robot Lévy Search . . . 95
5.8.3 Cluster Analysis . . . 95
5.8.4 Genetic Algorithm . . . 98
5.8.5 Effect of the Number of Searchers and the Configuration of Targets on the Optimal Fractal Dimension of Search . . . 99
5.9 Results . . . 101
5.9.1 The Hopkins Statistic . . . 101
5.9.2 The Genetic Algorithm . . . 102
5.9.3 Optimizing Fractal Dimension for the Number of Searchers and the Configuration of Targets . . . 104
5.9.4 Efficiency Scaling with Number of Robots . . . 111
5.10 Discussion . . . 112
5.11 Acknowledgements . . . 115
5.12 Conflicts of Interest . . . 116
5.13 Notes . . . 116
5.14 Software . . . 116

6 A Deterministic Swarm Search Algorithm 117
6.1 Author Contribution Statement . . . 117
6.2 Publication Notes . . . 118
6.3 Abstract . . . 118
6.4 Introduction . . . 118
6.5 Related Work . . . 120
6.6 Methods . . . 123
6.6.1 The Algorithm . . . 123
6.6.2 Robot Simulation . . . 126
6.6.3 Experimental Setup . . . 127
6.7 Results . . . 129
6.7.1 Performance . . . 129
6.7.2 Robustness . . . 131
6.7.3 Complete Collection . . . 132
6.7.4 Scaling with the Number of Robots . . . 134
6.7.5 Worst Case Performance . . . 136
6.7.6 Comparison to a Perfect Algorithm . . . 137

6.8 Conclusions . . . 138
6.9 Software . . . 140

7 Adaptation of Lévy Exponents to Error and Collisions 141
7.1 Author Contribution Statement . . . 141
7.2 Introduction . . . 141
7.3 Significance . . . 143
7.4 Related Work . . . 145
7.5 Methods . . . 145
7.5.1 Extending the search area boundaries beyond the area containing targets . . . 145
7.5.2 Going Big . . . 149
7.5.3 Mapping the target detection error rate per second to simulation iterations . . . 150
7.6 Results . . . 152
7.7 Discussion . . . 163
7.8 Software . . . 166

8 Conclusions 168
8.1 Swarm Search . . . 168
8.2 T cell Movement . . . 171
8.3 Connections . . . 172

Appendices 175

A T cell Analysis 176
A.1 Supporting Information . . . 176
A.1.1 Publication Notes . . . 176
A.1.2 Supporting Figures . . . 178
A.1.3 Supporting Tables . . . 190
A.2 Extended Materials and Methods . . . 191
A.2.1 Analysis of T cell Tracks . . . 191
A.2.2 Distribution Fitting . . . 193
A.2.3 Autocorrelation and Cross-Correlations . . . 195
A.2.4 Mean Squared Displacement . . . 196
A.2.5 Heterogeneity . . . 196
A.2.6 Search Efficiency Simulation . . . 196
A.2.7 Identifying Hotspots and Hot Tracks . . . 199
A.3 Modelling Software Interface . . . 199
A.4 Mutual Information Code . . . 200

B ALSA ANOVA Tables and Description **203**
B.1 Two-Way Analysis of Variance 203
B.2 ANOVA Tables 205

C Permission to Reproduce Previously Published Content **207**

List of Figures

1.1 T cell-Dendritic Cell Interaction Schematic 6

2.1 Stochastic Fractal Search Patterns 21

3.1 T cells Move in Lymph Nodes with Some Features of a Lévy Walk . 32
3.2 Distributions of T cell Speed and Step Lengths with MLE Fits . . . 33
3.3 T cell Search Balances Unique and Total Contacts with Targets . . . 34
3.4 T Cells Moving at Different Speeds Show Different Movement Patterns 35
3.5 T cells Visiting Hotspots Show a Different Distribution of Speeds Than T cells That Do Not Visit Hot Spots 36

4.1 Two-Photon Microscopy Movie Example Image 67
4.2 Two-Photon Movie Example Frames (Single Layer) 68
4.3 Volumetric Rendering of a Single 3D-Two-Photon Frame 69
4.4 Dependence of Mutual Information on Region Size 70
4.5 Mutual Information Over Change in Time (Individual) 73
4.6 Mutual Information Over Time 74

5.1 Observation of T cell Motion in Mouse Lymph Nodes 83
5.2 Stochastic Fractal Search Patterns 93
5.3 The iAnt Robot System. 100
5.4 Hopkins Index vs Number of Clusters 101
5.5 Convergence of the GA on an Optimum in the Fitness Landscape Defined by the Fractal Dimension 103
5.6 Dependence of Genetic Algorithm Selected Values of Fractal Dimension ($\mathcal{H}$) on Target Clusteredness 104
5.7 Dependence of Genetic Algorithm Selected Values of Fractal Dimension ($\mathcal{H}$) on the Hopkins Index 105
5.8 Full Factorial Heatmap 106

5.9 Multiple Comparison Plot of the Hausdorff fractal dimension ($\mathcal{H}$) and Swarm Size . . . 108
5.10 Efficiency Comparison for Search Patterns Evolved at the Extremes of the Hopkins Index . . . 109
5.11 Efficiency Comparison for Search Patterns Evolved at the Extremes of Swarm Size . . . 110
5.12 Search Efficiency Scales Linearly with the Number of Robots . . . 111

6.1 The DDSA Running in ARGoS, Overhead View . . . 124
6.2 The CPFA Running in ARGoS, Overhead View . . . 128
6.3 Comparison of DDSA and CPFA Performance for 3 Target Distributions . . . 130
6.4 Effect of Positional Noise on the DDSA and CPFA Search Patterns . 132
6.5 DDSA Performance Scaling with Increasing Error . . . 133
6.6 Time for the DDSA to Collect all Targets vs. Number of Targets . . 134
6.7 distributed deterministic spiral algorithm (DDSA) Performance Scaling with the Number of Searchers . . . 135
6.8 Crowding in the DDSA Degrades Performance . . . 136

7.1 Lévy Search Patterns are Naturally Adaptive to the Target Distribution 147
7.2 Dependence of Optimal μ on the Number of Searchers . . . 153
7.3 Dependence of Optimal μ on the Target Detection Error Rate . . . 154
7.4 Multiple Comparison of Optimal μ on Swarm Size and the Target Configuration . . . 154
7.5 Multiple Comparison of Optimal μ on the Error Rate and the Target Configuration . . . 155
7.6 Multiple Comparison Plots for Target Detection Errors . . . 156
7.7 Multiple Comparison Plots for Localisation Error . . . 156
7.8 DDSA and adaptable Lévy search algorithm (ALSA) efficiency for various swarm sizes for 30 min time limits . . . 157
7.9 Search Algorithm Comparison for Large Arenas . . . 158
7.10 Effect of Target Detection Error on Search Efficiency of μ - Uniform 159
7.11 Effect of Swarm Size on Efficiency as Function of μ . . . 160
7.12 Adding collisions causes μ to depend on swarm size . . . 161
7.13 The choice of μ affects efficiency in the presence of target detection error. . . . 161
7.14 Search strategy efficiency over long times. . . . 162

A.1 Example Tracks . . . 178
A.2 MSD Exponent Histograms for Various r^2 Filter Values . . . 179
A.14 Dependence of Speed on Frame Rate . . . 179

A.3 Prevalence of Tracks with Lévy Characteristics by Percentage of Data in the Power Law Tail . . . 180
A.4 Prevalence of Tracks with Lévy Characteristics by r^2 . . . 181
A.5 Comparison of the Gamma and Lognormal Distributions . . . 182
A.6 Sample DC Target Cluster Distributions in Simulation . . . 182
A.7 Mean Squared Displacement of Simulated Search Strategies . . . 183
A.8 Cluster Analysis of Tracks . . . 184
A.9 Dependency of Step Calculation on Angle Threshold . . . 185
A.10 Noise Dominates Mean Squared Displacement after 10 min . . . 186
A.11 Visualization of Search Patterns . . . 187
A.12 Hot Spot Visitor Empirical PDFs (Normalized Histograms) . . . 188
A.13 Visualization of Hotspots . . . 189
A.15 Screenshot of Track Efficiency Analyser . . . 200

List of Tables

3.1 MLE Fits to Step Lengths and Normalized Step Lengths 37
3.2 MLE Fits to Speeds and Normalized Speeds 38
3.3 Percent Change of Each Idealized Search Strategy for Unique Contacts 39
3.4 Percent Change of Each Idealized Search Strategy for Total Contacts 39
3.5 Best fit likelihood and MLE estimated parameters for the fastest and slowest cells 40
3.6 Hot and cold track step lengths show different MLE distribution fits 40

5.1 ANOVA of the Number of Searchers and Hopkins Index 106

7.1 Error Rate to Simulation Detection Rate 151

A.1 Extended Step Fit Statistics 190
A.2 Extended Speed Fit Statistics 191
A.3 Maximum likelihood estimated parameters and associated likelihood scores for steps calculated using a 30° threshold 191

B.1 ANOVA of Optimal μ on False Negative Target Detection Rate and the Target Configuration 205
B.2 ANOVA of Optimal μ on Swarm Size and the Target Configuration . 205
B.3 ANOVA of False Negatives and Localisation Error on Targets Collected for Fractal Dimension 0.7 206
B.4 ANOVA of False Negatives and Localisation Error on Targets Collected for Fractal Dimension 1.4 206
B.5 ANOVA of False Negatives and Localisation Error on Targets Collected for DDSA. 206

Glossary

$< ... >$	Angle brackets are used to denote the ensemble average.
$\mathcal{U}(0,1)$	Uniform probability variate generator on the unit interval, $\mathbb{I}$.
$\mathcal{N}(\mu,\sigma)$	Gaussian probability variate generator with mean = μ (not to be confused with the Lévy exponent μ) and standard deviation = σ.
$\mathcal{H}$	The Hausdorff fractal dimension. A measure of the fraction of points in some space that belong to a particular point set. We use this to measure the fraction of a search space that is visited by a particular search strategy.
H	The Hopkins index. A measure of how far from uniformly distributed points are.
μ	The exponent of a power law PDF of step lengths. Sometimes called the Lévy exponent. $\mathcal{H} = \mu - 1$.
α	Power law exponent of MSD.
$\mathbb{I}$	The unit interval $[0,1]$

Glossary

2PM — Two-photon microscopy. Technology that allows cells to be tagged with molecular dyes that fluoresce at narrow wavelengths when excited by laser light. 2PM is used to image T cells and DCs. We use wavelengths in the green and red spectrums to differentiate different populations of cells

95% CI $[a, b]$ — 95% confidence interval. A measure of the fraction of experiment repetitions that are expected to fall within the indicated range, [a,b], of values. Frequently used to measure the statistical significance of our results.

ACRW — Adaptive correlated random walk. A random movement model in which the degree of correlation between steps is influenced by information about the environment. The CPFA employs an ACRW.

Activated T cell — Once a T cell's TCR binds to antigen the T cell activates. Activated T cells undergo clonal expansion and seek out and destroy cells bearing the antigen that bound to the TCR.

Antigen — Chemical markers (typically proteins) that allow the immune system to distinguish cells belonging to self from non-self. The detection of antigen indicative of disease is the foundation of the immune response.

AICc — Corrected Akiake information criterion. An information theoretic *goodness of fit* measure based on MLE. The AICc reports the relative information lost when representing the observed data by statistical models. The corrected version penalises models with more

	parameters. The AICc is asymptotically equivalent to leave-one-out cross validation (Stone, 1977).
ANOVA	Analysis of variance. A statistical tool used to measure the amount of variance in an response variable that can be attributed to one or more input variables. We use this analysis to measure the contribution of experimental factors to the result. For example, ANOVA can measure the degree to which the best search strategy is influenced by the number of searchers.
April Tag	These visual codes are used as target markers. April tags provide information to the robot about target distance and orientation in space.
ARGoS	A swarm robot simulator. ARGoS is similar to Gazebo but runs much faster at the expense of realism. In this work we use the 2D physics engine dyn2d.
BRW	Biased random walk. A random movement model in which there is a global directional bias to turning angles. Not to be confused with a CRW.
CDF	Cumulative distribution function. The CDF for a value x is the sum of the PDF for all values less than or equal to x. We fit models to empirical CDFs rather than PDFs.
CCDF	Complementary cumulative distribution function. Also known as the hazard survival function and the first passage time density.

Glossary

Cognate	When antigen and TCR are cognate the antigen and TCR are capable of binding. When TCR and antigen bind cell signalling may be initiated that activates the T cell.
CPF	Central place foraging. A common task performed by groups of organisms is the discovery and transportation of food items to a central location. Transportation of resources, or other materials, to a depot is also of interest to developers of robot swarms.
CPFA	Central place foraging algorithm. A desert harvester ant inspired algorithm for the collection of resources. (Hecker and Moses, 2015)
CRW	Correlated random walk. A random movement pattern in which the turning angle between steps is not independent. This is a commonly used model of animal movement patterns that allows MSD to exceed that of Brownian motion. Also called a persistent random walk (PRW)
DC	Dendritic cell. Dentritic cells transport antigen to lymph nodes where it is displayed on their membranes. Dentritic cells are irregular in shape and are characterised by mobile antigen presenting protrusions. These cells are between 10 and 20 µm in diameter, though this measurement is complicated by their irregular and changeable shape.
DDSA	The deterministic spiral algorithm. A square spiral search strategy for swarms of robots.

Glossary

DMEM	Dulbecco's modified eagle medium. A commercial solution used to keep cells alive *ex vivo*.
ex vivo	Organs removed and kept viable for cells outside the host body are termed *ex vivo*.
EKF	Extended Kalman filter. A method of integrating sensor data that uses a dynamic noise covariance matrix to weight inputs.
FRC	Fibroblastic reticular cell. FRCs form a network within lymph nodes. The degree to which the FRC network influences T cell and DC motility and search patterns is under investigation by immunologists. As of this writing no consensus has emerged on the role of the FRC in guiding T cell-DC interaction.
GA	Genetic algorithm. An evolutionary optimisation technique in which a population of solutions navigate an optimality landscape using biologically inspired mechanisms such as gene crossover and mutation.
Gazebo	An environment closely tied to ROS that allows the simulation of robots. In this work we used Gazebo 2.0 and the Open Dynamics Physics Engine (ODE).
GPS	Global positioning system.
HEV	High endothelial venule. Vessels within the lymph node that allow migration of T cells in and out of lymph nodes.
Markovian	A Markovian process is memoryless. In the context of search Brownian Motion and Lévy search are characterised by turning angles that

	are Markovian, that is the turning angle is not related to the previous turning angles. In contrast correlated random walks have non-Markovian turning angles. We determine whether observed search patterns are Markovian using a the velocity autocorrelation function. Also known as the Markov property.
Lévy search	Lévy walks and Lévy flights are scale free fractal spacial distributions proposed by Mandelbrot (1983) and Shlesinger and Klafter (1986) and popularised as models of optimal search in an ecological context in Viswanathan et al. (1996). They were proposed as a description of T cell movement in the brain in Harris et al. (2012).
LN	Lymph node. Lymph nodes are where dendritic cells display antigen gathered from the host. Naïve T cells search lymph nodes until activated by cognate antigen. LNs are strategically placed throughout the body with tissues in each region of the body draining to particular lymph nodes. Lymph nodes are connected to one another via lymphatic vessels.
LogMCRW	Lognormal modulated correlated random walk. A type of random walk in which there is a correlation in turning angles and a lognormal distribution of step lengths. We use this model to describe T cell motility in Lymph Nodes.
MLE	Maximum likelihood estimation. A model parameter estimation technique especially well suited to selection of parameters for non-linear models such as power laws. Intuitively maximum likelihood

	estimation chooses the parameters that maximize the likelihood of drawing a particular sample. The likelihood function is the given by the sum of the probabilities of observing the input conditioned on the candidate parameter. Optimisation techniques are used to find the candidate parameter that maximises the likelihood function. Note that the likelihood function is not a probability since its sum over all input can exceed 1.
MSD	Mean squared displacement. We use MSD as a measure of how quickly a search pattern moves a searcher away from its starting point, and it is therefore one of the ways we measure the intensity-extent tradeoff.
MSE	Mean squared error. A common *goodness of fit* measure which uses the expected value of the square of the difference between model values and observed values. MSE methods are equivalent to MLE for linear fits when the error is normally distributed but are much faster to compute.
Naïve T cell	T cells leave the thymus and search for cognate antigen. Before a T cell's TRC binds to cognate antigen the T cell is called naïve.
QR	Quick response code. These visual codes are used as targets for the iAnt robot which does not have a gripper.
Reality gap	A common problem in robotics is the development of robot algorithms in simulation that do not translate well to real robots. Simulations enable the evaluation of algorithms many times faster

than would be possible with real robots, but cannot completely reproduce the complexity of real environments interacting with embodied robots. A robot interacting with its environment is in itself a complex system with potentially non-linear feedback between actuators, environment, and sensors. This makes modelling difficult, especially in the case of swarms of robots which have the added complication of robot-robot interactions.

TCR — T cell receptor. A binding site on the membranes of T cells which allows cognate antigen to dock and initiate a signal cascade within the T cell.

PDF — Probability density function. A function that maps outcome of an event to its probability, resulting in a probability distribution for a random variable. We use PDFs to model stochastic processes such as random search strategies.

ROS — The Robot Operating System. This is the software platform produced by the Open Robotics Software Foundation that we use to develop the software for the Swarmathon robots. In this work we use the Indigo Igloo release of ROS.

Search extent — The rate at which a searcher increases its distance from it's starting point independent of its velocity. For a given time interval search extent has an inverse relationship to search intensity. MSD is a measure of extent.

Search intensity The fraction of points in a space observed by a searcher in a fixed time period. For a given time interval search extent has an inverse relationship to search intensity. $\mathcal{H}$ is a measure of intensity.

Systems Biology Systems biology is the study of biology as a *complex system*, that is, as a system of interacting components that produce emergent behaviour. The system may be examined at the level of molecules, cells, organisms, species, or include multiple levels. The study of search in the immune system and central place foraging (CPF) by ant colonies are examples of systems biology.

Chapter 1

Introduction

Search is a fundamental process in a wide variety of natural and engineered systems. For example, animals search for food and mates, and computer programs are designed to search for solutions to problems. Some search processes involve multiple agents that cooperate to search for targets. For example, ants in a colony communicate as they search for seeds, the cells of the immune system collectively search for pathogens, and robots in a swarm may be designed to collectively search for targets.

We study systems that use search strategies that do not rely on prior knowledge of the location of targets. Such strategies fall into two types of search: stochastic and deterministic. For individual searchers, spiral search has been shown to be an optimal *deterministic* search strategy (Baeza-Yates et al., 1993; Skubch, 2012). Physicists have shown that a *stochastic* search strategy called Lévy search is optimal for systems given certain simplifying assumptions (Viswanathan et al., 1996). The optimal search strategy for swarms of searchers in finite areas with a limited number of consumable targets is not known.

Our goal in this dissertation is to develop a deeper understanding of swarm search processes using a combination of approaches: 1) quantitative analysis of the search patterns of swarms of T cells, and 2) development of new algorithms for robot swarms carrying out a resource collection task.

1.1 Lost Key Analogy

To build intuition about some of the issues involved with swarm search, consider a familiar problem: As you are leaving your house for work, you realise that you have misplaced your keys. As you search the house, you may choose to follow a deterministic search pattern that is guaranteed to locate the keys but will take a very long time to complete. Alternatively, you could follow a stochastic search pattern that maximises the probability of finding the keys quickly, but with no guarantee that the keys will be found at all.

You may have information about likely places to search; if an initial search of those locations is unsuccessful, you might choose new locations to search at random. Some may choose to search each room of the house thoroughly, perhaps deterministically, until the keys are found. Others may move quickly from room to room, giving each a cursory glance. The effectiveness of the search depends on the probability distribution of the location of the keys and the probability of detecting the keys if they are present.

If the probability of finding the keys at locations that are close to each another is highly correlated, then a thorough *intensive* search of those locations may be a good strategy. On the other hand, if the probability of finding a set of keys is uniform across all locations in the house, then a more *extensive* search may be warranted.

Because it is easy to miss a small and inconspicuous set of keys, an intense search in several areas, rather than a more extensive cursory search, could be effective.

When searching for something that is lost, we will often re-examine previously searched areas in case the object was missed on the first pass. Losing your keys can become a major event, with family members enlisted to help find them (attendant to a reward). The addition of multiple searchers introduces a new optimisation problem: How should the house be partitioned among the searchers? It would seem obvious that the searchers should be divided evenly across the house. However, young children may have a high error rate for key detection, so searching together in small groups could decrease the probability of a false negative.

Finally, in the enthusiasm of children trying to find the keys (and motivated by the promise of a reward), they may all decide to search the nearest areas first. Assigning multiple people to search the same area can be productive if the keys are hard to see, but may also cause problems due to crowding. Search with noisy sensors, limited time, little prior knowledge about target locations, and a high likelihood of collision with other searchers presents a complex optimisation problem.

This analogy highlights the factors that impact search that we address here: target detection error, multiple searchers, and varying target configurations. To be effective, robot swarms must be efficient in the presence of these factors. We show how these factors influence the search performance of the distributed deterministic spiral algorithm (DDSA) and the adaptable Lévy search algorithm (ALSA). A recurring theme in this work is the trade-off between *intensity* and *extent* following Méndez et al. (2013). Search extent and intensity lie on a continuum: more intensive patterns search thoroughly in the local area, while extensive patterns cover

more area, but may miss targets nearby. We use the fractal dimension of the search pattern as a convenient way to measure the intensity vs. extensiveness of search.

1.2 Relationship to Breadth-First and Depth-First search

There are similarities between the extent vs. intensity trade-off and the breadth-first search vs. depth-first search trade-off. When searching a tree, the trade-off depends on whether the best solutions tend to be localised to a few branches or distributed widely across many branches, and whether they are common or rare. Assuming that a solution-space tree is organised such that related solutions exist in the same subtree, deep nodes can be thought of as near in space to the current location. Choosing to traverse nodes using a depth first search is analogous to intensive spatial search, and breadth first search is analogous to extensive spatial search.

Each search technique has advantages, if the solutions are clustered in one branch depth-first search will be able to exploit the locality of solutions. If all of the solutions are close together, significant time may be wasted exploring subtrees that contain no solutions. If solutions are spread at random throughout the search tree breadth-first search may be more successful. The addition of a heuristic (the A* algorithm, for example) can reduce the time taken to discover the desired solution by choosing how often to visit related solutions (Hart et al., 1968). This is essentially an algorithm to balance the intensity vs. extent of search.

In spatial search, exploring areas far away from the start location incurs high travel time costs but the areas are unlikely to have been searched by other agents. Nearby areas are likely to have been visited by other members of the swarm but travel costs are small. This suggests a parallel between the exploit vs. explore trade-off and

the k-armed bandit problem (Robbins, 1985), in which a gambler must decide on the proper allocation of time to either explore the environment to gain information about regions in solution space with a potential high pay-off, or exploit resources already discovered. For the k-armed bandit problem the trade-off is between certainty and risk, in swarm spatial search the challenge is to balance travel cost vs. oversampling.

Decades of research have found ways to balance depth first vs breadth first search and the explore-exploit trade off. The explore-exploit trade-off governs whether to make use of known resources or spend time to discover unknown, but potentially more valuable resources. This trade-off is a recurring theme in complex systems science and provides a framework for approaching many disparate systems (Shalizi, 2006). This dissertation is an extension of that work applied to spatial search problems. Our goal is to reveal how the extent vs. intensity trade off depends upon the distribution of targets, the number of searchers, and the collision rate of searchers as swarms of agents conduct searches in physical space

1.3 Spatial Search in Two Swarm Systems

We study search in two swarm systems: T cells searching for pathogens in lymph nodes and robots searching for targets in a bounded 2-dimensional space. Though these systems seem unrelated, the underlying search patterns can be analysed in similar ways.

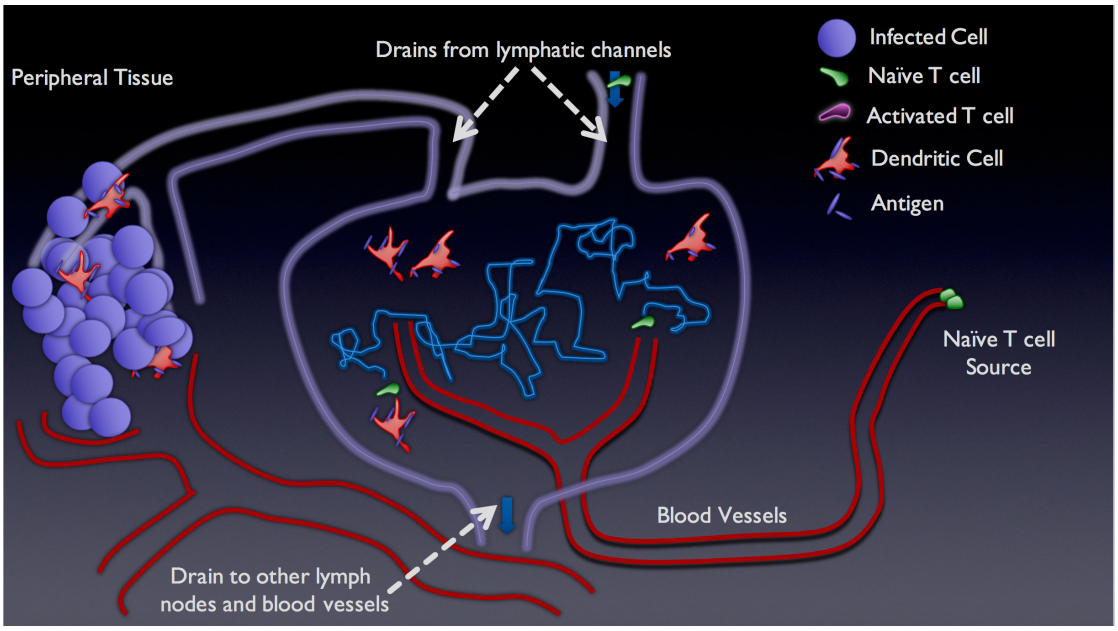


Figure 1.1: T cell-Dendritic Cell Interaction Schematic. DCs patrol tissues in search of antigen. Discovered antigen is brought to LNs, a confined space in which T cells search for antigen-bearing DCs. When T cells detect cognate antigen they activate and may eventually proliferate and move to peripheral tissue. These T cells will seek out and destroy cells displaying the antigen that their clonal ancestors detected in the lymph node.

1.3.1 Search in Immunology

We investigate the spatial patterns of T cells searching for dendritic cells (DCs) in lymph nodes (LNs) and we measure how quickly T cells find their DC targets given different distributions of targets and different requirements for how many and how often T cells need to contact DC targets to complete their search. We choose this system because we are interested in search where recruitment does not appear to guide searchers. Quantitatively characterising the stochastic search of T cells informs our understanding of a crucial part of the adaptive immune response, and provides a naturally-occurring model of swarm search.

This dissertation characterises the movement patterns of naïve T cell swarms searching LNs and proposes the *Effective Immune Search* hypothesis, analogous to Optimal Foraging Theory (Krebs, 1978; Stephens and Krebs, 1986). The Effective Immune Search hypothesis proposes that a fast immune response confers a fitness advantage on its host. Therefore, we expect T cells to have evolved patterns of search that minimise pathogen detection time. We specifically test whether T cell search patterns find targets quickly given different spatial distributions, densities and required contact rates with DCs.

1.3.2 Search and Foraging in Swarm Robotics

Robot swarms typically consist of many small, relatively simple and inexpensive robot agents that work collectively towards some goal. A major research challenge is the development of robot swarm systems that allow effective navigation through complex real-world environments without centralised control (Winfield et al., 2005; Hecker and Moses, 2015), but there is a considerable lack of adaptable, scalable, and robust search algorithms to meet this challenge (Winfield, 2009; Brambilla et al., 2013). Meeting this challenge has implications for other computing systems where agents work in parallel in the physical environment, for example embedded sensor networks and the Internet of Things.

Many real-world applications for robot swarms require the detection and collection of targets, including planetary surveys (Fink et al., 2005), land and sea mine clearance (Weber, 1995), pollution mapping (Hu et al., 2011), search and rescue (Birk and Carpin, 2006; Goodrich et al., 2008), military applications (Love et al., 2015), and agricultural pest control (Tamura and Naruse, 2014). Central-place foraging

applications, which require targets to be transported to a central collection point (Winfield, 2009), include crop harvesting (Bac et al., 2014) and planetary resource collection (Ramsey, 2015).

We study the effectiveness of novel search strategies for robot swarms in simulation. The simulations we use are specifically designed to model the behaviour and capabilities of the iAnt robots developed in our lab (Hecker and Moses, 2015). Using simulations allows the systematic manipulations of a variety of relevant parameters, including: target detection error rates, target distributions, and the presence or absence of collisions. Simulations have also allowed us to test search strategies in search areas that vary from $100\,\mathrm{m}^2$ to $1\,\mathrm{km}^2$ with many more robots than we have been able to build in hardware. We use two simulation testbeds: ARGoS which allows simulation using a physics engine, and our own simulation software that allows thousands of experiments to be conducted quickly in order to rapidly evolve parameters for ALSA.

We test the ALSA and DDSA search algorithms in hardware using the robot operating system (ROS). The robots we tested our algorithms in were built in our lab for the UNM-NASA swarm robotics challenge (Ramsey, 2015). These tests demonstrate that the algorithms we designed can run successfully in real robot swarms.

1.4 The Role of Simulation

Immune systems and robot swarms are both complex systems (Holland, 1992; Hofmeyr and Forrest, 2000; Winfield, 2009). Theoretical approaches to characterising search have been fruitful in generating bounds for ideal cases (Stone, 1975; Viswanathan

et al., 1999; Chechkin et al., 2008), but these formal mathematical models tend to be simplified, often assuming infinite 1-dimensional space and time.

Following Axelrod (1997), computer simulations provide an alternative 'third way of science' that allows more realistic finite-systems to be explored. Computer simulation also allows us to ask 'what if?' questions that may not be feasible to test in physical robots or biological systems. Accordingly, this work uses computational simulations to analyse the efficiency of different swarm search behaviours under a variety of conditions for both robots and T cells.

Swarms of agents searching for targets will be adaptive to the presence of other agents, through collisions, oversampling, and the distribution of undiscovered and discovered targets. These local interactions and feedback loops introduce non-linearities that make the application of purely analytical approaches difficult. The swarms we are interested in here are on the order of a few dozens to hundreds of agents. Analytical methods have been successful at describing the behaviour of a single agent,where there are no complex interactions with other agents modifying the search space, and for systems with very large numbers of agents, where interactions tend to cancel one another out. For smaller swarms individual interactions are significant for the behaviour of the whole and so must be modelled explicitly.

1.5 Organisation and Contributions

In Chapter 3, using computer simulations, we identify three distinct factors that contribute to increasing T cell search efficiency: 1) a lognormal distribution of step lengths, 2) motion that is directionally persistent over short time scales, and 3) heterogeneity in movement patterns. Furthermore, we show that T cells move dif-

ferently in specific frequently visited locations that we call hotspots within lymph nodes, suggesting that T cells change their movement in response to the lymph node environment.

Our analysis of random search strategies in lymph nodes is predicated on the assumption that nave T cells are not guided to target DCs through long range interactions. There are several proposed mechanisms for guidance of naïve T cells in LNs. In the presence of guidance by the environment we would expect there to be a co-localization of T cells and DCs to a degree greater than that expected by chance. In Chapter 4 we present a method using information theory to measure the degree of co-localization observed in lymph nodes in order to resolve this question (Munoz et al., 2014a; Donovan and Lythe, 2016). This method is applied to two-photon microscopy (2PM) images of T cells searching for DCs.

In Chapter 5 we use a simulated robot swarm to evaluate the effectiveness of a Lévy strategy applied to a *free search* task. In free search targets must be located, but there is no additional requirement that targets be transported to another location. We map the relationship between search parameters and target distributions in the absence of collisions. We show that the intensity-extent trade-off depends on the distribution of targets and but only weakly on the number of searchers.

We develop the distributed deterministic spiral algorithm (DDSA) in Chapter 6, as a benchmark algorithm that generalises a single searcher spiral to a swarm of robots. Biologically inspired search algorithms are difficult to analyse. Comparison with a deterministic algorithm provides an easily analysed baseline. We find that the DDSA is effective for smaller swarms but a stochastic ant-inspired algorithm, which reduces robot crowding, is more effective for larger swarms.

Surprisingly, localisation error does not impact the search efficiency of the deterministic spiral as much as the stochastic algorithms we tested. The explanation highlights the dependence of all the foraging algorithms that we have designed on site fidelity. Site fidelity causes searchers to return to areas where they last found targets, allowing them to exploit heterogeneity in the target distribution. When site fidelity fails due to localisation error, stochastic algorithms are adversely effected. The DDSA systematically sweeps an incrementally expanding area and so does not rely on site fidelity to rediscover clusters of targets.

In Chapter 7, we present the adaptable Lévy search algorithm (ALSA). ALSA allows the selection of the the Hausdorff fractal dimension ($\mathcal{H}$) through the Lévy exponent, μ. We analyse the performance of ALSA under a variety of experimental conditions, such as varying the boundary conditions, the size of the search area, and allowing collisions. We use ARGoS to provide accurate modelling of collisions and robot dynamics.

We find that in the presence of target-detection error, increasing search intensity makes search more efficient. Larger swarms benefit from increased extent, which reduces collisions and oversampling. Increasing the extent of search is also beneficial when targets are widely separated. The optimal intensity-extent trade-off is influenced by the configuration of targets, though the direction of the trade-off depends on the relationship between the search boundary and the region within which targets are placed.

We discover that search intensity increases with the local target density. This is a natural result of searcher reorientation when targets are encountered. This feedback leads to an emergent and dynamic coupling between the intensity-extent trade-off and the local target distribution, resulting in increased efficiency.

Together these chapters describe how the intensity-extent trade-off can be used to understand natural search systems and design effective search strategies for a variety of realistic conditions.

Chapter 2

Background

The dominant spatial search strategies can be grouped into three categories: random-search strategies such as correlated random-walks, Brownian motion, and Lévy flights; Bayesian approaches that use a prior model of target locations; and deterministic search strategies. The existing body of work on search processes can alternatively be divided up by the field in which they originated: physics, systems biology and ecology, and operations research. Hybrid search strategies also exist. Random environments can give rise to stochastic movement patterns even for deterministic agents. This aspect of movement is particularly important to our understanding of T cell movement in lymph nodes (LNs). (Hughes, 1996) provides a thorough analysis of movement in stochastic environments, as opposed to intrinsic random motion.

Search is well understood and characterised for single robots, and there are mathematical results for multiple searchers (primarily resulting from the statistical mechanics of reaction equations (Krapivsky et al., 2010), animals foraging in social groups (Giraldeau and Caraco, 2000), the intersection of these fields (Viswanathan et al., 2011; Méndez et al., 2013)), and in operations research (Stone, 1975). Stone laid the

foundation for optimal search following a Bayesian approach using Lagrange multipliers. The theory of optimal search laid out by Stone and his successors has two requirements: 1) the probability distribution of target placement is known to the searcher, 2) the conditional probability of finding a target in a particular location given the amount of time spent looking in that location is also known to the searcher. We confine our work to that in which the priors are unknown to the searcher, and so the Bayesian approach is not applicable.

Optimal search is a central problem in operations research. During World War II the Anti-Submarine Warfare Operations Research Group of the US Navy devoted significant resources to the development of optimal search patterns. The resulting algorithms are used to assist the US coastguard in planning search and rescue and interdiction efforts. Famous examples of the application of these methods to real-world problems include the search for an atomic bomb lost in the ocean in 1966 and the submarine Scorpion lost in 1968 (Stone, 1975). Trummel and Weisinger (1986) showed that optimising the probability of finding a target in limited time is NP-complete.

Ecologists in the 1960s hypothesised that natural selection acting on animals foraging for food would result in strategies that maximise caloric intake and therefore more efficient search strategies. This body of work became known as Optimal Foraging Theory (Krebs, 1978; Stephens and Krebs, 1986). For example, the marginal value theorem describes the optimal rate at which foragers leave a depleting patch of resources to search for a fresh patch (Charnov, 1976). Central place foraging theory extends optimal foraging theory to include organisms which move collected resources to a central location (Orians and Pearson, 1979). The ideas of optimal foraging theory have been extended to other search domains such as finding mates.

In the robotics community, spiral search has been used for robot search (Skubch, 2012) for single targets or uniformly distributed targets. Hoff et al. (2013) present an adaptive foraging algorithm that uses a combination of random walks and a circular sweep. Van Dartel et al. (2004) evolved neural controllers for agents searching a simulated world with targets drawn from a uniform distribution. Swarm robot simulations have used Lévy walks in combination with chemotaxis-inspired gradient sensing (Nurzaman et al., 2009) and artificial potential fields (Sutantyo et al., 2010) to search unmapped spaces efficiently with range-limited sensors. Keeter et al. (2012) use *underactuated* robots implementing Lévy walks in a 3D aquatic environment to search for four uniformly distributed targets. They sample various values of μ in 0.5 increments in the range $1.1 < \mu < 3$. Keeter et al. find that in simulation there is a monotonic improvement in search time as μ approaches their lower bound of 1.1, and ballistic motion.

Statistical physics, ecology, and operations research have all had a significant impact on the development of spatial search theory. Until recently the applications of search theory to engineered systems have been limited. However, with the rise of swarm robotics, there is an increasing need for computer scientists to take advantage of this disparate body of work in designing spatial search algorithms. The impact of error, target distributions other than uniform random, and swarm size may all be relevant for swarm search strategies. We explore these questions in our work.

2.1 Deterministic Search

Deterministic search methods have been developed in the operations research community, with particular emphasis on spiral algorithms. Spiral algorithms have been

studied analytically in two papers by Baeza-Yates. The first establishes the optimality properties of a spiral search pattern in the single searcher case (Baeza-Yates et al., 1993). Their second paper extends the spiral search pattern to multiple searchers but does so by having rovers move to a uniform random location before beginning the spiral. This n-searcher algorithm is no longer deterministic (Baeza-Yates and Schott, 1995). Burlington and Dudek (1999) extend the single searcher spiral search pattern to a complex environment.

Isbell (1957) described a target detection search pattern for individual ships in which it performs a continuous space-filling spiral. Ryan and Hedrick (2005) analyse a square search pattern carried out by a single helicopter. This search pattern is defined in Appendix H of the Coast Guard Operating Manual. (U.S. Coast Guard, 2002)

Short-lived spirals have also been observed to occur during search by ants by Müller and Wehner (1994) and by Elizabeth Esterly in our lab (private correspondence). These spirals apparently exist on very short time scales and are very noisy.

2.2 Stochastic Search

By the *central limit theory* almost all random search strategies become equivalent to a simple random walk (Brownian motion) (Aleksandr and Khinchin, 1949). This is because, asymptotically, all search patterns with finite moments converge on a Gaussian distribution of displacements. Lévy and Borel (1954) (cited in Viswanathan et al. (2011)) formulated the *generalised central limit theory*, which showed some distributions do not converge to a Gaussian distribution. Exceptions are *ballistic motion*, where the searcher always moves in a straight line and the *Lévy α-stable*

distributions of displacements. Méndez et al. (2013) provides a survey of the theory of such distributions as they relate to random search. Viswanathan et al. (1999) argues that since Lévy search patterns allow regions to be revisited while having a greater asymptotic mean squared displacement (MSD) than all other random search strategies, they are an optimal search strategy. In particular, Viswanathan showed that Lévy flights with a power law exponent of 2 maximise the searcher-target encounter rate for a single searcher in an infinite plane, with uniformly distributed targets that are replaced after detection. Zhao et al. (2015) and Levin (2016) find that in finite systems ballistic motion can be an optimal strategy.

Theoretically, systematic search outperforms random search (Stephens and Krebs, 1986), when searchers can localise themselves accurately (Bénichou et al., 2011; Keeter et al., 2012).

2.3 Lévy Search

Lévy walks consist of step lengths that fit a power law distribution, where step length is defined to be the displacement (shortest distance) between consecutive positions. Most step lengths are small, but with a heavy-tail, that is, a decreasing probability of larger steps and a non-zero probability of steps of any length. The infinite variance of this distribution prevents convergence to a Gaussian distribution of displacements and Brownian motion. Lévy walks assume that the direction of search at each step is drawn from a uniform distribution and is independent of previous steps (i.e. is isotropic and Markovian)(Mandelbrot, 1983; Viswanathan et al., 1996). Lévy search patterns are stochastic fractals. The probability density function (PDF) that governs the distribution of step lengths used to generate a particular Lévy pattern is a power

law:

$$\mathrm{L}(x) \propto x^{-\mu}, \tag{2.1}$$

where $\mathrm{L}(x)$ is the probability of a searcher moving in a straight line for distance x. The exponent μ that determines the shape of the PDF is known as the Lévy exponent. When search consists of a sequence of disconnected points, the motion is called a Lévy flight (the searcher is flying or jumping from point to point). Since we consider space between points to be part of the area searched, and therefore the intervening space is traversed with some finite velocity, we are working with a Lévy walk (Shlesinger and Klafter, 1986).

Lévy walks, as models of search, were first developed to explain the disparity between observed super-diffusive animal motion and simple random walk models. Animals searching for food tend to maintain relatively straight trajectories for longer distances than would be produced by a simple random walk. The Lévy walk foraging hypothesis is an explanation for this observation proposed by physicists and ecologists. Lévy walks have been used to explain the search patterns of numerous species including reindeer (Mårell et al., 2002), albatross (Viswanathan et al., 1996), and human foragers (Raichlen et al., 2014). James et al. (2011) provide a more comprehensive list along with criticism of Lévy walk analysis. Whether these search patterns are truly power law distributed is a matter of ongoing debate (Plank and James, 2008; Humphries et al., 2012). The issue is clouded, in part, because a true power law distribution of step lengths is impossible in a finite space. The question then is whether animals or cells use a truncated power law distribution of step lengths constrained by the environment in which they are searching.

The PDF governing Lévy flights as formulated in Eq. (2.2) not only describes the probability, $\mathrm{L}(x)$, of observing a step length of less than x, but also relates the resulting stochastic fractal to the Hausdorff fractal dimension ($\mathcal{H}$):

$$\mathrm{L}(x) = \frac{\gamma}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-1-\gamma} \tag{2.2}$$

where, x is the step length, $x_{\min}$ is the least possible step length, and γ determines the rate at which the probability of a particular step length occurring decays. The coefficient normalises the area under the curve to one and so enforces that the equation is a PDF. The fractal dimension of a Lévy flight is $\mu - 1$ in Eq. (2.1) on the previous page (Mandelbrot, 1983; Hughes, 1996).

Raposo et al. (2011) model the relationship between heterogeneity of searcher target distance and optimal μ values. Using a 1-dimensional analysis, they predict that decreasing μ will increase the success of target encounters in heterogeneous landscapes. They suggest that this theoretical result generalises to the 2-dimensional case.

2.4 Hausdorff Fractal Dimension

The Hausdorff fractal dimension is a compact measure of the trade-off between intensity and extent. Increasing extent, by decreasing $\mathcal{H}$, results in increased displacement of searchers from their start positions as a function of time, independent of searcher velocity. The fractal dimension of a search pattern is the asymptotic fraction of locations visited in a search space. For example, Brownian motion has fractal dimension $\mathcal{H}= 2$ (Taylor, 1953), meaning that, asymptotically, a Brownian search pattern visits

all positions in a 2-dimensional space. As far as we are aware the fractal dimension of Lévy walks, in which the searcher visits the intervening space between movement endpoints, has not been formalised. Therefore we calculate $\mathcal{H}$ for Lévy flights. Since Lévy flights visit only endpoints of steps, $\mathcal{H}$ is the dimension of the visited point-set (Seshadri and West, 1982; Mandelbrot, 1983). $\mathcal{H}$ for Lévy walks will be strictly greater than that for Lévy flights.

Substituting the $\mathcal{H}$ of Brownian search for γ results in a μ of 3. The resulting walk is maximally intense when embedded in a 2-dimensional space.

2.5 Mean Squared Displacement

The mean squared displacement, a measure of search extent, of a population of searchers is characterised by a power law:

$$\text{MSD} = \left\langle (\vec{r}(t + \Delta t) - \vec{r}(t))^2 \right\rangle \propto t^{\alpha} \tag{2.3}$$

where $\vec{r}(t)$ is the position vector of a searcher at time t and $\vec{r}(t + \Delta t)$ is the location of the searcher after some time increment, Δt. Angle brackets indicate the ensemble average over the population of searchers. The MSD exponent describes the rate of displacement over time and is related to the search extent. As α increases search extent increases. MSD is related to $\mathcal{H}$ by $2/\mathcal{H}$ (Zanette, 1999).

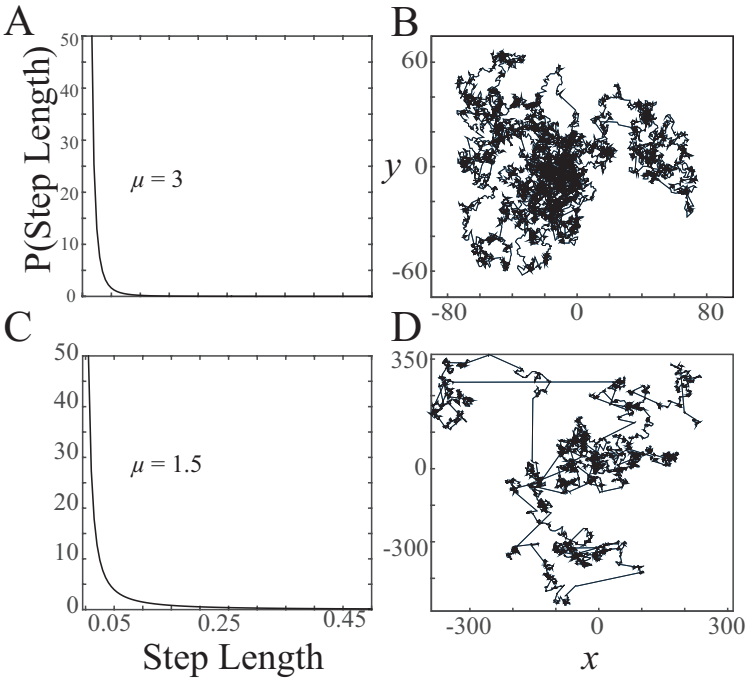


Figure 2.1: Stochastic fractal search patterns. A sample of 10,000 steps with uniform turning angles and a power law distribution of step lengths. The power law distributions of step lengths A) with exponent $\mu = 3$ ($\mathcal{H}$= 2) resulting in B) a Brownian pattern of search. C) A power law distribution of step lengths ($\mu = 1.5$, $\mathcal{H} = 0.5$), and D) the resulting pattern of search with lower dimensionality, and therefore lower intensity but greater extent. Reprinted from Fricke et al. (2016b).

2.6 Brownian Motion

The most common formalism used to generate random search is Brownian motion. This approach is so common, in fact, that the terms are often used interchangeably. The essential feature of Brownian motion is that the direction of movement at each time step is uncorrelated with the previous time step. The time between reorientations in direction is assumed to be small. In continuous time systems, this is taken to be the limit as time approaches zero. In discrete-time systems, the smallest time step is used between directional reorientations. In systems with continuous time and

space and finite velocities, as the time between steps goes to zero so does the distance covered between reorientations. In systems with continuous time but discrete space, the minimum distance covered between reorientations will have a discrete lower limit. This is also the case for discrete time and space systems. A simplifying assumption we make is that discrete time and space systems with sufficiently fine granulation are representative of a continuous system.

The displacement of a searcher from time t to $t+1$ is the sum of the magnitudes, M, of the movement vectors between time t and $t+1$. Given that the distribution of M has finite variance and is unbiased, the sum of M will be a Gaussian distribution by the central limit theorem.

Brownian motion has $\mathcal{H} = 2$ which can result in significant oversampling of the search space when performing a central place foraging task. When Brownian searchers start searching from the same location, the high fractal dimension of their movement results in locations being revisited by other robots in the swarm. For n searchers, the number of unique locations visited is proportional to $t \ln \frac{n}{\ln t}$. Only after t exceeds e^n, do searchers employing a simple random walk avoid search redundancy (Larralde et al., 1992).

Chapter 3

Analysis of T cell Search in Lymph Nodes

> *With four parameters I can fit an elephant, and with five I can make him wiggle his trunk.*
>
> — Jon von Neumann quoted in Dyson, F. (2004). A meeting with Enrico Fermi. *Nature*, 427(6972):297

3.1 Author Contribution Statement

I am the lead author of this chapter under the supervision of Melanie Moses (Associate Professor, UNM Computer Science, Biology, and Santa Fe Institute) and Judy Cannon (Associate Professor, UNM Pathology, Molecular Genetics and Microbiology). François Asperti-Bousin imaged T cells under the direction of Judy Cannon. Kenneth Letendre provided advice about statistical tests.

3.2 Publication Notes

Citation: Fricke G.M., Letendre K.A., Moses M.E., Cannon J.L., (2016) Persistence and Adaptation in Immunity: T Cells Balance the Extent and Thoroughness of Search. PLoS Comput Biol 12(3): e1004818. doi:10.1371/journal.pcbi.1004818

Editor: Martin Meier-Schellersheim, National Institutes of Health, United States

Received: October 6, 2015

Accepted: February 17, 2016

Published: March 18, 2016

Copyright: ©2016 Fricke et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Documentation is included in Appendix C on page 207.

Formatting: The original published text has been preserved as much as possible while still adhering to the formatting requirements of this dissertation. Any typographical errors have been left in place and marked with [*sic*].

Data Availability Statement: Data is available at the NIH funded Systems Biology Center website: http://stmc.health.unm.edu/tools-and-data/ under the Download Data Here link after Letendre K. et al. reference. Also available for download from: http:// stmc.health.unm.edu/tools-and-data/.

Funding: This work was supported by funding from the following: NSF EF 1038682 (MM), DARPA P1070-113237 (MM), NIH 1R01AI097202 (JLC), the Spatiotemporal Modeling Center (P50 GM085273), the Center for Evolution and Theoretical Immunology 5P20GM103452 (JLC), and a James S. McDonnell Foundation grant for the study of Complex Systems (MM). Also thanks to the UNM Cancer Center

Fluorescence Microscopy Facility (P30-CA118100) as well as the BRAIN Imaging Center (P30GM103400) for help with 2-photon microscopy. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Supporting Information: SI Figures, SI Tables, and the extended materials and methods are included in Appendix A on page 176, Section A.3 on page 199.

3.3 Author Summary

The immune system is responsible for clearing disease-causing infections, and T cells are an important immune cell type that helps eliminate viruses and bacteria. To become activated, T cells must encounter another type of immune cell called dendritic cells in the lymph node. T cell search for dendritic cells is similar to animal search for food. Here we precisely analyze how T cells move using search patterns originally developed to describe animals. We find that T cell motion is a complex combination of multiple strategies including moving in a persistent direction and using different step sizes. This allows T cells to balance the need to search both extensively throughout the lymph node and also to search some regions thoroughly for possible infection. Furthermore, we use a computer model to demonstrate that T cells are more likely to be found in specific locations in lymph nodes. We call these locations "hotspots". We find that T cells in hotspots move differently, apparently searching more thoroughly, suggesting that T cells can adapt to their environment, similar to

animals foraging for food. These results show that T cells share fundamental search strategies with foraging animals, exhibiting both persistence and adaptation.

3.4 Abstract

Effective search strategies have evolved in many biological systems, including the immune system. T cells are key effectors of the immune response, required for clearance of pathogenic infection. T cell activation requires that T cells encounter antigen-bearing dendritic cells within lymph nodes, thus, T cell search patterns within lymph nodes may be a crucial determinant of how quickly a T cell immune response can be initiated. Previous work suggests that T cell motion in the lymph node is similar to a Brownian random walk, however, no detailed analysis has definitively shown whether T cell movement is consistent with Brownian motion. Here, we provide a precise description of T cell motility in lymph nodes and a computational model that demonstrates how motility impacts T cell search efficiency. We find that both Brownian and Lévy walks fail to capture the complexity of T cell motion. Instead, T cell movement is better described as a correlated random walk with a heavy-tailed distribution of step lengths. Using computer simulations, we identify three distinct factors that contribute to increasing T cell search efficiency: 1) a lognormal distribution of step lengths, 2) motion that is directionally persistent over short time scales, and 3) heterogeneity in movement patterns. Furthermore, we show that T cells move differently in specific frequently visited locations that we call "hotspots" within lymph nodes, suggesting that T cells change their movement in response to the lymph node environment. Our results show that like foraging animals, T cells

adapt to environmental cues, suggesting that adaption is a fundamental feature of biological search.

3.5 Introduction

Search has been extensively studied in biology, particularly in ecology, to understand how animals search for food, mates and prey. The pattern of movement by searching agents affects search efficiency in a variety of biological contexts (Bartumeus et al., 2005; Viswanathan et al., 2002; Méndez et al., 2013). Optimal foraging theory suggests that animals, including social animals such as ants and bees, have evolved strategies to individually or collectively maximize food intake in minimal time (Pyke, 1984).

Similar to foraging animals, T cells of the immune system search for targets to mount an immune response. T cells are a critical immune effector, required to clear viral infections and to help B cells produce antibody. In order to initiate an effective immune response, naïve T cells must encounter and sample dendritic cells (DCs) bearing cognate antigen in lymph nodes (LNs). In the absence of infection, T cells continuously enter and exit LNs interacting with DCs. Upon infection, DCs present cognate antigen and provide stimulatory signals leading to T cell activation. T cell-DC interactions are required for naïve T cells to survive, activate and eventually clear infection as well as maintain immune memory (Munoz et al., 2014b; Mackay et al., 2000; Germain et al., 2012).

T cell activation is promoted by repeated sampling of nearby DCs (Textor et al., 2014), while at the same time T cells explore the entire population of DCs for rare antigen indicative of infection. This presents T cells with an optimization problem

in which T cells must balance thoroughness and extent of search. This requires that many T cells search across a broad extent, contacting many DCs quickly, a process similar to optimal foraging in animals. Simultaneously, T cell search is sometimes thorough, repeatedly sampling in a small area (Textor et al., 2014). Both of these factors contribute to the overall rate at which T cells encounter DCs within LNs, which is a critical component of organismal fitness impacting the overall timing of the immune response.

Relatively little quantitative analysis has been done to describe how T cells move in LNs or how that movement affects the rate at which T cells encounter DCs. Initial studies to understand the type of T cell motion in LNs from pioneering two-photon imaging of naïve T cells suggested that T cells move using a simple diffusive random walk, analogous to Brownian motion of molecules (Miller et al., 2004, 2003). Following these studies, computational modeling of T-DC interactions have often used simple diffusive random walks to represent T cell behavior (Donovan and Lythe, 2012; Preston et al., 2006). However, subsequent studies have not precisely described T cell motion in LNs, so it is unclear whether diffusive random walks are appropriate models for T cell movement.

Optimal random search strategies have been extensively studied in ecology, and ecological models of movement may be useful for characterizing T cell motility and search efficiency. Brownian motion, Lévy walks, and correlated random walks (CRWs, also called persistent random walks), have been proposed as idealized biological search models (Codling et al., 2008), but careful quantitative analysis is required to understand how well search models characterize T cell motility and search efficiency (Krummel et al., 2014). Brownian motion is often referred to as a simple random walk and is characterized by movement with uniformly distributed turning

angles and small fixed step sizes relative to the time resolution of observation (Miller et al., 2002, 2003; Mempel et al., 2004; Bousso and Robey, 2003; Celli et al., 2012). Qualitative similarities between Brownian motion and the movement of microorganisms resulted in simple random motion being used as a dominant model of cell motion (Przibram, 1913). Brownian motion results in diffusive movement in which distance travelled is proportional to the square root of time. In two dimensions this results in a normal distribution of speeds, and in three dimensions it results in a Maxwell distribution of speeds (Maxwell, 1860).

Lévy walks exist between ballistic (or straight directional) motion at one extreme and Brownian motion at the other. In contrast to Brownian motion, the step lengths of Lévy searchers fit a power law distribution with most step lengths being small, but with a heavy-tail, that is, a decreasing probability of larger steps and a non-zero probability of steps of any length (Viswanathan et al., 2002; Codling et al., 2008). Lévy walks have been used to model animal movement, for example, in albatross, ant, aphid and human foraging, and more recently, T cells in the brain (Viswanathan et al., 1996, 2002; Petrovskii et al., 2011; Raichlen et al., 2014; Harris et al., 2012). Both Brownian and Lévy walks assume that the direction of search at each step is drawn from a uniform distribution and is independent of previous steps (i.e. is isotropic and Markovian). CRWs on the other hand use fundamentally different mechanisms to model similar patterns of motion that tend to persist in direction over time. CRWs depend on the distribution of turning angles between successive steps leading to directional persistence. In search modelled by CRWs, the current direction of motion probabilistically influences future step directions (Codling et al., 2008). On relatively short time scales, Lévy walks and CRW may be difficult to distinguish since they both produce superdiffusive motion (Reynolds, 2010), that is, displacement

that increases faster than the square root of time. Compared to diffusive movement, superdiffusion increases search *extent* and decreases search *thoroughness*.

Despite the fact that many search strategies are well-characterized, there has been no systematic analysis of T cell motion in LNs. The lack of clarity in empirical studies has led to T cell motility being modelled using Brownian motion (Celli et al., 2012), Lévy walks (Harris et al., 2012), and correlated random walks (CRW) (Textor et al., 2014), or a combination of movement patterns (Banigan et al., 2015). Recently, Harris et al. showed that the movement of T cells in *Toxoplasma gondii* infected brain tissue fits a Lévy walk resulting in superdiffusion and efficient detection of protozoan targets (Harris et al., 2012). It is not clear if Lévy movement has not previously been found in LN because such movement does not occur there, or simply because it had not been looked for. The lack of precise quantitative understanding of T cell motion in LNs leads to inconsistent models and limits our ability to determine how T cell motility affects the efficiency with which T cells encounter DCs.

In this study, we analyze T cell search behavior in LNs using two-photon microscopy. We begin our analysis with traditional statistical methods that describe the velocities, step lengths, displacement, and turning angles taken by naïve T cells searching for DCs. We then extend these analyses to more accurately and comprehensively describe motility patterns, including using maximum likelihood estimates (MLE) to fit experimental data. Our study statistically analyzes T cell search strategies in LNs, and uses multiple efficiency metrics that measure the spatial thoroughness and extent of T cell search. We then directly quantify the contribution of different types of motion to the efficiency of T cell search. Additionally, by comparing T cell movement to the patterns generated by null models of random motion, interesting non-random interactions between T cells and their environment become

apparent, suggesting that T cells adapt movement in response to environmental cues. Our null models reveal hot spots that are visited more frequently than can be explained by chance. Our results suggest that even a precise characterization of T cell movement based on the assumption of random movement does not fully capture the complexity of T cell movement in the LN environment.

3.6 Results

3.6.1 Movement of naïve T cells in lymph nodes is superdiffusive, not Brownian

Two photon microscopy (2PM) has been used extensively to study the movement of T cells in intact lymph nodes (Stoll et al., 2002; Miller et al., 2002; Mempel et al., 2004; Celli et al., 2012; Cannon et al., 2013). We isolate bulk primary T cells from LNs of naïve C57Bl/6 animals, fluorescently label T cells with dyes, reintroduce labeled T cells into recipient mice, and then use 2PM to image labeled T cells in intact explanted LNs of recipients (see Materials and Methods for further details). We track cells for up to 10 minutes and include all motile cells in observation windows. We eliminate tracks with total track length shorter than 17 µm or that show squared displacement less than 300 μm^2 (= 17 µm × 17 µm) as described previously by Letendre et al. (Letendre et al., 2015). The data analyzed here are from 5,891 individual T cell tracks from 41 fields from 12 experiments. We group those 41 fields into 7 datasets, each dataset containing fields imaged using frame rates within one second of each other. This allows us to combine data across fields when performing analyses, such as velocity autocorrelation, that depend on the frame rate.

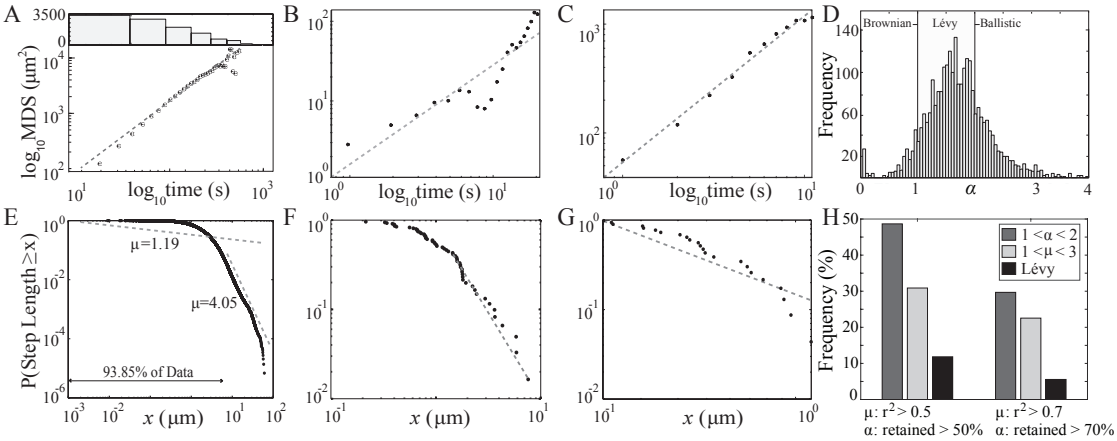


Figure 3.1: T cells move in lymph nodes with some features of a Lévy walk. Lévy walks are characterized by particular power law exponents of mean squared displacement (MSD) and step length distribution. (A, bottom) Observed T cell MSD vs. time. The dashed line is the linear regression with slope $\alpha = 1.41$ indicating superdiffusion. (A, top) The number of data points in the MSD calculation. (B) Example displacements for a single T cell track with $r^2 = 0.52$, and (C) with $r^2 = 0.93$. (D) Histogram of for individual tracks with $r^2 > 0.8$ (see Fig. S2 for other r^2 thresholds) with labels indicating the range of values of α consistent with Brownian, Lévy and ballistic motion. (E) Empirical complementary cumulative distribution function (CCDF) of all 145,731 step lengths for all 5,077 cells. The x-axis is all possible distances less than the maximum observed, the y-axis is the probability that an observed step length exceeds a particular value of x. The dashed line (offset for clarity) with slope 4.05 is the best fit to the power law tail of the CCDF which includes only 6.15% of the steps (Clauset et al., 2009). The line with slope 1.19 is the best fit to all data. (F,G) Examples of step length distributions and MLE fits for tracks with 49% and 93% of the track in the tail. (H) Percentage of tracks in the Lévy region for μ and power law exponents and their intersection. Data are included when the $r^2 > 0.5$ for and at least 50% (left histogram) or 70% (right histogram) of the track steps are retained in fitting the power law tail.

We observe T cell velocities and motility coefficients largely in agreement with those previously published (Miller et al., 2004; Mempel et al., 2004; Viswanathan et al., 2008; Letendre et al., 2015). [*sic*: (Miller et al., 2004; Mempel et al., 2004; Letendre et al., 2015)] We calculate the diffusion coefficient using the unweighted

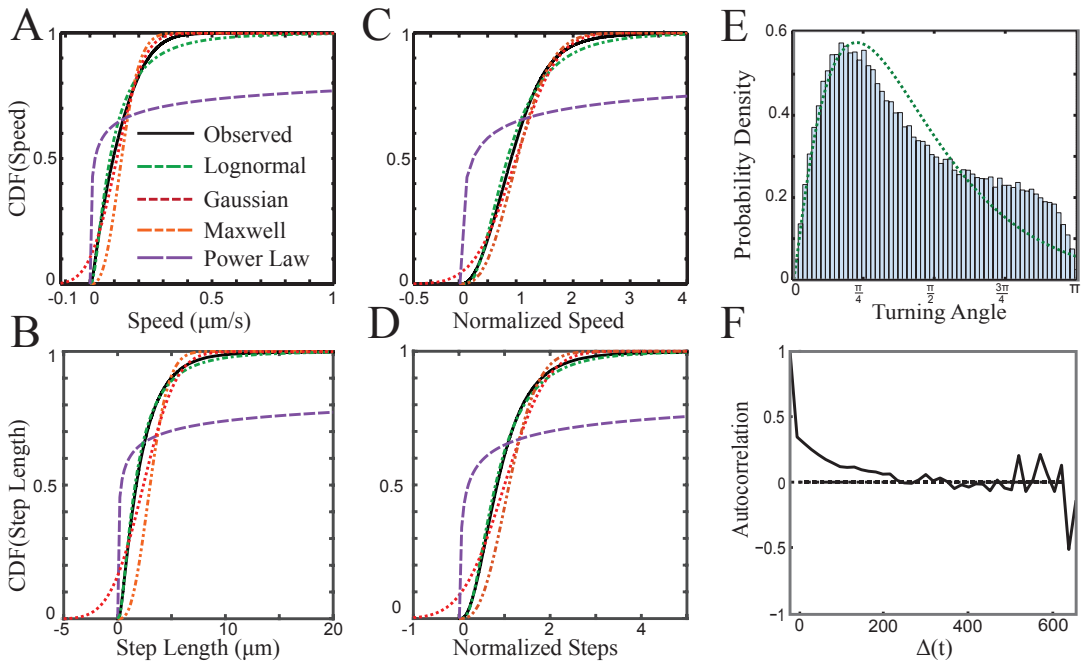


Figure 3.2: Distributions of T cell speed and step lengths with MLE fits. For (A) speed and (C) step length the lognormal function is the best fit (see Tables 3.1 on page 37 and 3.2 on page 38 for likelihood values and model parameters). Fits for normalized speed (B) and normalized step lengths (D) are divided by the mean speed or step length of the track from which they are drawn. (E) Histogram of all 149,592 observed turning angles. The green line is the maximum likelihood estimation of the gamma distribution used to model turning angles in the efficiency simulation. (F) Turning angle autocorrelation for 23,169 vectors from the 537 T cell tracks observed in one dataset. The correlation in movement direction decays until reaching zero at approximately 240 s.

average method (Michalet, 2010, 2011). T cells move with a mean speed with 95% confidence interval = 5.81 ± 0.024 µm/min, median speed = 4.22 µm/min, motility coefficient, D = 19.2±0.534µm^3/min [*sic*: µm^2/min], calculated from a linear fit MSD of 5,185 tracks (out of 5,891 tracks filtered for $r^2 > 0.8$). The motility coefficient is calculated using a linear model fit to the first 25% of each displacement curve and for positions not exceeding the 10 min track time.

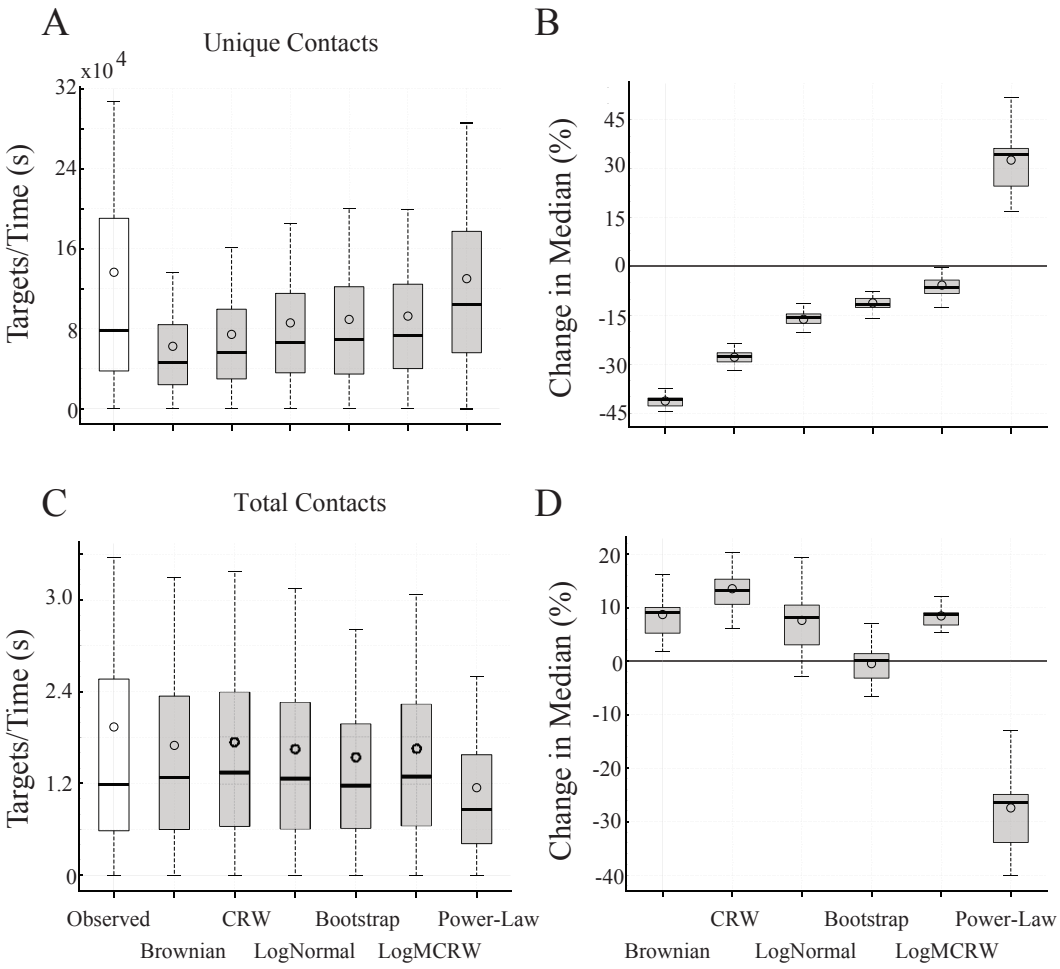


Figure 3.3: T cell search balances unique and total contacts with targets. Interquartile boxplots show search efficiency for DCs in 10 µm radius clusters. Panels (A) and (B) show unique contact efficiency; (C) and (D) show total contact efficiency. (A) and (C) show 1000 efficiency samples for each of the 41 fields. (B) and (D) compare the percent change in median search efficiency for each candidate search model relative to observed T cell search (indicated by the line at 0). See Tables 3.3 on page 39 and 3.4 on page 39 for other target distributions and significance values. Outliers are not shown for clarity.

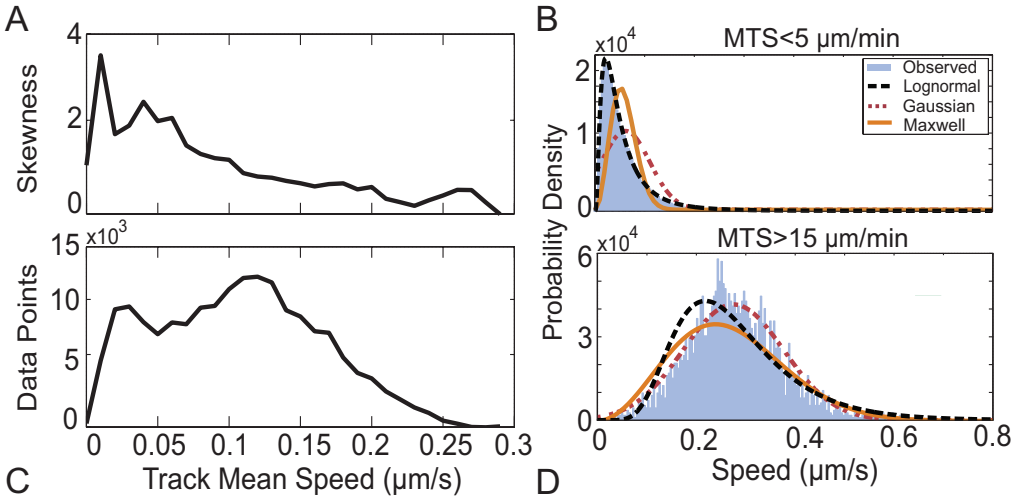


Figure 3.4: T cells moving at different speeds show different movement patterns. (A) Skew of step length distribution as a function of track mean speed and (C) the number of data points as a function of track mean speed. Tracks with mean track speeds (MTS) less than $5\,\mu m\,min^{-1}$ (B) and greater than $15\,\mu m\,min^{-1}$ (D) were selected to illustrate different MLE model fits for fast and slow tracks (for fits see Table 3.5 on page 40).

Displacement is commonly used as a first step to assess whether movement is consistent with a Lévy walk or Brownian motion (sample tracks in Fig. A.1 on page 178)(Viswanathan et al., 2008; Reynolds, 2010). We determine the displacement of individual T cells over time. Fig. 3.1A shows the mean squared displacement (MSD) of one of the 7 datasets, as well example tracks with lower (Fig. 3.1B) and higher (Fig. 3.1C) r^2 values. We then calculate the linear fit to the log-log-transformed data. Logarithmically transforming data before applying a linear regression is a common way to measure the exponent of a power-law relationship between depen-

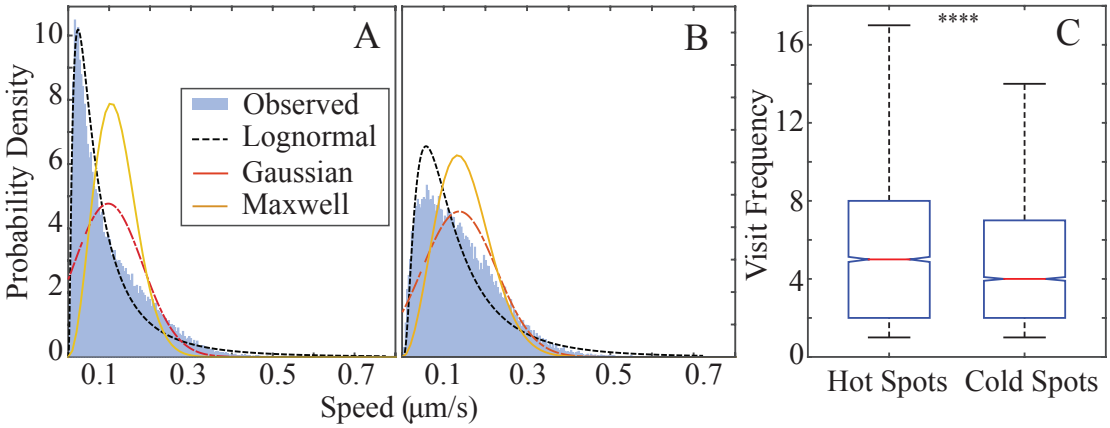


Figure 3.5: T cells visiting hotspots show a different distribution of speeds than T cells that do not visit hot spots. Cold tracks (A) have a speed distribution that is more peaked at low speeds with a more skewed, heavy-tailed distribution compared to hot tracks (B). For fits, see Table 3.6 on page 40. (C) Visit frequency, or number of observations of hot tracks in hot vs. cold spots. Hot tracks were observed to visit hotspots more than cold spots. The graph shows the distribution of average number of visits by hot tracks to hotspots versus cold spots. Interquartile box plot of the distribution with the red line indicating the median number of visits. Outliers are not shown. **** indicates $p << 10^{-3}$ using Mann-Whitney U test.

dent and independent variables (Michalski et al., 1986). Log-log-transformed Lévy walks produce displacement exponents, α, between 1 and 2 (Viswanathan et al., 2005). We calculate the distribution of α for all T cell tracks and find that 56% of T cells have a displacement exponent α falling in the expected window for a Lévy walk (Fig. 3.1D). Only 28.3% of cell tracks are subdiffusive ($\alpha < 1$), and the remaining tracks (15.6%) have a best-fit displacement exponent indicative of accelerating motion ($\alpha > 2$). Because low r^2 values of linear fits to log-log-transformed data may indicate that the data are not well-described by any displacement exponent, we

Step lengths				
Distribution	-log Likelihood $\times 10^5$	AICc$\times 10^5$	MLE Parameters	
Lognormal	**2.65**	**5.29**	$\boldsymbol{\mu = 0.4818}$	$\boldsymbol{\sigma = 0.9192}$
Gaussian	3.36	6.72	$\mu = 2.3895$	$\sigma = 2.4229$
Maxwell	4.02	8.04	a = 3.8497	
Power Law (Lévy)	4.58	9.16	$\alpha = 1.1921$	
Normalized Step lengths (step length/track mean step length)				
Lognormal	**1.20**	**2.40**	$\boldsymbol{\mu = -0.2217}$	$\boldsymbol{\sigma = 0.6896}$
Gaussian	1.61	3.23	$\mu = 1$	$\sigma = 0.7324$
Maxwell	1.69	3.37	a = 0.5117	
Power Law (Lévy)	3.32	6.63	$\alpha = 1.2245$	

Table 3.1: **MLE fits to step lengths and normalized step lengths (N = 145,731 steps)**. Negative log-likelihood measures the relative ability of candidate models to explain the observed data (For additional fits tested, see Tables A.1 on page 190 and A.2 on page 191). The corrected Akaike information criterion (AICc) and Bayesian information criterion (BIC) (Table A.2 on page 191) confirm that order of fit quality is not due to the number of model parameters. The most negative -log likelihood and AICc scores are the best fits; in this case that is the smallest positive score for the lognormal distribution. The last columns lists the distribution parameters that were selected by MLE. See Tables A.1 on page 190 and A.2 on page 191 for other distribution fits and goodness of fit statistics. doi:10.1371/journal.pcbi.1004818.t001

repeat the analysis on data sets restricted to r^2 values > 0.5, which discards 33% of all tracks, and $r^2 > 0.75$, discarding 50% of all tracks (see Fig A.2 on page 179 for figures with different r^2 filters). Increasing r^2filtering decreases the fraction of cells in the subdiffusive window, but the qualitative message remains the same: T cells demonstrate heterogeneous behavior, with some displacements consistent with subdiffusive, Brownian, ballistic and even accelerating motion, but the majority of T cells are superdiffusive but sub-ballistic. Fig 3.1D shows the histogram of α for tracks with an $r^2 > 0.8$, other r^2 thresholds are shown in Fig A.2 on page 179, including all tracks with no filtering in S2A Fig.

Speeds				
Distribution	-log Likelihood $\times 10^5$	AICc$\times 10^5$	MLE Parameters	
Lognormal	**-1.84**	**-3.68**	$\boldsymbol{\mu = -2.5027}$	$\boldsymbol{\sigma = 0.9329}$
Gaussian	-1.61	-3.23	$\mu = 0.1161$	$\sigma = 0.0881$
Maxwell	-1.12	-2.24	a = 0.0071	
Power Law (Lévy)	0.122	0.245	$\mu = 1.2069$	
Normalized Speeds (speed/track mean speed)				
Lognormal	**1.22**	**2.45**	$\boldsymbol{\mu = -0.1669}$	$\boldsymbol{\sigma = 0.5706}$
Gaussian	1.37	2.74	$\mu = 1$	$\sigma = 0.7324$
Maxwell	1.32	2.65	a = 0.4414	
Power Law (Lévy)	3.58	7.16	$\mu = 1.2446$	

Table 3.2: **MLE fits to speeds and normalized speeds (N = 159,746)**. The Lognormal distribution has the most -log-likelihoods and AICc score and therefore is the best fit. The parameters selected by MLE are shown for each distribution. See Tables A.1 on page 190 and A.2 on page 191 for other distribution fits and goodness of fit statistics. doi:10.1371/journal.pcbi.1004818.t002

3.6.2 Naïve T cell movement in LNs is not consistent with a Lévy walk

While displacement analysis suggests most T cells are consistent with a Lévy walk, another defining feature of Lévy walks is that the inverse power law complementary cumulative distribution function (CCDF) for step lengths has an exponent, μ, between 1 and 3. Therefore, we analyzed T cell step lengths for the μ exponent. We define a step to be the resultant of a velocity subsequence in which each T cell velocity vector deviates by no more than 15 from the previous vector and a step length is the distance covered by a step. Fig 3.1E shows that a power law fit to the population of T cell step lengths is only valid if almost 94% of the data are excluded from the analysis (see Materials and Methods: Distribution fitting). The resulting best-fit μ exponent for the remaining 6% of the power law tail is 4.05 (Fig 3.1E). The curvilinearity, the poor fit, as well as the μ value all indicate that a Lévy walk is not a good description of T cell motility. On average 51% of data must be excluded

	Search Strategy					
Target Distribution	Brownian	CRW	Lognormal	Bootstrap	LogMCRW	Power Law
10 µm (0.2)	-41.88 ± 0.82 $p < 10^{-4}, 10^{-4}$	-28.11 ± 1.09 $p < 10^{-4}, 10^{-4}$	-15.91 ± 1.69 $p < 10^{-4}, 10^{-4}$	-12.98 ± 0.99 $p < 10^{-4}, 10^{-4}$	-7.35 ± 1.92 $p < 10^{-4}, 10^{-4}$	27.63 ± 5.44 $p < 10^{-4}, 10^{-4}$
20 µm (0.32)	-39.828 ± 0.64 $p < 10^{-4}, 10^{-4}$	-25.87 ± 0.70 $p < 10^{-4}, 10^{-4}$	-13.39 ± 1.62 $p < 10^{-4}, 10^{-4}$	-9.926 ± 1.37 $p < 10^{-3}, 10^{-4}$	-3.98 ± 1.94 $p < 10^{-4}, 10^{-4}$	34.17 ± 8.49 $p < 10^{-4}, 10^{-4}$
40 µm (0.44)	-41.88 ± 0.81 $p < 10^{-4}, 10^{-4}$	-22.75 ± 0.55 $p < 10^{-4}, 10^{-4}$	-9.621 ± 1.97 $p < 10^{-4}, 10^{-4}$	-4.798 ± 1.37 $p < \mathbf{0.85}, 10^{-4}$	-0.218 ± 2.17 $p < 10^{-4}, 10^{-4}$	36.02 ± 6.41 $p < 10^{-4}, 10^{-4}$

Table 3.3: **Percent change of each idealized search strategy for unique contacts compared to the empirical search strategy across 3 different target distributions.** Table entries are percent change in median search efficiency from observed ± 95% confidence interval. Two p-values are shown: the first indicates the significance of the change in median efficiency between the observed and idealized runs (N=10 runs, each run consists of 4,100 samples, Fig. 3.3B). The second p-value tests whether all raw efficiency values differ between observed and idealized runs (N=41,000, Fig. 3.3A). All p-values are calculated using the Mann-Whitney U test. The values in parentheses are the Hopkins aggregation statistic. All search strategies are statistically different from observations except LogMCRW in the most diffuse 40 µm DC clusters (in bold).doi:10.1371/journal.pcbi.1004818.t003

	Search Strategy					
Target Distribution	Brownian	CRW	Lognormal	Bootstrap	LogMCRW	Power Law
10 µm (0.2)	8.7 ± 1.16 $p < 10^{-4}, \mathbf{0.29}$	12.94 ± 1.34 $p < 10^{-4}, 10^{-3}$	7.24 ± 3.25 $p < 0.01, 0.05$	7.24 ± 3.25 $p = \mathbf{0.63}, 10^{-4}$	8.4 ± 3.66 $p < 10^{-3}, \mathbf{0.73}$	-28.66 ± 2.43 $p < 10^{-4}, 10^{-4}$
20 µm (0.32)	12.71 ± 1.52 $p < 10^{-4}, \mathbf{0.87}$	15.67 ± 1.54 $p < 10^{-4}, 10^{-4}$	9.22 ± 2.72 $p < 0.05, 0.05$	2.29 ± 2.72 $p = \mathbf{0.19}, 10^{-4}$	12.18 ± 2.64 $p < 10^{-4}, \mathbf{0.8}$	-26.27 ± 4.14 $p < 10^{-4}, 10^{-4}$
40 µm (0.44)	17.71 ± 1.86 $p < 10^{-4}, 10^{-4}$	20.89 ± 1.58 $p < 10^{-4}, 10^{-4}$	13.07 ± 3.24 $p < 10^{-4}, 10^{-4}$	4.52 ± 2.51 $p < 0.05, 10^{-4}$	16.31 ± 2.69 $p < 10^{-4}, 10^{-4}$	-24.08 ± 4.4 $p < 10^{-4}, 10^{-4}$

Table 3.4: **Percent change of each simulated search strategy for total contacts compared to the empirical search strategy across 3 different target distributions.** Table entry format is identical to Table 3.3. These values correspond to Fig 3.3C and 3.3D. Brownian motion, bootstrap and LogMCRW are not significantly different from the observed distribution of efficiencies when targets are more clustered (in bold), but power law search underestimates the efficiency of search for total contacts. doi:10.1371/journal.pcbi.1004818.t004

in order to obtain a maximum likelihood estimated (MLE) power law fit (see Fig 3.1 on page 32F and G for example tracks with low and high percentage of steps in the power law tail; see Fig A.3A and D for histograms of μ using other GoF threshold values; and see Figs A.2 on page 179 and A.3 on page 180 for additional analysis.)

Distribution	Mean Speed < 5 µm/min -log Likelihood ($\times 10^5$)	MLE Parameters	Mean Speed > 15 µm/min -log Likelihood ($\times 10^3$)	MLE Parameters
Lognormal	-1.09	-3.35, 0.826	-2.60	-1.35, 0.387
Gaussian	-0.918	0.0482, 0.0407	-2.73	0.277, 0.0958
Maxwell	-0.789	0.0013	-2.66	0.0286
Power Law	-0.492	1.25	2.018	1.35
Skew	2.37		0.52	
Kurtosis	13.3		3.98	

Table 3.5: **Best fit likelihood and MLE estimated parameters for the fastest and slowest cells.** The Gaussian distribution better fits tracks with mean speed > 15 µm/min while lognormal better fits tracks with mean speed ¡ 5 µm/min. The step speed distribution for fast tracks has a shorter and lighter tail than the sample of tracks with slower mean speeds. Best negative loglikelihood scores are in bold. doi:10.1371/journal.pcbi.1004818.t005

Distribution	Hot Tracks -log Likelihood ($\times 10^5$)	MLE Parameters	Cold Tracks -log Likelihood ($\times 10^5$)	MLE Parameters
Lognormal	1.29	0.671, 0.752	1.85	0.819, 0.583
Gaussian	1.405	2.504, 1.72	4.22	4.53, 22.3
Maxwell	1.48	3.077	8.24	171
Power Law	2.96		2.65	1.21
Skew	1.45		11.12	
Kurtosis	6.74		136	

Table 3.6: **Hot and cold track step lengths show different MLE distribution fits.** Hot tracks tend to be faster than cold tracks and more Brownian in their movement pattern. The high kurtosis and skew is due to a long tail in the distribution of step lengths belonging to tracks that do not visit hotspots. doi:10.1371/journal.pcbi.1004818.t006

We determined the number of tracks that fit both $1 < \mu < 3$ and $1 < \alpha < 2$ parameters. Setting our goodness of fit (GoF) filtering criteria to require that at least 70% of the data per track is retained in calculating the exponent μ (Fig 3.1F), and the r^2 statistic for the power law exponent α is at least 0.7, we find that only 5.5% of all T cell tracks fit both criteria for Lévy walk (Fig 3.1 on page 32H). We note that the tracks excluded when filtering by r^2 and those filtered by the percent of track in the power law tail both tend to be subdiffusive. For any filtering criteria the vast majority of T cell tracks are not Lévy walks (Fig A.4 on page 181).

To further analyze T cell motion, we quantify speeds (T cell displacement between consecutive frames multiplied by the frame rate) of all T cell tracks (Fig 3.2A and 2C) and find that in LNs T cell speeds range from 6.5×10^{-4}µm/s to 0.9 µm/s (Fig 3.2A). We fit experimentally derived speeds (Fig 3.2A) and step lengths (Fig 3.2B) to idealized probability distributions. We use parametric distributions because they are associated with well-known generative processes; for example, the Gaussian distribution is produced by the cumulative effect of additive processes, the lognormal distribution is often associated with multiplicative or branching processes (Gunning and Wearing, 2013), and the Maxwell distribution is a product of Brownian motion in three dimensions. We use likelihood measures to rank how well different distributions explain the observed data (Tables 3.1 on page 37 and 3.2 on page 38).

The distribution of T cell step lengths and speeds are more consistent with a lognormal distribution than with Brownian motion (defined by a Gaussian or Maxwell distribution) or a Lévy walk (defined by a power law distribution of speeds (Shlesinger et al., 1999) as shown by the higher values in the MLE for power law fits in Tables 3.1 on page 37 and 3.2 on page 38. The variance of observed T cell speeds and lengths is high, and the distributions have a heavier tail (greater right skew) than both Gaussian and Maxwell distributions. The power law probability distribution overrepresents both very small steps and very large steps compared to observed T cells. The lognormal distribution shows the best statistical fit for both speed and step lengths. The gamma distribution also fits the observed speeds very well (S1 and S2 Tables, Fig A.5 on page 182). However since gamma and lognormal are often used to model the same phenomena, we present only lognormal here (Wiens, 1999).

It is possible that the right skew in the speed distribution arises from the variance between track mean speeds rather than from speed variance within tracks (Petrovskii

et al., 2011). To test for this possibility, we divide each speed drawn from within a cell track by the cell mean speed (called normalized) and ask whether the distribution becomes less heavy-tailed. We find that both normalized speed and step length distributions are still best fit by a lognormal distribution (compare Fig 3.2A with Fig 3.2C, 3.2B and 3.2D and the normalized vs. raw lengths and speeds in Tables 3.1 on page 37 and 3.2 on page 38), but the right skew is decreased. Our observations indicate that the heavy-tailed lognormal distribution is not simply due to distinct populations moving at different mean speeds, though heterogeneity in speed within the population is a factor.

Both Brownian motion and Lévy walks assume that the angle of each turn is drawn from a uniform random distribution. We analyze the turning angles of each T cell at each time step and find that T cell turning angles are not uniform, and that there is a bias toward turning angles of less than 90 (Fig 3.2E). The non-uniform distribution of turning angles suggests that T cells may move according to a CRW. We fit distributions to turning angles using MLE and find the gamma distribution to be the best fit, although it cannot capture all of the variation in the bi-modal distribution (Fig 3.2E green-dotted line). We then performed an autocorrelation analysis of directions over time to determine whether there is a dependency between the direction of T cells at one time step and the previous time steps (Fig 3.2F). We find that T cells show turning angle autocorrelation consistent with a CRW (indicated by positive values in Fig 3.2F). The correlation persists for approximately 4 minutes. Our cross-correlation analysis shows no drift in the observation fields (Materials and Methods: Eq. (A.2) on page 195).

3.6.3 T cells balance search for unique individual targets and interactions with multiple targets

A key function of naïve T cell search within LNs is to find and interact with antigen bearing DCs. To determine whether different types of search can affect T cell interaction with DCs, we use an agent-based model, using biologically informed parameters, to assess the degree to which different modes of random search predict the observed pattern of T cell search efficiency (i.e., the number of DCs encountered per unit time). We reproduce features of T cell movement by creating search tracks using the best distribution fit to speeds (Table 3.1 on page 37) and turning angles, limited by the total distance covered and time observed for empirical T cell tracks. We run simulations with DC targets placed with 3 different degrees of clustering: highly clustered (DC centers placed in 10 µm radius spheres), moderately clustered (in 20 µm radius spheres) to more evenly dispersed (in 40 µm radius spheres) (Fig A.6 on page 182). We confirm that these DC placements result in a range of clusteredness according to the Hopkins aggregation statistic that ranges from 0.44 for dispersed clusters (close to the 0.5 value expected for a uniform distribution) to 0.2 for compact clusters. We confirm that Brownian motion in our simulations results in diffusive movement (Fig A.7 on page 183). We then compare efficiency of modelled search with observed T cell tracks from the experimental data across this range of DC cluster sizes.

We calculate efficiency of T cell search in two ways. First, we determine how many unique "DC" targets were encountered by each T cell in a specific period of time. Previous studies suggest that naïve T cells have no a priori information about the location of DCs in LNs (Preston et al., 2006; Donovan and Lythe, 2012). Second, we determine how many total DC target encounters occur in the specified time. Total contacts count repeated contact with the same DC while unique contacts

counts only one contact per DC. Total contacts are important for T cell activation and potentially survival while unique contacts are a measure of how long it may take T cells to find rare DCs presenting cognate antigen. The simulation addresses two questions: do statistical descriptions of T cell movement produce search efficiencies that are similar to those of observed T cells; and, how do the relative efficiencies of the idealized models compare to each other and experimentally observed T cells.

Not surprisingly, the efficiency of observed T cells show a much wider range of variability compared with idealized models (Fig 3.3A), and we find clear differences in search efficiency between observed T cells and some idealized models. Brownian searchers are approximately 40% less efficient than observed T cells for unique DC contact (Fig 3.3A and B and Table 3.3 on page 39). In contrast, the power law (Lévy) fit was 30% more efficient than observed T cell tracks, and as expected, more efficient than any other model for unique contacts with DCs. We also modelled a correlated random walk (CRW) as well as a CRW with a lognormal distribution of step lengths (a lognormal modulated CRW, LogMCRW). We show that the idealized search that most closely fits the observed efficiency of experimentally derived T cell search in LNs is the LogMCRW (Fig 3.3B), in keeping with CCDF fits (Fig 3.2 on page 33). Efficiency is not dependent on placement of DC targets in the model: efficiency measures remain similar across multiple target distributions and degrees of clustering (Table 3.3 on page 39). Thus, LogMCRW is not only the best description of the step length distribution, but also the best efficiency match for unique contact T cell search in LNs.

Our simulation of unique target search also gives a quantitative estimate of the contribution of different types of T cell movement to search efficiency (Table 3.3 on page 39). Correlation in angles of T cells increases the search efficiency by 10%

(from -42% for Brownian without correlation to -28% for CRW; -17% for lognormal to -7% for LogMCRW). The heavy-tailed step lengths contributed a 20% increase in efficiency (-42% Brownian to -17% lognormal). These results show that T cell motion is a complex mix of multiple motility parameters that contribute to overall T cell search efficiency.

In addition to unique antigen search, multiple DC contacts by T cells contribute to T cell activation and may also be required for survival (Feuillet et al., 2005; Hochweller et al., 2010; Gérard et al., 2014). Interestingly, we find that the efficiency of total contacts is reversed from that seen for unique contacts (compare Fig 3.3B and D, Table 3.4 on page 39). Brownian searchers made the greatest number of total contacts, while power law (Lévy) searchers made the fewest total contacts (Fig 3.3D). Brownian searchers tend to resample the same locality and are therefore more thorough in their search at the cost of reduced search extent. In contrast, superdiffusive heavy-tailed searchers leave DC clusters more quickly and their total contact rate falls, increasing extent at the cost of thoroughness. Again, LogMCRW is closer to observed data than the other simulated patterns, and it successfully balances total contact rate with exploration of new DC clusters (Fig 3.3D).

We also performed a statistical bootstrap analysis in which search tracks were generated by sampling uniformly from all observed track speeds and turning angles (Feuillet et al., 2005). While the efficiency of total contacts for bootstrap tracks is statistically indistinguishable from observed T cells, bootstrap tracks are 12% less efficient than observed cells in unique contacts (Fig 3.3B and Table 3.3 on page 39). Thus, individual T cell tracks confer efficiency for unique DC target search that is lost when the steps within a track are randomized, suggesting that there is underlying heterogeneity in T cell tracks that increases T cell search efficiency.

3.6.4 Naïve T cells show heterogeneity in movement patterns

To assess potential variation in T cell motility, we analyzed differences in speeds across individual T cell tracks. We find that the distribution of speeds is highly skewed for cells with lower mean speeds, but there is less skew for cells with high mean speeds (Fig 3.4A). The fastest cells (mean speeds $> 15\,\mu m/min$, Fig 3.4D) produce more symmetric speed distributions as demonstrated by the low skew and kurtosis. Also, distribution fitting of speeds shows that the speeds are now best fit by Gaussian and Maxwell distributions (Table 3.5 on page 40). In contrast, slow cells (mean speed $< 5\,\mu m/min$, Fig 3.4C) have a heavier tailed distribution of speeds as shown by skew and kurtosis with lognormal remaining the best fit (Table 3.5). This is not due to the number of data points available at high speeds, as skew decreased even at the speeds with the highest number of data points (Fig 3.4B). However, slow and fast are not discrete populations, as a mixed Gaussian cluster analysis shows no evidence of discrete populations defined by mean speed and variance (Fig A.8 on page 184). These results suggest that T cells exhibit a continuum of movement patterns within LNs, leading to different types of searches: slow moving cells show a heavy-tailed distribution while faster moving cells are more Brownian.

3.6.5 "Hotspots" in the LN environment show differing patterns of T cell motion

The variation in movement shown in Fig 3.4 on page 35 suggests that T cells may alter their search pattern in response to environmental cues. Our previous work shows that altering movement in response to environmental cues can enhance search effi-

ciency (Letendre and Moses, 2013; Hecker and Moses, 2015). Extending our previous work in (Fricke et al., 2015), we analyze T cells in LN to identify whether T cells movement changes within local microenvironments of the LN. To do this, we identify whether there are locations in the LN that are visited by T cells more frequently than predicted by a null model. We analyzed each observation field separately; each field was discretized into cubes of 20 µm per side, which is approximately twice the diameter of a naïve T cell. We used the LogMCRW simulation we described earlier as a null model (for details of [*sic:* the] null model see Materials and Methods, Fig A.12 on page 188). We identified spots that were visited by T cells in the simulation null model and then identified spots that were visited by T cells from actual experimental data. We found that experimental T cells visited certain spots at significantly higher frequency than the null model (see Fig A.12 (Fricke et al., 2015)). Spots that were visited at a frequency 2σ higher than the null model were called hotspots (examples shown in Fig A.13 on page 189). Hotspots were observed in 37 of the 41 observation fields. The null model results in only 2.73% of visited locations being hotspots (as expected given that we identify hotspots as those visited 2 standard deviations above the mean, Fig A.12); in contrast, in empirical observations, 10.51% of locations from observed experimental data are hotspots. We also find that our null simulation predicts 32% tracks will visit hotspots but our observed tracks show that 51% of observed tracks visit hotspots. These data all support the hypothesis that hotspots exist in empirical observations.

We define *hot tracks* to be T cell tracks that intersect with hotspots and cold tracks to be those that do not. Hot tracks have median speeds that are significantly higher than cold tracks with hot track speed at 7.27 µm/min and cold tracks at 4.25 µm/min (median speed is 37.4% greater for hot tracks than for cold, p-values

$<< 10^{-3}$ Mann-Whitney U test). We also find that the step length distributions of hot tracks have a significantly lower skew and kurtosis compared to cold tracks (Table 3.6 on page 40), indicative of more Gaussian distributions in hot tracks. Furthermore, though the step lengths of hot tracks and cold tracks are both best fit by lognormal PDFs, the Gaussian and Maxwell distributions are nearly as good for hot tracks (Fig 3.5A and B and Table 3.6 on page 40). These results show that T cells that visit hotspots exhibit different, and more Brownian movement, suggesting that they search more thoroughly than T cells that do not visit hotspots.

The presence of hotspots suggests that a microenvironment within the LN might modify T cell behavior. To show T cell adaptation within LNs, we ask whether hot tracks (T cells that have visited hotspots) behave differently in hotspots versus other locations within the LN (cold spots). We find that T cells from hot tracks spend more time in hotspots than in other locations (cold spots), with T cells spending a median of 5.36 time steps in hotspots compared to 4.5 in cold spots (p-values $<< 10^{-3}$ Mann-Whitney U test, Fig 3.5C). T cells that visit hotspots are found in those hotspots between 13.3% and 23.2% (95% confidence interval) more often than they are in other LN locations, i.e. cold spots. Thus, hotspots are visited by more T cells than can be explained by chance, the T cells that visit those hotspots move differently than those that dont, and T cells spend more time in hotspots than in other locations; all suggesting that T cell movement changes in response to the LN environment.

3.7 Discussion

T cell activation depends on interactions between T cells and antigen-bearing DCs in secondary lymphoid organs including LNs (Miller et al., 2004; Bousso and Robey, 2003). In this study, we quantify the movement of T cells within LNs, and how efficiently they encounter DC targets (in terms of both unique and total contacts). We use quantitative analysis and computer simulations to show that a search strategy that employs both correlations in successive turning angles and a lognormal distribution of speeds is most representative of observed T cell motion, which we call a LogMCRW. However, T cell motion does not perfectly fit any simple parametric model, and different types of motility are observed depending on where the T cell is and how fast the T cell moves.

Accurate characterization of T cell movement is important because motility determines the timing of other immune processes downstream of T cell activation. Several groups have published models of how T cells interact with DCs in LNs. Mirsky at al. (Mirsky et al., 2011a) provide a review. Recent data also suggests that motility can affect both T cell recirculation (Textor et al., 2011) and T cell dwell time leading to activation especially when detecting rare antigen (Gérard et al., 2014; Textor et al., 2014). Different studies employ different models of T cell motion due to the lack of precise understanding of how T cells move. For example, some models assumed Brownian movement while another assumed a CRW with a Gaussian distribution of steps and speeds, and yet another uses tracks bootstrapped from empirical data (Beltman et al., 2007; Linderman et al., 2010; Celli et al., 2012; Gérard et al., 2014). Our results show that the LogMCRW pattern of motion not only fits the experi-

mental data, but also most faithfully reproduces the modelled search efficiency of observed T cell movement.

We use an agent-based model to compare empirical T cell movement to idealized simulations. These simulations demonstrate that simulated Lévy walks overestimate real T cell search efficiency (for unique DC contacts) while the Brownian walk, CRW, and bootstrap tracks underestimate it. The reverse is true for total contacts. A lognormal distribution of steps combined with correlation among steps (LogMCRW) best represents empirical T cell search efficiency for total and unique contacts.

We identify and quantify three mechanisms that increase T cell search efficiency for unique targets: 1) heavy-tailed step lengths (comparing lognormal versus Brownian search accounts for 20%); 2) directional correlation (comparing lognormal vs. LogMCRW accounts for 10%); and heterogeneity among T cells (comparing bootstrap to observed accounts for 10%) (Fig 3.4 on page 35 and Table 3.3 on page 39). Thus, computational models allow us to quantify the contribution of a variety of factors to T cell search efficiency.

In our study, we thoroughly analyze the motility of naïve T cells in LNs in the absence of antigenic stimulation. Our results largely agree with a recent study by Banigan et al. also showing persistent directional movement for 3-4 minutes by naïve T cells (Banigan et al., 2015). T cells have previously been shown to move in streams, which may correspond to the persistence in movement. Persistence may also reflect cells following a path of least resistance or intrinsic regulation of cell movement, for example, the time required to form a leading edge.

In contrast to Banigan, we find a lognormal distribution of T cell steps and show that the heavy tailed distribution of step lengths is important for search efficiency. Banigan et al. also suggested that modeling T cell movement using 2 subpopulations

may be a more faithful reproduction of T cell movement in LN (Banigan et al., 2015). Our data does not support the existence of 2 subpopulations of T cells. Rather, we find that there may be subregions (hotspots) within the LN that leads to differences in T cell search behavior. T cell motion near hotspots is less directionally persistent and more Brownian (Fig 3.5 on page 36). These results demonstrate that T cells react to their environment, and more specifically, they suggest that T cells that visit hotspots stay longer and thus search more thoroughly at those hotspots.

The identity of hotspots remains to be determined. It is possible that hotspots are locations of DCs or high endothelial venules from which T cells enter the LN. T cells that search areas with DCs more thoroughly may have more repeated contacts with the same DC as well as contacts with more DCs within the same area, enhancing the potential for productive T cell interaction with DCs presenting cognate antigen. One potential mechanism for hotspots is chemokine production by DCs, although there is no direct experimental evidence for this. Another possibility is that hotspots may reflect an underlying structure such as the fibroblastic reticular cells that may form a network that guides T cell movement (Bajénoff et al., 2006). However, the distribution of our hotspots does not obviously reflect any network structure. Others have tested the potential role of a network on T cell search efficiency (Donovan and Lythe, 2012; Graw and Regoes, 2012) and found that the presence of a network has little impact on T cell search efficiency.

Upon activation by cognate antigen, T cell motility within the LN changes, T cells slow down over a period of several hours and begin to form long lived interactions with DCs, essentially ending the search phase (Bousso and Robey, 2003; Mempel et al., 2004). Effector T cells then exit the LN and enter peripheral sites of inflammation. Effector T cell motion in the brains of *Toxoplasma gondii* infected animals was shown

to be a generalized Lévy walk based on displacement analysis (Harris et al., 2012). This differs from our findings that T cells in LNs do not fit a Lévy walk. The difference between our findings and those of Harris et al. may result from intrinsic differences between naïve and effector states. Another possibility is that differences between the tissues that the T cell resides in, for example, the LN for naïve T cells or the brain for effector T cells, contain structural and chemical variability leading to different motility.

As expected, our simulation shows that Lévy searchers are efficient at finding rare targets, but Brownian motion is more efficient when measuring total contacts. These results show that biological context may be important for T cell search efficiency: in the search for rare and unique antigens, the heavy-tailed search is more efficient. However, in situations where high numbers of DC contacts may be important for T cell activation and potentially survival, Brownian motion has an advantage. The observed T cell motion appears to combine the best properties of each, utilizing multiple modes of motility to achieve efficiency in different contexts.

Previous studies have used modeling to reproduce experimental results, and we use this approach to show that the LogMCRW statistical model captures immunologically important properties of T cell search. Similar to empirically observed T cell movement, combining multiple features of random search in the LogMCRW balances search over a wide spatial extent to find unique targets, with thorough search that allows repeated contacts within a cluster. In addition, we extend our use of modeling to identify novel features of the biology underlying T cell movement in LNs. Because the LogMCRW is a good estimate of search efficiency, it also provides a useful null model with which observed T cell motion can be compared, revealing that T cells move differently in different locations in the LN. Thus the statistical model

and search efficiency simulations not only characterize cell movement and provide estimates of search efficiency, they can also be used to reveal the complexity of T cell motility.

Indeed, comparison to our null model reveals non-random T cell movement which may indicate change in response to some feature of the LN. We find that T cells respond differently to specific microenvironments within the lymph nodes, which we call hotspots. The presence of hotspots suggest that, like foraging animals, T cells may respond to features of their environment in order to guide their search (Gordon, 1994; Viitala et al., 1995).

Prior work has characterized the movement of foraging animals using both CRW and Lévy walks. Lévy walks in particular have been suggested as optimal to maximize foraging rate (Viswanathan et al., 2002; Mempel et al., 2004) [*sic:* Memple is an incorrect citation for this]. Our work suggests that in order to balance maximizing repeated (total) contacts with maximizing new (unique) contacts, the LogMCRW may be more effective. More generally, walks with heavy tailed step length distributions and correlation among turning angles may be most effective at balancing the thoroughness and extent of search. In foraging animals as well as searching T cells, natural selection may opt for movement that is effective in a variety of circumstances, even if that movement is difficult to describe analytically.

T cells provide a unique window into biological search strategies because so many searchers can be visualized rapidly in relatively intact natural conditions. Such movement patterns can be included in agent-based models, even if they are not easy to present in closed form equations. Our data suggests that the LogMCRW strategy might be a better approach than either Brownian or Lévy walk in situations that need to balance repeated contacts with already-found targets and discovery of new

items. Additionally, T search for patchily distributed DCs (Mempel et al., 2004) in the LN may demonstrate response to cues, similar to other collective foragers such as ants collecting patchily distributed resources in natural habitats (Gordon, 1994).

In contrast to previous assumptions about simple random motion, our analysis shows that T cell movement in lymph nodes is complex, and involves correlation, variation in step lengths, and heterogeneity in response to local environments. The deviation from idealized models reflects the immunological need to balance the spatial extent and local thoroughness of search. The complex movements of T cells in LN provide a window into biological search strategies and how natural selection may balance multiple objectives in a variety of biological contexts.

3.8 Materials and Methods

3.8.1 Ethics statement

The protocol was approved by the IACUC at the University of New Mexico (protocol # 10- 100487). The breeding and maintenance of mice used in this research conform to the principles outlined by the Animal Welfare Act of the National Institutes of Health. All efforts were made to minimize suffering with use of ketamine and xylazine when appropriate. Euthanasia was performed by isofluorane overdose.

3.8.2 Mice

C57BL/6 mice were from Jackson Laboratories (Bar Harbor, ME). All mice were bred and/or maintained in a specific pathogen-free condition in barrier facilities

(Albuquerque, NM) and conform to the principles outlined by the Animal Welfare Act and the National Institutes of Health guidelines.

3.8.3 T cell observations using two-photon microscopy

Lymph nodes were prepared according to the protocol described previously (Allenspach et al., 2001; Bousso et al., 2002; Letendre et al., 2015; Matheu et al., 2007). T cells were purified by nylon wool or by negative selection using the pan-T cell kit (Miltenyi Biotec) as previously described by Cannon et al. (2013) and purified T cells labeled with either $1\,\mu mol\,dm^{-3}$ CFSE (Invitrogen) or $5\,\mu mol\,dm^{-3}$ CMTMR (Invitrogen, Carlsbad, CA). 5 to 10×10^6 labeled T cells were injected I.V. into recipient mice and inguinal lymph nodes were removed 15-/8 hours later and imaged using two photon-imaging.

Imaging experiments were performed using either a workstation with a Bio-Rad Radiance 2000 scanner mounted on an Olympus upright microscope with a chamber at 37 °Cor a 2-photon microscope in the Fluorescence Microscopy Facility in the UNM Cancer Center with a mode locked Ti:Sapphire infrared laser (Coherent Ultra II; tunable from 680-1080 nm; avg. power 3.5 W) for multiphoton fluorescence excitation on a Zeiss Axiovert 200 stand. For the Bio-Rad 2P, explanted lymph nodes were placed on a glass coverslip in the chamber. The sample is perfused with a 37 °C solution of DMEM (phenol red free, Gibco) bubbled with 95% O2 and 5% CO2. T cell motility within a lymph node was monitored in the T cell area at a minimum of 50-70 µm below the surface of the node. For the Zeiss 2P, the microscope stand is a Zeiss Axiovert 200 with motorized XY stage and IR-corrected long working distance objectives (25X:multi-immersion and 40X:water immersion) and image acquisition

via a Zeiss LSM510 scanhead. Ex-vivo tissue and organs are maintained during microscopic observation in a stage microincubator system (LCI-Live Cell Imaging) equipped with heating, humidity, CO2 atmosphere and perfusion. Explanted lymph nodes were placed on a glass coverslip in the chamber. The sample is perfused with a 37 °C solution of DMEM (phenol red free, Gibco) bubbled with 95% O2 and 5% CO2.

For 4D analysis of T cell motility, multiple stacks in the z-axis (z step = 3 µm) were acquired every 15-20 s (depending on the number of z stacks acquired) for 15-40 min, with an overall field thickness of 40-60 µm. Cell motility was analyzed with Imaris software (version 6; Bitplane). Tracks that lasted fewer than 3 time steps (duration filter in Imaris) were not taken into account in the analysis. Length filter (threshold of 17 µm= 3 times the diameter of the cell) Displacement^2 filter (threshold of 300 µm = 17 µm × 17 µm) were also used to discard tracks of non-motile cells. Videos were made by projecting the 4D information along the z-axis in a single plane.

The observation area covers approximately two thirds of the T cell zone of the lymph node. Cell motility was analyzed with Imaris 6.0 (Bitplane AG, Zurich, Switzerland). The point sequences generated by Imaris were used to create position vectors joining adjacent cell locations (sample tracks Fig A.1 on page 178). The Euclidean norm for each vector was calculated and divided by the time resolution to produce speeds.

3.8.4 Distribution fitting

Following Fisher (1925) we use maximum likelihood estimation (MLE) to parameterize candidate PDFs. We fit probability model parameters using cumulative distri-

bution functions (CDF), rather than by binning data which has been shown to bias conclusions about random walk distributions (Goldstein et al., 2004). We define a step as a vector of T cell motion that does not deviate beyond 15 from the original direction (see Fig A.9 on page 185 for analysis of threshold dependency).

Five PDF models (lognormal, Maxwell, Gaussian, exponential, and power law) for step length and speed were selected for analysis based on a combination of their negative log-likelihood scores, their importance in other biological processes, and their previous use in modeling T cell movement. Our selection of the relative goodness of fit (GoF) of each candidate PDF to empirical data was evaluated using likelihood functions, Anderson-Darling (AD), Bayesian information criterion (BIC), corrected Akaiki Information Criterion (AICc), and the Kolmogorov-Smirnov (KS) test.

Following (Clauset et al., 2009), we fit power laws using MLE and with the power law PDF: $\mathrm{P}(x) = \frac{\mu-1}{x_{\min}} \left(\frac{x}{x_{\min}}\right)^{\mu}$, where $x_{\min}$ is the smallest observed value, $\mathrm{P}(x)$ is the probability of x occurring , and μ is the estimated parameter. We used the $x_{\min}$ value with the best KS score of all possible choices as an estimator of the beginning of a power law tail. The percentage of positions in a track in the power law tail gives us a measure of the quality of the power law fit. Using this measure we show that a power law fit to the population of observed steps excludes 94% of the data (Fig 3.1F and H).

3.8.5 Autocorrelation and cross-correlations

Velocity autocorrelations were calculated following Qian et al. (1991) and Tarantino et al. (2014). The autocorrelation function is the ensemble mean for the $n-1$ possible delay times given the n vectors defining a T cell track. The result is a

measure of how much T cell direction depends on previous directions as a function of time delay. Our use of autocorrelation is distinct from the analysis of periodic velocity vector magnitudes by Beltman et al. (2007), but the [*sic*] similar to that done in Banigan et al. (2015). Letting $v(p_k(t))$ be the unit velocity vector at time t belonging to the kth path, we defined the cross-correlation function, $\mathrm{C_{cross}}$, to be: $\mathrm{C_{cross}} = \langle v(p_k(t)) \cdot v(p_m(t)) \rangle, \forall k, m$ where p_k and p_m are T cells paths. This measures the step angle dependence between T cell paths at the same moment in time, that is, a measure of drift due to global effects on the observation field.

3.8.6 Mean squared displacement

Mean squared displacement (MSD) coefficients, commonly called the α exponent (Viswanathan et al., 1996; Codling et al., 2008; Bartumeus et al., 2005), were calculated using least-squares polynomial fit by numerically solving the associated Vandermonde matrix (Von Mises and Geiringer, 1964) and fit quality assessed with the r^2 measure. Parametric and linear fits were also made to mean displacement. In Fig 3.1A we present only the first 10 minutes of observation (as was done in Mempel et al. (2004); Worbs et al. (2007) and Friedman et al. (2010)) at which point the curve reaches its first stationary inflection which in Tarantino et al. (2014) is indicative of unconstrained motion and therefore appropriate for determining α. In addition, in this study few tracks persist beyond 10 minutes and so the MSD signal also becomes dominated by noise (Fig 3.1A top).

3.8.7 Heterogeneity

We tested for heterogeneity by comparing track speed skew (Fig 3.4 on page 35) and AIC evidence ratios as a function of mean speed. The sample skew of the distribution of speeds was calculated using the method of moments applied to a mean speed sliding window of width 0.125 µm/s progressing in 0.1 µm/s increments.

3.8.8 Search efficiency simulation

The simulation to test T-DC interaction efficiency was implemented as a continuous (floating-point) 3D model written in C++. Boost libraries (Boo, 2013) were used to generate variates drawn from model PDFs. Because the clustering and density of targets can influence which movement types are most efficient, we replicated the estimated density of DCs and varied the degree of clustering in our simulations.

We use LN DC density of 2-5% as determined in (Hochweller et al., 2010) to calculate a target DC density of 3.17×10^{-5} targets/µm^3. Our observed fields have an average volume of 6.3×10^6µm^3. We scale the number of targets as a function of field volume in order to maintain the same target density between simulation fields. DCs were clustered into groups of 10 and were uniformly distributed within spheres defining a cluster. By varying the sphere radius, we controlled the degree of clustering from uniform to highly clustered. A 3D version of the Hopkins statistic (Hopkins and Skellam, 1954) was used to measure the resulting non-uniformity of target placement (Tables 3.3 on page 39 and 3.4 on page 39). In the Hopkins statistic scores range from 0 to 0.5 where 0 is highly clustered and 0.5 indicates no clustering (Fig A.6 on page 182).

T cell tracks were observed and recorded as 3D coordinate sequences within a bounding box defined by the visible section of the ex vivo lymph node. Idealized models (Brownian, CRW, Power Law, etc.) of search were parameterized by the speeds and turning angles estimated from observation (see Distribution fitting). Searchers in the idealized model start at the same initial positions as the observed T cells, and exist in a volume equal to the observed field volume. Candidate search patterns were generated for each of the 41 observation fields.

Our efficiency measure is the number of targets found divided by the sum of the time used by searchers. Since we modelled walks rather than flights (i.e. speeds are finite) the sum of D(k) for all simulated tracks k was limited to the total distance travelled by observed T cells. Therefore the average velocity of the population of searchers is kept within the observed range. Based on an assumed radii of 5 µm for DCs and T cells, targets were marked as discovered if a searcher track passed within 10 µm of a target point. We define two versions of the efficiency measure, one that increments its output value only when a target was not previously detected by that searcher, and another that increments for all targets found. These two versions allow us to record unique contacts and total contacts (Fig 3.3 on page 34).

The simulation measures the target encounter rate and determines, using the Mann-Whiney test, whether the candidate search models search efficiency is significantly different from that observed in T cells. We use the Mann-Whitney test because the observed and simulated distribution of efficiencies is non-Gaussian. Simulations were replicated 100 times per field, producing 4,100 efficiency data points for each search model. The entire process was repeated 10 times in order to generate confidence intervals for the simulation; in all this results in 41,000 efficiency samples.

3.8.9 Identifying hotspots and hot tracks

In order to test whether the environment within LNs influences T cell movement we extend an analysis begun in (Fricke et al., 2015). Fields were discretized into 8000 μm^3 cubes (the length of a cube is 20 µm, approximately twice the diameter of a T cell). We use the LogMCRW simulation as a null model and record the number of times a location is visited by unique T cells in simulation (repeated 10 times). We use a 2σ (two-standard deviation) threshold for determining which locations are visited more frequently in the observed fields than expected and call these hotspots. This is repeated for each of the 41 individual observational fields. All other visited locations are called cold spots. A comparison of the number of hotspots in simulation and in the observed data gives an indication of how much behavior is not captured by the simulation.

We define hot tracks to be T cell tracks that visit hotspots and cold tracks to be T cell tracks that do not. We also examine the number of visits by hot tracks to cold spots and hotspots. We also examine the distribution of step lengths and speeds for hot and cold tracks.

For additional information on methods, see supplementary materials and methods (Appendix A).

3.9 Acknowledgments

Thanks to François Asperti-Boursin for data collection, T cell tracking, Imaris analysis and discussion. Thanks to Christian Gunning, Helen Wearing, David Ackley, Vasudev Kenkre, Stephanie Forrest, Aaron Neumann, Deborah Gordon, Brianna Mulligan, and Grant Lythe for helpful discussion and reviews of this paper. Thanks

to Genevieve Phillips and Becky Lee of the UNM Cancer Center Fluorescence Microscopy Facility as well as John Connor and Denis Bragin of the BRAIN Imaging Center for help with 2-photon microscopy.

3.10 Software

The software used in this chapter is available at:

`https://github.com/BCLab-UNM/TcellAnalysis`

`https://github.com/BCLab-UNM/TcellSearchSim/releases/tag/18`

Chapter 4

Spatial Association of T cells and Dendritic Cells

> *Von Neumann told me, 'You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one really knows what entropy really is, so in a debate you will always have the advantage.'*
> — Claude Shannon quoted in Tribus, M. and McIrvine, E. C. (1971). Energy and information. *Scientific American*, 225(3):179–188

4.1 Author Contribution Statement

I am the lead author of this chapter under the supervision of Melanie Moses (Associate Professor, UNM Computer Science, Biology, and Santa Fe Institute) and Judy Cannon (Associate Professor, UNM Pathology, Molecular Genetics and Microbiology), with major contributions from Justyna Tafoya (Undergraduate, UNM Mathematics and Statistics) and Janie Byrum (UNM Molecular Genetics and Microbiology). Janie produced the two-photon movies of T cells and dendritic cells

under the direction of Judy Cannon. Justyna wrote software to apply the mutual information calculation to the two-photon movies in various combinations.

4.2 Overview

Through the efforts of our collaborators at the UNM-HSC Department of Molecular Genetics and Microbiology, we now have access to two-photon recordings of T cells and their dendritic cell (DC) targets in the same experiment. This allows us to analyse the information theoretic relationship of T cells to DCs and so directly measure interactions without having to model DC placement using estimates from data published by other labs.

Though recruitment of T cells by DCs may seem to have obvious benefit, this is not always the case. Mechanisms that concentrate T cells and DCs together may serve to decrease interaction rates due to crowding effects. We have observed something similar in our simulations of ant foraging. The use of information to guide ants to seeds can lead to congestion that hinders resource collection unless the signal strength is balanced.

We measure the mutual information between images of T cells and DCs to determine to what degree the location of DCs and T cells influence each other. This begins to address directly the current assumption that naïve T cells perform an unguided search without regard to DC placement. An alternate hypothesis is that T cells are guided to DCs by structures within lymph nodes (LNs). Two candidate structures are the fibroblastic reticular cell (FRC) network (Miyasaka and Tanaka, 2004) and high endothelial venules (HEVs) (Donovan and Lythe, 2012).

An alternative strategy would be to determine the locations of T cells and DCs, and directly measure the distance between the cells. Many algorithms exist for measuring the clustering of discrete objects in a plane that use variations on this approach, and they would provide a characteristic distance. A barrier to this approach, however, is that the two-photon microscopy (2PM) data does not provide reliable location data for DCs. That is, due to the irregular shape of DCs (from which they derive their name), and their tendency to form large clusters it is difficult to discretise the fluorescence into individual cells. Conceivably a technique could be applied that measures the distance between pixels instead of cells, but the fluorescence varies in intensity between DCs and T cells. While investigating the spatial association between T cells and DCs we initially tried applying a threshold to pixels. We discovered that our choice of the pixel colour intensity threshold entirely determined the degree of association that resulted. The benefit of mutual information is that it measures the association of pixel values associated with the presence of T cells and DC parsimoniously. Mutual information is also attractive because it provides a natural baseline at 0 for which the interaction of pixel colours is uniformly random. Deviation from that baseline exactly targets our question of interest: is the association of T cells and DCs random or non-random, or in other words, is information about the location of one cell type available to the other.

4.3 Significance

Our analysis of random search strategies in lymph nodes is predicated on the assumption that nave T cells are not guided to target DCs through long-range interactions.

(Germain et al., 2012) This is in contrast to activated T cells which commonly home to their targets by following chemical gradients.

Our findings in Chapter 3 suggest that there is an over abundance of T cells in some locations in LNs when compared to our LogMCRW model. We have dubbed these locations hotspots (Fig. 3.5 on page 36). This is not surprising given the heterogeneous physical environment within the LN however an alternative explanation is that this over-abundance is due to chemical or morphological guidance by DCs to particular locations.

It has also been suggested that T cells may use structures, as opposed to chemical gradients, within lymph nodes to guide T cells to DCs. The two candidate structures are networks of FRC which, it is hypothesised, T cells follow to find DCs that congregate at certain points on the network. Alternatively, observations of DCs congregating at HEVs, which serve as entry points into lymph nodes for T cells, have led to a hypothesis that T cells sample DCs on entry into the lymph node. This is analogous to a greeting line at a wedding where T cells the guests and the DCs are the wedding couple coming in contact with each guest in turn. Empirical observations of T cells following the FRC network are qualitative and inconclusive. Beyond informing our analysis of T cells movement patterns, i.e. as a purely stochastic process, how they find DCs has medical implications. For example, HIV is known to suppress FRC and T cell interactions (Zeng et al., 2012; Donovan and Lythe, 2016). If T cells use the FRC network for DC localisation this would further suppress the immune system.

In all the above cases we would expect there to be a co-localization of T cells and DCs to a degree greater than that expected by chance. We use information theory to measure the co-localization observed, as described in the next section.

4.4 Methods

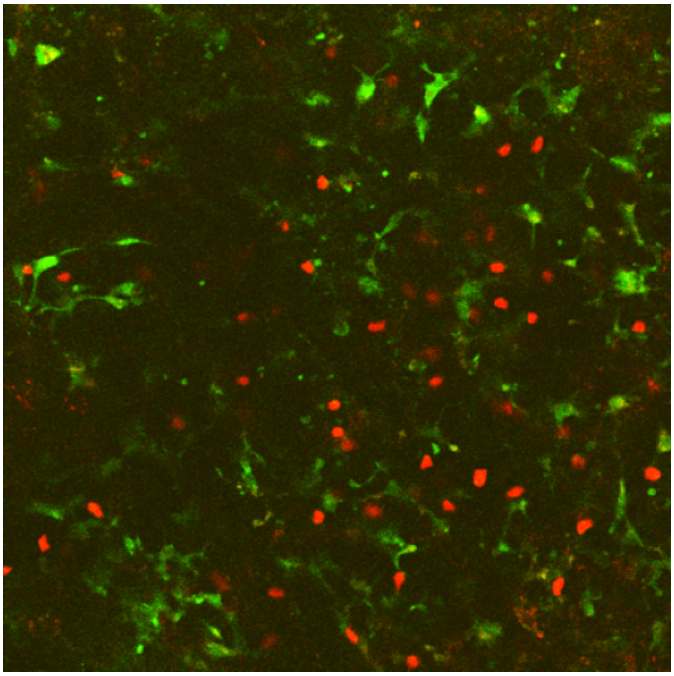


Figure 4.1: 2PM example image. T cells fluoresce in the red channel, DCs in green. Credit: Janie Byrum.

Dendritic cells are tagged with CFSE dye that has excitation/emission wavelengths of 492/517 nm (green). T cells are tagged with CMTMR dye which has an excitation/emission wave lengths of 541/565 nm (red). For the remainder of this chapter, we refer to the red fluorescence as T cells and the green fluorescence as DCs. As shown in Fig 4.3 on page 69 this association does not always appear to hold. Some images contain objects that fluorescence in both the red and green channels. The association between T cells and DCs and the red and green channels for these objects is unknown. These objects persist over time and will tend to increase the observed

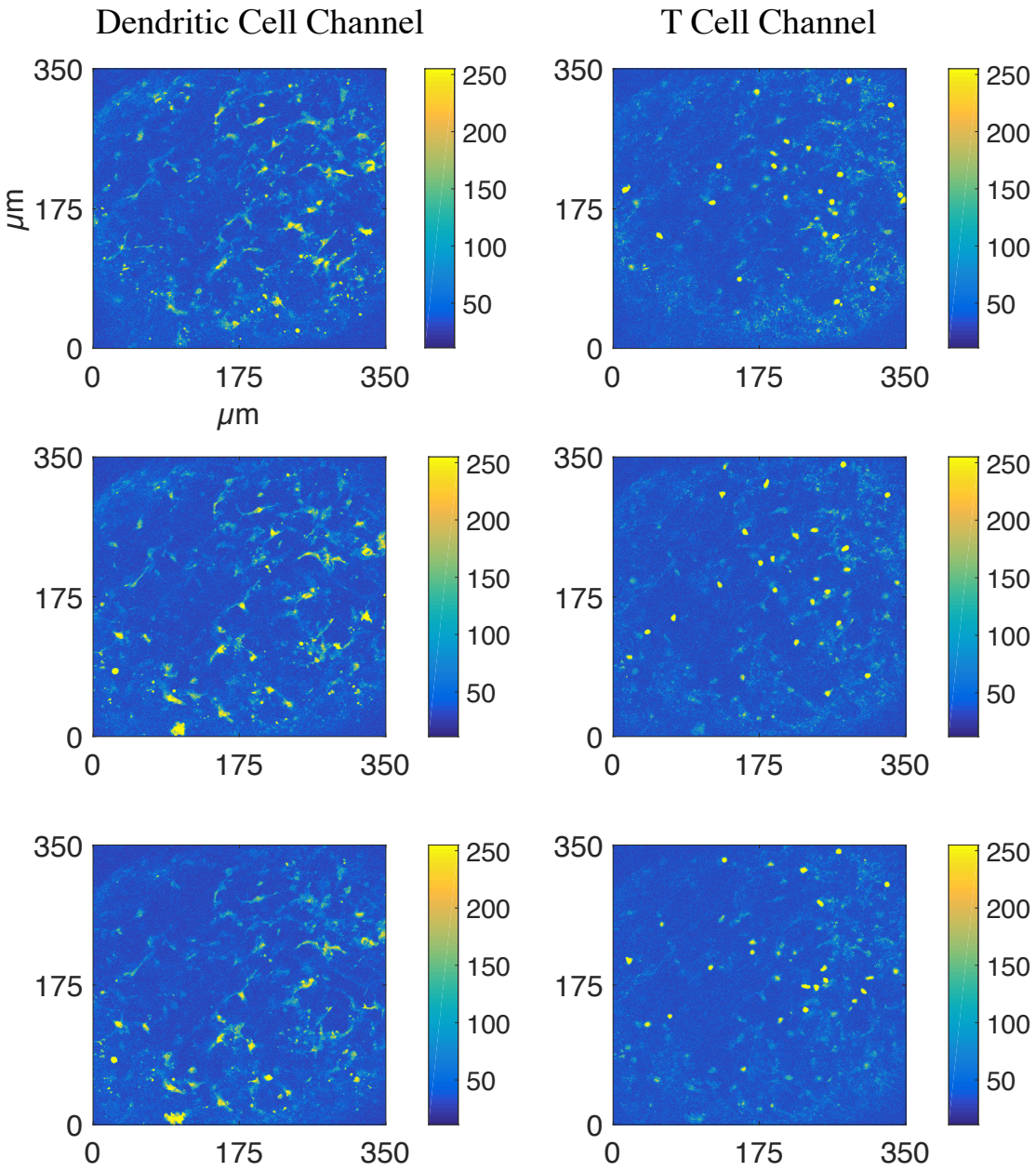
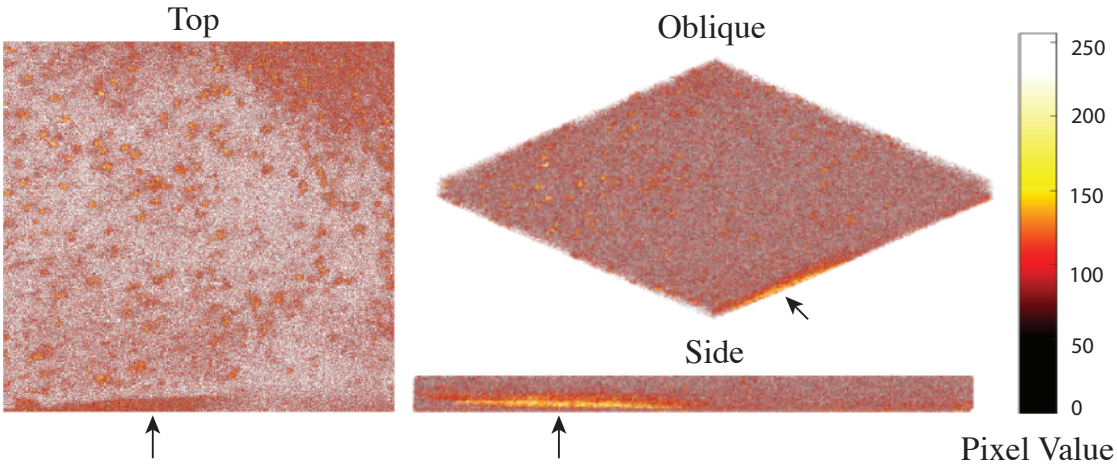
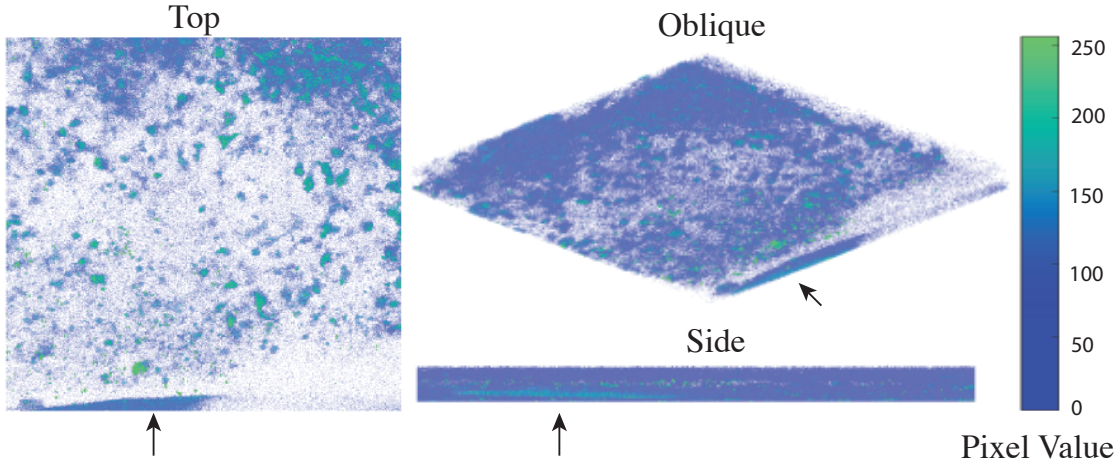


Figure 4.2: 2PM movie example frames (single layer). Left column shows green (DC) channel fluorescence, the right column shows red (T cell) channel fluorescence. T cells are labelled with red fluorescing dye, DCs are labelled with green fluorescing dye. Color bars indicate the intensity of fluorescence.



(a) T Cell Channel



(b) Dendritic Cell Channel

Figure 4.3: **Volumetric rendering of a single 3D 2PM movie frame.** Mouse 2, LN 1, Field 2, Frame 1. (a) Red (T cell) colour channel intensity and (b) green (DC) colour channel intensity. The image contains regions of fluorescence (arrows) that do not appear to be associated with T cells or DCs, for example, the region of intensity in both red (T cell) and green (DC) channels in the lower left hand corner of the 3D frame, possibly indicating a physical object in the frame. Photoelectric noise is visible as a scattering of isolated non-zero pixels. The frame height in the oblique view is not to the same scale as the width and depth since the frame is lower resolution in the z-direction. Pixels with a colour intensity below 50 are transparent in this figure. We use all pixels in our analysis.

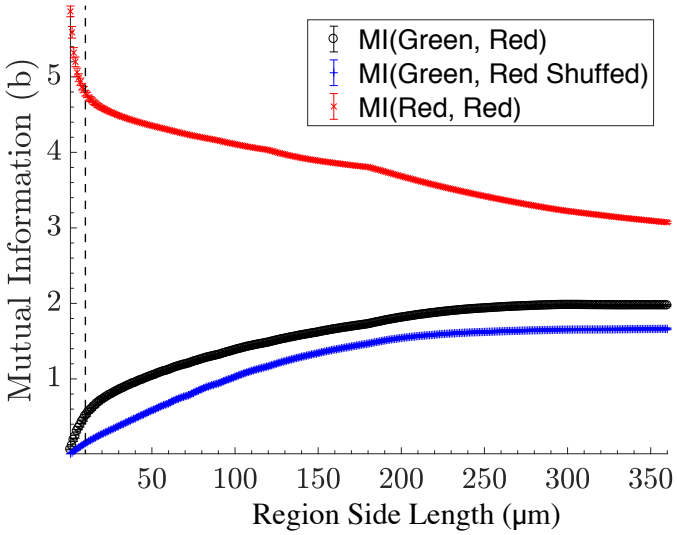


Figure 4.4: **Dependence of Mutual Information on Region Size.** Region size impacts the mutual information calculation. We use square (2D) 10 µm×10 µm regions because this is about the same size as a T cell. The mutual information below this region size decreased, with the limiting case being zero information for regions of 1 pixel. The vertical dashed line is the region size we use (10 µm per side) Mean of 30 frames sampled from 7 experiments. Bars are the 95% CI.

mutual information. Also, the images contain noise with an unknown association to physical objects. This noise follows a Poisson distribution and is due to the photon arrival times at the photodetector device, and variance in photon emission mediated by thermal fluctuations. (Pawley and Masters, 2008). The effect of this noise is to reduce the mutual information between the DC and T cell channels. Fluorescence in the red channel is therefore strongly correlated with the presence of T cells, and green with DCs but the association is not strict. While we refer to the red as T cells and green to DCs these caveats should be kept in mind.

Three-dimensional images are recorded over time. The distribution of DCs in lymph nodes is poorly understood but, as we have shown, the configuration of targets is important to predicting the optimal pattern of search for T cells. We measure the

aggregation of DCs using the Hopkins index (Hopkins and Skellam, 1954). The Hopkins index takes the ratio of k-neighbour distances between targets to k-neighbour distances between targets and a uniform distribution of locations.

In order to determine the extent to which DCs and naïve T cells associate we calculate the mutual information of the green (DC) channel image, G, and the red (T cell) channel image, R, in each frame. Where *entropy* is defined to be:

$$\mathrm{H}(X) = -\sum_{x} p(x) \log_2 p(x) \tag{4.1}$$

and the joint entropy:

$$\mathrm{H}(X, Y) = -\sum_{x} p(x, y) \log_2 p(x, y) \tag{4.2}$$

Entropy is a measure of uncertainty in terms of bits of information. The maximum entropy for a given system occurs when the probability of any particular state value is uniformly distributed. The mutual information $\mathrm{MI}(T, D)$ is,

$$\mathrm{MI}(T, D) = \mathrm{H}(T) + \mathrm{H}(D) - \mathrm{H}(T, D), \tag{4.3}$$

in which, $\mathrm{H}(T)$ and $\mathrm{H}(D)$ are the entropies of the red (T cell) and green images, H(T,D) and is their joint entropy. This technique has been extensively used in medical image registration Pluim et al. (2003). Mutual information provides a measure of the reduction in uncertainty about the red (T cell) pixel values given the values of the green (DC) pixels, and vice versa. The code for calculating these values using Matlab is included in Appendix A on page 176.

We are interested in the spatial association of cells rather than their exact colocation. The mutual information of each pixel in the source image is necessarily zero since DCs and T cells cannot occupy the same physical space. There are pixels that have positive values in the T cell and DC channel in the 2PM images, but this is due to artefacts in the image (Fig 4.3 on page 69). Rather, we map pixel-space into 10 µm *regions*. The regions are large enough to capture pixels that would correspond to a DC and a T cell if they are adjacent, but not small enough to be contained entirely within a T cell. Regions are assigned the average value of the pixels they contain.

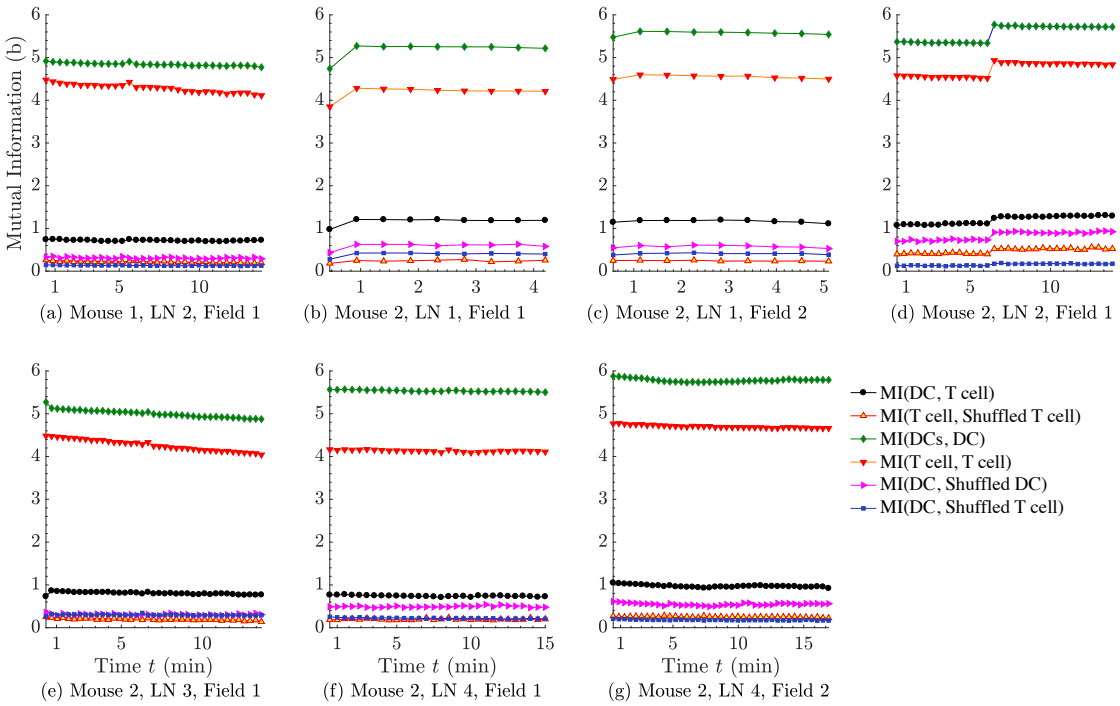
Since $\mathrm{MI}(T, D)$ depends strongly on the internal entropies, $\mathrm{H}(T)$ and $\mathrm{H}(D)$, of the images we compare the value of $\mathrm{MI}(T, D)$ to that of $\mathrm{MI}(T, D')$, where $\mathrm{H}(D')$ is the green signal from a different experiment. $\mathrm{H}(D')$ should be unrelated to T and so provide a reasonable baseline while preserving the DC signals internal entropy. The internal entropies are determined by the distribution of pixel values in the frame.

The calculation of mutual information includes the calculation of the joint entropy (Eq. Eq. (4.2) on the previous page). The calculation of the joint entropy involves building a table containing a count of the number of regions with the same intensity in the T cell and DC frames. Since we are averaging over pixel values the number of possible values increases and the probability of observing the same values in the same region from both channels decreases. This may cause our method to underestimate the association of T cells and DC. In order to qualitatively determine the sensitivity of our method we observe the rate at which $\langle \mathrm{MI}(T_t, T_{t+\Delta t}) \rangle$ and $\langle \mathrm{MI}(D_t, D_{t+\Delta t}) \rangle, \forall \Delta t > 0$, fall as Δt increases, here t is time and Δt is separation in time. That is we plot the disassociation if T cells and DCs from their start location over time, as measured by average mutual information (Fig. 4.6 on page 74). The mutual information falls

as frames are separated by time providing evidence that our method can distinguish between cells in proximity and those that are not.

The analysis of DC clustering has been performed using the Hopkins index. The result is that DCs appear highly clustered indicated by the Hopkins index approaching 1.

4.5 Results



(a) Mouse 1, LN 2, Field 1 (b) Mouse 2, LN 1, Field 1 (c) Mouse 2, LN 1, Field 2 (d) Mouse 2, LN 2, Field 1

(e) Mouse 2, LN 3, Field 1 (f) Mouse 2, LN 4, Field 1 (g) Mouse 2, LN 4, Field 2

Figure 4.5: Mutual information over change in time (individual). 190 images from 7 experiments. 2 mice 4 lymph nodes. Region level mutual information with baseline mutual information between different experiments. T cell and DC co-location is higher than for any of the shuffled baselines and so is non-random.

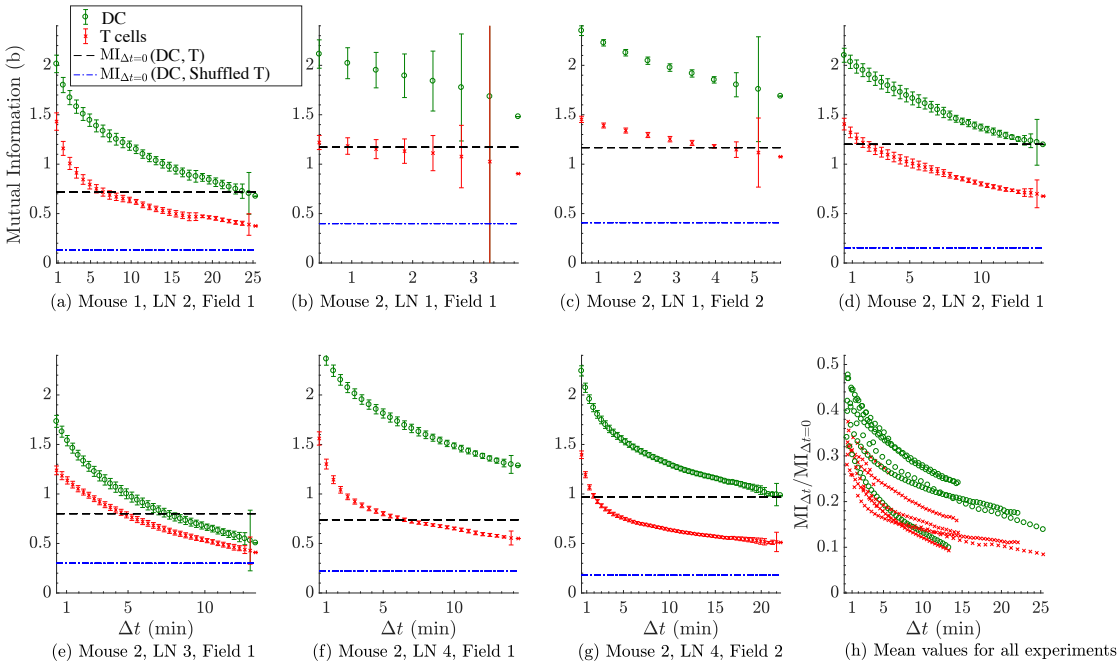


Figure 4.6: Mutual information over time. 190 images from 7 experiments. 2 mice 4 lymph nodes. Region level mutual information with baseline mutual information between different experiments for comparison. Both DCs and T cell mutual information decay over time. The black dashed and blue dot-dashed line correspond to Fig. 4.5.

We use mutual information to measure the correspondence in space between T cells and DCs. However, mutual information depends not only on the spacial correspondence of colour values in the images, but it also depends on the internal entropy of the images being compared. This is because the mutual information between two images cannot exceed the information contained in either of the images. Mutual information provides only a relative measure of similarity and requires contextualization. We measure the mutual information of four reference points in order to provide this context: 1) the green (DC) channel with itself, 2) the red (T cell) channel with itself, 3) the shuffled red (T cell) and green (DC) channels, and 4) the green (DC) channel with the red (T cell) channel from a different experiment. The mutual

information of colour channels with themselves provides maximum values given the internal entropy of the channels. The shuffled and inter-experiment mutual information provide minimum values given the internal channel entropy.

Knowing locations and colour intensity of the red (T cell) fluorescence channel communicates 0.9147 95% CI [0.8870, 0.9424] bits of information about the location of fluorescence in the green (DC) channel produced by dyes associated with DCs. In comparison, the mutual information of disassociated red (T cell) and green (DC) channels, either because they are taken from different experiments or because the position data was randomised through shuffling are 0.2617 95% CI [0.2617, 0.2697] and 0.3789 95% CI [0.3363, 0.4214] respectively. The mutual information of the green (DC) and red (T cell) channels with themselves are 5.3506 95% CI [5.2981, 5.4032] and 4.4224 95% CI [4.3849, 4.4599] bits respectively (Fig 4.5 on page 73).

4.6 Conclusions

Mutual information analysis indicates that T cells and DCs are associated in space to a greater degree than uniformly random associations predict. However, the degree of association appears to be weak, only communicating 0.54 bits of information more than randomly shuffled channels, and 0.65 bits more information than the mutual information of T cells and DCs from different fields. We describe this as a weak association in comparison to the upper reference point in which the green (DC) and red (T cell) channels are compared to themselves (5.3 and 4.4 bits).

Through we characterise the interaction as weak, we have only one example system. Measurement of the mutual information between T cells and antigen-bearing cognate DCs would provide more context. The activation process is known to cause

the association of T cells and DCs and would presumably result in higher mutual information between T cells and DCs. Such an experiment would validate or invalidate the technique and provide mutual information values for a system in which cells are known to associate.

We are also analysing the association of CCR7-knockout T cells. CCR7 is known to be important for T cell signalling, a change in the measured association of CCR7-deficient T cells and DCs may indicate CCR7-mediated naïve T cell-DC interactions.

4.7 Software

The software used in this chapter is available at:

`https://github.com/BCLab-UNM/TcellAnalysis`

Chapter 5

Adaptation of Lévy Exponents to Target Configuration

5.1 Author Contribution Statement

I am the first author of this chapter under the supervision of Melanie Moses (Associate Professor, UNM Computer Science, Biology, and Santa Fe Institute). Judy Cannon (Associate Professor, UNM Pathology, Molecular Genetics and Microbiology) co-authored the immunology sections. Joshua Hecker (Postdoc, UNM Computer Science) provided the iAnt simulator and co-authored the related methods section.

5.2 Publication Notes

Citation: G.M.Fricke, J.P.Hecker, J.L.Cannon and M. E. Moses, "Immune-inspired search strategies for robot swarms," *Robotica* (2016) volume 00, pp. 119. Cambridge University Press 2016 doi:10.1017/S0263574716000382

Copyright: ©Cambridge University Press. License for use in this dissertation is included in Appendix C on page 207.

5.3 Abstract

Detection of targets distributed randomly in space is a task common to both robotic and biological systems. Lévy flights have previously been used to characterize T cell search in the immune system. We use a robot swarm to evaluate the effectiveness of a Lévy search strategy and map the relationship between search parameters and target distributions. We show that the fractal dimension of the Lévy search which optimises search efficiency depends strongly on the distribution of targets but only weakly on the number of agents involved in search. Lévy search can therefore be tuned according to the target distribution while also being scalable. Implementing search behaviours observed in T cells in a robotic swarm provides insight into swarm robotic search strategies; additionally, the flexibility and scalability of Lévy walks may explain why Lévy-like movement has been observed in T cells in different immunological contexts.

5.4 Summary

Detection of targets distributed randomly in space is a task common to both robotic and biological systems. Lévy flights have previously been used to characterize T cell search in the immune system. We use a robot swarm to evaluate the effectiveness of a Lévy search strategy and map the relationship between search parameters and target distributions. We show that the fractal dimension of the Lévy search which

optimises search efficiency depends strongly on the distribution of targets but only weakly on the number of agents involved in search. Lévy search can therefore be tuned according to the target distribution while also being scalable. Implementing search behaviours observed in T cells in a robotic swarm provides insight into swarm robotic search strategies; additionally, the flexibility and scalability of Lévy walks may explain why Lévy-like movement has been observed in T cells in different immunological contexts.

5.5 Introduction

Robot swarms typically consist of many small, relatively simple and inexpensive robotic agents that work collectively toward some common goal (Brambilla et al., 2013). Swarm robotic algorithms are often inspired by biological behaviours that generate emergent collective behaviour from the interactions of robots and their environment (Sahin, 2005). A major research challenge is the development of swarm robotic algorithms that allow effective navigation through complex real-world environments without centralized control (Winfield et al., 2005; Hecker and Moses, 2015).

Foraging is a canonical problem in swarm robotics in which robots have to locate targets distributed in space, and often to transport those targets to some specified location (Winfield et al., 2005). Hecker and Moses (2015) demonstrate an ant-inspired algorithm (the central place foraging algorithm (CPFA)) that is error tolerant, adaptable to different resource distributions, and scalable across swarms sizes. Here we present an immune-inspired search pattern that can be adapted to replace the current, more complex, search subtask in the CPFA. An advantage of a simpler pattern

is that it can be more easily analysed and tuned for maximum search performance given different resource distributions and swarm sizes. The search pattern discussed in this work is simple, efficiently scales with the number of searchers, is robust to error, is adaptable to the distribution of targets, and requires no centralized control.

As robots have become smaller, cheaper, and are increasingly expected to operate in natural environments, designing flexible and error tolerant algorithms for robot swarms has become more important. Swarms tend to be made up of robots built from cheaper components than monolithic robots, which increases the chance of component failure and decreases the accuracy of actuation and sensor input. The small size of swarm robots and operation in natural environments also leads to robot loss. Swarms must be designed to be resilient to individual robot loss and to the effects of robots producing erroneous data.

Robustness is the ability to cope with the loss or malfunction of individuals. In biological systems robustness is promoted by redundancy and decentralized control that avoids single points of failure. *Scalability* is the ability to perform well with different group sizes. Search strategies that can be tuned to optimise performance for different target configurations we call *adaptable*.

The success of robot swarms searching for targets in an unknown environment depends on the adaptability and robustness of the search strategy employed. Such tasks include surveying planetary surfaces and interplanetary space (Fink et al., 2005), land and sea mine clearance (Weber, 1995), pollution mapping by subsurface robots (Hu et al., 2011), and survivor location in hazardous environments (Birk and Carpin, 2006) as well as military applications.

Sensor and actuation error are crucial motivation for using a stochastic search strategy. Though most robot swarms remain confined to simulation (Brambilla et al.,

2013), error models can only be realistically derived from an embodied physical system. This leads us to use a swarm of iAnt robots for which we have a realistic error model to explore the stochastic search strategies observed in T cell search. The hardware error model was developed from search trials using 6 hardware robots. We modify the simulation described in Hecker and Moses (2015) that allows us scale the immno-inspired search pattern to 32 robots. The simulation was written in tandem with development of the physical robots, and tested to make sure that the simulation reproduces the performance of the physical iAnts (Hecker et al., 2013).

An argument that particularly applies to swarm robotics and motivates our use of a random search strategy has recently been made by Ackley et al. (2012). Computational processes that guarantee correctness are no longer tenable as systems increase in size and complexity; rather, large distributed systems should sacrifice determinism for robustness. Von Neumann (1951), shortly after the development of the ENIAC, recognized that computing systems would need to take on aspects of biological systems that are both robust to the failure of individual parts, and which are inherently stochastic, if computing systems were ever to become truly scalable.

5.5.1 Search in Immunology

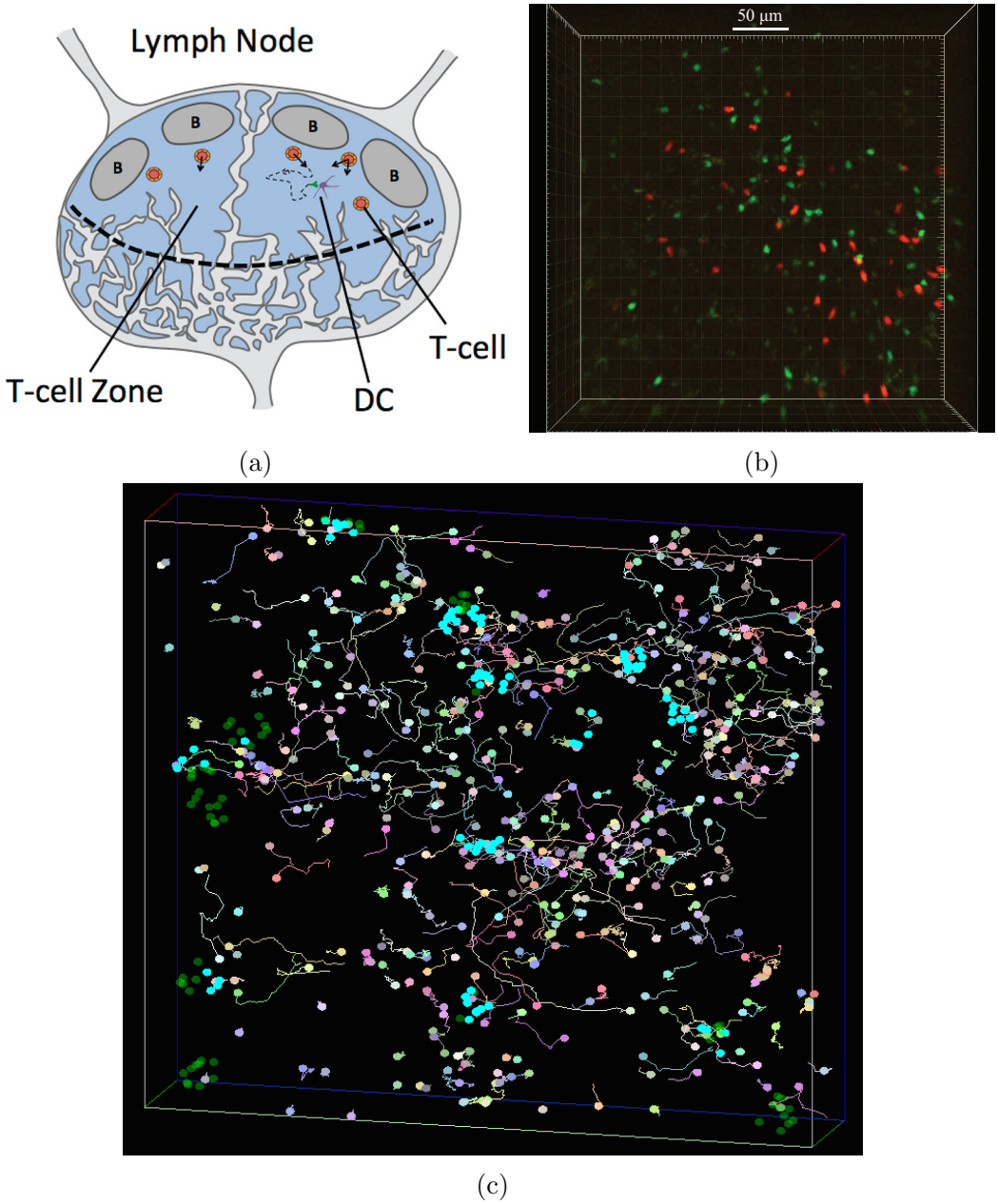
The requirements for search by robotic swarms are similar to those encountered by biological searchers, such as T cells in search of pathogens. Cells of the immune system are under selective pressure to rapidly find and eradicate pathogens. T cells search collectively but without centralized control (Groom et al., 2012; Sung et al., 2012; Textor et al., 2014). Cells have limited ability to sense and navigate, and they do so based on stochastic processes that are prone to error.

As in swarm robotics, the search patterns of T cells and other cells in the immune system must be robust, efficient, and scalable. Detecting and destroying pathogens early during an infection, before they become unmanageable, relies on efficient search patterns. The immune system must also be robust to the loss of individual cells and to malfunctioning cells, which in the worst case do not just result in failure to eradicate disease but in the destruction of the host's own cells.

T cells are crucial to the adaptive immune response: activation of T cells leads to B cell production of antibody as well as antiviral immune responses. T cells search for dendritic cells (DCs) carrying parts of pathogens (antigens) which they display in the T cell zone (Figure 5.1a on the next page) of lymph nodes (LNs). Interaction of T cells with DCs displaying cognate antigen leads to T cell activation. Activated T cells move to the site of infection, where they search for and destroy cells displaying the antigen that activated them (Alberts et al., 2002). Within LNs each T cell searches a space approximately 10^6 times its own volume. If DCs are randomly located throughout the T cell zone of a LN, T-DC interactions impose a requirement for T cells to search efficiently. Efficient search is crucial to mount a timely and effective immune response against pathogen populations that often expand geometrically (Mirsky et al., 2011b).

When an immune response is triggered, activated T cells proliferate geometrically (De Boer et al., 2001), going in a matter of days from a population on the order of 10 cells to millions. The search strategy must be both adaptable to given different target distributions, and scalable to vast changes in the number of T cells.

Infection can occur anywhere in the body, including but not exclusive to gut, lung, skin, vaginal tract tissue, and the brain. In order to promote an effective immune



(a) (b)

(c)

Figure 5.1: T cell motion was observed in mouse lymph nodes (LNs). The resulting data was translated into a model of search using a Lévy walk. (a) Schematic diagram of an LN in which T cells search for dendritic cells (DCs) (Fricke et al., 2013), (b) Two-photon image of T cell zone. T cells fluorescing in red and green, (c) T cell tracks in a single observation field (1 of 7).

response T cells must be capable of efficient search in many different tissues. Each of these tissues may have different properties, including their effective dimensionality. T cells search an essentially 2-dimensional space in the epithelium (Ariotti et al., 2012), while searchers in the lungs and brain must navigate higher dimensional search spaces. Tissues also differ dramatically in the cell types that are in each tissue, the physical structures that T cells need to migrate through, and the extracellular matrix proteins. T cells show remarkable ability to search and move through multiple environments. Pathogens also exhibit an array of growth strategies and patterns (Lindquist et al., 2004; Miller et al., 2004). T cell search strategies must be capable of effectively finding targets in various spatial configurations and in a variety of environments.

Cells infected with bacteria, parasites, or viruses can all be search targets of T cells in peripheral tissue. In LNs the situation is more complex since only when a T cell finds a matching DC does activation occur. The size of the subset of T cells that match each DC varies by pathogen (Regner, 2001), so it is not clear what proportion of T cells will match target DCs. In addition, the number of DCs bearing cognate antigen is highly dynamic with a small initial population early in an infection but quickly growing (Alberts et al., 2002). While it is important that as many DCs are scanned by each T cell as possible, it is less clear to what degree the total number of DCs scanned by the T cell population affects the immune response. The robot search problem is simpler in that each target is treated as being equally valuable to all robots, more similar to T cell search for pathogens in periphery.

T cell motility is an active area of research and multiple models of search have been proposed, for example, in peripheral tissues (Potdar et al., 2008; Harris et al.,

2012) and in LNs (Banerjee et al., 2011; Donovan and Lythe, 2012; Fricke et al., 2013; Gérard et al., 2014; Banigan et al., 2015).

5.5.2 Lévy Search

Lévy walks are described by step lengths that fit a power law distribution Most step lengths are small, but with a heavy-tail, that is, a decreasing probability of larger steps and a non-zero probability of steps of any length. Lévy walks assume that the direction of search at each step is drawn from a uniform distribution and is independent of previous steps (i.e. is isotropic and Markovian) (Mandelbrot, 1983; Viswanathan et al., 1996).

Lévy search patterns are stochastic fractals. The probability density function (PDF) that governs the distribution of step lengths used to generate a particular Lévy pattern is a power law:

$$\mathrm{L}(x) = x^{-\mu} \tag{5.1}$$

where $\mathrm{L}(x)$ is the probability of a searcher moving in a straight line for distance x. The exponent μ that determines the shape of the PDF is known as the Lévy exponent.

When search consists of a sequence of disconnected points, the motion is called a Lévy flight (the searcher is flying or jumping from point to point). Since we consider space between points to be part of the area searched, and therefore the intervening space is traversed with some finite velocity, we are working with a Lévy walk (Shlesinger and Klafter, 1986). We refer to the displacement (shortest distance) between consecutive positions as the *step length*.

Stochastic search has been studied extensively by biologists, especially mathematical ecologists. Ecologists typically state the problem of efficiency in the interaction of searchers and targets as one of foragers and food items. Foraging problems and immunological search have much in common.

Lévy flights as models of search were first developed in order to explain the disparity between observed super-diffusive animal motion and simple random walk models. Animals searching for food tend to maintain relatively straight trajectories for longer distances than would be produced by a simple random walk. The Lévy walk foraging hypothesis is an explanation for this observation proposed by physicists and ecologists. Lévy walks have been used to explain the search patterns of numerous species including reindeer (Mårell et al., 2002), albatross (Viswanathan et al., 1996), and human foragers (Raichlen et al., 2014). James et al. (2011) provides a more comprehensive list along with criticism of Lévy walk analysis.

Whether these search patterns are truly power law distributed is a matter of ongoing debate (Bartumeus et al., 2005; Edwards, 2011; Humphries et al., 2012). The issue is clouded in part because a true power law distribution of step lengths is clearly impossible in a finite space. The question then is whether animals or cells use a truncated power law distribution of step lengths constrained by the environment in which they are searching.

Similarly, immunologists have recently considered Levy walks as models that capture superdiffusive movement of T cells both during search of peripheral tissues and LN (Harris et al., 2012; Fricke et al., 2013).

In this work we examine the properties of Lévy search patterns because they provide a model of search that captures heavy-tailed movement patterns observed in T cells, are simple enough to be translated into robots, and exhibit the properties

of robustness, scalability, and adaptability that are required of T cells and robot swarms.

We explore the relationship between target distributions, the number of robot searchers, and μ in Equation Eq. (5.1) on page 85 optimised by a genetic algorithm (GA). We confine our robots to a $100\,\mathrm{m}^2$ arena, which places an upper limit on the distance searchers can travel without turning. Similarly, since power laws diverge as step lengths approach zero, we define a minimum step length for our robots to be 8 cm. Despite these constraints, we find Lévy walks to be a very useful abstraction of search patterns. We show that the mathematical properties of this search pattern that ecologists and physicists found so compelling can be used to engineer a search pattern with desirable properties.

5.6 Related Work

5.6.1 Robotic Lévy Search

In work related to our own, Van Dartel et al. (2004) evolved neural controllers for agents searching a simulated world with targets drawn from a uniform distribution. The authors observed convergence of the best performing robots to a Lévy walk pattern defined by a power law PDF exponent of 2 ($\mu = 2$ in Eq. (2.1) on page 18), consistent with optimal foraging behaviour described by Viswanathan et al. (1999).

Swarm robot simulations have used Lévy walks in combination with chemotaxis-inspired gradient sensing (Nurzaman et al., 2009) and artificial potential fields (Sutantyo et al., 2010) to efficiently search unmapped spaces with range-limited sensors. These papers fix the Lévy exponent to 2 the result from Viswanathan et al. (1999)

and explore a uniformly random distribution of targets. In contrast, in this work we evolve the Lévy walk exponent for combinations of target configurations and numbers of searchers.

Sutantyo et al. (2010) find that when collisions increase significantly with the number of searchers, performance scales sub-linearly. The rate of collision between searchers is determined by a number of factors, including the effectiveness and cost of collision avoidance algorithms and size of searchers relative to the rate of displacement. Hecker and Moses (2015) find that even when collisions are ignored, efficiency does not scale linearly with the number of robots. Our simulation similarly ignores collisions, allowing us to isolate the effect of oversampling on search performance.

Keeter et al. (2012) use underactuated robots implementing Lévy walks in a 3D aquatic environment to search for four uniformly distributed targets. They sample various values of μ in 0.5 increments in the range $1.1 < \mu < 3$. Keeter et al. find that in simulation there is a monotonic improvement in search time as μ approaches their lower bound of 1.1.

All the related work above examines search for uniformly distributed targets, but naturally occurring resources tend to be spatially heterogeneous (Dunning et al., 1992; Hecker and Moses, 2015). Here we are interested in how the configuration of targets with varying degrees of clusteredness influences stochastic search. We relate the size of the swarm to the optimal Lévy exponent, μ rather than using a fixed swarm size.

Beal (2015), examine the effectiveness of Lévy flights as a mechanism for positioning robots such that they provide even coverage of a given area with obstacles. Beal identifies a trade-off between *aggressiveness* and *evenness* in the pattern of dispersal. These concepts have interesting connections to the ideas of intensity and

extent in our own work. However, the formulation of the coverage problem and the search problem are significantly different (the lack of discrete targets in the coverage problem, for example), and so our approach does not map directly to theirs.

5.6.2 Lévy Search in Immunology

Harris et al. (2012) examined T cell search patterns in the brains of *Toxoplasma gondii* infected animals. They found the pattern of motion to be superdiffusive, and consistent with a generalized Lévy walk with $\mathrm{L}(x) \propto x^{2.15}$, $\mu = 2.15$ in Eq. (5.3) on page 92. In supplemental material the authors also describe a computer simulation of Brownian motion and the generalized Lévy walk search in a sphere. They report that the Lévy walk was able to detect targets an order of magnitude more efficiently than Brownian motion.

In our own work (Fricke et al., 2013) we also find an order of magnitude decrease in first contact times when using a Lévy flight vs. Brownian search. We have recently investigated the search patterns of T cells searching for DCs in LNs and implications for efficiency of search (Fricke et al., unpublished data). The difference between Lévy walks and Brownian walks decreases, though is still significant, when the efficiency metric is normalized by the total distance, and time, searchers cover. We find that T cells use strategies that have aspects of Lévy walks, but that Lévy walks by themselves are unable to explain the full range of search behaviour. In particular we find that T cells balance extent and intensity of search. This is not surprising given that immunological search is the result of extremely complex interactions between numerous cell types communicating via a range of chemical signals. Despite this, we find the Lévy walk foraging hypothesis (Viswanathan et al., 1999) to be a useful

model of T cell search both in peripheral tissue (i.e. the brain (Harris et al., 2012)). Lévy walks also provides a simplified model of of T cell search in LNs (Fricke et al., 2013). In addition, providing a useful engineering approach for designing searchers with desirable properties.

5.6.3 Lévy Search with Heterogeneous Target Distributions

Raposo et al. (2011) model the relationship between heterogeneity of searcher target distance and optimal μ values. Using a 1-dimensional analysis they predict that decreasing μ will increase the success of target encounters in heterogeneous landscapes. They suggest that this theoretical result generalizes to the 2-dimensional case. They do not test this hypothesis as we do here.

5.7 Stochastic Fractal Search

Deterministic search strategies may be effective in relatively fixed environments and when localization is error free. Theoretically, systematic raster scan search outperforms random search (Bénichou et al., 2011; Keeter et al., 2012). However, in environments where target distributions are unknown or change over time, randomized search strategies are more effective (Stephens and Krebs, 1986; Acar et al., 2003). Deterministic strategies depend on accurate information about the searchers' current location, and the ability to move from the current location to the next without error, which is difficult in practice even with global positioning systems (Humphreys et al., 2008; United States Department of Defense, 2008; Maier and Kleiner, 2010).

Errors in sensor input and in actuation introduce randomness into even apparently deterministic processes. iAnt robots are underactuated (i.e. are unable to follow arbitrary trajectories without error) and therefore the problem of error free navigation from one point to another is especially difficult. In contrast T cell motility mechanisms are holonomic with motion governed by cytoplasmic flow into membrane protrusions and an actin cytoskeleton that can be rapidly reoriented in any direction. However, observation of T cells demonstrate that their motion is stochastic, which may be due to environmental or intrinsic factors (Linderman et al., 2010; Celli et al., 2012; Harris et al., 2012; Textor et al., 2014). Like many swarm robots T cells sensors are short range, being limited to molecular interactions at the cell surface (Alberts et al., 2002).

If the search area or volume is greater than that which can be explored completely in the time allowed, an efficient solution must involve a trade-off between *intensity*, or the degree to which the local area has been searched and *extent*, the displacement of searchers from their starting locations (Méndez et al., 2013). The Hausdorff fractal dimension ($\mathcal{H}$) is a compact measure of this trade-off between intensity and extent. Increasing extent, by decreasing the Hausdorff fractal dimension ($\mathcal{H}$), results in increased displacement of searchers from their start positions as a function of time. The fractal dimension of a search pattern is a measure of the fraction of locations visited in a search space. For example, Brownian motion has fractal dimension $\mathcal{H} = 2$, meaning that, asymptotically, a Brownian search pattern visits all positions in a 2-dimensional space. As far as we are aware the fractal dimension of Lévy walks has not been formalized, therefore we calculate $\mathcal{H}$ for Lévy flights. Since Lévy flights visit only endpoints of steps, $\mathcal{H}$ is the dimension of the visited point-

set (Seshadri and West, 1982; Mandelbrot, 1983). $\mathcal{H}$ for Lévy walks will be strictly greater than that for Lévy flights. We hope to investigate this in future work.

The PDF governing Lévy flights as formulated in Eq. (5.2) not only describes the probability, $\mathrm{L}(x)$, of observing a step length of x, but also relates the resulting stochastic fractal to $\mathcal{H}$:

$$\mathrm{L}(x) = \frac{\gamma}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-1-\gamma} \tag{5.2}$$

where, x is the step length, x_{min} is the least possible step length, and γ determines the decay rate of the step length probability distribution. $\mathcal{H} = \gamma$ in Eq. (5.2) (Mandelbrot, 1983; Hughes, 1996). The coefficient $\frac{x}{x_{\min}}$ normalises the area under the curve to be one and so enforces that Eq. (5.2) is a PDF. Since $x_{\min}$ is a constant, and labelling the exponent μ results in Eq. (5.1) on page 85.

$$\mathcal{H} = \mu - 1 \tag{5.3}$$

Brownian motion has $\mathcal{H} = 2$ (Taylor, 1961) and $\mu = 3$. The resulting walk is maximally intense when embedded in a 2-dimensional space.

The mean squared displacement (MSD), a measure of search extent, of a population of searchers is also characterized by a power law:

$$\mathrm{MSD} = \left\langle \left(\vec{r}(t + \Delta t) - \vec{r}(t) \right)^2 \right\rangle \propto t^{\alpha} \tag{5.4}$$

where $\vec{r}(t)$ is the position vector of an agent at time t and $\vec{r}(t+\Delta t)$ is the location of an agent after some time increment Δt. The difference between $\vec{r}(t)$ and $\vec{r}(\Delta t)$ is the displacement between searcher positions at t and Δt. Angle brackets indicate the

ensemble average over the population of searchers. The MSD exponent α describes the rate of displacement over time and is related to search extent. As α increases search extent increases.

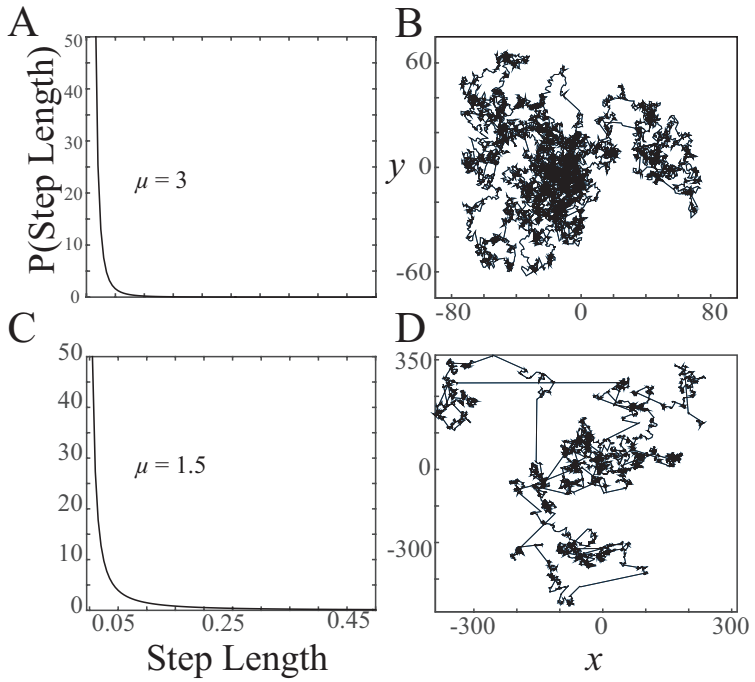


Figure 5.2: Stochastic fractal search patterns resulting from a sample of 10,000 steps and uniform turning angles with a power law distribution of steps. The power law distributions of step lengths A) $\mu = 3$, the Hausdorff fractal dimension ($\mathcal{H}$) = 2, $\alpha = 1$ B) a Brownian pattern of search. C) $\mu = 1.5$, $\mathcal{H} = 0.5$, $\alpha > 1$, and D) a Lévy walk with lower dimensionality and greater extent.

The walks produced by Lévy exponents between 1 and 3 are more extensive than Brownian motion, but have lower $\mathcal{H}$ and so are less intensive. As μ approaches the limiting value of 1 the MSD and extent of search becomes infinite.

Figure 5.2 shows two search patterns. In panel (A) $\mu = 3$ resulting in (B), a Brownian pattern of search with diffusive motion; $\mathcal{H} = 2$ and $\alpha = 1$ in Eq. (5.4) on the preceding page. The displacement of a Brownian random walk grows as the

square root of time. Panel (C) shows a power law distribution resulting in a Lévy walk with $\mu = 0.5$, and a superdiffusive pattern of motion with $\alpha > 1$. MSD describes how far searchers are likely to travel from their starting locations over time. The fractal dimension determines how thoroughly an area is searched. Both MSD and the fractal dimension are governed by a single parameter μ. Changing μ allows control over the extent and intensity of search, a property we exploit in order to adapt the robot swarm to the distribution of targets.

5.8 Methods

5.8.1 iAnt Robot Platform

The iAnts implement a scalable and robust central place foraging algorithm (CPFA) inspired by desert harvester ants (Hecker and Moses, 2015).

The iAnt simulator replicates the movement and sensing capabilities of iAnts, small autonomous robots that sense and compute using onboard iPods (Figure 5.3a on page 100). iAnts performing the CPFA have several phases, including a random search phase implemented as an adaptive correlated random walk. Here we implement the Lévy walk as an alternative random search strategy. The parameters for this search are determined by a genetic algorithm (GA) which evolves simulated iAnts and produces a strategy for the physical robots to use in the search task.

Hecker et al. (2013) observed that iAnt robots fail to observe targets within their viewing range 50% of the time. We incorporate this error in our simulations.

5.8.2 Robot Lévy Search

Lévy walks are defined by the Hausdorff fractal dimension ($\mathcal{H}$) of the random path they generate. In order to analyse the relationship between Lévy walk efficiency and the distribution of targets, we have the iAnts generate a Lévy walk with parameter μ (Eq. (5.1) on page 85) and corresponding $\mathcal{H}$ (Eq. (5.3) on page 92).

The GA explores the fitness landscape and optimises μ to produce a search pattern with $\mathcal{H}$ that most efficiently solves the problem of detecting a particular distribution of targets.

iAnts draw a random variate, t, from a power law probability density function (PDF) (with units 0.5 s):

$$t = t_{\min}(\mathcal{U}(0,1))^{\frac{-1}{\mu-1}} \tag{5.5}$$

where $t_{\min}$ is the minimum time (0.5 s) and μ is evolved by the GA to maximize the discovery of targets. The GA is described in more detail in section 5.8.4 on page 98.

Each robot chooses a direction from a uniform distribution and moves in that direction for t time steps. At the end of this movement a new t is drawn and the process repeats. Collisions with the edge of the search area require that the robot draw a new direction.

5.8.3 Cluster Analysis

We explore the optimised $\mathcal{H}$ evolved by the GA in response to various target configurations. To accomplish this we use 256 targets in every simulation but distribute

them into varying numbers of clusters. For generality we measure the patchiness (distance from uniform) of the distribution of targets using the Hopkins index.

Analysing search $\mathcal{H}$ across a varying number of clusters explores the space between a completely organized configuration in which all targets are in a single pile, and a completely disordered configuration in which the targets are uniformly distributed into 256 piles of one target each. Cluster centres are positioned by drawing x and y-coordinates from a uniform PDF.

We use two cluster progressions amongst which the targets are divided to map the Hopkins index to different configurations of targets: A linear progression, 1, 10, 20, ..., 100, and a power law progression, 1, 2, 4, ..., 256, of clusters. This map allows the exploration of the Hopkins index's resolution in distinguishing target configurations.

Zhang et al. (2006) describe experiments comparing cluster analysis algorithms. They report that the Hopkins index is the most sensitive test for distinguishing fine scale clustering. They identify this fine scale analysis as particularly important in biological systems.

The Hopkins statistic tests spatial randomness by comparing nearest-neighbour distances from uniformly distributed points and randomly chosen targets (Jain and Dubes, 1988). If there are n targets in the set T, then let $m \ll n$ and choose m sampling points $s_j = (x_j, y_j) \in$ uniform distribution on the interval [a,b] ($\mathcal{U}(a, b)$), where x and y are coordinates in the arena. Also choose m targets $\tilde{p}_k = p_{i_k} \in \mathcal{U}(a, b)$. If we define U and W to be,

$$\mathrm{U} = \sum_{j=1}^{m} ||s_j - t||, \forall t \in \mathrm{T} \qquad \mathrm{W} = \sum_{k=1}^{m} ||\tilde{p}_k - t||, \forall t \in \mathrm{T} \tag{5.6}$$

Then the normalized Hopkins index is,

$$\mathrm{H} = \frac{\mathrm{U}}{\mathrm{U} + \mathrm{W}} \tag{5.7}$$

Intuitively, the Hopkins index compares (as a normalized ratio) W, the distribution of distances between targets (S and s_j), and the distribution of distances between targets and, U a set of uniformly distributed points (S and $\tilde{p}_k$). Yielding a dimensionless statistic that does not depend on the units used to measure distance. The values of the Hopkins index lie in the interval $[\frac{1}{2}, 1]$. For randomly generated points the expected value is $\frac{1}{2}$. As targets become more highly clustered the value of the Hopkins index approaches 1.

Use of the Hopkins index requires the use of multiple datasets, S, and that the Hopkins index is computed for each S. We are able to generate multiple datasets (N=10) for each configuration of targets. This meets the requirements for using the Hopkins index and allows us to generate a 95% confidence interval around H.

For this work we choose $m = 50$ and we have $n = 256$ targets. To generate the Hopkins index we repeat the analysis 100 times and take the mean value. Since each configuration has 10 samples there are 1000 Hopkins index samples contributing to each data point in Figure 5.4 on page 101. This allows us to confirm that the empirical PDF is Gaussian distributed about the mean and suggests that variation in the Hopkins index is due to random rather than systematic effects. We examine a total of 18 different distributions (10 linear and 8 power law progressions).

Using the Hopkins index has the advantage of being a generalized statistic that measures clusteredness rather than being specific to our experimental design.

5.8.4 Genetic Algorithm

The GA evaluates the fitness of various search strategies by simulating robots that search for targets. We vary the number of robots from 1 to 32 and the Hopkins index for target configurations from $\frac{1}{2}$ to 1 and evolve an $\mathcal{H}$ that maximizes the search efficiency of the robotic swarm. Fitness is defined as the number of targets detected by the robot swarm in one hour of simulated time. Because the fitness function is evaluated many times, the simulation must run quickly. Thus, we use a parsimonious simulation that uses a gridded, discrete world without explicitly modelling sensors or collision detection. This simple fitness function also helps to mitigate condition-specific idiosyncrasies and avoid over-fitted solutions (Hecker and Moses, 2015).

We evolve a population of 100 simulated robot swarms for 100 generations, though convergence consistently occurred in fewer generations. We used the recombination and mutation described in Hecker and Moses (2015). The GA evolves μ in order to govern Eq. (5.5) on page 95 and therefore the resulting $\mathcal{H}$. From Eq. (5.3) on page 92, $\mathcal{H}+1 = \mu$. Parameter μ is randomly initialized using independent samples from $\mathcal{U}(1, 10)$ for each swarm. Robots within a swarm use identical parameters throughout the hour-long experiment. During each generation, all 100 swarms undergo 8 fitness evaluations, each with different random placements drawn from the specified target distribution.

At the end of each generation, the fitness of each swarm is evaluated as the sum total of targets collected in the 8 runs of a generation. Deterministic tournament selection with replacement (tournament size = 2) is used to select 99 candidate swarm pairs. Each pair is recombined using uniform crossover and 10% Gaussian mutation with fixed standard deviation (0.05) to produce a new swarm population.

We use elitism to copy the swarm with the highest fitness, unaltered, to the new population – the resulting 100 swarms make up the next generation. After only 10 to 20 generations the evolutionary process converges on a $\mathcal{H}$ for the Lévy search.

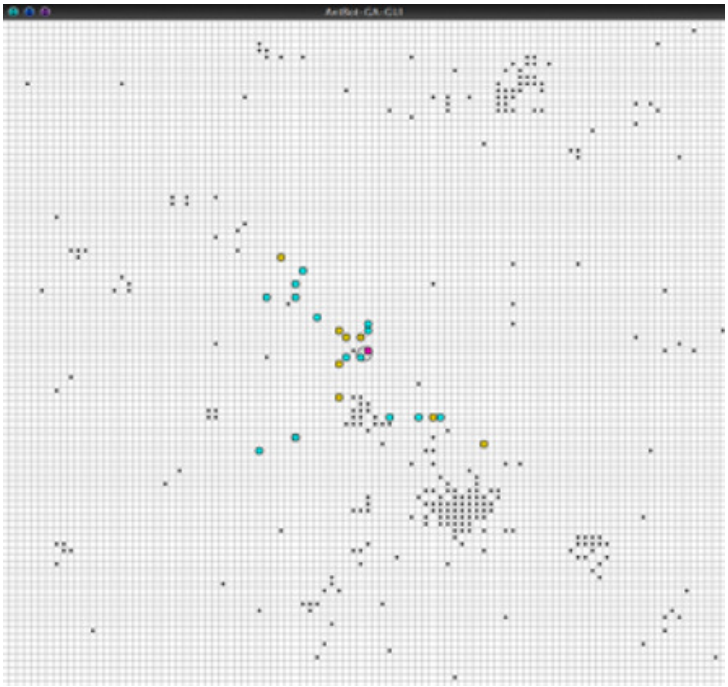
5.8.5 Effect of the Number of Searchers and the Configuration of Targets on the Optimal Fractal Dimension of Search

We use an analysis of variance (ANOVA) to determine whether the relationships between factors (number of searchers and the Hopkins index) and the observed $\mathcal{H}$ selected by the GA are statistically significant. The ANOVA also allows the quantification of the relative contributions of factors to the resulting $\mathcal{H}$.

The *error* factor measures the amount of variation in evolved $\mathcal{H}$ that results from the combination of changes in the number of searchers and the number of target clusters. SS is the sum of differences used in calculating the the mean squared error (MSE). MSE measures the variance within and between factor groups.



(a)



(b)

Figure 5.3: The iAnt Robot System. Physical robots perform a search task informing the error model in simulation. The resulting simulation reproduces the performance of the physical robots (Hecker et al., 2013). (a) Robots searching for targets. The targets are QR tags attached to poker chips. The lamp corresponds to the robot start position. (b) Simulation of the physical iAnt robots in a $100\,\mathrm{m}^2$ search arena. Grey dots are QR tags which are the target of search. Circles are robot locations. Blue circles indicate a robot that has just found a tag. Green circles are robots engaged in search. The red robot is at the start position.

5.9 Results

5.9.1 The Hopkins Statistic

We use the normalized Hopkins index as a measure of the clusteredness of target distributions. In line with (Zhang et al., 2006) we find that the Hopkins index is able to capture changes in target distribution over a wide range of target configurations (Figure 5.4). However, the relationship between the number of clusters and the Hopkins index is non-linear, and it becomes more difficult to distinguish differences in target configurations as the number of clusters approaches 1. For example, the

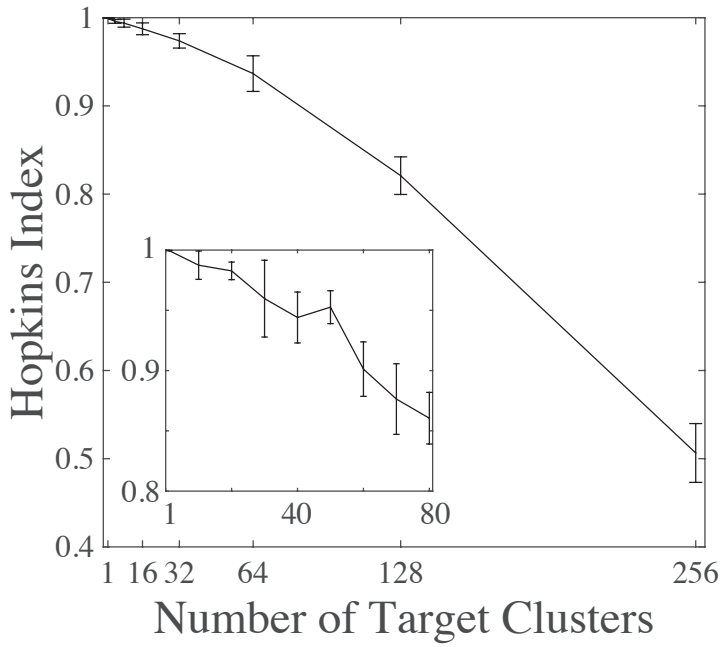


Figure 5.4: Hopkins index vs number of clusters. The Hopkins index is able to distinguish between target configurations ranging from a single cluster to a uniform distribution of 256 singleton clusters. Inset: Hopkins index for a linear progression of between 1 and 80 clusters.

difference between 1 cluster and 40 clusters only maps to a 0.05 decrease in the Hopkins statistic. Despite this we are able to distinguish all target configurations that we considered from one another using the Hopkins index. We also find that the progression of target configurations, from 256 clusters of 1 target each to 1 cluster of 256 targets, maps to the full range of the Hopkins index.

5.9.2 The Genetic Algorithm

We use the genetic algorithm (GA) to explore the relationship between the number of searchers, the Hopkins index of targets and the optimal Hausdorff fractal dimension ($\mathcal{H}$). The GA evolves values of $\mathcal{H}$ that provide a fitness advantage, where the fitness is the efficiency of target detection. We define efficiency to be the number of targets detected in one hour of search.

The Hausdorff fractal dimension ($\mathcal{H}$) is used as a measure of the trade-off between extent and intensity selected by the GA. The Hopkins index provides a metric for how disorganized the targets are.

The fitness landscape for $\mathcal{H}$ was explored over 100 generations for one combination of factors, and visualized in Figure 5.5 on the next page. A linear regression was performed and plotted on log-log transformed data (dashed line). The slope with 95% confidence interval (CI) is -0.501 ± 0.024 and intercept with 95% CI is 2.141 ± 0.008. ($R^2 = 0.948$, p-value $< 10^{-4}$).

The rapidity of GA convergence (10-25 generations), the convergence to disjoint $\mathcal{H}$ depending on the factors defining the search environment, and the power law fitness

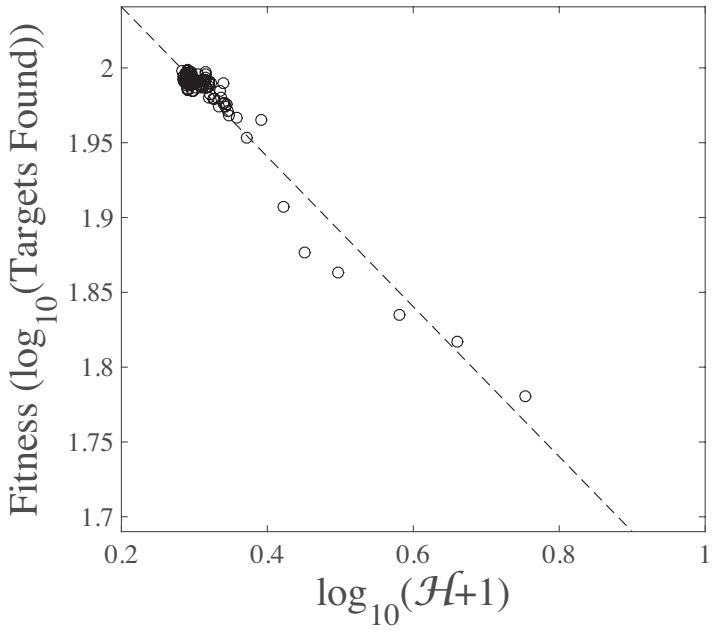


Figure 5.5: Convergence of the GA on an optimum in the fitness landscape defined by $\mathcal{H}$. Fitness is defined to be the number of targets detected in 1 h. This example run has 6 searchers and 100 clusters. The mean fitness values for each of 100 generations are shown as open circles. The fitness peak is reached after only 25 generations. The increase in fitness as the population approaches the optimal fractal dimension ($\mathcal{H}$) is fit well by a power law ($R^2 = 0.948$), indicated by the dashed line.

relationship as a function of $\mathcal{H}$ all suggest that the fitness landscape for optimizing $\mathcal{H}$ is simple and therefore particularly amenable to evolutionary optimization.

The example fitness landscape revealed by the GA is based on 6 searchers and 100 clusters. The landscape has a clear slope from a randomly assigned starting value of $\mu = 6.3$ ($10^{0.8}$) to a peak at approximately 1.86 ($10^{0.27}$) (Figure 5.5). We note that below Hopkins index 0.9 the GA is unable to discover values of $\mathcal{H}$ that provide significant advantage over other values. The efficiency of Lévy walks for distributions with Hopkins index between 0.75 and 0.9 are statistically similar and converge to $\mathcal{H}$

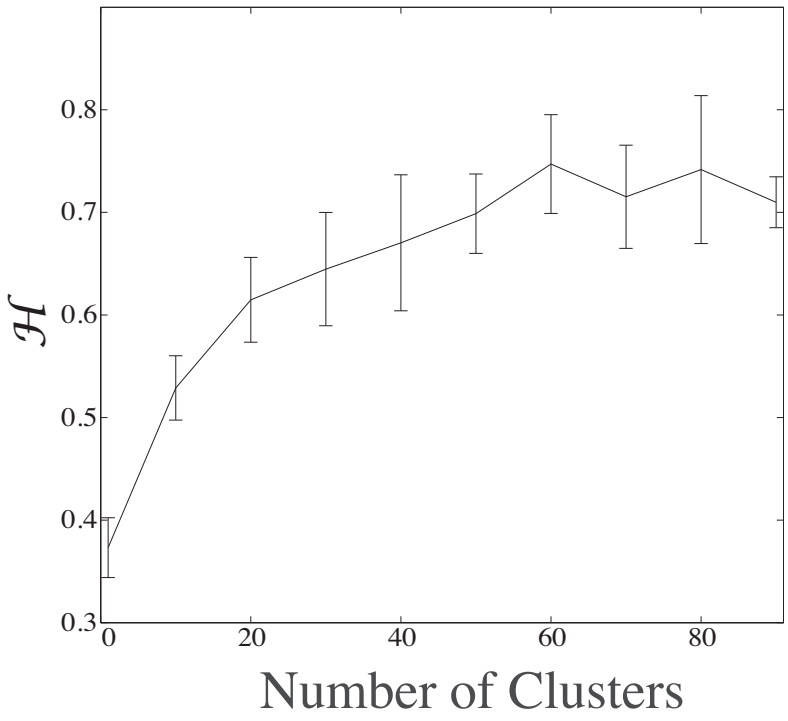


Figure 5.6: Dependence of genetic algorithm (GA) selected values of fractal dimension ($\mathcal{H}$) on target clusteredness. Bars indicate the 95% confidence interval (CI).

≈ 0.71. For Hopkins index between 0.9 and 1, that is, highly clustered, $\mathcal{H}$ falls from 0.7 to ≈ 0.38 (Figure 5.7 on the next page).

5.9.3 Optimizing Fractal Dimension for the Number of Searchers and the Configuration of Targets

We used a full factorial experimental design to explore the relationship between $\mathcal{H}$ as optimised by our GA, the number of searchers, and the configuration of targets. The results are displayed in Figure 5.8 on page 106. For the target configuration closest to the uniform distribution (with Hopkins index $= \frac{1}{2}$, 256 clusters), the optimal evolved

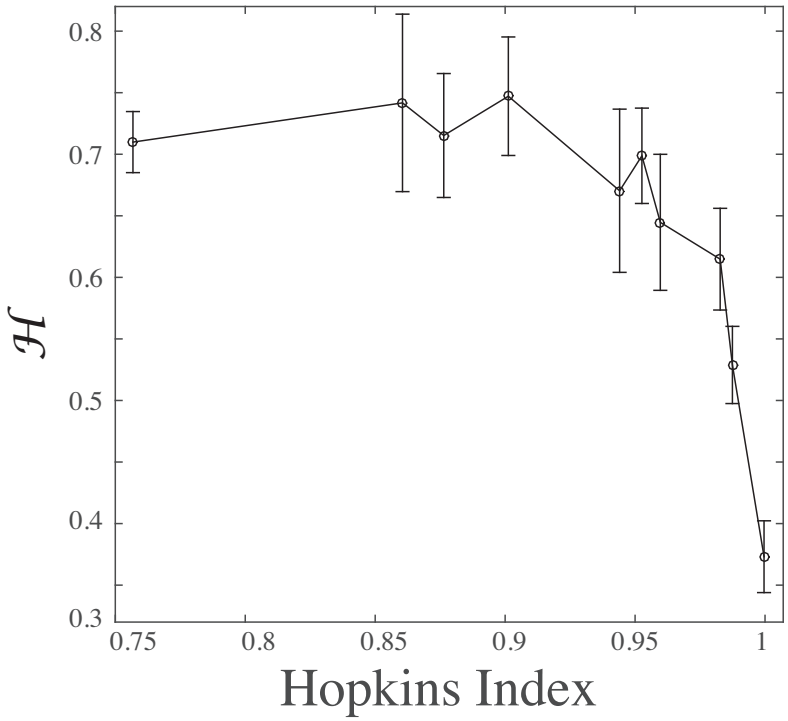


Figure 5.7: Dependence of genetic algorithm (GA) selected values of fractal dimension ($\mathcal{H}$) on the Hopkins index. Bars indicate the 95% confidence interval (CI).

$\mu = 1.8$ ($\mathcal{H} = 0.8$). This is in general agreement with Viswanathan et al. (1999) who predict an optimal μ of 2 ($\mathcal{H} = 1$) for the uniformly distributed target configuration.

A two-way ANOVA analysis (Hogg and Ledolter, 1987) shows a statistically significant correlation between fractal dimensions of the Lévy walk evolved by the GA, the number of searchers, and the Hopkins index. The p-values for the number of searchers and the Hopkins index of target clustering are less than 10^{-5} and 10^{-18} respectively, indicating a statistically significant influence of the number of searchers and the Hopkins index on optimal values of $\mathcal{H}$.

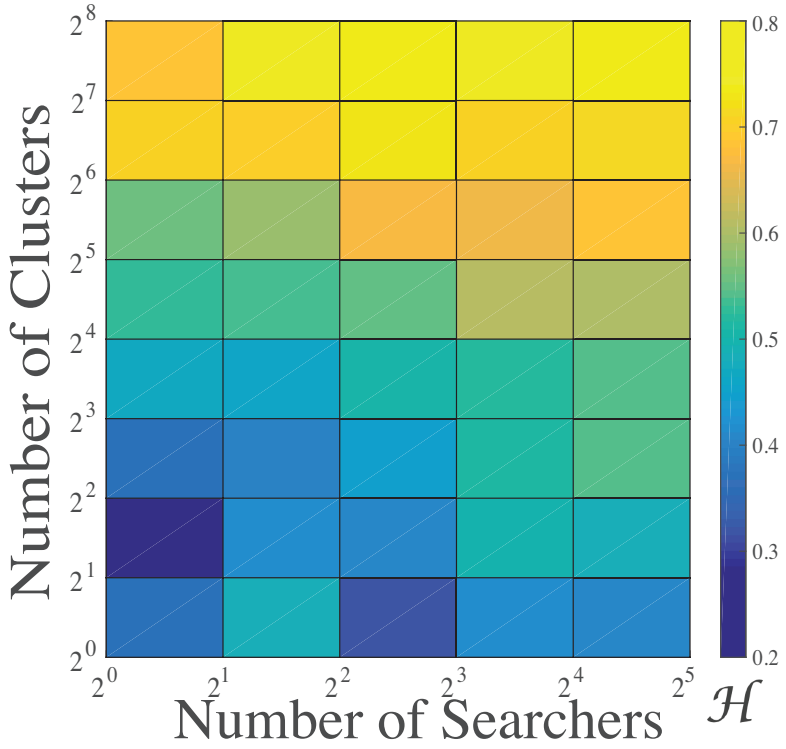


Figure 5.8: Full factorial heatmap. Colors indicate GA optimised values of $\mathcal{H}$ for number of searchers and target configuration.

The statistical tests indicate only that the GA selects different optimal values of $\mathcal{H}$ for different target configurations and different numbers of searchers. We measure the resulting change in search efficiency using simulations.

Factor	SS	df	MSE	F	p-value
N Searchers	0.072	5	0.014	8.93	$< 10^{-4}$
Hopkins	0.682	8	0.085	53.15	$< 10^{-4}$
Error	0.064	40	0.002		
Total	0.818	53			

Table 5.1: ANOVA Results. The number of searchers and and the Hopkins index both have statistically significant effects on the GA optimised $\mathcal{H}$. The error term is the remaining variance in $\mathcal{H}$ not explained by the factors.

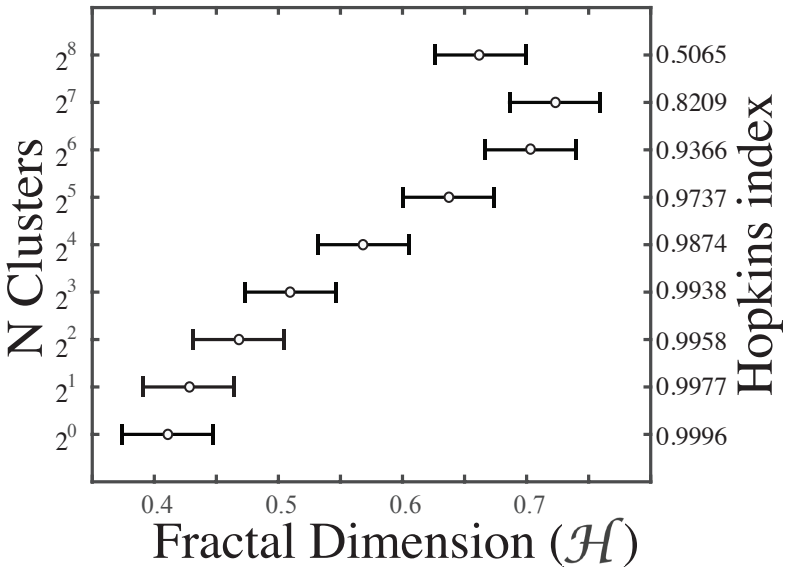
We create a multiple comparison plot (Figure 5.9 on the next page) to further explore the relationships between the number of searchers, and $\mathcal{H}$. The majority of samples of cluster sizes in (Figure 5.9a on the following page can be arranged into significantly different groups. However, Figure 5.9b on the next page shows that there are no statistical differences between the number of robots for $N > 1$.

In order to investigate the magnitude of the effect of $\mathcal{H}$ on search efficiency, we analyse four Lévy walks characterized by different $\mathcal{H}$ evolved for different numbers of searchers and for target configurations with different Hopkins indices.

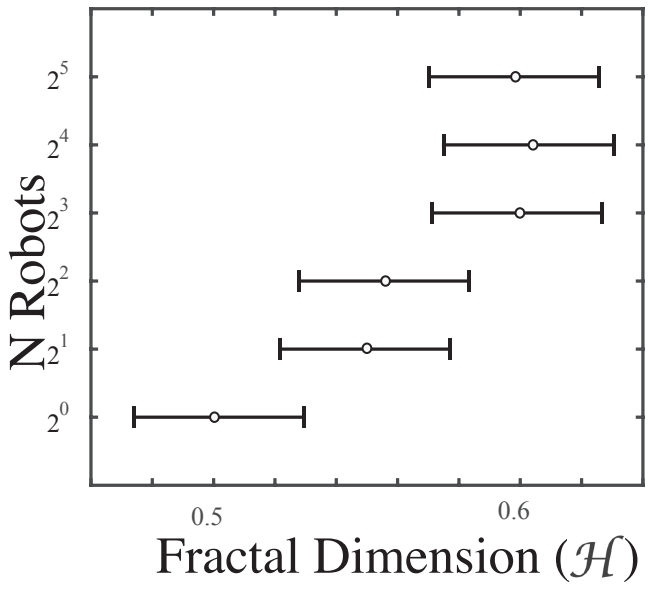
Figures 5.10 on page 109 and 5.11 on page 110 show the practical impact of the number of target clusters and the number of robots on efficiency. Each inter-quartile box is the result of 1000 searches. We plot the efficiency of a simple random walker for comparison.

For Figure 5.10 on page 109 we evolve Lévy walks on different target configurations, 1 target cluster and 256 target clusters, while holding the number of robots fixed at 8. Lévy searchers perform much better when applied to the target configuration for which they were evolved than when applied to a target configuration for which they were not evolved. The change in median efficiency is 19 and 26%, with p-value $< 10^{-4}$ for both the Student's t-test and the Mann-Whitney U test.

In Figure 5.11 on page 110 we repeat the experiment but this time we hold the target configuration constant and vary the number of robots. This time, despite the statistically significant difference reported by the ANOVA, the change in efficiency is negligible. This suggests that the fitness landscape defined by the interaction of the number of searchers and $\mathcal{H}$ is relatively flat.

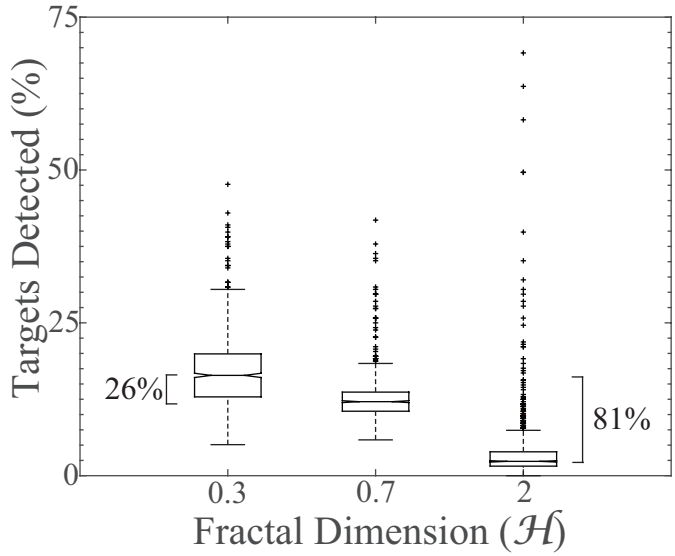


(a) 8 Robots

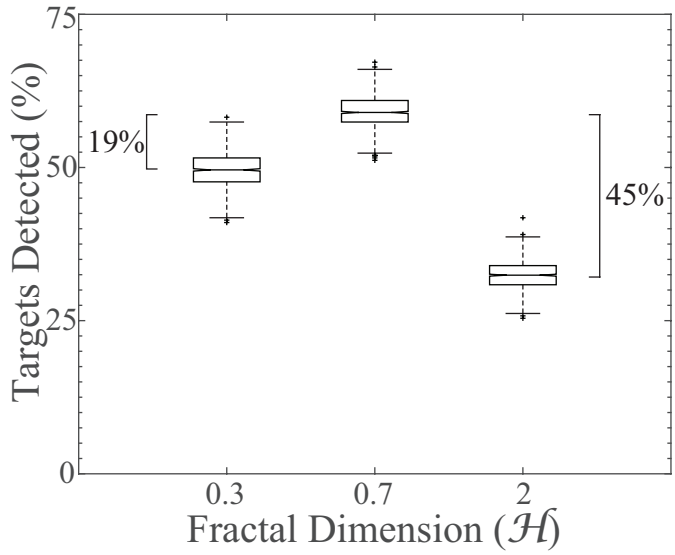


(b) Hopkins index = 0.85

Figure 5.9: Multiple comparison plots (Montgomery, 2012) showing the statistical separation of mean $\mathcal{H}$ grouped by the number of searchers and the number of clusters. Open circles are sample means and the bars indicate the Tukey range test at 95% confidence (Tukey, 1949). Non-overlapping bars indicate populations with means that are significantly different. In (a) most genetic algorithm (GA) optimised values for the fractal dimesion ($\mathcal{H}$) are statistically different from those evolved for a different target configuration. In (b) only $\mathcal{H}$ evolved for $N = 1$ is statistically different from $\mathcal{H}$ evolved for multiple robots.

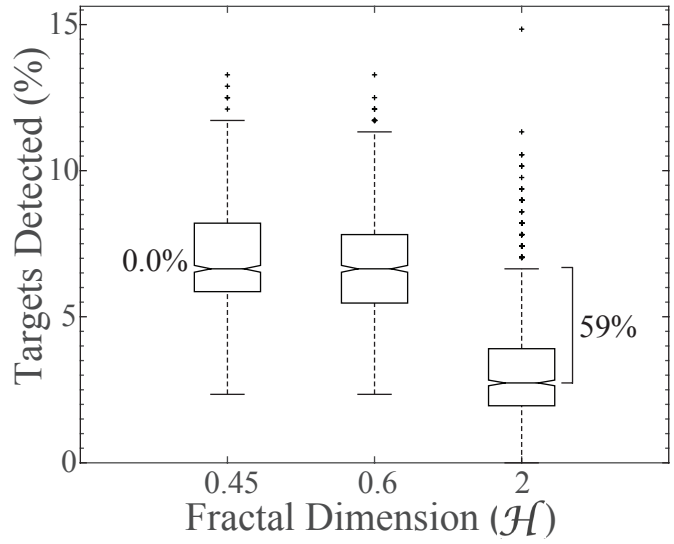


(a) 8 Robots, 1 Target Cluster

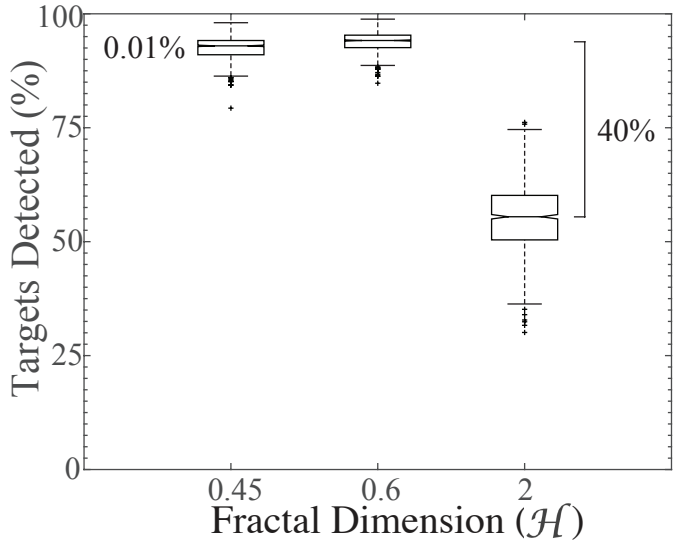


(b) 8 Robots, 256 Target Clusters

Figure 5.10: Efficiency comparison for search patterns evolved at the extremes of the Hopkins index. The search pattern with $\mathcal{H} = 0.7$ was evolved for a target configurations with Hopkins index = $\frac{1}{2}$ (uniformly distributed targets). The search pattern with $\mathcal{H} = 0.3$ was evolved for Hopkins index = 0.9996 (1 target cluster). A simple random search ($\mathcal{H} = 2$) is given for comparison. The efficiency gain for optimised values of $\mathcal{H}$ are significant (19 and 26%). As expected the difference in efficiency between a simple random walk and either Lévy search is significant (45 and 81%). (a) 8 Robots and Hopkins index = 0.9996 (1 cluster). (b) 8 Robots and Hopkins index =$\frac{1}{2}$ (256 clusters).



(a) 1 Robot, 16 Target Clusters



(b) 32 Robots, 16 Target Clusters

Figure 5.11: Efficiency comparison for search patterns evolved at the extremes of swarm size. The search pattern with $\mathcal{H} = 0.6$ was evolved for a swarm with 32 robots. The search pattern with $\mathcal{H} = 0.45$ was evolved for a swarm with 1 robots. A simple random search ($\mathcal{H} = 2$) is given for comparison. The efficiency gain for optimised values of $\mathcal{H}$ are minimal. As expected the difference in efficiency between a simple random walk and either Lévy search is significant (40 and 59%). (a) 1 Robot and Hopkins index 0.98 (16 target clusters). (b) 32 Robots and Hopkins index 0.98 (16 target clusters). The percentage change between the search pattern evolved for 1 and 32 robots are negligable (0.0% and 0.01%).

Our results show that the configuration of targets as measured by the Hopkins index of clusteredness influences the $\mathcal{H}$ of optimal Lévy search. The optimal $\mathcal{H}$ does not depend significantly on the number of searchers (N) for $N > 1$.

5.9.4 Efficiency Scaling with Number of Robots

When target discovery is difficult, for example when there is only a single cluster of targets (Hopkins index ≈ 1), the evolved $\mathcal{H}$ is approximately 0.4. In this case

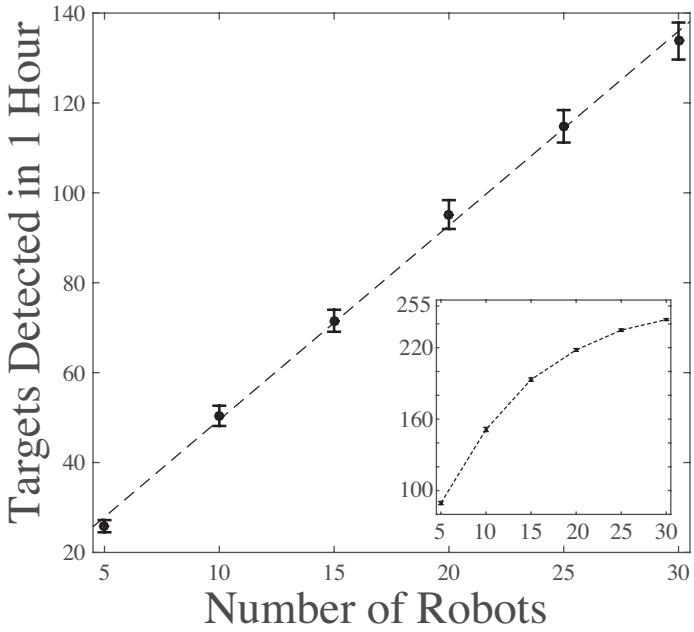


Figure 5.12: Search efficiency scales linearly with the number of robots. Closed circles are mean values and bars are the 95% confidence interval for 100 samples. Targets are clustered with Hopkins index ≈ 1.0 (1 cluster). The dashed line is the linear regression ($R^2 = 0.924$, p-value $< 5 \times 10^{-3}$). Inset: search efficiency saturates when targets have Hopkins index ≈ 0.93 (64 clusters). Closed circles are mean values and bars are the 95% confidence interval for 100 samples. The dashed line is a linear interpolation.

the relationship between the number of searchers and search efficiency is linear (Figure 5.12 on the preceding page). Doubling the number of robots doubles the number of detected targets. This is in contrast with prior studies, where per robot search efficiency declines with number of searchers (Winfield et al., 2005; Hecker and Moses, 2015).

However, we note that the linear scaling does not hold in all cases: when the number of target clusters is large, adding more searchers does not result in a proportional increase in efficiency, because as targets become rare, they become more difficult to find. We term this searcher saturation. The saturation effect is seen in the Figure 5.12 on the previous page inset, where we see saturation after more than half the targets are collected. This effect is discussed in Hecker et al. (2015).

5.10 Discussion

In the absence of accurate localization, randomized search provides a useful alternative strategy to deterministic search. Harris et al. (2012) found that T cells searching in peripheral tissue can be modelled using a Lévy walk. In previous work we observed the stochastic search patterns of T cells searching for dendritic cells (DCs) in lymph nodes (LNs) and characterized them as being reasonably approximated by a Lévy flight (Fricke et al., 2013). Here we explore the ability to optimise the efficiency of Lévy search in a robot swarm.

A significant advantage of Lévy search is its simplicity and adaptability: a range of search behaviours can be defined using just one parameter. Evolving solutions is fast since the fitness landscape is simple with a well defined optima (Figure 5.5 on page 103).

We show that the Hausdorff fractal dimension ($\mathcal{H}$) of the Lévy walk can easily be adapted to the configuration of targets (Figure 5.10 on page 109), and that the optimal Lévy walk is insensitive to the number of searchers (Figure 5.11 on page 110). Specifically the $\mathcal{H}$ selected by the genetic algorithm (GA) decreases as the clusteredness and the Hopkins index of the targets increase. This supports a prediction made by (Raposo et al., 2011) based on their theoretical analysis of the one dimensional case. The insensitivity of Lévy search efficiency to the number of searchers is a clear benefit to swarm robotics. As robots fail or get lost the optimal $\mathcal{H}$ for a given target configuration does not change.

The effect of the number of searchers on individual efficiency is extremely small for the cases we tested (Figure 5.11 on page 110). We also find a linear relationship between the number of searchers and the efficiency of search, up until so many targets have been found that search itself becomes more difficult (Figure 5.12 on page 111).

A possible explanation for the linear scale-up in efficiency with the number of Lévy searchers, as opposed to simple random searchers, is that when Brownian searchers start searching from the same location the high fractal dimension of their movement results in locations being revisited by other robots in the swarm. For n searchers the number of unique locations visited is proportional to $t \ln \frac{n}{\ln t}$. Only after t exceeds e^n do searchers employing a simple random walk avoid search redundancy (Larralde et al., 1992). For a swarm of 256 robots performing a simple random search this implies wasted effort for more than 10^{100} time steps.

In contrast, for n searchers employing the evolved Lévy search pattern, the number of unique locations visited is proportional to nt, implying relatively little search redundancy (Viswanathan et al., 1996). Therefore, given a search pattern with $\mathcal{H}$ tuned to a particular configuration of targets, adding searchers increases search ef-

ficiency linearly. There is little interaction between the number of searchers and the configuration of targets which could have potentially complicated the proper selection of $\mathcal{H}$.

Our analysis shows that there is a systematic relationship between $\mathcal{H}$ optimised by a GA, the number of searchers, and the configuration of targets as measured by the Hopkins index. Using this empirical relationship (Figure 5.8 on page 106) we are able to predict values of $\mathcal{H}$ that result in improved search performance given various target configurations (Figure 5.10 on page 109).

Lévy walks have been used as a model of biological search including immunological search. The properties of Lévy search we explore in this paper have implications for the biological systems in which Lévy flight models have been suggested. For example, in the immune system different numbers of T cells are required to find a wide variety of targets distributed in different ways in different tissues. The distribution of DC targets in lymph nodes may be very different than virus-infected cells in the lung or brain, and in each case the number and distribution of targets may change over time. Because targets of immunological search are heterogeneously distributed, and the number of searchers range by many orders of magnitude during the search process, the flexibility and scalability of Lévy walks may be particularly useful in immune search. The Lévy model provides an example of how a simple movement pattern could be adapted by natural selection to search efficiently given a variety of target distributions over a wide range of T cell numbers.

Search for clustered targets is an important problem in swarm robotics because it generalizes to many real-world applications, such as collecting hazardous materials, natural resources, search and rescue, and environmental monitoring (Liu et al., 2007; Parker, 2009; Winfield, 2009; Brambilla et al., 2013). A particularly exciting

application is space exploration. For example, NASA recently announced the Swarmathon challenge in which robots operate in concert to autonomously search for, retrieve, and map patchy natural resources, such as water ice. These robots swarms are intended to be used for resource exploration on other planets (Ramsey, 2015). In this arena particularly, robot searchers must be robust to hardware failure and sensor limitations, and adaptable to heterogeneous target distributions. Scalability is desirable because it allows flexibility in robot allocation. For example, rich resource areas may be assigned more robots without loss of efficiency. Our work demonstrates a mechanism for tuning the fractal dimension of a search pattern to most efficiently encounter targets given a measure of their clusteredness, while also being scalable and robust.

5.11 Acknowledgements

The authors gratefully acknowledge financial support from NSF EF 1038682 (MM), DARPA P-1070-113237 (MM), NIH 1R01AI097202 (JLC), the Spatiotemporal Modeling Center 1 P50 GM085273 (FAB), the Center for Evolution and Theoretical Immunology 5P20GM103452 (JLC), and a James S. McDonnell Foundation grant for the study of Complex Systems (MM). We also thank the reviewers, each of whom provided insightful comments and corrections that have greatly improved this manuscript.

5.12 Conflicts of Interest

5.13 Notes

For more information about the iAnt platform including instructions on how to build them visit: http://iant.cs.unm.edu

5.14 Software

The software used in this chapter is available at:
`https://github.com/BCLab-UNM/LevyBots-AntBots/releases/tag/0.1-beta`

Chapter 6

A Deterministic Swarm Search Algorithm

Capt. Kirk: "I was not aware, Mr. Baris, that 12 Klingons constitutes a swarm."

— The Trouble with Tribbles

6.1 Author Contribution Statement

I am the lead author of this chapter under the supervision of Melanie Moses (Associate Professor, UNM Computer Science, Biology, and Santa Fe Institute). Antonio Griego (Undergraduate Student, UNM Computer Science), Josh Hecker (Postdoctoral Fellow, UNM Computer Science), and Linh Tran (Undergraduate Student, UNM Computer Science) and I co-authored the CPFA and DDSA ARGoS libraries used in this paper.

6.2 Publication Notes

This chapter is in press. ©2016, IEEE, reprinted with permission from the authors. Fricke, G. M., Hecker, J. P., Griego, A. D., Tran, L., T., & Moses, M. E. (2016). "A Distributed Deterministic Spiral Search Algorithm for Swarms." In *Proceedings of the 29th IEEE/RSJ 2016 International Conference on Intelligent Robots and Systems (IROS)*. The paper is reproduced here in accordance with IEEE guidelines on reuse in thesis and dissertations (Appendix C).

6.3 Abstract

As robot swarms become more viable, efficient solutions to fundamental tasks such as swarm search and collection are required. We propose the distributed deterministic spiral algorithm (distributed deterministic spiral algorithm (DDSA)) which generalises a spiral search pattern to robot swarms. While being an effective search strategy in its own right, the DDSA is also a useful point of comparison for other swarm search strategies. Such a benchmark for robot swarm search is currently needed but missing. As a case study, we compare the DDSA to a biologically-inspired central-place foraging algorithm that uses stochastic search, memory, and communication to efficiently collect resources in a variety of different resource distributions.

6.4 Introduction

Many swarm robot applications require the detection and collection of targets by teams of robots. These tasks include planetary surveys (Fink et al., 2005), land and sea mine clearance (Weber, 1995), pollution mapping by subsurface robots (Hu et al.,

2011), environmental monitoring, survivor location in hazardous environments (Birk and Carpin, 2006; Goodrich et al., 2008), military applications (Love et al., 2015), and agricultural pest control (Tamura and Naruse, 2014). When the requirement that targets be transported to a single collection point is included in a swarm search problem, it becomes a central place foraging task (Winfield, 2009). Central place foraging tasks include crop harvesting (Bac et al., 2014) and planetary resource collection (Ramsey, 2015).

Spiral search patterns for single searchers with stationary targets have been studied extensively and found to have desirable optimality properties (Goodrich et al., 2008; Bentley et al., 1980; Baeza-Yates et al., 1993; Burlington and Dudek, 1999; Langetepe, 2010; ElHadidy, 2015; Gabal and El-Hadidy, 2015). These properties include detection of the nearest targets first, complete coverage of the area within the spiral, and minimal oversampling. Detection of the nearest targets first is particularly important for central place foraging because it minimises the per-target trip time to the collection point.

Here we present the distributed deterministic spiral algorithm (DDSA). The DDSA generalises a single robot square spiral to any number of robots. The generated spirals are interlocking paths that preserve the determinism of the single robot case and the consequent optimality guarantees. We implement the DDSA using the ARGoS swarm simulator (Pinciroli et al., 2012) in order to observe how foraging efficiency scales with the number of searchers and targets. The physical robots we simulate are called iAnts (Hecker et al., 2013). Groups of iAnts are designed to meet the fundamental properties of a swarm described by Brambilia et al (Brambilla et al., 2013). We compare our results to the central place foraging algorithm (CPFA) developed by Hecker and Moses (Hecker and Moses, 2015; Hecker et al., 2015).

In both swarm robotics (Keeter et al., 2012) and ecology (Reynolds et al., 2007; Papi, 2012) the performance of systematic search strategies, including spiral search, is assumed to degrade significantly in the presence of error. Stochastic methods which are more resilient to noise are therefore favored (Sebbane, 2011). We investigate the impact of positional error on the performance of the DDSA and compare our results to the CPFA.

Central place foraging algorithms do not have a baseline of comparison. This makes it difficult to evaluate the effectiveness of foraging algorithms. Swarm algorithms tend to be probabilistic and contingent on the hardware or simulation in which they are implemented. This makes improvements in performance difficult to compare across systems and hard to describe analytically. We propose the use of the DDSA as a point of comparison for other central place foraging algorithms. The DDSA has two essential properties that make it a good candidate as a baseline algorithm: 1) it is simple from a theoretical point of view, being deterministic and having behaviour definable using a simple recurrence relation, and 2) in the error free case, it guarantees collection of the nearest targets first, complete coverage, and minimal repeated sampling.

6.5 Related Work

In addition to the single searcher work listed in the introduction, Ryan and Hedrick describe a square search pattern carried out with a single helicopter (Ryan and Hedrick, 2005). This search pattern is defined in Appendix H of the Coast Guard Operating Manual and is similar to the DDSA single searcher base case.

Baeza-Yates and Schott describe a multi-agent spiral search algorithm in which agents begin at a central point (Baeza-Yates and Schott, 1995). However, because they use a circular spiral the searchers diverge from one another over time. As a result, the approach is only able to reliably detect lines rather than point targets placed at arbitrary locations in the plane.

Parallel spiral search approaches have also been implemented in which each searcher performs an independent single agent spiral spatially removed from the other members of the swarm (Hayes et al., 2001; Feinerman et al., 2012), a behaviour observed in ants in our own lab.[2] Stocastic spiral search patterns have also been observed as a central place foraging strategy of desert ants (*Cataglyphis fortis*) (Müller and Wehner, 1994). These ants inhabit salt pans, which are flat and obstacle free compared to most natural landscapes.

The Multiple Robots Internal Spiral Coverage algorithm is a solution that guarantees complete coverage of a environment by partitioning the space equally among multiple robots (Hao et al., 2008). This approach differs from the DDSA because the robot search paths are discretized by a grid structure and robots are assigned to regions within the grid.

In work closely related to our own, Skubch describes a "proof of concept"' approach for generating a circular distributed spiral for multiple robots Skubch (2012). A dynamic constraint optimisation function uses stateful-feedback and a shared datastructure to coordinate the movement of robots in the swarm. Intriguingly, this dynamic constraint update at each time step allows the redistribution of robots in the event of robot failure, but also results in robots randomly switching between

[2]Spiralling Ant Video: `youtu.be/N46u0xLl56o`

each other's spiral paths. In our own approach the search pattern is predetermined and robots do not communicate with one another during search or maintain a shared datastructure.

López and Maftuleac describe a deterministic search strategy for idealized searchers that in the case of 2 and 4 searchers results in an interlocking spiral (López-Ortiz and Maftuleac, 2016). When the number of searchers exceeds 4, the algorithm partitions the space into expanding wedges. This strategy requires searchers to perform a right angle turn every step. They find this approach to be robust to error in target detection and searchers with differing speeds.

The CPFA with which we compare the DDSA is an ant-inspired algorithm (Hecker and Moses, 2015; Hecker et al., 2015). As robots search the environment they probabilistically place waypoints at locations with high target density. These waypoints influence other members of the swarm towards searching areas where more targets are expected to be found. When not using waypoints robots perform a random walk with decaying correlation. This strategy ensures that areas near where a target was previously found are searched intensively but that if nothing is found the robot quickly moves to a new area. The parameters that govern the CFPA, such as the probability of placing waypoints as a function of local target density, are optimised with a genetic algorithm.

6.6 Methods

6.6.1 The Algorithm

The DDSA specifies the interlocking spirals for a group of robots by calculating how far each robot must travel along each edge of their particular spiral (Figure 6.1 on the next page). Calculating the spiral paths requires knowing 1) how many robots there are so enough room is left for all of them, 2) the target detection range of the robots so the gap between spirals is eliminated, 3) how far into the spiral the robot is, since the spiral expands over time, and 4) the index of each robot in a predefined order.

Let D_H be a function that determines how far the current robot should travel in a given cardinal direction, H, for a particular circuit count, c. The circuit count is the number of times a robot has completed movement in all four directions in N, E, S, W order. Robots move away from the central location and order themselves on the 0th circuit. Formally, $D_H(i, c, R) : I \times \mathbb{N}^0 \times \mathbb{N}^0$ where $I = \{i \mid 1 \leq i \leq R\}$ is the index of the current robot, c is the current circuit, and R is the total number of robots. The north (N) and east (E), and south (S) and west (W) movement cases are symmetric, so $D_{\mathrm{N}} = D_{\mathrm{E}}$ and $D_{\mathrm{S}} = D_{\mathrm{W}}$.

$$D_{\mathrm{N}}(i, c, R) = \begin{cases} i & c = 0 \\ D_{\mathrm{N}}(i, 0, R) + R + i & c = 1 \\ D_{\mathrm{N}}(i, c-1, R) + 2R & c > 1 \end{cases} \tag{6.1a}$$

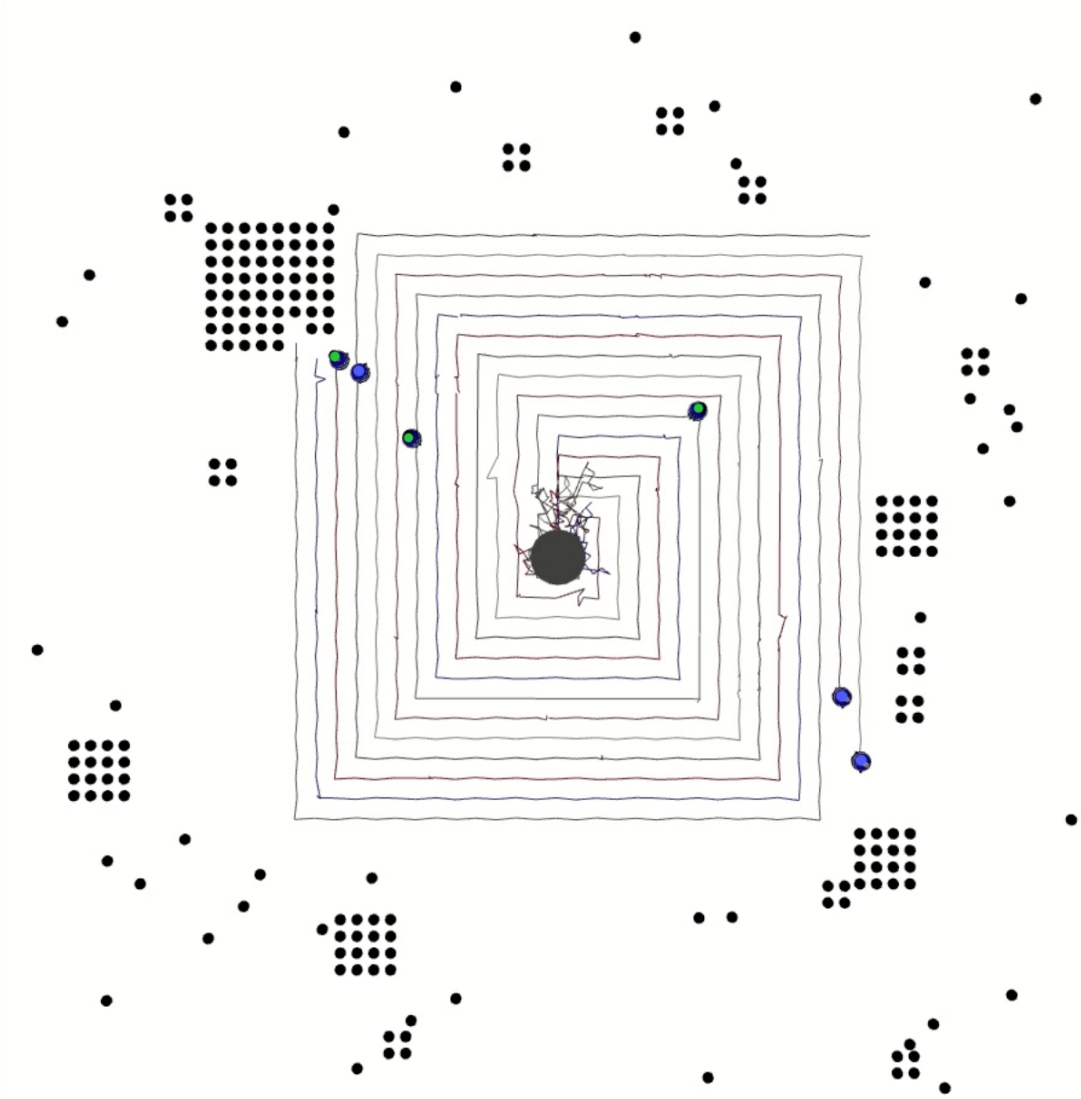


Figure 6.1: The DDSA running in ARGoS, overhead view. The robots search a continuous plane employing a spiral search pattern beginning at a central collection point. Targets are shown as black dots arranged in a partially clustered distribution. Robots are marked with a blue or green dot. Robots with a green dot are carrying targets, those with a blue dot are searching. Coloured lines are the paths of the various robots. Paths taken to and from the central collection point are not displayed. The simulation is performed in continuous space and robot paths and target placements are not lattice bound.

$$D_{\mathrm{S}}(i,c,R) = \begin{cases} D_{\mathrm{N}}(i,c,R) + i & c = 0 \\ D_{\mathrm{N}}(i,c,R) + R & c > 0 \end{cases} \tag{6.1b}$$

Requiring each robot to know the swarm size and its index implies global knowledge which would theoretically violate a principle of swarm design. Fortunately there are a number of consensus addressing algorithms for multi-robot systems that allow the swarm to determine its size and introduce an ordering (Thoppian and Prakash, 2006).

The gap between adjacent search paths, g, must be narrow enough so that the target detection ranges of robots on adjacent paths overlap to guarantee complete coverage; however, overlap should be minimized to avoid resampling of the same location. The target detection range, r, is 13 cm in our robots, suggesting a gap of 26 cm. However, at the corners of the square spirals the distance between paths increases to $\sqrt{2g^2}$. Therefore, in order to guarantee complete collection within the spiral, we set $g = \sqrt{(2r)^2/2} \approx 18\,\mathrm{cm}$ to compensate.

Let S be the set of searcher positions, along with D_H and g we can define interlocking square spirals for each robot and the state machine given in Algorithm 1 on the following page.

Algorithm 1 DDSA

```
1: for all robots i ← 1 to R do
2:     for c ← 0 to NCircuits do
3:         Q.enqueue(⟨0, gD_N(i, c, R)⟩)
4:         Q.enqueue(⟨gD_E(i, c, R), 0⟩)
5:         Q.enqueue(⟨0, −gD_S(i, c, R)⟩)
6:         Q.enqueue(⟨−gD_W(i, c, R), 0⟩)
7:     end for
8:     while ¬Q.empty() do
9:         w ← s+Q.dequeue()
10:        Move toward w
11:        if target found at current location s then
12:            Return to collection point with target
13:            if at collection point then
14:                Deposit target
15:                Return to location s
16:            end if
17:        end if
18:    end while
19: end for
```

6.6.2 Robot Simulation

We implement the DDSA and central place foraging algorithm (CPFA) using the ARGoS swarm robot simulator (Pinciroli et al., 2012).[3] ARGoS supports high fidelity ODE physics engines which allow the accurate detection of robot collisions. We use the 2D physics solver provided with ARGoS running at 320 updates per second for our simulations.

The parameters for the robots are informed by those of the physical iAnt robots designed and built in our lab (Hecker and Moses, 2015). Basing the simulated robots on physical robots allows us to choose positional error that closely matches our

[3]The software used in this work is available on GitHub:
`github.com/BCLab-UNM/DDSA-ARGoS/release/0.2-beta`
`github.com/BCLab-UNM/CPFA-ARGoS/release/0.1-beta`

experience with real hardware. In order to represent iAnt hardware, the simulated robots are 8 cm in radius and have a downward facing camera capable of seeing directly below the robot. Targets have a radius of 5 cm, together with the 8 cm robot viewing area this gives a value of 13 cm for parameter r in the DDSA. Robots have a forward movement rate of $16\,\mathrm{cm\,s^{-1}}$ and a rotation rate of $8\,\mathrm{cm\,s^{-1}}$ or approximately $1\,\mathrm{rad\,s^{-1}}$. Robots move 8 cm towards their goal locations between reorientations.

We simulate error by applying Gaussian noise to robot destinations. To replicate our observations that the iAnt robots accumulate error linearly with distance (Hecker et al., 2013), we make the standard deviation increase with the distance from the robot's current position to its destination position, x. We multiply the standard deviation by a noise coefficient, e, in order to change noise severity. Therefore, our noise variates, v, are generated by:

$$v \sim \mathcal{N}(0, \sigma^2) \text{ where } \sigma = d(s, x) \times e \tag{6.2}$$

We do not apply positional noise to robots returning to the central location point for the DDSA or CPFA. We assume that the collection point is marked by a beacon following our previous studies with iAnt hardware (Hecker and Moses, 2015).

6.6.3 Experimental Setup

The problem domain for central place foraging algorithms is target placements within a plane. Each problem instance consists of a set of coordinates representing the locations of targets. To be useful a search algorithm must work effectively across a wide variety of potential target configurations. We measure the performance of the

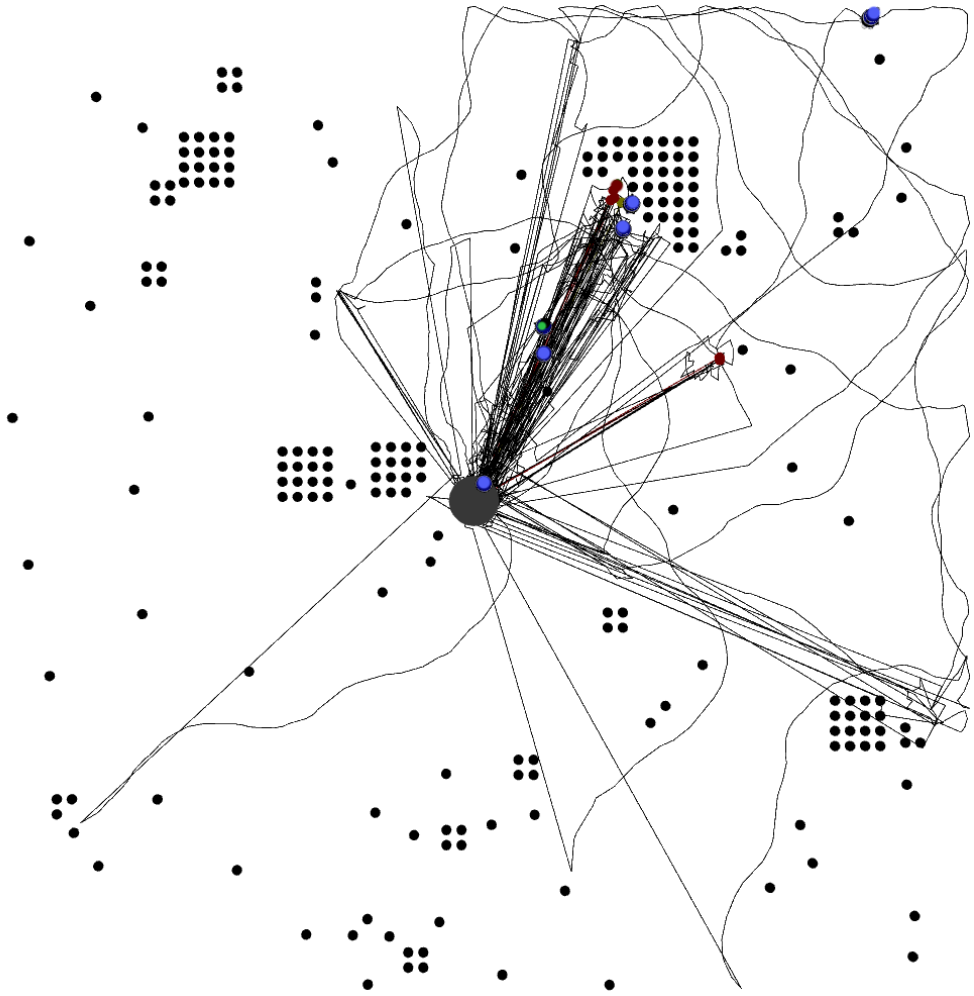


Figure 6.2: The CPFA running in ARGoS, overhead view. The partially clustered distribution of targets is shown as black dots, robots green or blue dots, lines indicate the paths taken by searchers during the experiment. Red dots indicate waypoints used by searchers to communicate the location of dense target clusters to the swarm.

DDSA and CPFA by measuring target collection times over many randomly chosen problem instances. While the DDSA is a deterministic algorithm, its performance on any particular problem instance varies according to the particular placement of targets. Collisions between robots can also introduce non-deterministic effects into the search process.

In our experiments, we place 256 targets in a $100\,\mathrm{m}^2$ arena according to three random distributions 1) uniform, 2) partially clustered, and 3) clustered. The uniform distribution places targets at all locations in the search arena with equal probability. The partially clustered distribution follows a power law with 1 cluster of 16, 4 clusters of size 64, 16 of size 4, and 64 single targets. The clustered distribution consists of 4 target clusters with 64 targets per cluster. If a cluster or target location is occupied a new uniform random location is chosen. We choose the partially clustered distribution of targets as our default target distribution because naturally occurring targets tend to occur in a variety of cluster sizes (Dunning et al., 1992).

All experiments last 30 min, except experiments measuring complete collection time, which run indefinitely. In all figures 25 experimental runs contribute to each data point, except in Figure 6.3 on the next page where we use 50 replicates.

6.7 Results

6.7.1 Performance

Our experiments show that the DDSA performs at least as well as the CPFA in the 6 robot, no error case (Figure 6.3 on the following page). This confirms our expectation that the DDSA has desirable search and collection properties. Beyond those already discussed (closest targets first, complete collection and minimal oversampling), the DDSA always returns to the location at which it last found a target. This is so that the search spiral can be rejoined at the point it was interrupted, but it has the useful side effect of sending robots back to clusters of targets. The process of returning to the location of the last-found target is called site fidelity, and is a common search

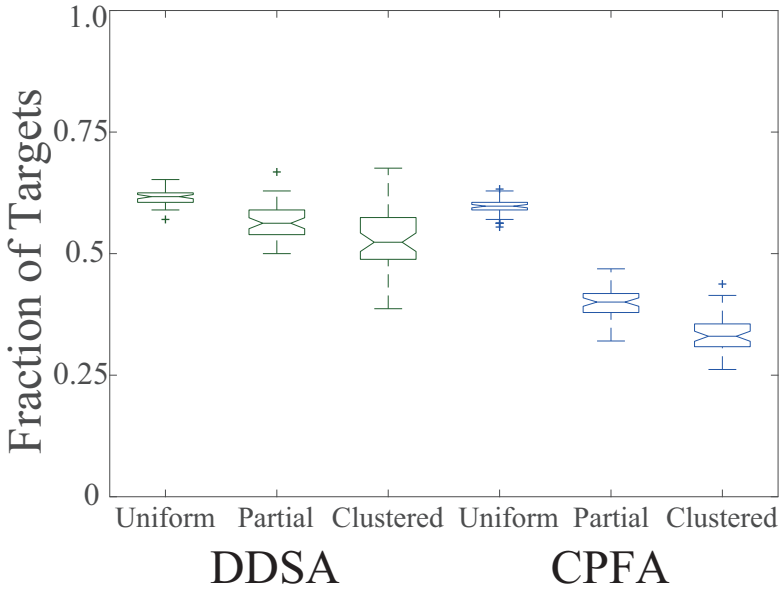


Figure 6.3: Comparison of DDSA and CPFA performance for 3 target distributions. Experiments are 30 min in a 100 m^2 search arena with 6 robots and without noise.

strategy in ants (Beverly et al., 2009; Flanagan et al., 2012). This strategy is also used by the CPFA.

The DDSA partitions the search space among searchers equally after the 0th circuit. This is reflected in the order of performance we observed in Figure 6.3. The uniform target distribution results in an equal allocation of robots to the target collection task, partially clustered less so, and the clustered case least of all. The unequal allocation of robots to targets results in a decrease in performance. Additionally, when a cluster of targets is encountered collisions between robots increase near the cluster.

Similarly, in the CPFA experiments uniform targets are collected fastest, followed by partially clustered targets, and clustered targets are collected slowest. This pattern is reversed from Hecker and Moses (Hecker and Moses, 2015), likely because that prior work did not consider collisions. Once a cluster is detected the CPFA

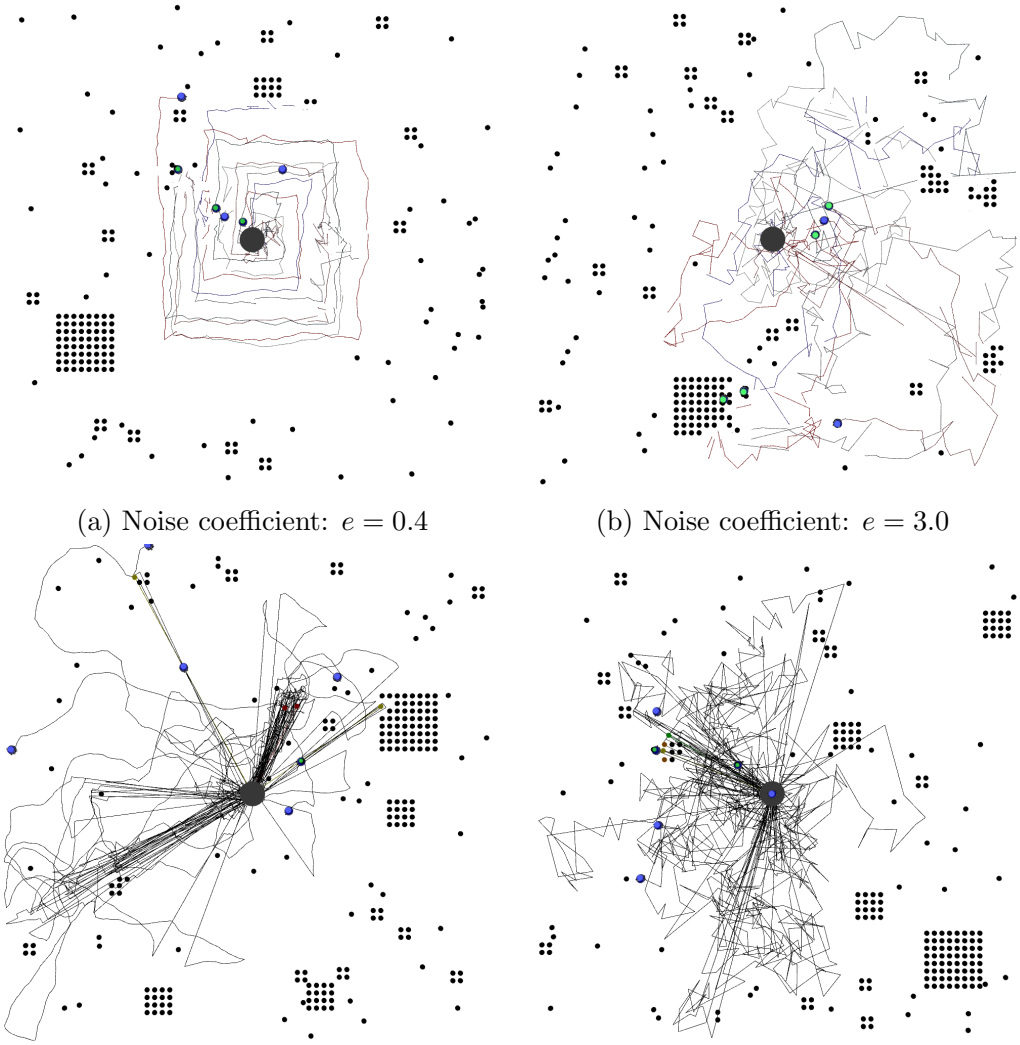
can take advantage of it by recruiting robots to that location, but the initial time to discover a cluster, and the increase in collisions at clusters, offsets this advantage.

6.7.2 Robustness

Degradation of the search pattern under positional noise results in only modest decreases in performance. The number of collected targets is reduced by only 15% between the error free case and our maximum error case (Figure 6.5 on page 133).

In the maximum error case the positional noise is substantial. For example, a robot travelling 10 cm has destination positional error with standard deviation 30 cm (Figure 6.4b on the next page). This robustness to error is due to robots progressively searching locations close to the collection point even with positional noise. Additionally, positional error in one robot may be compensated for by noise in adjacent robots. When the swarm is large, tags that are missed by one robot that is out of alignment with its spiral are likely to be picked up by robots noisily following adjacent paths (Figure 6.4a on the following page).

The CPFA is also robust to error. Compared to the CPFA without positional noise shown in Figure 6.2 on page 128, search paths with $e = 0.4$ are qualitatively unchanged in Figure 6.4c on the following page. In the maximum error case, $e = 3.0$, shown in Figure 6.4d on the next page an increase in path tortuosity is apparent. Noise in the CPFA does not systematically decrease performance (Figure 6.5 on page 133).



(a) Noise coefficient: $e = 0.4$

(b) Noise coefficient: $e = 3.0$

(c) Noise coefficient: $e = 0.4$

(d) Noise coefficient: $e = 3.0$

Figure 6.4: Effect of positional noise on the DDSA and CPFA search patterns. The DDSA is shown in panels (a) and (b) and the CPFA in panels (c) and (d). Black dots are targets, green and blue dots indicate the current location of robots and lines are the paths taken by robots.

6.7.3 Complete Collection

The DDSA guarantees complete collection of targets within the search spiral in the noise free case and as a result performs complete collection tasks faster than the

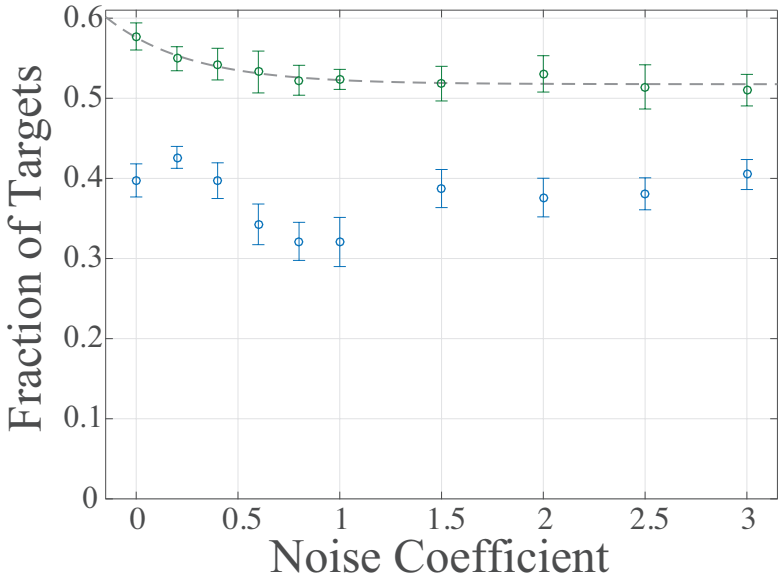


Figure 6.5: In green, DDSA performance scaling with increasing error. Experiment time is 30 min. Dashed line is an exponential decay fit with $R^2 = 0.926$. In blue, CPFA performance scaling with increasing error. The best linear fit has $R^2 = 0.004$ indicating that the noise coefficient, e, explains very little of the variance in performance. Circles are means and bars are the 95% confidence intervals. We use a partially clustered distribution of targets in a 100 m^2 arena.

CPFA (mean decrease in collection time is 59.2%). In the CPFA the time to find uniform targets increases exponentially as the number of remaining targets decreases (Hecker et al., 2015). However, the time to complete collection scales linearly with the number of targets. For each additional target the time for DDSA collection increases by 10.67 s compared to 23.4 s per additional target with the CPFA (Figure 6.6 on the next page).

The 95% confidence interval is tight relative to the stochastic CPFA, as expected for a deterministic algorithm.

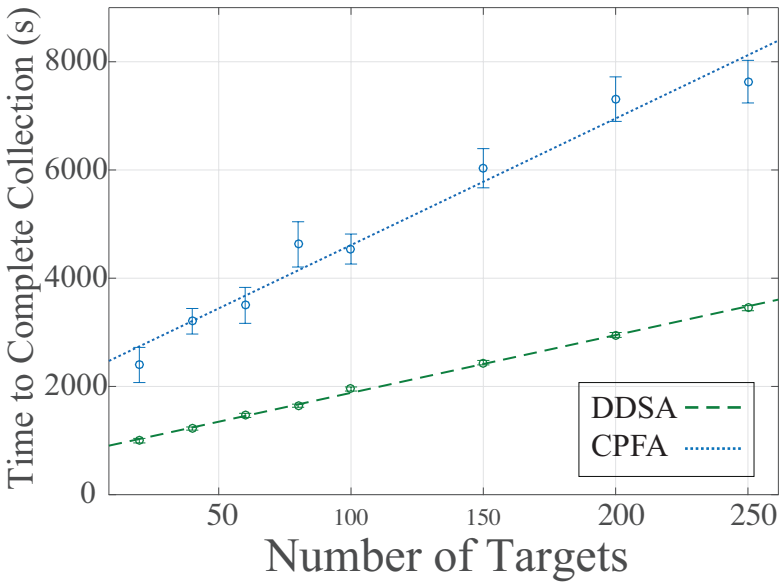


Figure 6.6: In green, with dashed fit line, time for the DDSA to collect all targets vs. number of targets. The dashed line is the linear fit with slope 10.67 seconds per target, $R^2 = 0.998$. In blue, with dotted fit line, time for the CPFA to collect all targets vs. number of targets. Dotted line is a linear fit, slope 23.4 seconds per target, $R^2 = 0.968$. The circles are means and bars are the 95% confidence intervals. We use a uniform distribution of targets in a $100\,\mathrm{m}^2$ arena with 6 robots.

6.7.4 Scaling with the Number of Robots

The DDSA outperforms the CPFA for swarms consisting of between 1 and 15 robots. For swarms with between 20 and 30 robots DDSA performance drops below that of the CPFA (Figure 6.7 on the following page). The DDSA performance curve follows a parabola reaching its maximum between 15 and 20 searchers. Degradation of performance is due to crowding at the collection point (Figure 6.8 on page 136). That crowding at the collection point is the main driver for degradation in performance is supported by our observation of linear scaleup when robots are not required to return to the collection point (data not shown) and in previous work (Fricke et al., 2016c).

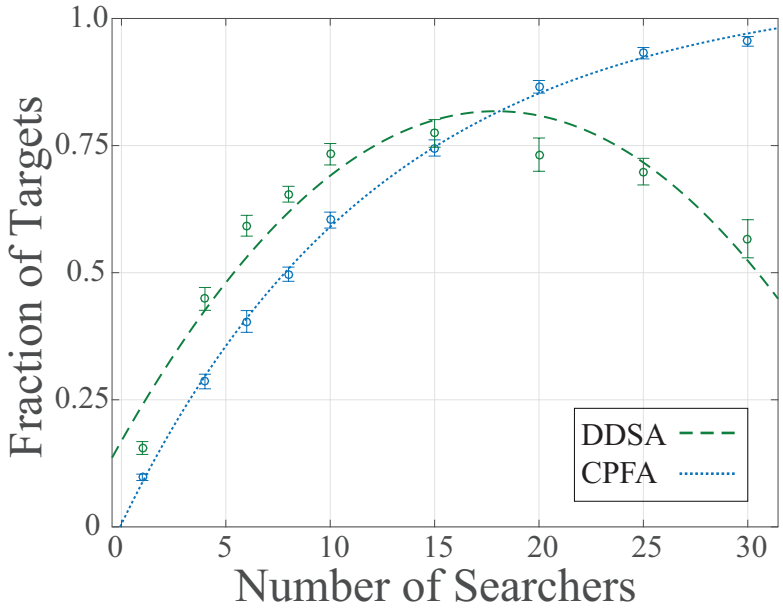


Figure 6.7: In green, with dashed fit line, DDSA performance scaling with the number of searchers. The dashed line is a parabolic fit with $R^2 = 0.998$. In blue, with dotted fit line, CPFA performance scaling. The dotted line is a negative exponential fit with slope with $R^2 = 0.922$. Circles are means and bars are the 95% confidence intervals. We use the partially clustered distribution of targets in a 100 m^2 arena with a 30 min time limit.

In the CPFA we observe a negative exponential increase in performance (Figure 6.7). It is possible that the CPFA is also following a parabolic curve with an inflection point at a much higher number of robots than in the DDSA. Lower levels of congestion at the collection point are likely due to the stochastic nature of the CPFA which reduces the likelihood of robots contending for the same location at the same point in time.

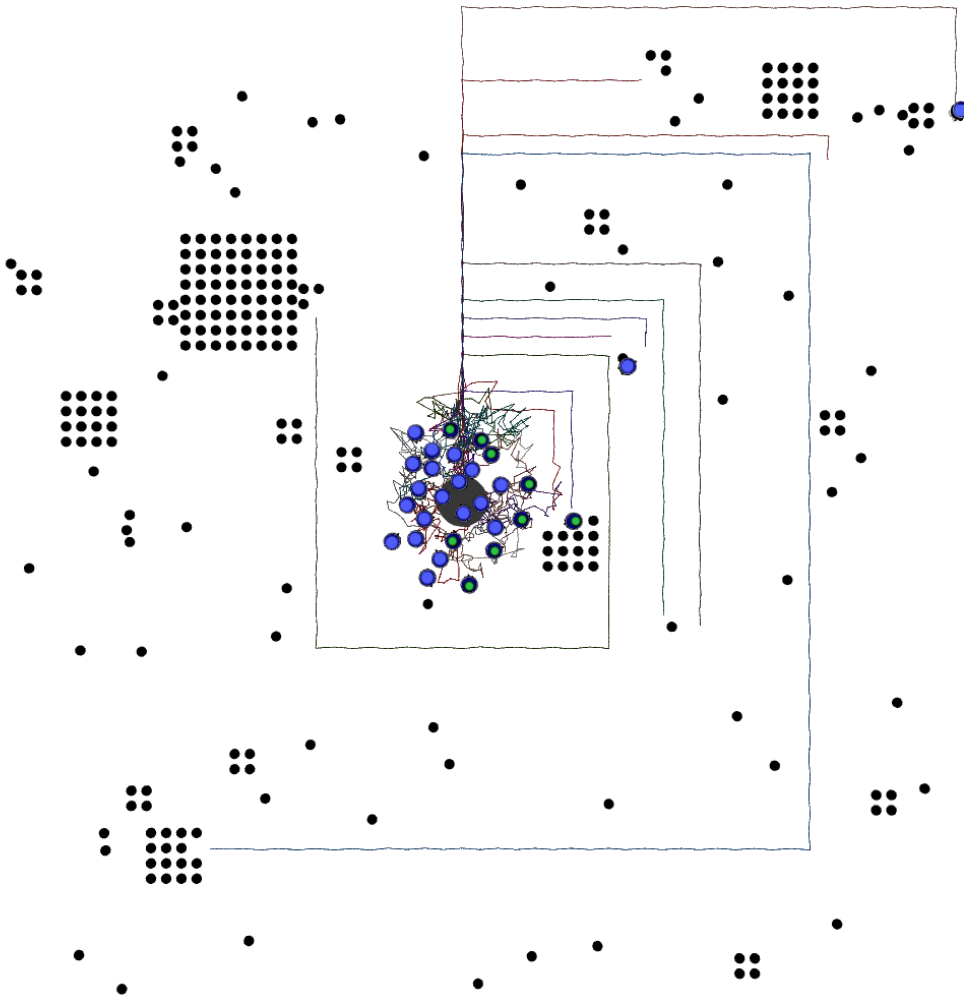


Figure 6.8: Crowding in the DDSA degrades performance. State after 30 min with 30 robots.

6.7.5 Worst Case Performance

Since the DDSA always collects items close to the collection point first an adversary that placed all targets at the edge of the search arena would force the DDSA to maximise its search and collection time. For our examples with a 100 m^2 arena, 6 searchers, 256 targets, a gap of 18 cm and a robot moving at 16 $cm\,s^{-1}$ it takes approximately 4 h to collect all the targets, compared to the observed mean time

of approximately 1 h. Since the DDSA is a preplanned deterministic algorithm, an adversary could force the DDSA to maximise its collection time by placing all targets along the spiral path of a single searcher. The CPFA, in contrast, is a stochastic algorithm that adapts to the distribution of targets. As such, design of a worst case scenario for the CPFA is much more difficult.

6.7.6 Comparison to a Perfect Algorithm

We can calculate the performance of a hypothetical perfect algorithm in which the location of all targets are known to the swarm *a priori*. The expected distance from the collection point $p = (0, 0)$, to a uniformly distributed target, within the $100\,\mathrm{m}^2$ square arena is, without loss of generality, the expected distance to a target, $t = (x, y)$, in the arena's positive quadrant. Numerically solving the double integral over the probability of each Euclidean distance, $\frac{1}{5^2}\int_0^5\int_0^5 d(p,t)dxdy$, gives an expected distance of 3.826 m to each target. For 250 targets the perfect algorithm, neglecting collisions, will collect all targets using a single robot in expected time $\mathbf{E}[C_{\mathrm{time}}] = \frac{250\times2\times3.826\,\mathrm{m}}{0.16\,\mathrm{m\,s^{-1}}} + 250\left(\pi + \frac{\pi}{2}\right)\mathrm{s} = 13\,241\,\mathrm{s}$, where the first term is the linear travel time and the second is the turning time. Therefore, in the 6 robot case complete collection will take an expected 2207 s. The mean time taken by the DDSA to collect 250 targets with 6 robots is 3447, 95% CI [3399, 3493] s, an increase of approximately 56.2% over the theoretical minimum.

The mean distance from the centre of a square of width 10 to its perimeter is 5.74 (Johnson, 1907). Substituting this value for 3.826 in the expected time formula above yields 19 116 s in the single robot case. So for 6 robots we have a collection

time of approximately 53 min. This gives a worst case increase in DDSA collection time over the perfect algorithm of 353%.

6.8 Conclusions

We show that desirable properties of the single agent square spiral, extensively demonstrated in previous work, can be extended to multiple robots. The DDSA has optimality properties which make it ideal for use as a central place foraging benchmark: guaranteed collection of nearest objects first, complete collection, and minimal oversampling. Benchmark algorithms should provide an efficient and theoretically tractable point of comparison for more complex algorithms. By comparing the CPFA to the DDSA we have a better understanding of the CPFA's strengths and weaknesses.

Adaptive search patterns such as the CPFA take advantage of information about target clusters. Doing so increases target detection rates but does not minimise trip time. This is highlighted by the relatively good performance of the DDSA (Figure 6.3 on page 130). This suggests that a modification to the CPFA, the use of distance information when deciding whether a robot should use a waypoint, could be beneficial.

While the DDSA is surprisingly resilient to error the CPFA is even less affected (Figures 6.4c on page 132 and 6.5 on page 133). This suggests that the DDSA can be an effective strategy even for robots with limited ability to localise.

The DDSA solves the complete collection problem optimally, in that there is no redundancy in the search pattern and it guarantees collection of all targets within the spiral. In the 6 robot case, the DDSA mean complete collection time is only 56.2%

greater than the theoretical perfect algorithm, which has perfect prior knowledge of all target locations and no collisions. In comparison the CPFA's stochastic strategy takes much longer to collect all targets; using the DDSA as a benchmark provides a point of comparison that allows us to quantify this difference (Figure 6.6 on page 134).

Two factors make finding scalable solutions to central place foraging difficult. Congestion at the central collection point and the mean distance to the targets both grow with the rate of target collection, which in turn grows with the number of robots. However, the stochastic nature of the CPFA means that it does not suffer as much as the DDSA from the central-point collision bottleneck. This results in the CPFA outperforming the DDSA when the swarm size exceeds 20 robots (Figure 6.7 on page 135). The congestion in the beginning setup phase of the DDSA (Figure 6.8 on page 136) could be mitigated by staggering a time delay when robots begin the spiral, or by moving directly to their positions in circuit 0 without travelling to the centre of the map first. Generalisation to multiple collection points allows for more scalable solutions such as the multiple-place foraging algorithm (MPFA) (Lu et al., 2016).

The DDSA provides both theoretical and practical advantages as a general search algorithm, central place foraging strategy, and a benchmark. We expect this approach will find applications in a wide variety of robot swarm tasks.

Acknowledgments

We gratefully acknowledge funding from a James S. McDonnell Foundation Complex Systems Scholar Award and NASA MUREP #NNX15AM14A for the UNM NASA

Swarmathon. Thanks to Elizabeth Esterly for her video of an ant performing a spiral. Reviewer comments and corrections were invaluable.

6.9 Software

The software used in this chapter is available at:

`github.com/BCLab-UNM/DDSA-ARGoS/release/0.2-beta`

`github.com/BCLab-UNM/CPFA-ARGoS/release/0.1-beta`

Chapter 7

Adaptation of Lévy Exponents to Error and Collisions

> *It is time to study and manage incorrectness in the interest of robustness. We should not shun the trade-off, but rather, we should understand, engineer, and teach computation beyond correctness and efficiency only.*
> -David Ackley (Ackley, 2013)

7.1 Author Contribution Statement

I am the sole author of this chapter under the supervision of Melanie Moses (Associate Professor, UNM Computer Science, Biology, and Santa Fe Institute).

7.2 Introduction

As shown in Chapter 5, the exponent of the power law that characterises the Hausdorff fractal dimension ($\mathcal{H}$) of a Lévy search can be profitably adapted to the config-

uration of targets. We found that the number of searchers only weakly influenced the optimal Lévy exponent. However, we did not model the role of searcher collisions.

In this chapter, we investigate the relationship between the Lévy exponent, swarm size in the presence of collisions, target detection error, and localisation error. To examine these relationships, we require robot simulator that can model collisions. The ARGoS simulator (Pinciroli et al., 2012) coupled with a 2D physics engine (dyn2d) provides the ability to model collisions for swarms of robots while also running very quickly in comparison to other robot simulators. ARGoS is also able to model robots in continuous space rather than as a grid (as in Chapter 5) which allows more accurate modelling of localisation error. We reuse the localisation error models developed for testing the distributed deterministic spiral algorithm (DDSA) in Chapter 6 and implement sensor error simulation to investigate the impact of error on the optimal Lévy exponent. We also move the simulation boundaries beyond the region in which targets are placed. This this prevents searchers from staying in the target area by simply bouncing off the simulation boundaries. In Chapter 5 we evaluated a *free search* task in which targets are consumed as they are encountered. In this Chapter, we apply adaptable Lévy search algorithm (ALSA) to the central place foraging (CPF) task.

Ultimately ALSA is intended to adapt the $\mathcal{H}$ of its search pattern in response to environmental factors. Currently ALSA is able to estimate the target configuration and choose an appropriate search $\mathcal{H}$ according to our findings in Chapter 5. Robots running ALSA individually calculate the Hopkins index of the target distribution, updating the index as each new target is encountered, However, as we show in this chapter the relationship between $\mathcal{H}$ and the target configuration is dominated by the target detection error rate and swarm size. Incorporating detection of these para-

meters is under development. Swarm size can be determined directly, or dynamically by measuring the robot-to-robot collision rate.

In this chapter we define *target detection error rate* to be the false negative rate, we do not consider false positives.

7.3 Significance

Our work so far on swarm search patterns (Fricke et al. (2013) and Chapters 3 and 5) shows that selecting a random search strategy with the appropriate balance of intensity and extent is critical to maximising the efficiency of stochastic search. Here we examine the role of error in choosing the Lévy exponent that maximises search efficiency. Engineering swarms in real hardware is dominated by the need to control for sensor error. Sensor error impacts search patterns in two ways: First, sensor error interferes with the ability of a robot to know its current location and navigate to a new location, a fundamental component of any search strategy. Secondly, sensor error may result in target detection error when targets are encountered.

Selection of Search Lévy Exponent in Robot Swarms

The work done so far demonstrates a useful mapping from the distribution of target locations to the efficiency of a Lévy search pattern given a particular exponent. Our genetic algorithm can optimise the Lévy exponent given multiple presentations of a particular target clustering. Ultimately, we would like each robot to learn the appropriate Lévy exponent for the distribution of targets in its environment online. We assume the distribution of targets, swarm size, false negative rate, and

degree of localisation error are global properties of the search environment. This simplifying assumption allows there to be a single optimal Lévy exponent for each search problem rather than depending on the robot position or time. This assumption is unrealistic but provides an entry point for exploring the relationships between these parameters and the best Lévy exponent, and allows the creation of a static lookup table that maps these input parameters to the ideal, empirically determined $\mathcal{H}$. We have performed experiments in which each member of the swarm calculates the As we will show these parameters interact with one another necessitating the consideration of all these parameters when selecting $\mathcal{H}$.

A critical task for online Lévy exponent selection is a measurement of these parameters by robots in real time. We have performed experiments which demonstrate that individual robots calculating the Hopkins index as they encounter targets are very quickly able to differentiate between uniform and power law/clustered distributions, but not reliably between power law and clustered distributions. The target detection error rate can be measured by having a rover approach a location with a known target and record how often the target is discovered. Swarm size can be calculated via a *gossip* algorithm that performs only local communication. Localisation error depends on an unknown array of sensors, such as wheel encoders, global positioning system (GPS), inertial measurement units (IMUs), magnetometers, etc. Each of which has different noise properties that can be time and space dependent. Some sensors, such as GPS, include noise covariance matrices in their output, others do not. Localisation is a significant challenge in robotics and as such has received a great deal of attention (Cox, 1991; Borenstein et al., 1996; Weingarten and Siegwart, 2005; Moore and Stouch, 2016). Our primary method of measuring localisation error is the extended Kalman filter (EKF). EKFs have been important in our implementa-

tion of the ALSA and DDSA strategies in the Gazebo simulator and physical robots. In this chapter, we examine the swarm size question again in the presence of collisions. We are effectively measuring the influence of collisions as a function of swarm size in comparison to that of oversampling which also depends on swarm size.

A strength of the Lévy approach is that it does not require communication between members of the swarm. We endeavour to maintain those properties while enabling online estimation of the distribution while recognising that measuring parameters such as swarm size requires local communication at the start of the search process.

ALSA is currently not capable of estimating all these parameters. Instead, in this chapter, we focus on mapping the relationships between these parameters and the best Lévy exponent.

7.4 Related Work

Chapter 5, Section 5.6 on page 87 summarises work related to Lévy search by robots.

7.5 Methods

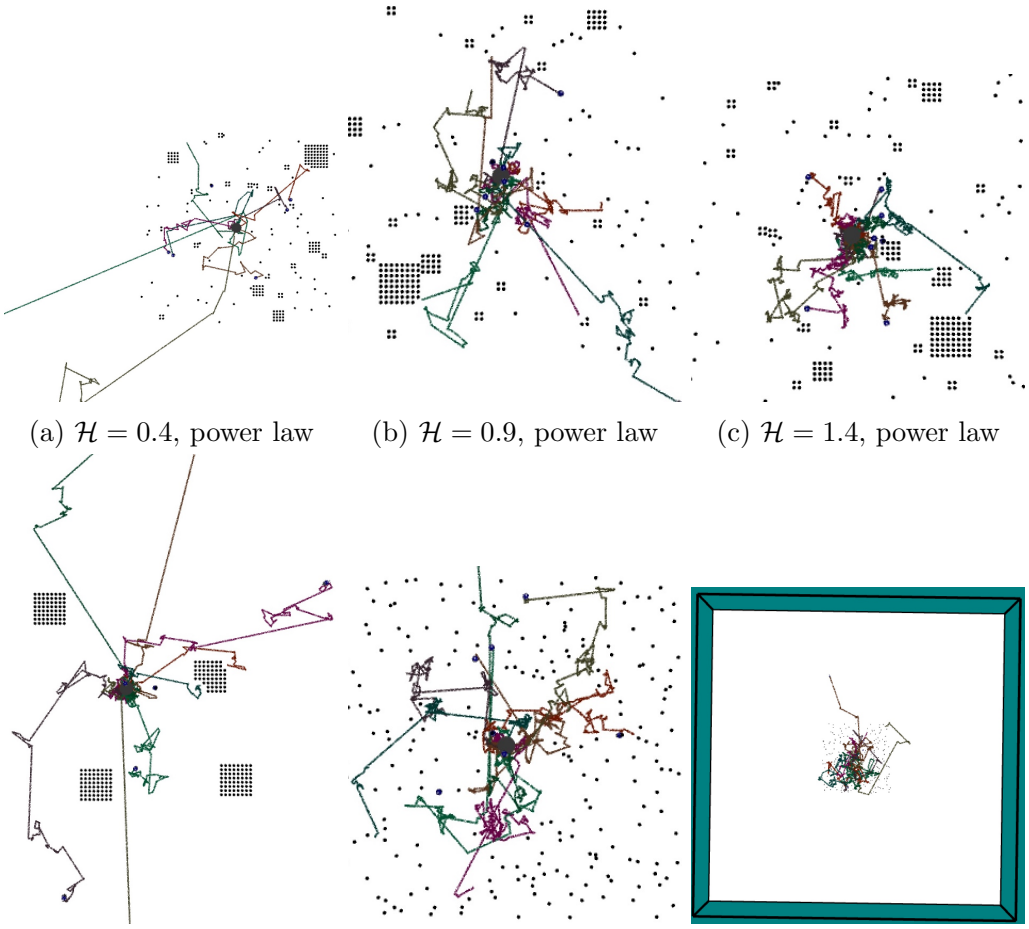
7.5.1 Extending the search area boundaries beyond the area containing targets

In this work, we make a change in the size of the arena. We, and others (Levin, 2016), have observed that search agents in search areas with reflective boundaries may exploit ballistic motion to cover the space efficiently. Reflective boundaries in the short time scales of our experiments allow searchers to move in a straight

line without being in danger of leaving the search area. This observation follows a prediction of Viswanathan et al. (2000) for a uniform distribution of targets in an infinite plane without target replacement.

Figure 7.10 on page 159 shows the penalty that a swarm of searchers would pay if a Lévy search pattern were employed that assumed there was no target detection error. The best μ, in that case, is 1.4, but if robots experience a detection error rate of $\frac{7}{8}$ per second, then a μ of 1.4 would result in a search strategy that collects 81% fewer targets than if a μ of 2.3 were used instead. The most efficient strategy in the error-free case under-performs in the cases with error, Lévy walks provide a simple mechanism for adjusting and maximising search efficiency even in the presence of target detection error. Intuitively the $\mathcal{H}$ of a search pattern trades efficiency for robustness to error but probabilistically re-sampling. This has connections to the efficiency-robustness trade-off discussed by Ackley et al. (2012).

Our Lévy search implementation has two related properties that emerge naturally from the power-law distribution of steps: avoiding collisions and remaining in target rich environments. As searchers move, their steps are interrupted by the discovery of targets and by collisions with other rovers. In target rich areas even very long step lengths will be interrupted by the discovery of a target. Since most step lengths will be short, the searcher is strongly biased to remain in that area. The probability of escaping increases as targets are locally depleted. Since the probability exists of very long steps, once an escape path becomes available the searcher may travel a long way into previously unsearched areas. The probability of moving out of the region with targets as a function of the local target density decreases is mediated by the value of μ, that governs the skewness of steps taken (Figure 7.1 on the next page).



(a) $\mathcal{H} = 0.4$, power law (b) $\mathcal{H} = 0.9$, power law (c) $\mathcal{H} = 1.4$, power law

(d) $\mathcal{H} = 0.9$, clustered (e) $\mathcal{H} = 0.9$, uniform (f) Wall positions.

Figure 7.1: **Lévy Search Patterns are Naturally Adaptive**Search patterns depend on the intrinsic movement of searchers, and also on interactions with targets. The distribution of targets governs, in part, the fractal dimension of the search pattern. Searchers are corralled in high resource areas by the targets themselves. The $\mathcal{H}$ determines the weight searchers give to being constrained to an area by the targets that are there. Compare the locality maintained by searches in regions with a uniform distribution of targets (d) vs. those in highly clustered environments (e). The fraction of available paths that end with a targeted encounter is greater in the uniform case. In the clustered case targets tend to shadow one another resulting in a much smaller target silhouette. This is not necessarily detrimental since searches should move to new areas if locally targets are hard to locate. $\mathcal{H}$ balances the tradeoff between remaining in a region with targets and moving on.

A similar process occurs with robot collisions. Long steps are preferentially interrupted preventing searchers from escaping a region dense with other searchers. Too high a density of searchers results in inefficiencies as robots waste time in collision avoidance behaviour instead of searching for targets. It also results in oversampling, though oversampling can be a benefit if targets are difficult to detect. In this case then decreasing $\mathcal{H}$ to produce enough long steps that searchers can escape areas with many searchers dominates the optimal selection of the Lévy exponent μ. This effect can be seen in the decrease in $\mathcal{H}$ as swarm size increases in Figure 7.4a on page 154.

We use the ARGoS swarm robot simulator with simulated iAnts as described in Section 6.6.2 on page 126. The experimental setup is largely as described in Section 6.6.3 on page 127, that is we distribute 256 targets in uniform, power law, and clustered configurations.

We place targets within the same 10×10 m area described in Section 6.6.3 on page 127, but place the walls 20 m from the centre so they enclose a 40×40 m region.

This was unnecessary in our study of the DDSA in Chapter 6 since in the error-free case searchers follow the boundary rather than being reflected by it and so gain no advantage. However, the central place foraging algorithm (CPFA) is constrained by the arena barriers and so does not suffer the possible downside of stochastic search strategies in which searchers leave the search area entirely. The CPFA contains a mechanism, the *give-up-rate*, that was included based on observations that ants limit the distance they travel from the nest. In our experiments with the CPFA in environments where the area containing targets is the same as the entire searchable space the CPFA's genetic algorithm (GA) consistently selects values for this parameter that approach zero. Small values imply that there is no penalty for extensive search strategies approaching ballistic motion in an effectively unbounded target space, but

in which the maximum distance from the nest is limited. Increasing the boundary distance in this chapter is intended to prevent this artefact from causing ALSA to trivially select ballistic motion as the best solution.

Fig 7.2 on page 153 shows the dependence of the best search $\mathcal{H}$ on the number of searchers. Despite the large variance a relationship between $\mathcal{H}$ and the number of searchers is discernible. As the swarm size increases the most effective $\mathcal{H}$ decreases. The influence of the target distribution on the best $\mathcal{H}$ is also apparent, with the 1.2 being a better choice for the uniform distribution than for the power law and clustered cases. This may be the result of uniform distributions corralling searchers with higher $\mathcal{H}$ more effectively than clustered configurations (compare Figures 7.1a and 7.1d with 7.1e on page 147).

7.5.2 Going Big

Our simulations have all been designed so that they can be compared to our physical swarms. We can then validate that our results hold outside the simulated world. For that reason, we limit our simulations to relatively small search spaces with swarm sizes that are close to the number of physical rovers that we have built.

However, a benefit of simulation is that we can go beyond the bounds of the number of robots we can build and the space we can reserve for experiments. Ultimately, with real applications in mind the areas swarms of robots must search will be larger than those we have discussed so far. The battery life of the UNM-NASA swarmie robots we have developed is currently 10 h. To that end, Figure 7.9 on page 158 shows the results of increasing the target arena from 100 m to 250 000 m and 1 km^2. We also increase the search time to 8 h and the swarm size to 256 robots. The num-

ber of targets remains 256 in a power law configuration. The time required to run simulations this large is significant. As a result we are able to run 30 replicas for the 0.25 km case but only 5 for the 1 km experiment. We include the CPFA evolved for a power law distribution of targets as a point of comparison. It should be noted however that the CPFA we use was evolved for a small arena not ones of the sizes we use here.

7.5.3 Mapping the target detection error rate per second to simulation iterations

The target detection error rate is calculated by solving for the probability in the Bernoulli equation for 32 failures with the desired probability. We use 32 failures because there are 32 simulation ticks per second. $t = 32$ ticks per second. These are Bernoulli trials. With the probability of t failures being equal to $d = (1 - p)^t$. (d is the probability of no detections in 1 second, p is the per tick detection probability).

Eqn. (Eq. (7.1)) Solve for p in terms of d:

$$d = p^t \implies p = d^{-t} \tag{7.1}$$

This simple conversion results in the detection rates shown in Table 7.1 on the next page.

We perform analysis of variance (ANOVA)s to map the relationship between the optimal mu that results from a parameter scan over μ. The mean $\mathcal{H}$ that resulted in the most targets being collected is recorded and entered into theANOVA as the output. The influence of each of the factors in the selection of μ can be estimated.

Error Rate (s)	0	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1
Detection Rate (1/32 s)	1	0.0629	0.0424	0.0302	0.0214	0.0146	0.0089	0.0042	0

Table 7.1: **Error rate to simulation detection rate.** We calculate values for 0 probability of target detection errors per second, and progress in increments of $\frac{1}{8}$ until we get to a target detection error rate of 1.

The ANOVA tables can be found in Appendix B on page 203 along with description of the ANOVA and caveats.

7.6 Results

Figures 7.1 on page 147 shows search patterns associated with $\mathcal{H}$ and various target configurations. Appropriately chosen $\mathcal{H}$ keeps the searchers within the target area while higher $\mathcal{H}$ cause searchers to exit the target area reducing search efficiency. The target distribution also has an impact on containment of searchers in the area with targets. Comparing Figure 7.1b, d, and e the probability of encountering a target when taking a large step in a straight line is greatest for the uniform distribution of the three target configurations. In contrast, searchers in the clustered case tend to escape the area containing targets. The $\mathcal{H}$ interacts with the target distribution, causing the effective $\mathcal{H}$ to be higher in the presence of a uniform distribution of targets than in the clustered case. This is in contrast to our findings in Chapter 5, Figure 5.9a on page 108 where we saw the reverse trend. However, in that case, there was no penalty for choosing a low $\mathcal{H}$ in the presence of target clusters since the arena boundaries prevented the searchers from leaving the area containing targets.

Figure 7.2 on the next page demonstrates that the mean value for the best performing μ is near 2 for the single searcher case for all distributions. The most efficient choice for μ falls as the number of searchers increases. This is consistent with higher $\mathcal{H}$ resulting in less unnecessary oversampling by spreading out searchers, and also with reducing the number of collisions between searchers.

The increase in the choice of μ that maximises target collection as detection error increases is shown in Figure 7.3 on page 154. The swarm increases oversampling as a way to compensate for the lower probability of an individual robot detecting targets. The trend is less pronounced in the clustered case than in the uniform case because the probability of a searcher finding a target, even with a high detection error rate,

is still good when targets are clustered together. When a robot encounters a cluster of 64 targets the probability of seeing at least one of them compensates for the low target detection rate, leading to the effective use of site fidelity. This is not the case when targets tend to be isolated, as in the uniform random distribution.

Figures 7.5a on page 155 and 7.4a on the following page summarise the dependence of μ on detection error and on swarm size respectively. These figures show the effect of target detection error and swarm size on the best performing values of μ over all target configurations. Figures 7.5b and 7.4b differ from one another because 7.4b contains data for a range of swarm sizes with zero error, whereas Figure 7.5b has a range of error but a fixed swarm size of 8.

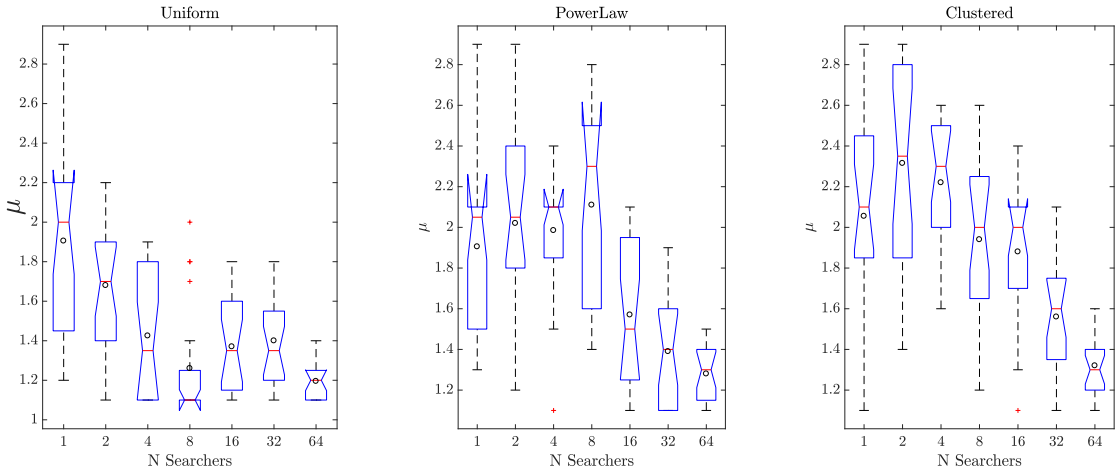


Figure 7.2: Dependence of optimal μ on the number of searchers by target configuration. Interquartile boxplot. The selection of optimal $\mathcal{H}$ depends on the number of searchers and the target configuration. 20 experiments per box.

Figure 7.10 on page 159 shows the effect on the mapping between the Lévy exponent μ and the number of targets collected. The maximum in the curves shifts towards patterns with increased search intensity as target detection error increases.

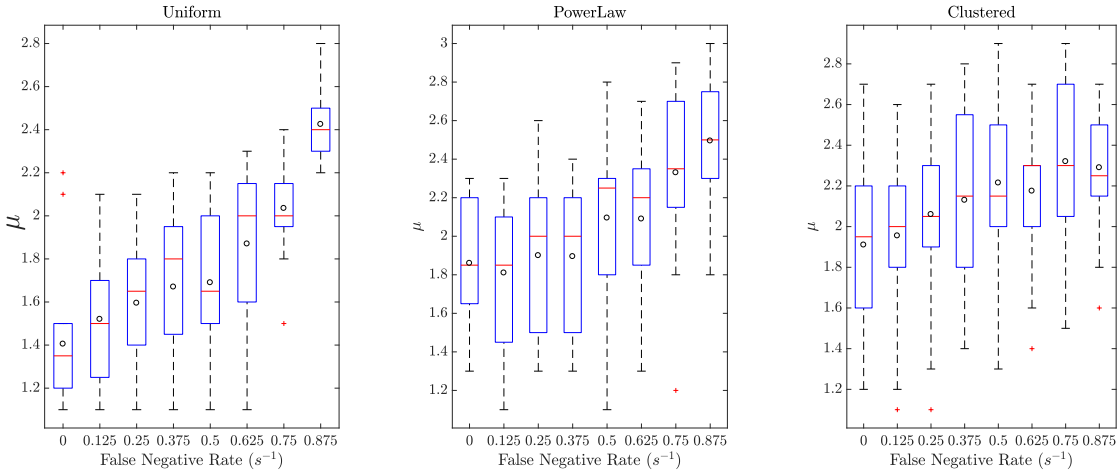
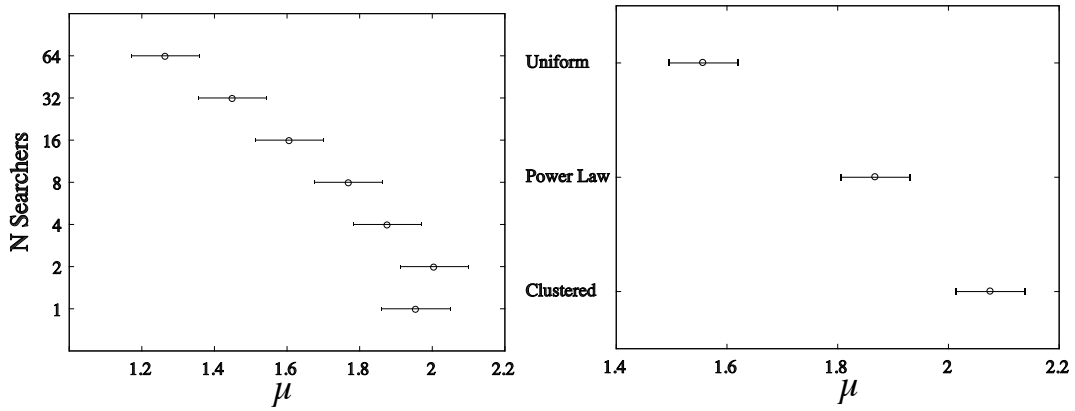


Figure 7.3: Dependence of Optimal μ on the Target Detection Error Rate. Interquartile boxplot. The selection of optimal $\mathcal{H}$ depends on the target detection error rate and the target configuration. 20 experiments per box



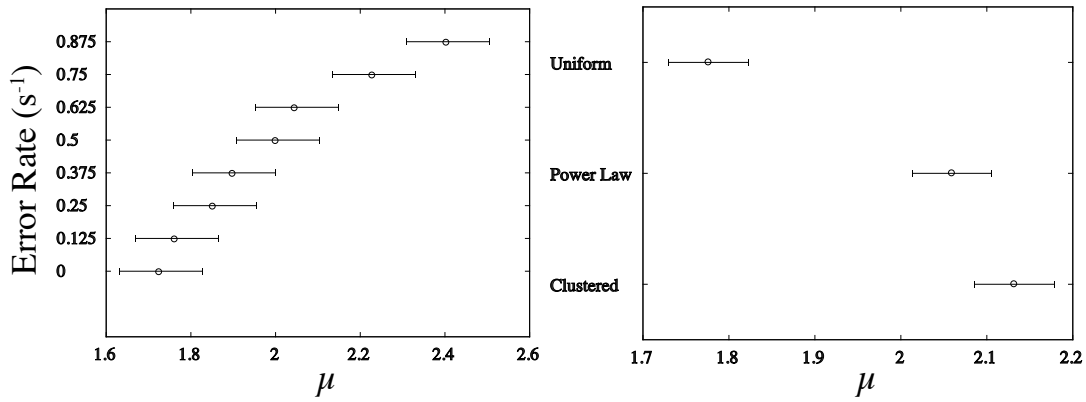
(a) Optimal μ by swarm size.

(b) Optimal μ by target configuration.

Figure 7.4: Dependence of optimal μ on swarm size and the target configuration. No localisation or target detection error. 256 Targets. 40×40 m arena. Targets placed in a 10×10 m region.

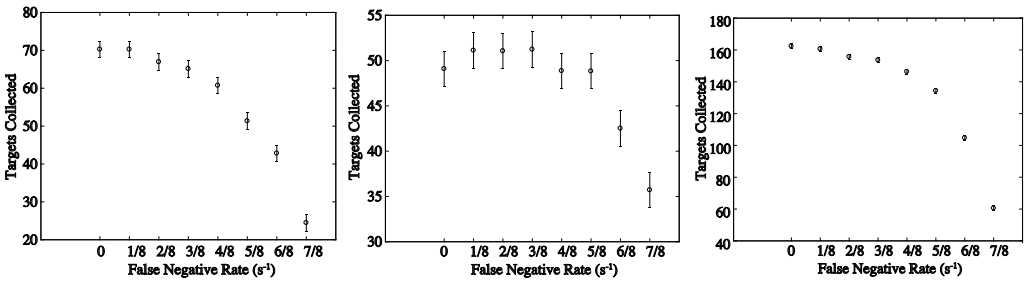
This may be explained by the compensation for increased target detection errors by additional oversampling, and so providing more opportunities for a target to be discovered. This experiment establishes that error in detection of targets makes the selection of corresponding $\mathcal{H}$ worthwhile in at least some cases, with an 81%

difference in efficiency in Figure 7.10. In order to explore this relationship further with different target distributions we further analyse the data by performing ANOVA (Table B.1 on page 205) and multiple comparisons (Figure 7.3 on the previous page) over a range of target configurations and target detection error rates. This shifts the best performing μ to higher values for all distributions to compensate for the target detection errors, and to the lower values when compensating for the oversampling and collisions that result from larger swarms.



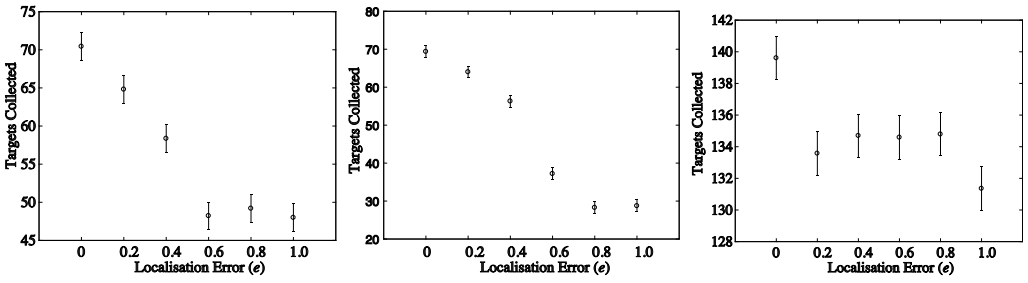
(a) Optimal μ by error rate. (b) Optimal μ by target configuration.

Figure 7.5: Dependence of optimal μ on the error rate and the target configuration. No localisation error. 8 searchers. 256 targets. 30 min time limit. 40×40 m arena. Targets placed in a 10×10 m region. 60 repetitions per bar.



(a) ALSA $\mathcal{H} = 0.7$ (b) ALSA $\mathcal{H} = 1.4$ (c) DDSA

Figure 7.6: Multiple Comparison Plots for target detection errors. 30 experiments for each value of the target detection error rate. ALSA with $\mu = 1.7$ shows a decrease in efficiency from 0 error to a 7/8 target detection error rate of 71.4%, with $\mu = 2.4$ the reduction is 30%. For the DDSA the reduction is 62.5%. Power law configuration of targets. No localisation error. 256 targets. 40×40 m arena. Targets placed in a 10×10 m region. 30 min time limit. 60 repetitions per bar.



(a) ALSA $\mathcal{H} = 0.7$ (b) ALSA $\mathcal{H} = 1.4$ (c) DDSA

Figure 7.7: Multiple Comparison Plots for Localisation Error. ALSA is more susceptible to performance degradation than the DDSA as localisation error increases. 30 experiments for each value of localisation error. Power law configuration of targets. No localisation error 256 targets. 40×40 m arena. Targets placed in a 10×10 m region. 20 repetitions 40×40 m arena. 30 min time limit. 60 repetitions per bar.

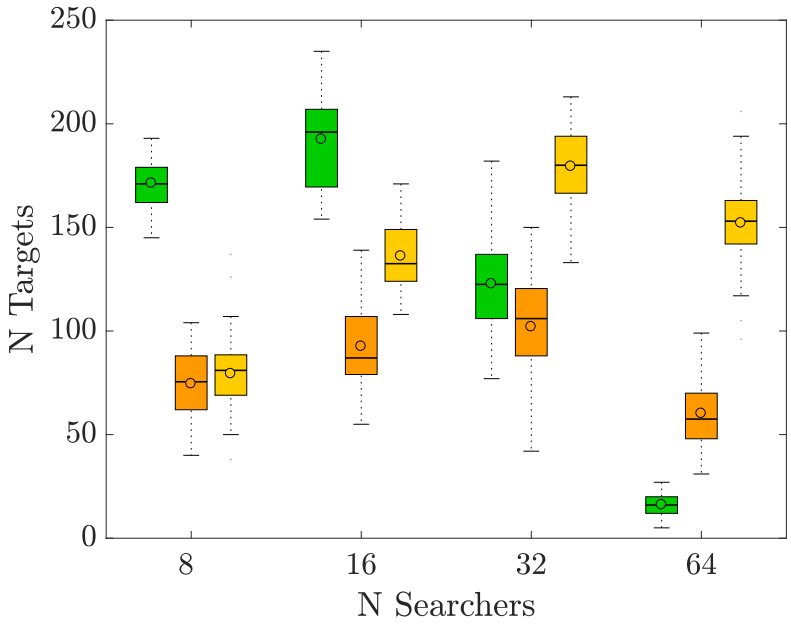


Figure 7.8: DDSA and ALSA efficiency for various swarm sizes for 30 min time limits The DDSA is more efficient than ALSA for smaller swarms but performance degrades drastically for larger swarms. ALSA with higher $\mathcal{H}$ (1.4) follows a similar pattern to the DDSA but the degradation due to swarm size is delayed. ALSA with a lower $\mathcal{H}$ (0.7) is able to sustain performance increases to the largest swarms we consider. 30 experiments per box. Power law configuration of targets. No localisation error and no target detection errors. Time limit is 30 min. 256 targets. The DDSA has a longer setup time than ALSA caused by increased congestion at the start point.

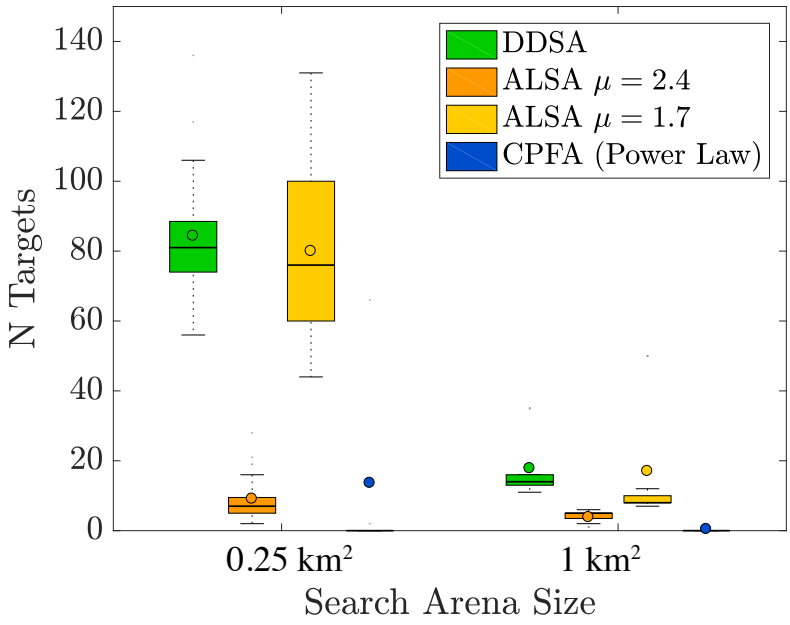


Figure 7.9: **Search Algorithm Comparison for Large Arenas.** Experiment setup: 256 searchers, 256 targets in a power law distribution 8 h. Both the DDSA and ALSA with $\mu = 1.7$ are able to find and collect targets in very large areas with low density. In the 1 km^2 case the target density is 1:10,000 that of our previous experiments. The CPFA is presented for comparison, however this is the CPFA with parameters described in Hecker and Moses (2015), which were evolved for a 10×10 m arena.

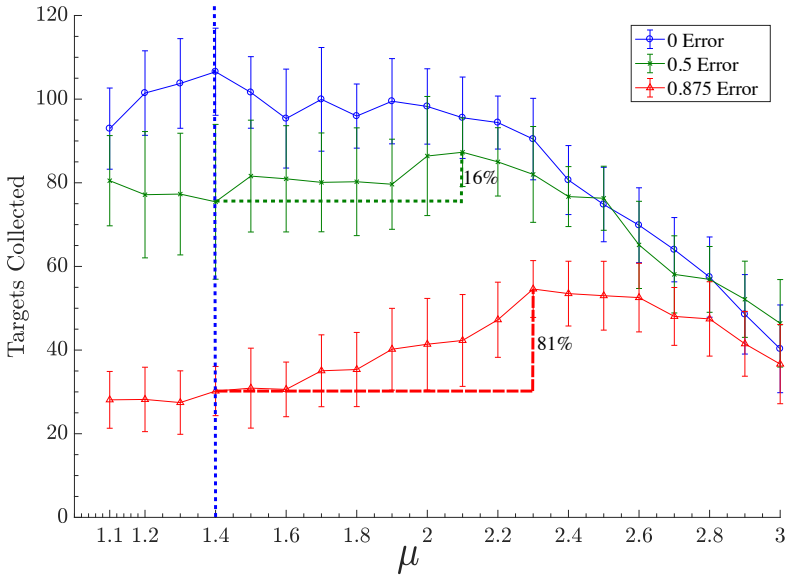
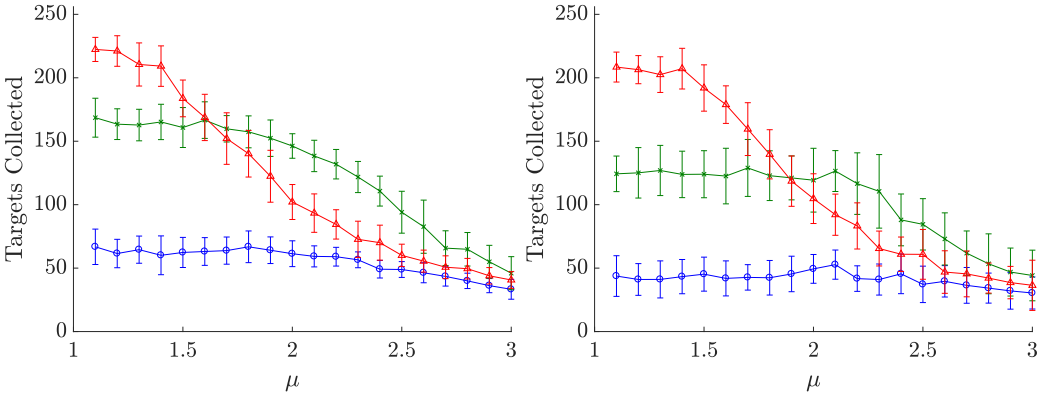
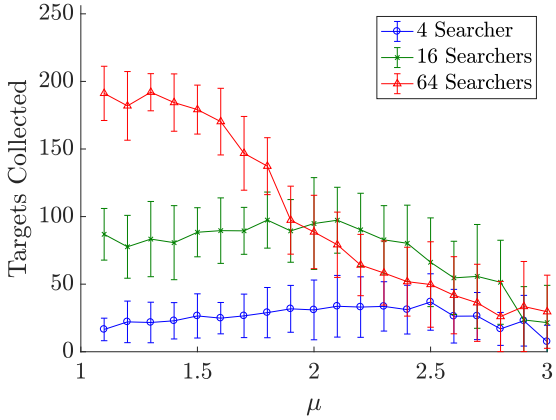


Figure 7.10: **Effect of Target Detection Error on Search Efficiency of μ for the uniform distribution of targets.** Vertical lines indicate the μ that results in the best performance. Dashed lines and associated percentages show penalty in efficiency of choosing the best μ for the error free case and applying the resulting search pattern to the 0.5 target error case (green dotted) and 0.875 target error case (red dashed). Experiment setup: 8 searchers, 30 min, uniform distribution of targets.



(a) Effect of Swarm Size on Search Efficiency of μ - Uniform

(b) Effect of Swarm Size on Search Efficiency of μ - Power Law



(c) Effect of Swarm Size on Search Efficiency of μ - Clustered

Figure 7.11: Dependence of optimal μ on swarm size and the target configuration. No sensor error. 40×40 m arena, targets restrivted to a 10×10 m area.

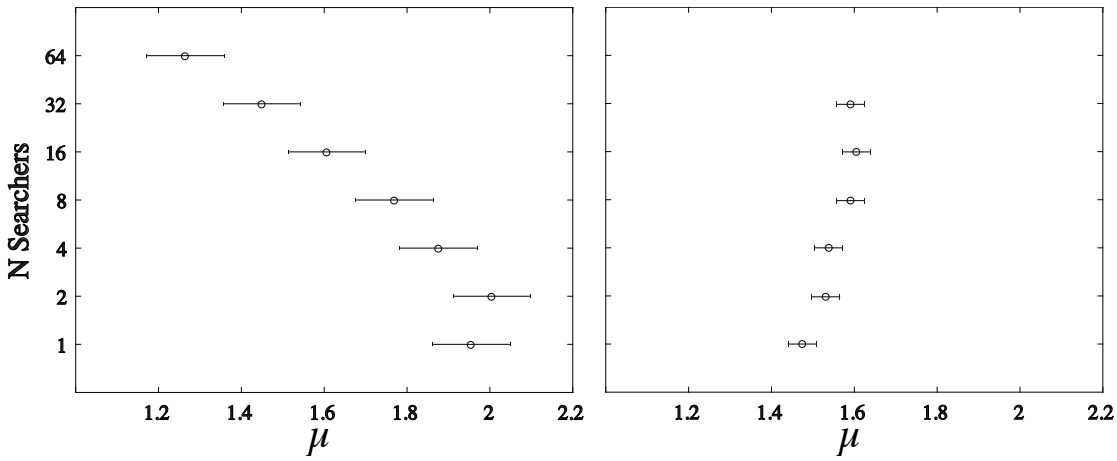


Figure 7.12: Adding collisions causes μ to depend on swarm size. Left panel is a reproduction of Figure 7.4a. The right panel is a representation of Figure 5.9b, showing that the optimality of μ in this range does not depend on the number of searchers.

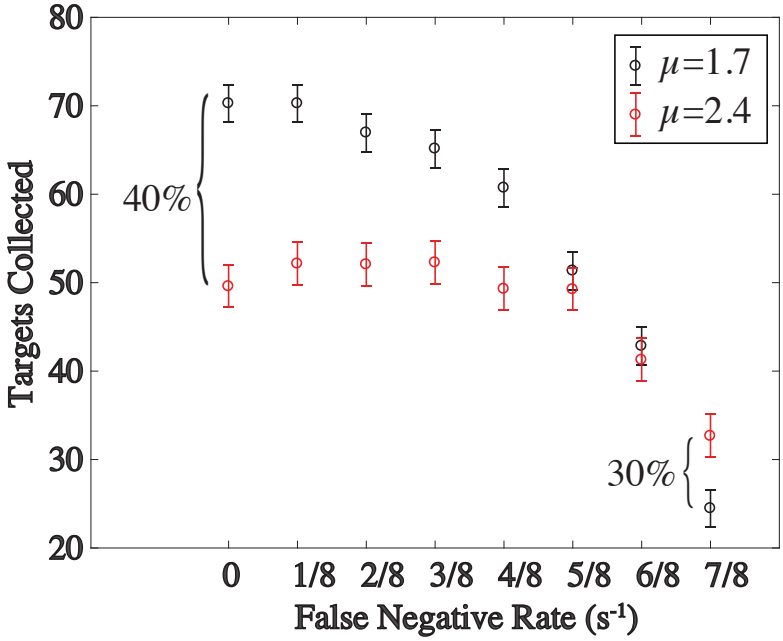


Figure 7.13: The choice of μ affects efficiency in the presence of target detection error. Selecting values of μ that result in lower fractal dimension results in more efficient search when the target detection error rate is low. Conversely, low fractal dimensions perform better as target detection error rates increase.

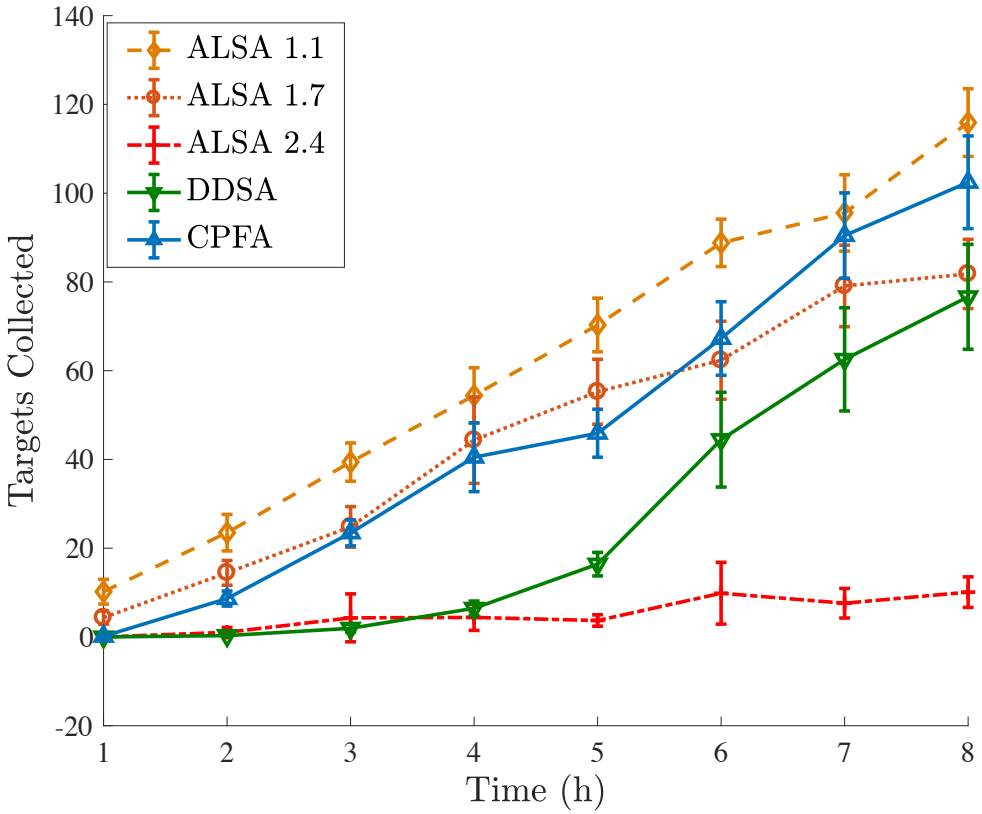


Figure 7.14: Search strategy efficiency over long times. The distributed deterministic spiral algorithm (DDSA) increases in efficiency over time as the swarm is able to resolve initial crowding. Stochastic algorithms such as adaptable Lévy search algorithm (ALSA) and central place foraging algorithm (CPFA) are relatively linear in efficiency over time, since crowding is less of a problem. Swarm size is 256, 256 targets in a power-law distribution, 0.25 km^2 arena size.

7.7 Discussion

Tables B.1 and B.2 on page 205 indicate that swarm size, target detection errors, and target configuration are significant influences on the value of $\mathcal{H}$ that optimises search efficiency.

ANOVAs analysis suggests that the target detection error rate has a greater influence than localisation on foraging efficiency for ALSA with $\mathcal{H} = 0.7$ (Table B.3) and the DDSA (Table B.5 on page 206). In the DDSA case this confirms our surprising result from Chapter 6 that localisation has very little influence on the efficiency of the algorithm. This is reversed for ALSA with $\mathcal{H} = 2.4$ (Table B.4 on page 206). Revisiting sites by increasing $\mathcal{H}$ may help to compensate for lossy sensors that increase the probability that targets will not be detected. The significant difference in performance due to selecting the appropriate $\mathcal{H}$ is shown in Figure 7.13 on page 161.

The number of searchers influences the optimal $\mathcal{H}$ but only in the presence of error. In Chapter 5 we observed the influence of target distribution on the optimal $\mathcal{H}$ but in that case, the simulated robots had target detection error rate of 50% per second as determined from observation of the physical iAnt robot.

Multiple comparison analysis shows that all three target distributions influence $\mathcal{H}$ in the absence of error but that when error is introduced both power law and clustered distributions result in statistically indistinguishable values of $\mathcal{H}$, however, a significantly different value of $\mathcal{H}$ is optimal for the uniform case.

The distribution of target detection errors may serve to make the power law distribution appear more clustered. We have observed that single targets are easily missed when target detection errors are possible. Clustered resources, on the other

hand, tend not to be missed because there are more opportunities to see at least one target. Once one target has been detected site fidelity ensures that other targets in the cluster may be detected.

The reduction in efficiency due to localisation error in ALSA may be a consequence of not reliably returning to the last position that a target was found (site fidelity). This gives an indication of how much of the ALSA success depends on site fidelity as compared to that of the DDSA, assuming the DDSA reduction in efficiency is also due to interference with site fidelity, though this remains unclear.

Figure 7.8 on page 157 shows an advantage of ALSA over the DDSA; it can maintain efficiency for larger swarms where the DDSA performance degrades due to the increased number of collisions produced by a larger swarm. Since the DDSA is deterministic, there are particular points in the search space that must be reached for the algorithm to progress. If those points become unreachable due to occlusion by robots in the swarm the performance of the DDSA degrades drastically. In contrast, due to the stochastic nature of ALSA no points other than the central collection point can become bottlenecks. ALSA by decreasing $\mathcal{H}$ can increase the extent of search and more effectively disperse searchers away from one anther, reducing the number of robot collisions. Also as the number of searchers goes up, oversampling of the search area near the central collection point increases. Increasing $\mathcal{H}$ reduces oversampling and increases performance. This is clearly apparent in our observations of ant colonies. Larger colonies disperse their foragers over a much wider area, in part to avoid oversampling near the nest (Moses, 2005; Flanagan et al., 2012).

Robots are increasingly expected to operate in environments that have not been designed to accommodate them. This requires that robot swarms be capable of

performing search tasks in the presence of error. We examine the role of localisation and target detection error in selecting an appropriate $\mathcal{H}$ for ALSA.

We find that localisation error has a surprisingly strong impact on ALSA compared to the DDSA. The target detection error rate impacts both the ALSA and DDSA, but can be compensated for by ALSA by increasing the $\mathcal{H}$. (Tables B.3, B.4 and B.5 on page 206, and Figure 7.7 on page 156).

ALSA outperforms our benchmark DDSA as swarms increase in size. However, as Figure 7.14 on page 162 shows, the DDSA closes the gap as over longer time periods as it overcomes the initial crowding of searchers at the collection point. Decreasing the $\mathcal{H}$ of ALSA allows it to avoid collisions by disbursing the searchers. Since this is in opposition to how ALSA would compensate for a high target detection error rate the interaction of these two factors becomes particularly interesting.

We have shown that the impact of error on Lévy search can be quantified and that choosing appropriate Lévy exponents can be used to, at least in part, compensate for degradation due to the target detection error rate and swarm size. Swarm size is easily determined and could be used to increase the performance of the swarm by increasing searcher dispersal.

Notice the reversal of fractal dimension as a function of the target distribution. Here lower fractal dimension is associated with more clustered target configurations. This is the opposite trend compared to Chapter 5. A possible explanation comes from the change in boundary placement. In Chapter 5, the boundary closely surrounded the targets, so a high fractal dimension was not penalised by leaving the target placement region. However, in this chapter higher fractal dimensions do result in leaving the area and so result in unproductive searchers. The distribution of targets impacts the probability that a large step will take the searcher out of the arena. This

is because the probability of encountering a target and returning home is greatest for the uniform distribution. Therefore in the clustered case, where escape from the productive region of the arena is most likely, the extent of search is reduced to compensate.

This chapter highlights the interaction of Lévy search with the distribution of targets. Lévy search is characterised by a power law distribution of movements. Critically, those movements are not made independently of the target distribution. When a searcher following a long trajectory encounters a target it's movement is stopped, the target collected, and a new movement vector and angle selected. The angle will be uniformly chosen and but the step length is likely to be near the current location. This increases the probability of finding another target if the positions of targets are correlated. Even if a long step is selected, there is a chance that it will be truncated by the discovery of another target nearby. If the region is depleted of targets, these long steps will not be intercepted, and the searcher will move to a new area where there may be targets. Figure 7.1f on page 147 shows how Lévy searchers tend to be corralled in regions with targets. This also highlights the importance of site fidelity in the CPF task, without which Lévy searchers would not be able to take advantage of target locality. This also provides an explanation of why localisation error is so detrimental to Lévy searchers in the CPF task since they are unable to reliably return to the site where they last found a target.

7.8 Software

The software used in this chapter is available at:

`https://github.com/BCLab-UNM/CPFA-ARGoS/releases/tag/v0.`

32-beta-21-g9a2a80b

https://github.com/BCLab-UNM/DDSA-ARGoS/releases/tag/0.3-beta

https://github.com/BCLab-UNM/ALSA-ARGoS/releases/tag/0.1-beta

Chapter 8

Conclusions

8.1 Swarm Search

Swarm search is a fundamental task in biological and robotic systems. The work presented in this dissertation leads to a better understanding of Lévy search in the context of robot swarms, includes the development of the distributed deterministic spiral algorithm (DDSA), and advances our knowledge of a critical component of the immune system.

Stone (1975) and Krebs (1978) provide analysis for single searchers that have information about the location of targets and knowledge of the conditional probability of finding a target at a location over time. Efficient search strategies for uninformed searchers have been studied by operations researchers who have designed geometric solutions and by statistical physicists who have modelled search on the stochastic propagation of particles. The latter has been widely used to describe the movement of organisms in ecology and microbiology. Viswanathan et al. (1999) established

that Lévy search with an exponent of 2 is an optimal stochastic strategy for foraging by a single searcher with perfect sensors on an infinite plane, and where targets are uniformly distributed and reappear after collection.

In this dissertation, we study realistic search problems in which agents with imperfect sensors search for targets that are in finite spaces, far from equilibrium, and heterogeneously distributed (Ritchie, 2009). We focus on search by multiple agents that do not have prior information about the distribution of targets.

We analyse the effectiveness of two search strategies, Lévy search and spiral search, for which optimality has been claimed in the single searcher case. By generalising these approaches to swarms and measuring the effectiveness of the resulting search patterns in computer models, we find that the intensity-extent trade-off, formalised as the fractal dimension of the search pattern, can be used to adapt search to common challenges in swarm search.

Increasing the intensity of search mitigates the effects of error in target detection. Increasing extent allows for more effective search when targets are distributed uniformly, but not when targets are more clustered. However, this relationship depends on the boundary conditions of the search area. When searchers are confined to a target rich search area extent can profitably be increased even when targets are clustered.

We find that swarm size has little impact on the intensity-extent trade-off for searchers in the absence of collisions. When we evolve the intensity-extent trade-off, the resulting search patterns have high enough extent that oversampling is not a significant factor. In the presence of collisions, however, more extensive search patterns are clearly advantageous.

We explain the increase in search intensity as a reaction to false negatives due to sensor error. Searchers can utilise the oversampling that results from increased search intensity to integrate the probability of target detection.

In the T cell case, the trade-off between intensity and extent can also be explained by the need to balance detection of rare and common antigen. In our robot study, we find the same trade-off, with increased extent being an effective way to increase performance when targets are widely separated. Increasing intensity is more effective when targets are common.

We also find that feedback between targets and search patterns emerges naturally. Searchers become confined in target rich areas as the discovery of targets causes reorientation. The result is that searchers in target clusters increase the intensity of search and so find more targets in the cluster. Those in less dense regions increase the extent of their search pattern, which tends to result in the discovery of new target rich areas. This emergent property is dynamic and reacts to the changes in target distribution as targets are collected.

Collisions between searchers tend to cause them to stay in the same region, leading to more collisions. In contrast to target encounters where the resulting increase in search intensity tends to improve performance, the opposite is the case with searcher collisions.

Localisation error has been suggested as a reason not to use deterministic spiral search in the single searcher case. When a single searcher misses a target, it cannot be recovered. However, we find that localisation error introduces a stochastic element that allows missed targets to be rediscovered. This combined with the averaging effect of localisation error for swarms of searchers make the DDSA resilient to

localisation error. In contrast, adaptable Lévy search algorithm (ALSA) is relatively sensitive to localisation error, due to its reliance on site fidelity.

Site fidelity returns searchers to the site where a target was last discovered. The result is that searchers that discover clusters tend to exploit that cluster. Site fidelity is a major component in the success of the foraging strategies we have examined when targets are distributed heterogeneously. The DDSA also uses site fidelity as a by-product of its search pattern, but here the expanding spiral forms a front that allows searchers to rediscover target clusters.

8.2 T cell Movement

Search by T cells for dendritic cells (DCs) in lymph nodes (LNs) is a critical part of the adaptive immune response, and as such has important health consequences. (Zeng et al., 2012; Donovan and Lythe, 2016) Understanding T cell search informs out understanding of disease processes that target T cell movement.

When studying swarms of T cells, we test whether previously proposed search models fit our observations and the effect of those models on search effectiveness in lymph nodes. We find that T cell movement is well characterised by a log-normal distribution of step lengths and speeds. T cells also exhibit correlation in turning angles. We model T cell-DC interactions with the observed probability distribution of step sizes and correlation between turning angles. Using this model, we find that T cells balance the intensity-extent trade-off. Using our observations of swarm search and published observations of T cells revisiting DCs we hypothesise that T cells choose a search that allows for antigen signal integration. T cells must be able to discover rare antigen widely separated in space and rapidly find nearby common

antigen. We find that the search pattern T cells employ is as good as Brownian motion at finding nearby antigen but more rapidly finds distant targets.

Using our agent based T cell model we provide evidence that there are regions of the lymph node, hotspots, that are preferentially visited by T cells, and that the search patterns of T cells that visit hotspots differs from those that do not. The nature of these regions is unknown, but fibroblastic reticular cell (FRC) and high endothelial venule (HEV) have both been proposed as possible mechanisms (Mirsky et al., 2011a; Munoz et al., 2014a).

We also use mutual information to measure non-uniform random interactions between naïve T cells and DC targets. Our findings support the hypothesis that T cells and DCs are in contact for longer than purely random models of movement can explain. This suggests that chemical signalling or morphological features such as the FRC network or HEV have a roll in bringing naïve T cells into contact with DCs. This interaction appears to be weak, however. The application of our method to other systems is required to determine the relative strength of these interactions. Measurement of T cell-DC interactions in the presence of antigen, where chemical recruitment is known to occur, would provide an example of a system with strong interactions. Measurement of the strength of interaction between T cells with various signalling receptors knocked-out and DCs would allow us to test which signalling molecules, if any, are involved.

8.3 Connections

In nature, stochastic search processes reminiscent of ALSA are prevalent, while deterministic movement patterns like the DDSA virtually unknown. Possibly for-

aging organisms are not, in fact, optimising target encounters. This objection is the primary reason that Optimal Foraging Theory became controversial (Pyke, 1984). For example, the desert harvester ants we study may not be optimising the collection rate of seeds, rather, it may be more important to establish territorial boundaries to prevent other ants from collecting seeds in their territory. Search patterns may have evolved to reduce swarm loss to predation. In the immune system, it may be that antigen transport to the LNs is optimised placing less pressure on the optimisation of the search processes within LNs. Our work reinforces that stochastic search processes are more tunable and can balance competing requirements better than search processes like the DDSA.

Though the DDSA is effective in our experiments, we do not test complex landscapes with obstacles. Complex search environments may interfere with deterministic algorithms and prevent their execution. It may be telling that the only naturally occurring spiral search pattern observed in a biological system is performed by ants that live on salt flats devoid of obstacles (Müller and Wehner, 1994).

As we have shown, in large arenas with many searchers ALSA can be as effective as DDSA but only with the proper selection of an appropriate the Hausdorff fractal dimension ($\mathcal{H}$). Search time coordination may be more difficult for many species than optimisation of a stochastic strategy over evolutionary time.

Finally, the DDSA may not be a good strategy for continual target collection in the case where targets are replenished over time. In both foraging and the immune system search targets can appear during the search process. Indeed, the search process does not have a well-defined beginning and end. The DDSA as currently designed would have difficulty finding targets that appear near the collection point after the DDSA search frontier has moved far from the starting point.

Despite these caveats, our experiments suggest that the DDSA will prove to be an effective resource collection strategy for engineered swarms. Foraging by artificial swarms often operate under fixed time constraints, do not have conflicting priorities (such as avoiding predation), can communicate with one another and coordinate their behaviours at search time.

We describe search patterns that enable robot swarms to effectively search for targets given realistic search parameters. This work provides a bridge between the existing body of search theory and real-world robotic search problems. Also, we have shown that T cell search in lymph nodes cannot be explained by simple stochastic or guided search models. Using simulations of these systems as a 'third way of science' has allowed us to explore a wide variety of relevant search conditions and their impact on the search strategies we have designed.

Our findings suggest several future directions, such as, using the tools we have developed for analysing cell search patterns in other areas of the immune system, and application of the DDSA and ALSA to *in-situ* resource utilisation tasks.

When targets are sparsely distributed over large areas ALSA is at least as effective as established information based search strategies such as the central place foraging algorithm (CPFA). ALSA is a remarkably simple search strategy that requires no co-ordination between searchers and no information storage. Viswanathan et al. (1999) and others have shown that Lévy search is efficient for individual searchers and that the optimal fractal dimension can depend on the sparsity of targets, we have shown that the intensity-extent trade-off applies to many other critical aspects of swarm search such as sensor error, swarm size, and the configuration of targets. In the future, as swarms of agents are applied to more search problems, this understanding of the intensity-extent trade-offs will allow more effective search strategies.

Appendices

Appendix A

T cell Analysis

A.1 Supporting Information

A.1.1 Publication Notes

Citation: Fricke G.M., Letendre K.A., Moses M.E., Cannon J.L., (2016) Persistence and Adaptation in Immunity: T Cells Balance the Extent and Thoroughness of Search. PLoS Comput Biol 12(3): e1004818. doi:10.1371/journal.pcbi.1004818

Editor: Martin Meier-Schellersheim, National Institutes of Health, United States

Received: October 6, 2015

Accepted: February 17, 2016

Published: March 18, 2016

Copyright: ©2016 Fricke et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted

use, distribution, and reproduction in any medium, provided the original author and source are credited. Documentation is included in Appendix C on page 207.

Formatting: The original published text has been preserved as much as possible while still adhering to the formatting requirements of this dissertation. Any typographical errors have been left in place and marked with [*sic*].

A.1.2 Supporting Figures

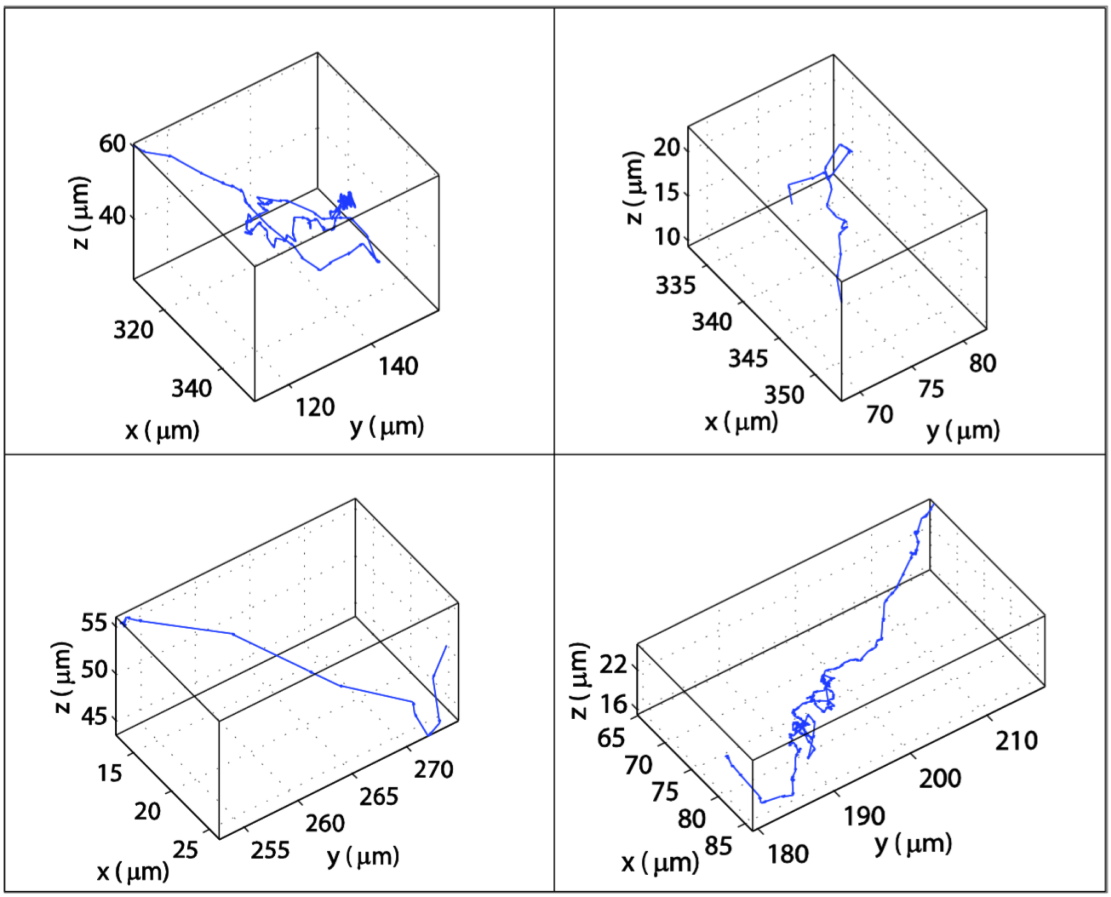


Figure A.1: **Example Tracks.**

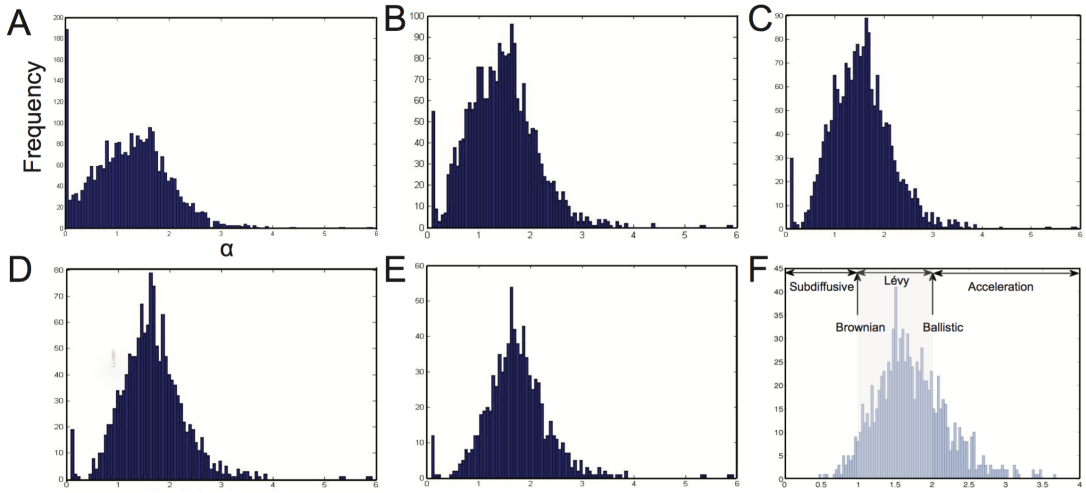
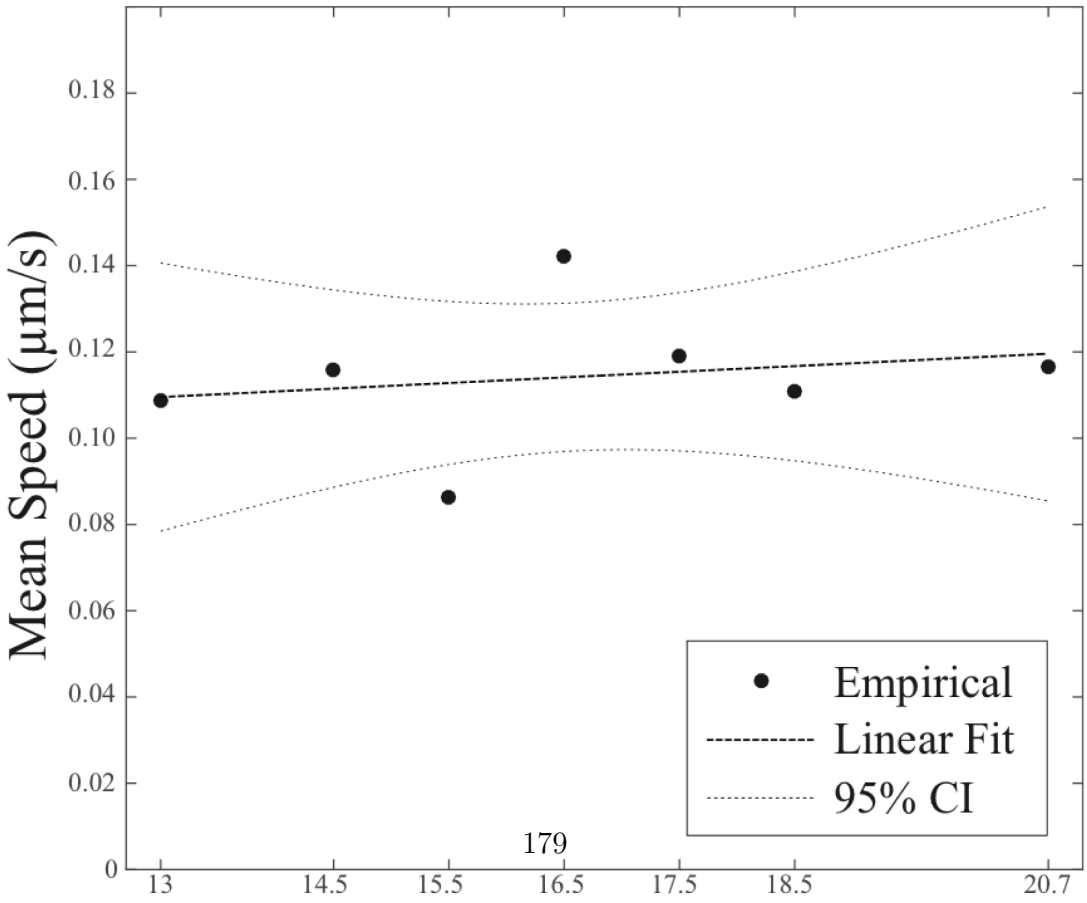


Figure A.2: **MSD Exponent Histograms for Various r^2 Filter Values.** As the linear regression slopes are filtered by the r^2 statistic, the histogram narrows but maintains its mean value. (A) $r^2 > 0$, 3.5% of tracks filtered, (B) $r^2 > 0.25$, 21% filtered, (C) $r^2 > 0.5$, 33%, (D) $r^2 > 0.75$, 50%, and $r^2 > 0.9$, 69% of tracks filtered out. (E) $r^2 > 0.8$ with regions of interest marked.



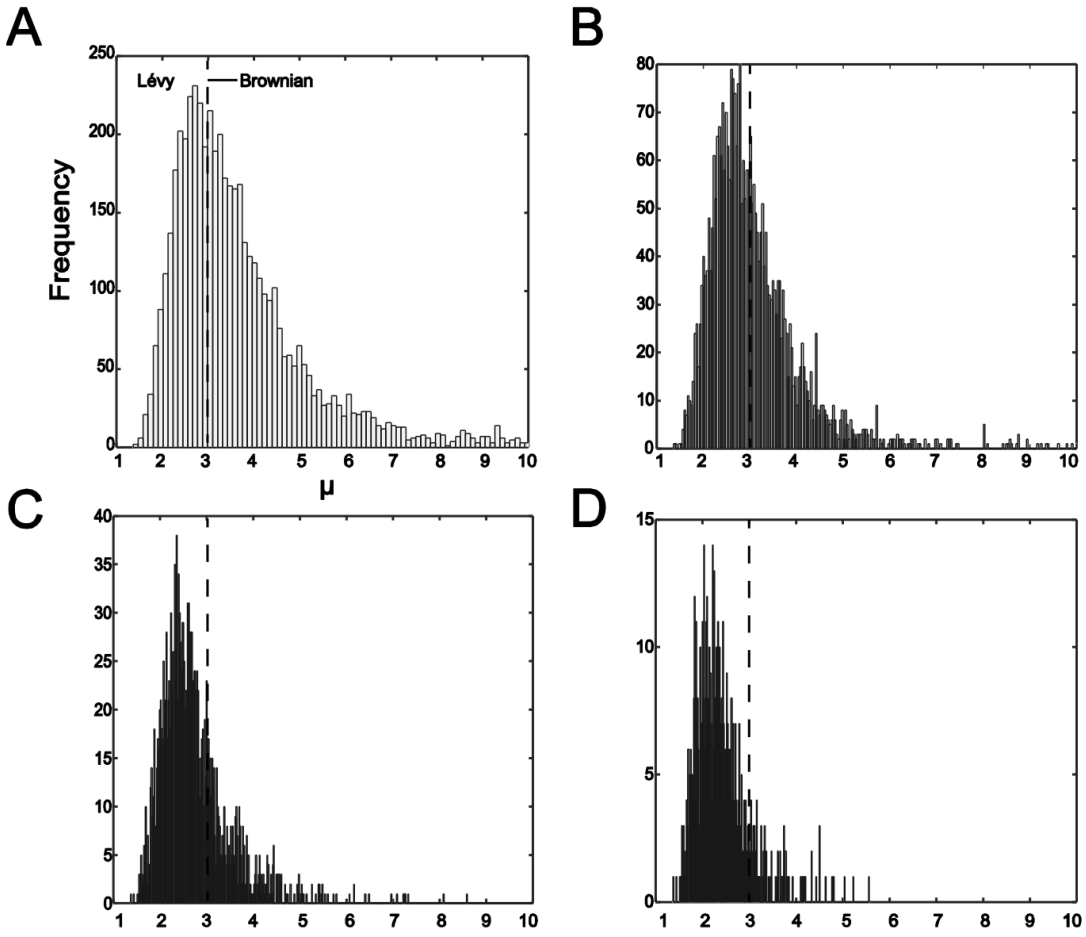


Figure A.3: **Histogram of power law exponents fit to the CCDF of step length for tracks with varying percentages of their steps in the power law tail:** (A) all tracks, (B) tracks with at least 50%, (C) 70%, and (D) 90% of steps in the power law tail. An increasing fraction of steps in the tail results in values being more likely to be between 1 and 3 but as a total fraction of all tracks those well fit by a power law falls rapidly, for (A) 35%, (C) 31%, (D) 24%, and (E) 7% of total tracks are represented. (E) Fraction of Tracks with Lévy characteristics. Power law exponents, μ, for step length and α, for displacement. Tracks are grouped by fit quality (GoF). Retained percentage refers to the amount of data discarded in order to obtain a power law fit (see methods for μ fitting). Displacement α, values are filtered by r^2.

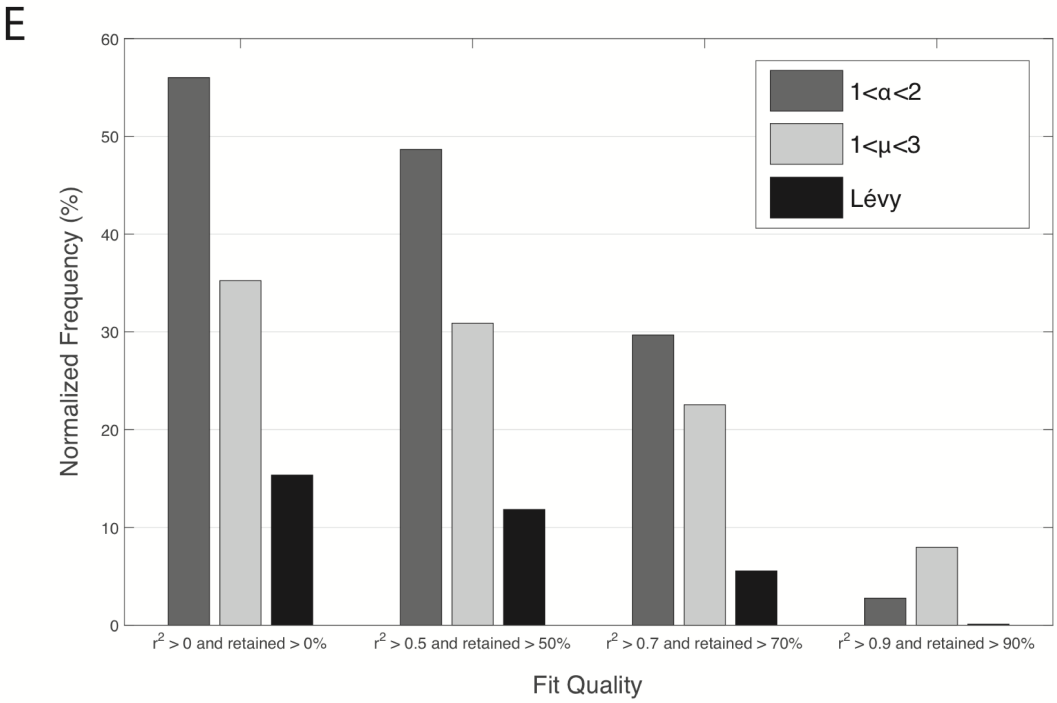


Figure A.4: **Prevalence of Tracks with Lévy Characteristics.**

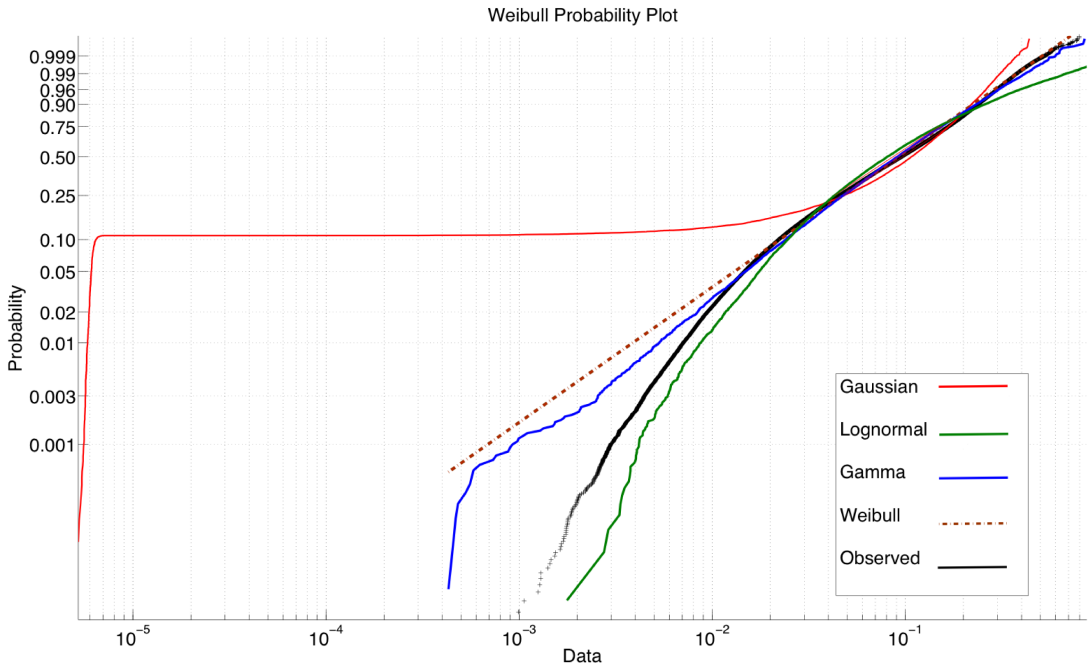


Figure A.5: **Weibull probability plot.** The gamma probability distribution has comparable negative log-likelihood scores to the lognormal distribution (speeds shown here). The lognormal model overestimates the probability of high speeds at the tail of the distribution while the gamma distribution over estimates the probability of very low speeds.

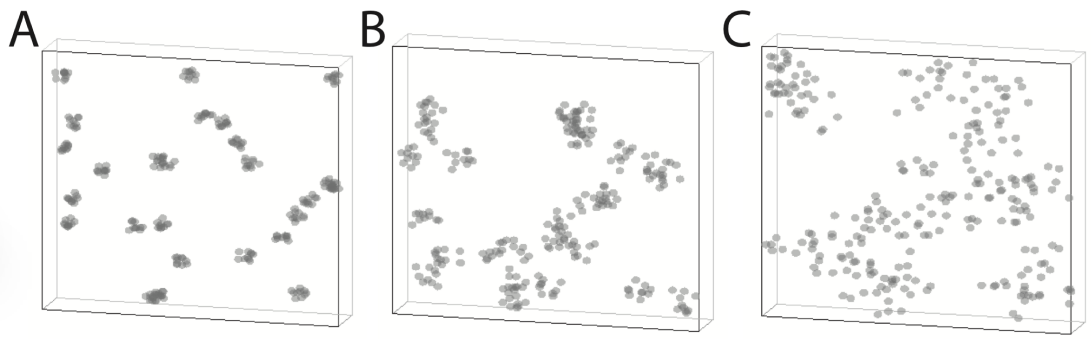


Figure A.6: **Sample DC target cluster distributions in simulation.** Panel A: 10 µm radius clusters with Hopkins index = 0.2. Panel B: 20 µm [*sic*: 20 µm] radius clusters with Hopkins index = 0.32. Panel C: 40 µm radius clusters with Hopkins index = 0.44.

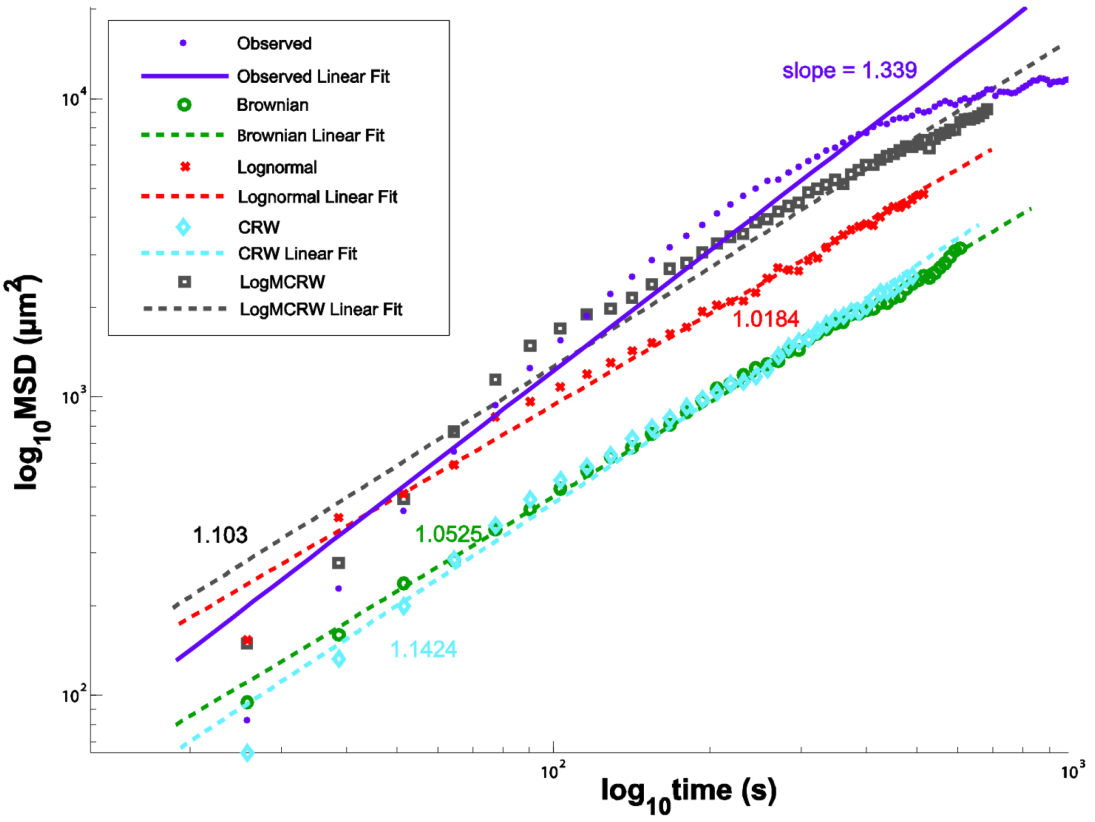


Figure A.7: **Mean squared displacement for simulated search models.** Numbers in color indicate the slope of the mean-squared linear fit to the log-log transformed displacement curve. As expected, Brownian motion has a slope close to one, as does the lognormal step distribution model. All other models produce superdiffusive motion.

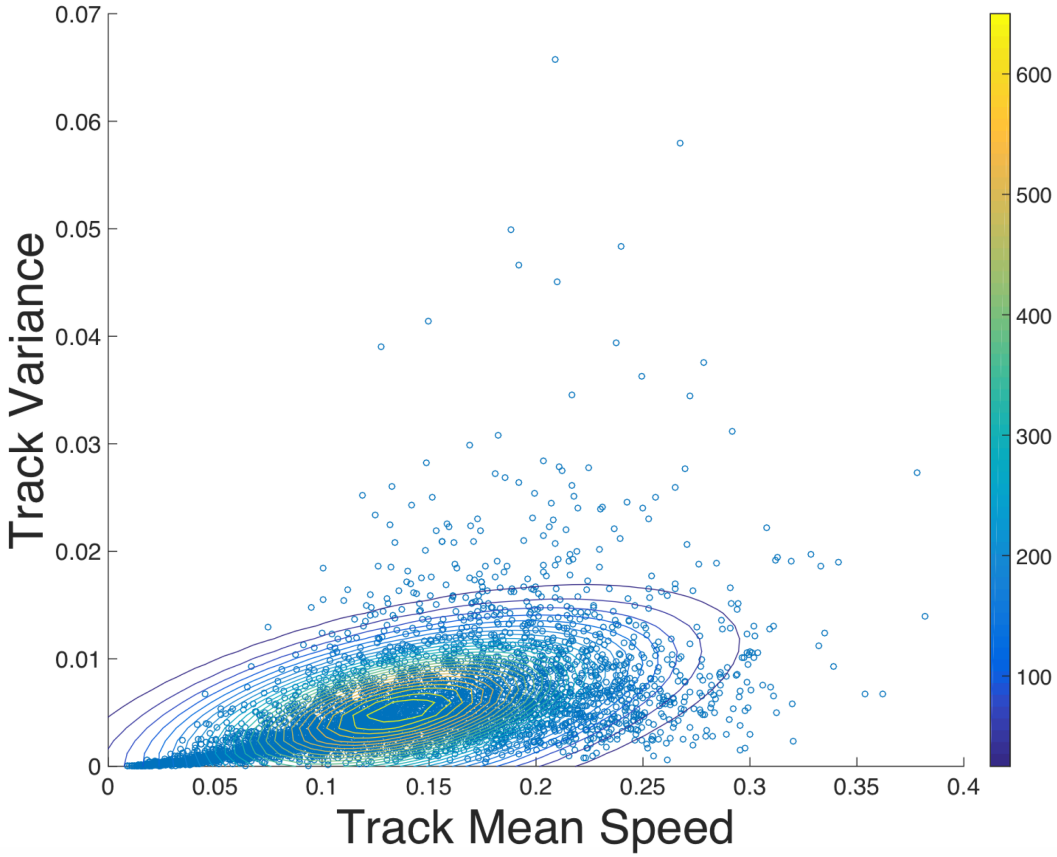


Figure A.8: **We found no evidence of distinct subpopulations defined by variance and mean speed.** An expectation maximization Gaussian mixture model finds that clustering tracks according to track speed and track variance results in a single grouping. The color bar and contour map indicate the height of the best-fit Gaussian model. Increasing the number of Gaussians to fit incrementally up to 16 does not reveal any natural clusters. This figure supports the skew plot Fig 3.4 on page 35C. Example field (1 of 41).

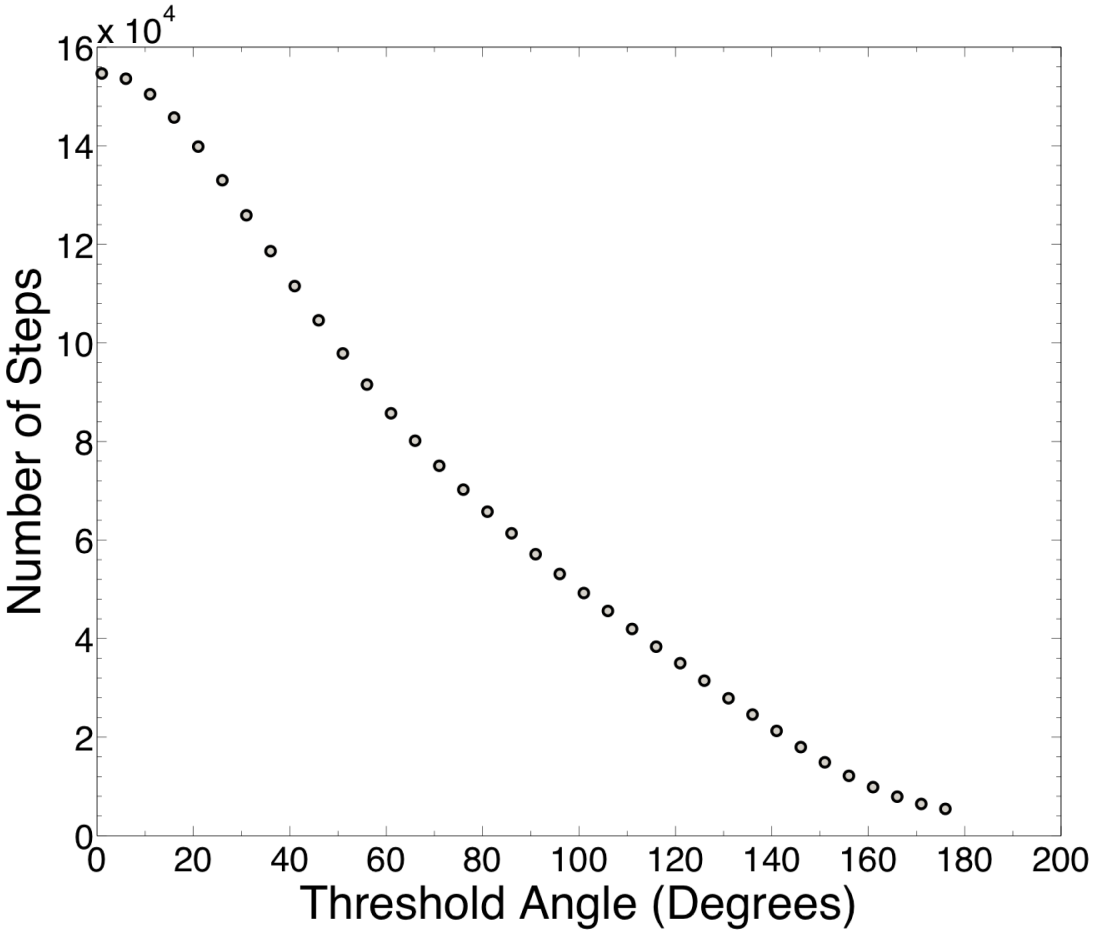


Figure A.9: **The dependency between the angle used to calculate steps from T cell positions and the number of steps resulting.** For example at threshold of 180° all steps in each track are combined and the resulting number of steps in the population is small. The influence of the angle threshold on the number of combined positions is smooth. No natural choice of thresh- old angle is apparent.

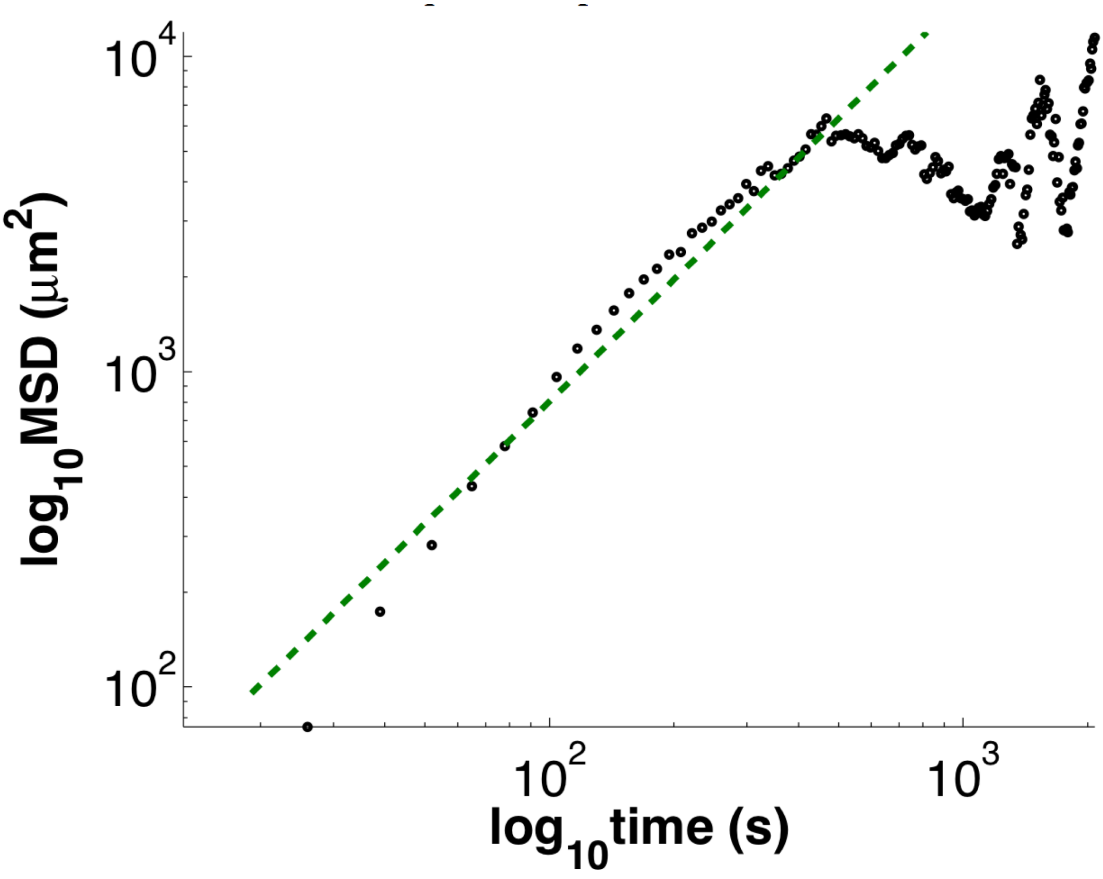


Figure A.10: **As the number of data points in tracks lasting more than 10 minutes drops, MSD becomes dominated by noise.** As a result we perform linear regression only on the first 10 minutes [*sic*: 10 min] of each track (green line). (1 of 7 datasets)

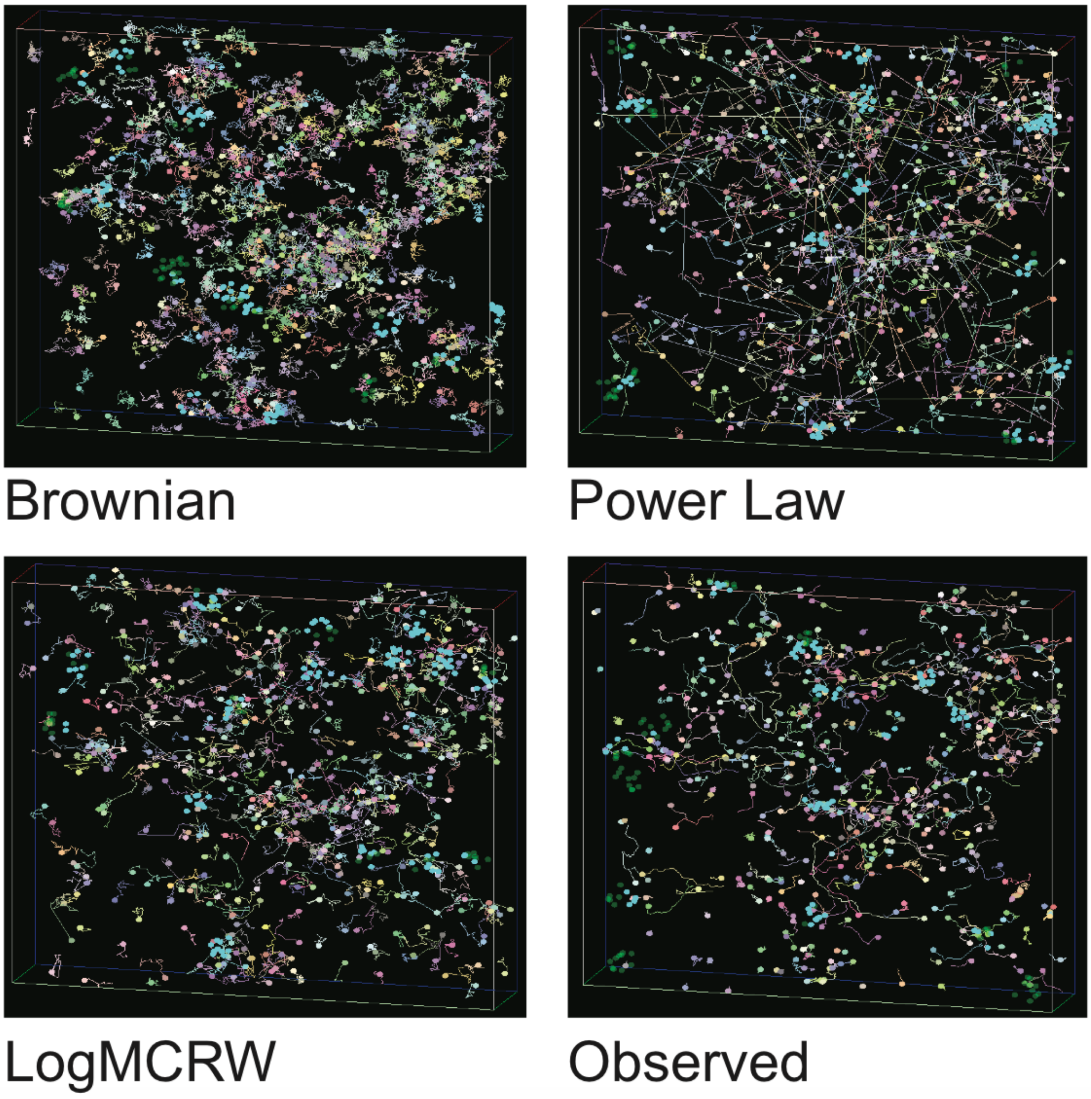


Figure A.11: **Visualization of search tracks.** Dark green targets are undiscovered. Targets become cyan if they are within the search volume of a T cell track (detected). In this example targets are grouped into clusters of 10 with radius 10 µm. Each T cell track is assigned a random color to help distinguish them from one another. Example field (1 of 41).

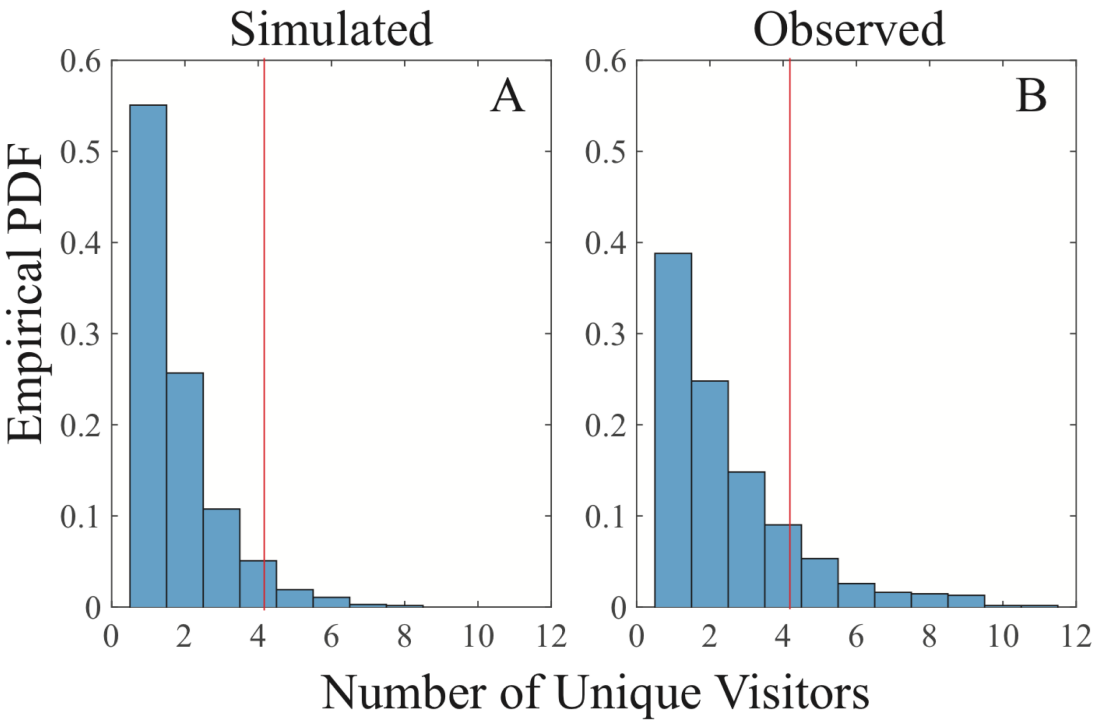


Figure A.12: **Distribution of hotspot visitor counts.** Spot counts for (A) simulated locations over 10 repetitions, and (B) observed locations. Example plot of observed field and the corresponding simulation (1 of 41). The red lines correspond to the hotspot threshold for this field (μ+2σ of the simulated location visitor counts). For this field the threshold is 4.047. Of the 498 locations in the simulated field 17 (3.41%) are hotspots (mean of 10 simulations). The observed field had 621 locations, of which 78 (12.5%) are hotspots, an increase of 258% over simulation.

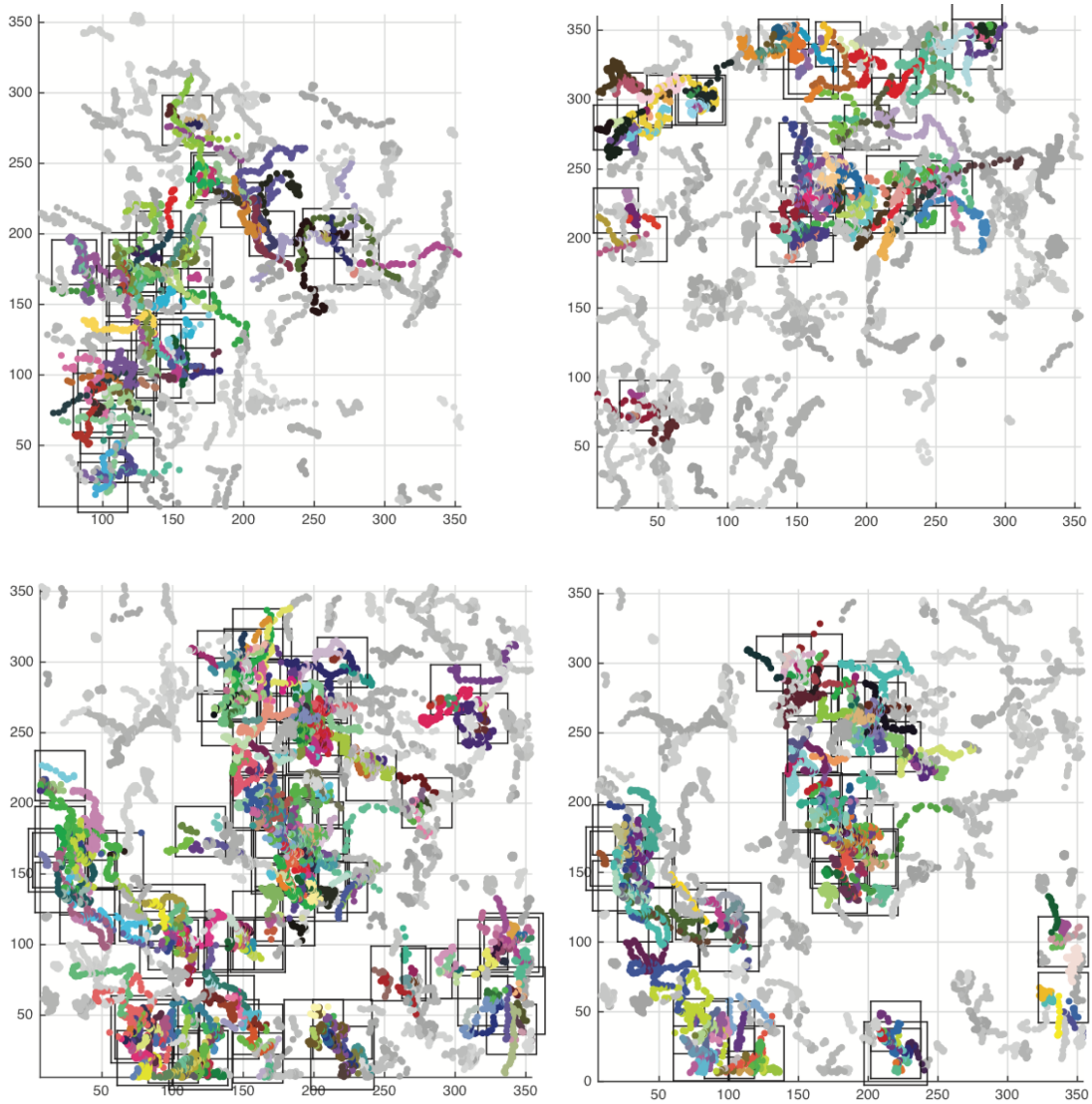


Figure A.13: **Visualization of hotspots and hot tracks in 4 of 41 observed fields.** Hotspots are indicated by black rectangles where the area is proportional to the number of unique visitors. Hot tracks are displayed in color with each color corresponding to a track. Tracks that do not visit a hot spot are shown in grey with the shades corresponding to individual tracks. Plots are a projection of a 3D space into the xy-plane. Overlapping hotspots indicate distinct z-coordinates.

A.1.3 Supporting Tables

Distribution	**AIC E**	**AICc** ($\times 10^5$)	**nlogl** ($\times \mathbf{10^5}$)	**KS**	**AD** ($\times \mathbf{10^3}$)	$\boldsymbol{\chi^2}$ ($\times \mathbf{10^3}$)	**BIC**
Lognormal	1	2.65 (2)	2.65 (2)	0.06 (4)	1.50 (2)	8.06 (7)	5.29 (2)
Gamma	0	5.34 (3)	2.67 (3)	0.04 (2)	0.57 (1)	4.15 (4)	5.34 (3)
Gaussian	0	6.72 (10)	3.36 (10)	0.10 (5)	3.57 (4)	26.56 (9)	6.72 (10)
Power Law	0	9.16 (15)	4.58 (15)	0.33 (8)	38.02 (6)	154.68 (11)	9.16 (15)
Maxwell	0	8.04 (14)	4.02 (13)				8.04 (14)
Exponential	0	7.34 (11)	3.67 (11)	0.09 (7)	2.38 (3)	13.13 (8)	7.34 (11)
Gen. Pareto	0	5.48 (5)	2.74 (5)	0.04 (2)	19.89 (5)		5.48 (5)
Fatigue		4.98 (1)	2.49 (1)	0.05 (3)	597.0 (9)	2.95 (2)	4.98 (1)
Nakagami		5.41 (4)	2.7 (4)				5.41 (4)
Weibull		5.51 (6)	2.76 (6)	0.05 (3)	1117.1 (11)	6.57 (6)	5.51 (6)
Loglogistic		5.56 (7)	2.78 (7)	0.04 (2)	442.86 (8)	4.01 (3)	5.56 (7)
T-location scale		5.71 (8)	2.86 (8)				5.71 (8)
Extreme Value		7.80 (13)	3.9 (13)	0.03 (1)	278.92 (7)	2.69 (1)	7.80 (13)
Inv. Gaussian		6.01 (9)	3.04 (10)	0.03 (1)	696.67 (10)	4.96 (5)	6.01 (9)
Logistic		7.73 (12)	3.87 (12)				7.73 (12)
Rayleigh		10.06 (16)	5.29 (16)	0.17 (6)	14565.0 (12)	56.65 (10)	10.06 (16)
Rician		21.40 (17)	10.71 (17)				21.40 (17)

Table A.1: **Extended Step Fit Statistics.** Table shows the Akaike information criterion evidence ratio (AIC E), applied to first 7 rows only; the corrected Akaike information criterion (AICc); negative log-likelihood (nlogl), Kolmogorov-Smirnov (KS), Anderson-Darling (AD), chi- squared (χ^2), and Bayesian information criterion (BIC). Score ranking is in parentheses. Differences in BIC and AICc scores are less than 1:103 of the AICc score.

Distribution	AIC E	AICc ($\times 10^5$)	nlogl ($\times 10^5$)	KS	AD ($\times 10^3$)	χ^2 ($\times 10^3$)	BIC
Lognormal	0	-3.68 (5)	-1.84 (4)	0.04 (5)	784.54 (3)	8.06 (4)	-3.68 (5)
Gamma	1	-3.87 (1)	-1.93 (1)	0.03 (3)	579.35 (2)	4.15 (2)	-3.87 (1)
Gaussian	0	-3.23 (11)	-1.61 (9)	0.09 (9)	3578.20 (9)	26.56 (9)	-3.23 (11)
Power Law	0	0.245 (14)	0.122 (12)	0.33 (12)	28021 (13)	154.68 (12)	0.245 (14)
Maxwell	0	-2.24 (13)	-1.12 (11)				-2.24 (13)
Exponential	0	-3.68 (5)	-1.84 (4)	0.07 (7)	3122.80 (8)	15.79 (6)	-3.68 (5)
Gen. Pareto	0	-3.78 (4)	-1.89 (3)	0.01 (1)	6645.50 (10)		-3.78 (4)
Fatigue		-3.67 (6)	-1.83 (5)	0.09 (9)	2940.10 (7)	16.57 (7)	-3.67 (6)
Nakagami		-3.84 (3)	-1.92 (2)				-3.84 (3)
Weibull		-3.86 (2)	-1.93 (1)	0.03 (2)	395.88 (1)	2.94 (1)	-3.86 (2)
Loglogistic		-3.67 (6)	-1.83 (5)	-1.83 (5)	1817.1 (6)	20.05 (8)	-3.67 (6)
T-location scale		-3.27 (9)	-1.63 (8)				-3.27 (9)
Extreme Value		-2.37 (12)	-1.18 (10)	-1.18 (10)	863.43 (4)	7.40 (3)	-2.37 (12)
Inv. Gaussian		-3.53 (7)	-1.76 (6)	0.11 (10)	8931.40 (11)	46.51 (10)	-3.53 (7)
Logistic		-3.26 (10)	-1.63 (8)	0.06 (6)	1468.30 (5)	11.74 (5)	-3.26 (10)
Rayleigh		-3.34 (8)	-1.67 (7)	0.14 (11)	10948.0 (12)	61.20 (11)	-3.34 (8)
Rician		-3.34 (8)	-1.67 (7)				-3.34 (8)

Table A.2: **Extended Speed Fit Statistics.** Table shows the Akaike information criterion evidence ratio (AIC E), applied to first 7 rows only; the corrected Akaike information criterion (AICc); negative log-likelihood (nlogl), Kolmogorov-Smirnov (KS), Anderson-Darling (AD), chi-squared (χ^2), and Bayesian information criterion (BIC). Score ranking is in parentheses. Differences in BIC and AICc scores are less than 1:103 of the AICc score.

Distribution	-log Likelihood ($\times 10^5$)	MLE Parameters
Lognormal	4.89	$\mu = 0.52$, $\sigma = 1.00$
Gaussian	6.53	$\mu = 2.75$, $\sigma = 3.24$
Maxwell	8.40	$a = 6.00$
Power Law	8.05	$\alpha = 1.19$

Table A.3: **Maximum likelihood estimated parameters and associated likelihood scores for steps calculated using a 30° threshold.** The lognormal probability distribution is still the best fit when steps are calculated using a 30° rather than 15° threshold. Compare to Table 1 in the main text.

A.2 Extended Materials and Methods

A.2.1 Analysis of T cell Tracks

Cell motility was analyzed with Imaris 6.0 (Bitplane AG, Zurich, Switzerland). Tracks with fewer than 3 time steps were removed from consideration. Tracks with

total length or displacement from the start location less than 17 µm over the course of the observation were assumed to be non-motile and not included in analysis. The point sequences generated by Imaris were used to create position vectors joining adjacent cell locations (sample tracks Fig. A.1 on page 178). The Euclidean norm for each vector was calculated and divided by the time resolution to produce speeds. A nested ANOVA analysis (Letendre et al., 2015) showed no differences between experiments replicated 41 times, using 17 mice, resulting in 159,746 positions over 5,077 T cell tracks. The combined track length for all cells was 34.8 cm and total observation time was 17 hours and 12 minutes. The mean track length is 34 positions with a median length of 21 and a max of 200. The maximum T cell velocity over all observations was 0.9 µm/s with a mean of 0.11 µm/s.

Microscopy fields with time resolution differences of less than 1 second were combined into groups so that mean squared displacement and vector autocorrelation could be calculated separately for groups with similar time resolution. Seven datasets were generated from the 41 observation fields. The autocorrelation plot, shown in Fig. 3.2 on page 33E as a representative example, contains data from fields with time resolutions between 15 s and 16 s, and consists of 23,169 vectors from 537 tracks.

Speed measurements can be sensitive to the frame rate of observation. To determine whether the observed speed was influenced by the frame rate we fit a linear response model to frame delay and mean speed. The speed is not linearly correlated to the frame delay; as can be seen in Figure A.14 on page 179.

A.2.2 Distribution Fitting

Following Fisher (1925) we use maximum likelihood estimation (MLE) to parameterize candidate PDFs. De Jager et al. provide details of the MLE method in their analysis of the motility patterns of mussels (de Jager et al., 2014). We fit probability model parameters using cumulative distribution functions (CDF), rather than by binning data which has been shown to bias conclusions about random walk distributions [8,9]. Software for fitting distributions was written in Matlab (Mat, 2014). We examined a further 57 commonly used PDFs which were narrowed down to 17 for which we calculated negative log-likelihood scores (Tables A.1 on page 190 and A.2 on page 191) as candidate models.

A variety of probability distributions have relatively good statistical fit scores. Those that do are heavy-tailed, such as gamma and Weibull. Lognormal, while not the best fit is consistently high in the rankings for all goodness of fit measures (GoF), except the chi-squared test for step sizes. We choose to use the lognormal distribution for our simulations because it is relatively well understood mathematically and intuitively. See Fig. A.5 on page 182 for a comparison of the lognormal CDF to the gamma and Weibull CDFs. The lognormal distribution shows up repeatedly in biological contexts (Beltman et al., 2007; Furusawa et al., 2005)[*sic:* Beltman et al. (2007) is not an appropriate reference] so we have made it our representative for the heavy-tailed family of distributions.

Five PDF models for step length and speed were selected for analysis based on a combination of their negative log-likelihood scores, their importance in other biological processes, and their previous use in modeling T cell movement. Those models are lognormal, Maxwell, Gaussian, exponential, and power law PDFs. Our

selection of the relative goodness of fit (GoF) of each candidate PDF to empirical data was evaluated using likelihood functions, Anderson-Darling (AD), Bayesian information criterion (BIC), corrected Akaiki Information Criterion (AICc), and the Kolmogorov-Smirnov (KS) test. We found all these measures to be in rank agreement and therefore present only the likelihood measure. We observe that for all GoF measures lognormal is ranked as a better fit to our observed step and speeds than the alternatives. In addition to track speeds, which we measure directly, we determined the distribution of step lengths, which is an essential element in the definition of Lévy walks. Since we image T cell motility frame by frame, we cannot track each cell continuously to ascertain step length. Instead, we define a step as a vector of T cell motion that does not deviate beyond 15° from the original direction (see Fig. A.9 on page 185 for analysis of threshold dependency). We did not find any discrete angle thresholds and when we used 30° as a cutoff angle to determine step length, we saw the same ranking in MLE fits (Table A.3 on page 191).

It is difficult to fit power laws in their simplest formulation, $f(x) = Cx^y$, with its infinite right variance, to necessarily finite biological systems. Additionally, power law behavior is often found in combination with other processes. Therefore, it is common to fit only a portion of the data to a power law, using formulations such as the generalized Pareto distribution, exponential cutoff, or generalized Lévy walk. In previous work (Fricke et al., 2013) we modelled T cell search as a Lévy flight and found a good fit to T cell motion using the generalized Pareto distribution. Following Clauset et al. (2009), we fit power laws using MLE and with the power law PDF: $\mathrm{P}(x) = \frac{\mu-1}{x_{\min}} \left(\frac{x}{x_{\min}}\right)^{\mu}$, where $x_{\min}$ is the smallest observed value, $\mathrm{P}(x)$ is the probability of x occurring, and μ is the estimated parameter. We used the $x_{\min}$ value with the best KS score of all possible choices as an estimator of the beginning

of a power law tail. The percentage of positions in a track in the power law tail gives us a measure of the quality of the power law fit. Using this measure we show that a power law fit to the population of observed steps excludes 94% of the data (Fig. 3.1 on page 32F and H). This measure is also used to filter tracks in [Fig.] A.3 on page 180.

A.2.3 Autocorrelation and Cross-Correlations

Velocity autocorrelations were calculated following Qian et al. (1991) and Tarantino et al. (2014). The autocorrelation function, C_{auto} , is the ensemble mean for the $n-1$ possible delay times given the n vectors defining a T cell track.

$$C_{\text{auto}}(\Delta t) = \langle v(t_i) \cdot v(t_j) \rangle, \forall i, j : \Delta t = t_i - t_j \text{ where } t_i, t_j \text{ are times.} \tag{A.1}$$

The result is a measure of how much T cell direction depends on previous directions as a function of time delay. Letting $v(p_k(t))$ be the unit velocity vector at time t belonging to the kth path, we defined the cross-correlation function, C_{cross}, to be:

$$C_{\text{cross}}(p) = v(p_k(t)) \cdot v(p_m(t)), \forall k, m \text{ where } p_k, p_m \text{ are T cell paths.} \tag{A.2}$$

This measures the step angle dependence between T cell paths at the same moment in time, that is, a measure of drift due to global effects on the observation field.

A.2.4 Mean Squared Displacement

Mean squared displacement (MSD) coefficients, commonly called the α exponent were calculated using least-squares polynomial fit by numerically solving the associated Vandermonde matrix (Von Mises and Geiringer, 1964) and fit quality assessed with the r^2 measure. Parametric and linear fits were also made to mean displacement. In Fig. 3.1A we present only the first 10 minutes of observation at which point the curve reaches its first stationary inflection which is indicative of unconstrained motion and therefore appropriate for determining α. In addition, in this study few tracks persist beyond 10 minutes and so the MSD signal also becomes dominated by noise (Fig 3.1A top and Fig. A.10 on page 186).

A.2.5 Heterogeneity

We used mixed Gaussian clustering (McLachlan and Peel, 2004) to investigate whether there is heterogeneity the distribution of track speeds among T cell tracks with different track mean speeds. The mean speed and variance in tracks is shown in both Fig. 3.2 on page 33 and in Fig. A.8 on page 184. We further tested for heterogeneity by comparing track speed skew (Fig. 3.4 on page 35) and AIC evidence ratios as a function of mean speed (data not shown).

A.2.6 Search Efficiency Simulation

We built a simulation to test how different movement patterns affect the efficiency with which T cells encounter DC targets, implemented as a continuous (floating-point) 3D model written in C++. Boost libraries (59) were used to generate variates

drawn from model PDFs. Because the clustering and density of targets can influence which movement types are most efficient, we replicated the estimated density of DCs and varied the degree of clustering in our simulations.

Beltman et al. report a DC density in lymph node T cell zones of 2-5% [11,18]. We use this to calculate a target DC density of 3.17×10^{-5} targets/μm^3. Our observed fields have an average volume of $6.3 \times 10^6 \mu m^3$. We scale the number of targets as a function of field volume in order to maintain the same target density between simulation fields. DCs were clustered into groups of 10 and were uniformly distributed within spheres defining a cluster. By varying the sphere radius, we controlled the degree of clustering from uniform to highly clustered. A 3D version of the Hopkins statistic (Hopkins and Skellam, 1954) was used to measure the resulting non-uniformity of target placement (Tables 3.1 on page 37 and 3.2 on page 38). In the Hopkins statistic scores range from 0 to 0.5 where 0 is highly clustered and 0.5 indicates no clustering.

T cell tracks were observed and recorded as 3D coordinate sequences within a bounding box defined by the visible section of the *ex vivo* lymph node. Idealized models (Brownian, CRW, Power Law, etc.) of search were parameterized by the speeds and turning angles estimated from observation (see Distribution Fitting). Searchers in the idealized model start at the same initial positions as the observed T cells, and exist in a volume equal to the observed field volume. Candidate search patterns were generated for each of the 41 observation fields. For sample visualization of each idealized model search as well as the observed see Fig. A.11 on page 187. Similar to (James et al., 2010) our measure of efficiency, $E(k, \text{Obs})$, for a combination of n observed tracks (Obs), search strategy, and target distribution is the number of targets discovered, $F(k)$,by each searcher, k, divided by the time expended, $T(k)$, by

the kth track. The distance covered by track k is D(k), and the total distance covered by observed tracks is D(Obs). L is the property that the total distance expended by searchers does not exceed the distance expended by observed T cells.

$$\mathrm{E}(k,\mathrm{Obs}) = \frac{\sum_{k=1}^{\mathrm{L(k)}} \mathrm{F}(k)}{\sum_{k=1}^{\mathrm{L(k)}} \mathrm{T}(k)}, \text{ where } \mathrm{L}(k) = \sum_{i=1}^{k} \mathrm{D}(i) < \mathrm{D(Obs)}. \tag{A.3}$$

Efficiency measure E is the number of targets found divided by the sum of the time used by searchers. Since we modelled walks rather than flights (i.e. speeds are finite) the sum of D(k) for all simulated tracks k was limited to the total distance travelled by observed T cells. Therefore the average velocity of the population of searchers is kept within the observed range. In the limit where the field is saturated with targets, the efficiency of unique contacts would be the swept volume of each track. Based on an assumed radii of 5 µm for DCs and T cell, targets were marked as discovered if a searcher track passed within 10 µm of a target point. We define two versions of F(k), one that increments its output value only when a target was not previously detected by searcher, k, and another that increments for all targets found. These two versions of F(k) allow us to record unique contacts and total contacts (Fig. 3.4 on page 35).

The simulation measures the target encounter rate and determines, using the Mann-Whiney test, whether the candidate search models search efficiency is significantly different from that observed in T cells. We use the Mann-Whitney test because the observed and simulated distribution of efficiencies is non-Gaussian. Bonferoni correction (Bland and Altman, 1995) was used to adjust for the minimum p-value that can be called significant. Simulations were replicated 100 times per field, pro-

ducing 4,200 efficiency data points for each search model. The entire process was repeated 10 times in order to generate confidence intervals for the simulation; in all this results in 41,000 efficiency samples.

A.2.7 Identifying Hotspots and Hot Tracks

In order to test whether the environment within LNs influences T cell movement we extend an analysis begun in (Fricke et al., 2015). To determine hotspots, we use the LogMCRW simulation as a null model. We discretize the LN into cubes with 20 µm edges (about twice the diameter of a T cell). We record the number of times a location is visited by unique T cells in simulation (repeated 10 times). We use a 2σ (standard deviation) threshold for determining which locations are visited particularly frequently than expected and call these hotspots (threshold indicated by red line in Fig. A.12 on page 188). This is repeated for each of the 41 individual observational fields. We then determine hotspots for observed data as locations visited more frequently than the threshold set by the null model that corresponds to the individual observation field (graphically shown in Fig. A.13 on page 189). All other locations are called cold spots.

We define hot tracks to be T cell tracks that visit hotspots and cold tracks to be T cell tracks that do not. We also examine the number of visits by hot tracks to cold spots and hotspots.

A.3 Modelling Software Interface

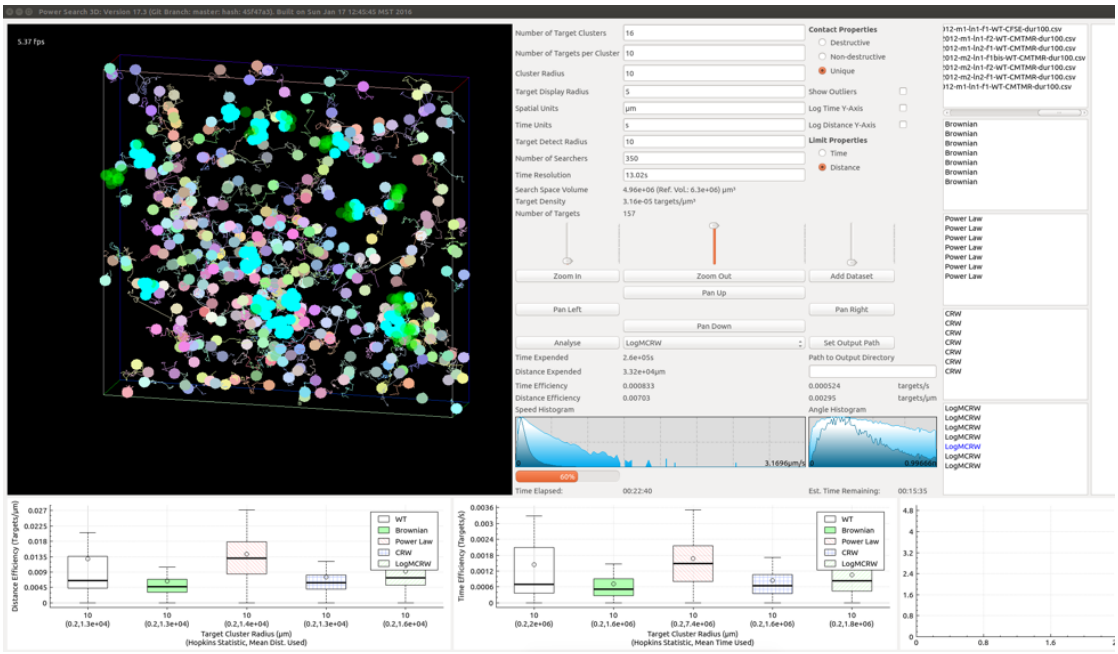


Figure A.15: Track Efficiency Analyser. Screen capture of the PowerSearch3D interface. The user selects files containing observed track data (exported from Imaris for example), selects idealized search patterns to compare with the empirical data, and chooses parameters such as the density of targets and whether to limit searchers by time or distance. In the example above 7 empirical experiments have been selected and 4 idealized search patterns are being generated. The user has elected to limit search by distance covered. The software measures the number of target contacts for each experiment and generates idealized patterns that correspond to the observed search volumes, number of searchers, and the total distance and time spent by the searchers. Statistics, such as the Mann Whitney test, are used to compare the efficiency of the search patterns and displayed to the user at the end of the run. This tool was written in C++ with Qt.

A.4 Mutual Information Code

```
1 % Mutual information
2 % This function takes two 3D images and returns
3 % the mutual information.
```

```
% Usage: h = mutual_info( image_a, image_b, #bits of color )
function [h,joint_prob]=mutualinfo(image_a,image_b,color_depth)
    a = jointinfo( image_a, image_b , color_depth );
    hab = 0;
    % Matlab provides entropy functions
    ha = entropy(image_a);
    hb = entropy(image_b);
    joint_prob = a./ numel(image_a);
    for i = 1:color_depth
        for j = 1:color_depth
            if joint_prob(i,j) == 0
                continue;
            end
            %JOINT ENTROPY
            hab = hab + joint_prob(i,j) * log2(joint_prob(i,j));
        end
    end
    hab = -hab;
    h = ha + hb - hab;
end

function h=joint_info3D( image_a, image_b , color_depth )
    rows=size( image_a, 1 );
    cols=size( image_a, 2 );
    layers = size( image_a ,3 );
    N=color_depth;
    h = zeros( N , N );
    for k = 1:layers
        for i = 1:rows
            for j = 1:cols
```

```
36                  h( image_a(i,j,k)+1,image_b(i,j,k)+1 )...
39                        =h(image_a(i,j,k)+1,image_b(i,j,k)+1)+1;
40              end
41          end
42      end
43  end
```

Appendix B

ALSA ANOVA Tables and Description

B.1 Two-Way Analysis of Variance

We make extensive use of 2-way analysis of variance (ANOVA) which performs linear regression tests on the variation within experiments grouped by factors of interest (such as swarm size) vs variation between those groups. In other words ANOVA measures the correspondence between changes in a factor and the output in comparison to variance not attributable to the factors of interest. We break our analysis into 2-way ANOVAs so that we can make pairwise comparisons of various factors. This pairing also makes plotting the associated multiple comparison plots easier. A full factor analysis and n-ANOVA would be prohibitively expensive in terms of computing resources. The table headings are as follows: Factor, the parameter being varied in order to determine its influence on the output. *SS* is the Sum of Squares, quantifies the variability between groups within the corresponding factor (such as

the groups of swarm sizes for the swarm size factor)). The SS indicates the relative influence of the factor on the output. *MSE* is the mean squared error of the linear regression associated with the corresponding factor. *F* is the Fisher statistic and is the ratio of inter-group variation to intra-group variation, informally is is a measure of how much variation in the output is explainable by the factor of interest as compared to other sources of variation in the experiments. *p-value*, informally this is a score indicating whether the hypothesis that the corresponding factor influences the output is statistically supportable. The *interaction* row indicates the degree to which the two factors under consideration influence the output in combination. The *error* row indicates how much variation in the the data set is not explained by variation in the two factors under consideration.

It should be noted that ANOVA assumes Gaussian distributed samples, or for large sample sizes a distribution of samples with a common skew direction. These assumptions are violated for Tables B.1 on the following page, B.2 on the next page, and B.3 on page 206 which contain factor samples which are not Gaussian or skewed in the same direction. For the remaining ANOVA datasets the skew is consistently to the left. While technically violating the assumptions of the ANOVA, Glass et al. (1972) provide evidence that, in practice, ANOVA is robust to false positives even when the Gaussian assumption is violated. However the p-values based on the F-statistic will often be reduced, overestimating the statistical significance of factor influence. We provide Tukey multiple comparison plots for each ANOVA in order to support the conclusions we draw from ANOVA Tukey (1949).

B.2 ANOVA Tables

Factor	SS	df	MSE	F	p-value
False Negative Rate (s^{-1})	22.84	7	3.263	25.38	$< 10^{-4}$
Target Distribution	11.3005	2	5.6502	43.95	$< 10^{-4}$
Interaction	3.682	14	0.26301	2.05	0.0136
Error	58.62	456	0.1286		
Total	96.44	479			

Table B.1: ANOVA Results. Dependence of optimal μ on false negative target detection rate (per second) and the target configuration. The target configuration and false negative rates are statistically significant factors in the empirically optimal $\mathcal{H}$. 20 experiments for each combination of factors and each tested $\mathcal{H}$ (N = 9600, of which 480 experiments corresponding to the best $\mathcal{H}$ are analysed). 8 robots.

Factor	SS	df	MSE	F	p-value
N Searchers	6.555	4	1.639	11.17	$< 10^{-4}$
Target Distribution	13.64	2	6.822	46.51	$< 10^{-4}$
Interaction	3.105	8	0.3881	2.65	0.0082
Error	41.805	285	0.1467		
Total	65.11	299			

Table B.2: ANOVA Results. Dependence of optimal μ on swarm size and the target configuration. 20 experiments for each combination of factors and each tested $\mathcal{H}$ (N = 6000, of which 300 experiments corresponding to the best $\mathcal{H}$ are analysed). Power law configuration of targets. No localisation or target detection error.

Factor	SS	df	MSE	F	p-value
Localisation Error	110632.2	5	22126.4	114.1	$< 10^{-4}$
False Negatives	327572.6	7	46796.1	241.31	$< 10^{-4}$
Interaction	18215.1	35	520.4	2.68	$< 10^{-4}$
Error	269944	1392	193.9		
Total	726364	1439			

Table B.3: ANOVA of False Negatives and Localisation Error on Targets Collected for Fractal Dimension 0.7. 30 experiments for each combination of factors (N = 10080). Power law configuration of targets. 8 robots.

Factor	SS	df	MSE	F	p-value
Localisation Error	397570.4	5	79514.1	517.58	$< 10^{-4}$
False Negatives	37679.8	7	5382.8	35.04	$< 10^{-4}$
Interaction	20817.5	35	594.8	3.87	$< 10^{-4}$
Error	213848.3	1392	153.6		
Total	669916	1439			

Table B.4: ANOVA of False Negatives and Localisation Error on Targets Collected for Fractal Dimension 1.4. 30 experiments for each combination of factors (N = 10080). Power law configuration of targets. 8 robots.

Factor	SS	df	MSE	F	p-value
Localisation Error	8761	5	1752.2	15.61	$< 10^{-4}$
False Negatives	1579457.9	7	225636.8	2009.74	$< 10^{-4}$
Interaction	20900.3	35	597.2	5.32	$< 10^{-4}$
Error	156281.9	1392	112.3		
Total	1765401.1	1765401.1	1439		

Table B.5: ANOVA of False Negatives and Localisation Error on Targets Collected for DDSA. 30 experiments for each combination of factors (N = 10080). Power law configuration of targets. 8 robots.

Appendix C

Permission to Reproduce Previously Published Content

The University of New Mexico requires that previously published works included in this dissertation be licensed by their respective copyright holders. Copies of those licences must be included in the dissertation. If the material was published under an open access license the guidelines for use must be included in lieu of a license. Accordingly this appendix includes the following documents:

- A license to reproduce material in Chapter 5 for which Cambridge University Press is the copyright holder.
- The guidelines for open access use of material in Chapter 3, previously published by the Public Library of Science and the governing Creative Commons legal code.

- Instructions for reproduction of IEEE publications in a dissertation. The contents of Chapter 6 previously appeared in the proceedings of the International Conference on Intelligent Robots and Systems (IROS), published by the IEEE.

CAMBRIDGE UNIVERSITY PRESS LICENSE TERMS AND CONDITIONS

Oct 03, 2016

This Agreement between George M Fricke ("You") and Cambridge University Press ("Cambridge University Press") consists of your license details and the terms and conditions provided by Cambridge University Press and Copyright Clearance Center.

License Number	3961370302687
License date	Oct 03, 2016
Licensed Content Publisher	Cambridge University Press
Licensed Content Publication	Robotica
Licensed Content Title	Immune-inspired search strategies for robot swarms
Licensed Content Author	G. M. Fricke, J. P. Hecker, J. L. Cannon, M. E. Moses
Licensed Content Date	Jul 11, 2016
Licensed Content Volume Number	34
Licensed Content Issue Number	8
Start page	1791
End page	1810
Type of Use	Dissertation/Thesis
Requestor type	Author
Portion	Full article
Author of this Cambridge University Press article	Yes
Author / editor of the new work	Yes
Order reference number	
Territory for reuse	World
Title of your thesis / dissertation	Search in T cell and Robot Swarms: Balancing Extent and Intensity
Expected completion date	Nov 2016
Estimated size(pages)	130
Requestor Location	George M Fricke 4412 Inspiration Dr SE ALBUQUERQUE, NM 87108 United States Attn: George M Fricke
Publisher Tax ID	GB823847609
Billing Type	Invoice
Billing Address	George M Fricke 4412 Inspiration Dr SE

ALBUQUERQUE, NM 87108
United States
Attn: George M Fricke

Total 0.00 USD

Terms and Conditions

TERMS & CONDITIONS

Cambridge University Press grants the Licensee permission on a non-exclusive non-transferable basis to reproduce, make available or otherwise use the Licensed content 'Content' in the named territory 'Territory' for the purpose listed 'the Use' on Page 1 of this Agreement subject to the following terms and conditions.

1. The License is limited to the permission granted and the Content detailed herein and does not extend to any other permission or content.
2. Cambridge gives no warranty or indemnity in respect of any third-party copyright material included in the Content, for which the Licensee should seek separate permission clearance.
3. The integrity of the Content must be ensured.
4. The License does extend to any edition published specifically for the use of handicapped or reading-impaired individuals.
5. The Licensee shall provide a prominent acknowledgement in the following format: author/s, title of article, name of journal, volume number, issue number, page references, , reproduced with permission.

Other terms and conditions:
v1.0

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

Licenses and Copyright

The following policy applies to all PLOS journals, unless otherwise noted.

What Can Others Do with My Original Article Content?

PLOS applies the Creative Commons Attribution (CC BY) license to articles and other works we publish. If you submit your paper for publication by PLOS, you agree to have the CC BY license applied to your work. Under this Open Access license, you as the author agree that anyone can reuse your article in whole or part for any purpose, for free, even for commercial purposes. Anyone may copy, distribute, or reuse the content as long as the author and original source are properly cited. This facilitates freedom in re-use and also ensures that PLOS content can be mined without barriers for the needs of research.

May I Use Content Owned by Someone Else in My Article?

If you have written permission to do so, yes. If your manuscript contains content such as photos, images, figures, tables, audio files, videos, etc., that you or your co-authors do not own, we will require you to provide us with proof that the owner of that content (a) has given you written permission to use it, and (b) has approved of the CC BY license being applied to their content. We provide a form you can use to ask for and obtain permission from the owner. Download the form (PDF).

If you do not have owner permission, we will ask you to remove that content and/or replace it with other content that you own or have such permission to use.

Don't assume that you can use any content you find on the Internet, or that the content is fair game just because it isn't clear who the owner is or what license applies. It's up to you to ascertain what rights you have—if any—to use that content.

May I Use Article Content I Previously Published in Another Journal?

Many authors assume that if they previously published a paper through another publisher, they own the rights to that content and they can freely use that content in their PLOS paper, but that’s not necessarily the case – it depends on the license that covers the other paper. Some publishers allow free and unrestricted re-use of article content they own, such as under the CC BY license. Other publishers use licenses that allow re-use only if the same license is applied by the person or publisher re-using the content.

If the paper was published under a CC BY license or another license that allows free and unrestricted use, you may use the content in your PLOS paper provided that you give proper attribution, as explained above.

If the content was published under a more restrictive license, you must ascertain what rights you have under that license. At a minimum, review the license to make sure you can use the content. Contact that publisher if you have any questions about the license terms – PLOS staff cannot give you legal advice about your rights to use third-party content. If the license does not permit you to use the content in a paper that will be covered by an unrestricted license, you must obtain written permission from the publisher to use the content in your PLOS paper. Please do not include any content in your PLOS paper which you do not have rights to use, and always give proper attribution.

What Are Acceptable Licenses for Data Repositories?

If any relevant accompanying data is submitted to repositories with stated licensing policies, the policies should not be more restrictive than CC BY.

Removal of Content Used Without Clear Rights

PLOS reserves the right to remove any photos, captures, images, figures, tables, illustrations, audio and video files, and the like, from any paper, whether before or after publication, if we have reason to believe that the content was included in your paper without permission from the owner of the content.

How Does One Give Proper Attribution for Use of Content?

When citing a PLOS research article, use the “Vancouver style”, as outlined in our Submission Guidelines. For example:

Kaltenbach LS et al. (2007) Huntingtin Interacting Proteins Are Genetic Modifiers of Neurodegeneration. *PLOS Genet* 3(5): e82. doi:10.1371/journal.pgen.0030082.

When citing non-article content from a PLOS website (e.g., blog content), provide a link to the content, and cite the title and author(s) of that content.

For examples of proper attribution to other types of content, see websites such as Open.Michigan.

Creative Commons Legal Code

Attribution 4.0 International

Official translations of this license are available in other languages.



Creative Commons Corporation (“Creative Commons”) is not a law firm and does not provide legal services or legal advice. Distribution of Creative Commons public licenses does not create a lawyer-client or other relationship. Creative Commons makes its licenses and related information available on an “as-is” basis. Creative Commons gives no warranties regarding its licenses, any material licensed under their terms and conditions, or any related information. Creative Commons disclaims all liability for damages resulting from their use to the fullest extent possible.

Using Creative Commons Public Licenses

Creative Commons public licenses provide a standard set of terms and conditions that creators and other rights holders may use to share original works of authorship and other material subject to copyright and certain other rights specified in the public license below. The following considerations are for informational purposes only, are not exhaustive, and do not form part of our licenses.

> ***Considerations for licensors:*** *Our public licenses are intended for use by those authorized to give the public permission to use material in ways otherwise restricted by copyright and certain other rights. Our licenses are irrevocable. Licensors should read and understand the terms and conditions of the license they choose before applying it. Licensors should also secure all rights necessary before applying our licenses so that the public can reuse the material as expected. Licensors should clearly mark any material not subject to the license. This includes other CC-licensed material, or material used under an exception or limitation to copyright. More considerations for licensors.*

> ***Considerations for the public:*** *By using one of our public licenses, a licensor grants the public permission to use the licensed material under specified terms and conditions. If the licensor’s permission is not necessary for any reason–for example, because of any applicable exception or limitation to copyright–then that use is not regulated by the license. Our licenses grant only permissions under copyright and certain other rights that a licensor has authority to grant. Use of the licensed material may still be restricted for other reasons, including because others have copyright or other rights in the material. A licensor may make special requests, such as asking that all changes be marked or described. Although not required by our licenses, you are encouraged to respect those requests where reasonable. More considerations for the public.*

Creative Commons Attribution 4.0 International Public License

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution 4.0 International Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits the Licensor receives from making the Licensed Material available under these terms and conditions.

Section 1 – Definitions.

a. **Adapted Material** means material subject to Copyright and Similar Rights that is derived from or based upon the Licensed Material and in which the Licensed Material is translated, altered, arranged, transformed, or otherwise modified in a manner requiring permission under the Copyright and Similar Rights held by the Licensor. For purposes of this Public License, where the Licensed Material is a musical work, performance, or sound recording, Adapted Material is always produced where the Licensed Material is synched in timed relation with a moving image.
b. **Adapter's License** means the license You apply to Your Copyright and Similar Rights in Your contributions to Adapted Material in accordance with the terms and conditions of this Public License.
c. **Copyright and Similar Rights** means copyright and/or similar rights closely related to copyright including, without limitation, performance, broadcast, sound recording, and Sui Generis Database Rights, without regard to how the rights are labeled or categorized. For purposes of this Public License, the rights specified in Section 2(b)(1)-(2) are not Copyright and Similar Rights.
d. **Effective Technological Measures** means those measures that, in the absence of proper authority, may not be circumvented under laws fulfilling obligations under Article 11 of the WIPO Copyright Treaty adopted on December 20, 1996, and/or similar international agreements.
e. **Exceptions and Limitations** means fair use, fair dealing, and/or any other exception or limitation to Copyright and Similar Rights that applies to Your use of the Licensed Material.
f. **Licensed Material** means the artistic or literary work, database, or other material to which the Licensor applied this Public License.
g. **Licensed Rights** means the rights granted to You subject to the terms and conditions of this Public License, which are limited to all Copyright and Similar Rights that apply to Your use of the Licensed Material and that the Licensor has authority to license.
h. **Licensor** means the individual(s) or entity(ies) granting rights under this Public License.
i. **Share** means to provide material to the public by any means or process that requires permission under the Licensed Rights, such as reproduction, public display, public performance, distribution, dissemination, communication, or importation, and to make material available to the public including in ways that members of the public may access the material from a place and at a time individually chosen by them.
j. **Sui Generis Database Rights** means rights other than copyright resulting from Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, as amended and/or succeeded, as well as other essentially equivalent rights anywhere in the world.
k. **You** means the individual or entity exercising the Licensed Rights under this Public License. **Your** has a corresponding meaning.

Section 2 – Scope.

a. **License grant**.
    1. Subject to the terms and conditions of this Public License, the Licensor hereby grants You a worldwide, royalty-free, non-sublicensable, non-exclusive, irrevocable license to exercise the Licensed Rights in the Licensed Material to:
        A. reproduce and Share the Licensed Material, in whole or in part; and
        B. produce, reproduce, and Share Adapted Material.
    2. Exceptions and Limitations. For the avoidance of doubt, where Exceptions and Limitations apply to Your use, this Public License does not apply, and You do not need to comply with its terms and conditions.
    3. Term. The term of this Public License is specified in Section 6(a).
    4. Media and formats; technical modifications allowed. The Licensor authorizes You to exercise the Licensed Rights in all media and formats whether now known or hereafter created, and to make technical modifications necessary to do so. The Licensor waives and/or agrees not to assert any right or authority to forbid You from making technical modifications necessary to exercise the Licensed Rights, including technical modifications necessary to circumvent Effective Technological Measures. For purposes of this Public License, simply making modifications authorized by this Section 2(a)(4) never produces Adapted Material.
    5. Downstream recipients.
        A. Offer from the Licensor – Licensed Material. Every recipient of the Licensed Material automatically receives an offer from the Licensor to exercise the Licensed Rights

under the terms and conditions of this Public License.

B. No downstream restrictions. You may not offer or impose any additional or different terms or conditions on, or apply any Effective Technological Measures to, the Licensed Material if doing so restricts exercise of the Licensed Rights by any recipient of the Licensed Material.

6. No endorsement. Nothing in this Public License constitutes or may be construed as permission to assert or imply that You are, or that Your use of the Licensed Material is, connected with, or sponsored, endorsed, or granted official status by, the Licensor or others designated to receive attribution as provided in Section 3(a)(1)(A)(i).

b. **Other rights**.

1. Moral rights, such as the right of integrity, are not licensed under this Public License, nor are publicity, privacy, and/or other similar personality rights; however, to the extent possible, the Licensor waives and/or agrees not to assert any such rights held by the Licensor to the limited extent necessary to allow You to exercise the Licensed Rights, but not otherwise.
2. Patent and trademark rights are not licensed under this Public License.
3. To the extent possible, the Licensor waives any right to collect royalties from You for the exercise of the Licensed Rights, whether directly or through a collecting society under any voluntary or waivable statutory or compulsory licensing scheme. In all other cases the Licensor expressly reserves any right to collect such royalties.

Section 3 – License Conditions.

Your exercise of the Licensed Rights is expressly made subject to the following conditions.

a. **Attribution**.

1. If You Share the Licensed Material (including in modified form), You must:

A. retain the following if it is supplied by the Licensor with the Licensed Material:
i. identification of the creator(s) of the Licensed Material and any others designated to receive attribution, in any reasonable manner requested by the Licensor (including by pseudonym if designated);
ii. a copyright notice;
iii. a notice that refers to this Public License;
iv. a notice that refers to the disclaimer of warranties;
v. a URI or hyperlink to the Licensed Material to the extent reasonably practicable;

B. indicate if You modified the Licensed Material and retain an indication of any previous modifications; and

C. indicate the Licensed Material is licensed under this Public License, and include the text of, or the URI or hyperlink to, this Public License.

2. You may satisfy the conditions in Section 3(a)(1) in any reasonable manner based on the medium, means, and context in which You Share the Licensed Material. For example, it may be reasonable to satisfy the conditions by providing a URI or hyperlink to a resource that includes the required information.
3. If requested by the Licensor, You must remove any of the information required by Section 3(a)(1)(A) to the extent reasonably practicable.
4. If You Share Adapted Material You produce, the Adapter's License You apply must not prevent recipients of the Adapted Material from complying with this Public License.

Section 4 – Sui Generis Database Rights.

Where the Licensed Rights include Sui Generis Database Rights that apply to Your use of the Licensed Material:

a. for the avoidance of doubt, Section 2(a)(1) grants You the right to extract, reuse, reproduce, and Share all or a substantial portion of the contents of the database;
b. if You include all or a substantial portion of the database contents in a database in which You have Sui Generis Database Rights, then the database in which You have Sui Generis Database Rights (but not its individual contents) is Adapted Material; and

c. You must comply with the conditions in Section 3(a) if You Share all or a substantial portion of the contents of the database.

For the avoidance of doubt, this Section 4 supplements and does not replace Your obligations under this Public License where the Licensed Rights include other Copyright and Similar Rights.

Section 5 – Disclaimer of Warranties and Limitation of Liability.

a. **Unless otherwise separately undertaken by the Licensor, to the extent possible, the Licensor offers the Licensed Material as-is and as-available, and makes no representations or warranties of any kind concerning the Licensed Material, whether express, implied, statutory, or other. This includes, without limitation, warranties of title, merchantability, fitness for a particular purpose, non-infringement, absence of latent or other defects, accuracy, or the presence or absence of errors, whether or not known or discoverable. Where disclaimers of warranties are not allowed in full or in part, this disclaimer may not apply to You.**
b. **To the extent possible, in no event will the Licensor be liable to You on any legal theory (including, without limitation, negligence) or otherwise for any direct, special, indirect, incidental, consequential, punitive, exemplary, or other losses, costs, expenses, or damages arising out of this Public License or use of the Licensed Material, even if the Licensor has been advised of the possibility of such losses, costs, expenses, or damages. Where a limitation of liability is not allowed in full or in part, this limitation may not apply to You.**

c. The disclaimer of warranties and limitation of liability provided above shall be interpreted in a manner that, to the extent possible, most closely approximates an absolute disclaimer and waiver of all liability.

Section 6 – Term and Termination.

a. This Public License applies for the term of the Copyright and Similar Rights licensed here. However, if You fail to comply with this Public License, then Your rights under this Public License terminate automatically.

b. Where Your right to use the Licensed Material has terminated under Section 6(a), it reinstates:

    1. automatically as of the date the violation is cured, provided it is cured within 30 days of Your discovery of the violation; or
    2. upon express reinstatement by the Licensor.

    For the avoidance of doubt, this Section 6(b) does not affect any right the Licensor may have to seek remedies for Your violations of this Public License.
c. For the avoidance of doubt, the Licensor may also offer the Licensed Material under separate terms or conditions or stop distributing the Licensed Material at any time; however, doing so will not terminate this Public License.
d. Sections 1, 5, 6, 7, and 8 survive termination of this Public License.

Section 7 – Other Terms and Conditions.

a. The Licensor shall not be bound by any additional or different terms or conditions communicated by You unless expressly agreed.
b. Any arrangements, understandings, or agreements regarding the Licensed Material not stated herein are separate from and independent of the terms and conditions of this Public License.

Section 8 – Interpretation.

a. For the avoidance of doubt, this Public License does not, and shall not be interpreted to, reduce, limit, restrict, or impose conditions on any use of the Licensed Material that could lawfully be made without permission under this Public License.
b. To the extent possible, if any provision of this Public License is deemed unenforceable, it shall be automatically reformed to the minimum extent necessary to make it enforceable. If the provision cannot be reformed, it shall be severed from this Public License without affecting the enforceability of the remaining terms and conditions.

c. No term or condition of this Public License will be waived and no failure to comply consented to unless expressly agreed to by the Licensor.
d. Nothing in this Public License constitutes or may be interpreted as a limitation upon, or waiver of, any privileges and immunities that apply to the Licensor or You, including from the legal processes of any jurisdiction or authority.

Creative Commons is not a party to its public licenses. Notwithstanding, Creative Commons may elect to apply one of its public licenses to material it publishes and in those instances will be considered the "Licensor." The text of the Creative Commons public licenses is dedicated to the public domain under the CC0 Public Domain Dedication. Except for the limited purpose of indicating that material is shared under a Creative Commons public license or as otherwise permitted by the Creative Commons policies published at creativecommons.org/policies, Creative Commons does not authorize the use of the trademark "Creative Commons" or any other trademark or logo of Creative Commons without its prior written consent including, without limitation, in connection with any unauthorized modifications to any of its public licenses or any other arrangements, understandings, or agreements concerning use of licensed material. For the avoidance of doubt, this paragraph does not form part of the public licenses.

Creative Commons may be contacted at creativecommons.org.

Additional languages available: Bahasa Indonesia, Nederlands, norsk, suomeksi, te reo Māori, українська, 日本語. Please read the FAQ for more information about official translations.

Frequently Asked Questions Regarding IEEE Permissions

- When is permission to reuse IEEE required?
- From whom do I need permission?
- What if I do not see the "Request Permission" link on either the Table of Contents or the Abstract Page in Xplore?
- Does IEEE require individuals working on a thesis or dissertation to obtain formal permission for reuse?
- If I want to republish an article in another language do I still need to obtain a license from IEEE?
- How do I obtain permission to use photographs or illustrations?
- Do I need to obtain permission to use IEEE material posted on its website?
- Does IEEE require certain rights when requesting permission to use material in an IEEE work?
- What is Rightslink®?
- Is IEEE an STM signatory publisher?

- **When is permission to reuse IEEE required?**

As a general rule, IEEE requires permission be sought to reproduce any substantial part of its intellectual property, including any text, illustrations, charts, tables, photographs, or other material from previously published sources used. IEEE also requires that all references or sources used be credited, whether or not permission is required. For further guidance, please contact pubs-permissions@ieee.org.

- **From whom do I need permission?**

Permission must be sought from IEEE to reuse its intellectual property. In most cases this will mean locating the material you wish to reuse in IEEE Xplore, where you will find a "request permission" link either on the Table of Contents or on the Article Abstract Page.

- **What if I do not see the "Request Permission" link on either the Table of Contents or the Abstract Page in Xplore?**

If you do not see a permission link on the Abstract Page, we recommend you review the front cover and/or the copyright page in the document itself (often, these pages are freely available for viewing in Xplore) in order to determine copyright owner. If you are unsure, please contact pubs-permissions@ieee.org.

- **Does IEEE require individuals working on a thesis or dissertation to obtain formal permission for reuse?**

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you must follow the requirements listed below:

Textual Material

Using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.

In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.

If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Full-Text Article

If you are using the entire IEEE copyright owned article, the following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]

Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.

In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

- **If I want to republish an article in another language do I still need to obtain a license from IEEE?**

If you are republishing IEEE intellectual property, we do require you obtain a license that includes any translations. The required translation disclaimer and other translation guidelines are available in the IEEE Terms and Conditions contained in the license provided by the Copyright Clearance Center (RightsLink service).

- **How do I obtain permission to use photographs or illustrations?**

IEEE does not always own reproduction rights to photographs or illustrations; rather, such rights may have been retained by the photographer or illustrator. If the source from which the material is borrowed does not indicate who owns reproduction rights, users of these photographs or illustrations are required to locate the rightsholder, directly.

- **Do I need to obtain permission to use IEEE material posted on its website?**

Yes. As a general rule, most material found on the internet is protected by copyright law even if a notice is not displayed. IEEE does require that you inquire about such permission before using any material found on all IEEE copyright owned websites.

- **Does IEEE require certain rights when requesting permission to use material in an IEEE work?**

IEEE does allow permission to reuse small portions of text in another IEEE copyright owned document only (e.g., the equivalent of several paragraphs only) and figures. Our only requirement is that you 1) provide full credit information pertaining to the original IEEE publications (e.g., author name, paper title, publication title, month and year of original publication). Requests for permission to reuse larger portions should be sent to pubs-permissions@ieee.org .

- **What is Rightslink®?**

Rightslink® is the Copyright Clearance Center's automated permissions granting service, which is used by IEEE along with many other STM publishers such as Springer, Elsevier, and Taylor & Francis. Through this permission service, customers can request permission for IEEE Periodical and Conference content from the point of access; (normally found on the abstract page of the individual article, in IEEE Xplore).

- **Is IEEE an STM signatory publisher?**

No, IEEE is not a signatory to the STM (International Association of Scientific, Technical & Medical Publishers) Permissions Guidelines, last updated February 2012.

April 2013, nbd

Bibliography

(2013). BOOST C++ Libraries 1.53.0.

(2014). MATLAB with Statistics Toolbox Release 2014a.

Acar, E. U., Choset, H., Zhang, Y., and Schervish, M. (2003). Path planning for robotic demining: Robust sensor-based coverage of unstructured environments and probabilistic methods. *The International Journal of Robotics Research*, 22(7-8):441–466.

Ackley, D. H. (2013). Beyond efficiency. *Communications of the ACM*, 56(10):38–40.

Ackley, D. H., Cannon, D. C., and Williams, L. R. (2012). A movable architecture for robust spatial computing. *The Computer Journal*, page bxs129.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). *Molecular Biology of the Cell*. Garland Science, New York, 4th edition.

Aleksandr, I. and Khinchin, A. (1949). *Mathematical foundations of statistical mechanics*. Courier Corporation.

Allenspach, E. J., Cullinan, P., Tong, J., Tang, Q., Tesciuba, A. G., Cannon, J. L., Takahashi, S. M., Morgan, R., Burkhardt, J. K., and Sperling, A. I. (2001). ERM-dependent movement of CD43 defines a novel protein complex distal to the immunological synapse. *Immunity*, 15(5):739–750.

Ariotti, S., Beltman, J. B., Chodaczek, G., Hoekstra, M. E., van Beek, A. E., Gomez-Eerland, R., Ritsma, L., van Rheenen, J., Marée, A. F. M., and Zal, T. (2012). Tissue-resident memory CD8+ T cells continuously patrol skin epithelia to quickly recognize local antigen. *Proceedings of the National Academy of Sciences*, 109(48):19739–19744.

Axelrod, R. (1997). Advancing the art of simulation in the social sciences. In *Simulating social phenomena*, pages 21–40. Springer.

Bac, C. W., Henten, E. J., Hemming, J., and Edan, Y. (2014). Harvesting Robots for Highvalue Crops: Stateoftheart Review and Challenges Ahead. *Journal of Field Robotics*, 31(6):888–911.

Baeza-Yates, R. and Schott, R. (1995). Parallel searching in the plane. *Computational Geometry*, 5(3):143–154.

Baeza-Yates, R. A., Culberson, J. C., and Rawlins, G. J. E. (1993). Searching in the plane. *Information and Computation*, 106(2):234–252.

Bajénoff, M., Egen, J. G., Koo, L. Y., Laugier, J. P., Brau, F., Glaichenhaus, N., and Germain, R. N. (2006). Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes. *Immunity*, 25(6):989–1001.

Banerjee, S., Levin, D., Moses, M., Koster, F., and Forrest, S. (2011). The value of inflammatory signals in adaptive immune responses. In *Artificial Immune Systems*, pages 1–14. Springer.

Banigan, E. J., Harris, T. H., Christian, D. A., Hunter, C. A., Liu, A. J., and Asquith, B. (2015). Heterogeneous CD8+ T Cell Migration in the Lymph Node in the Absence of Inflammation Revealed by Quantitative Migration Analysis. *PLoS computational biology*, 11(2):e1004058–e1004058.

Bartumeus, F., da Luz, M. G. E., Viswanathan, G. M., and Catalan, J. (2005). Animal search strategies: a quantitative random-walk analysis. *Ecology*, 86(11):3078–3087.

Beal, J. (2015). Superdiffusive Dispersion and Mixing of Swarms. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 10(2):10.

Beltman, J. B., Marée, A. F. M., Lynch, J. N., Miller, M. J., and de Boer, R. J. (2007). Lymph node topology dictates T cell migration behavior. *The Journal of experimental medicine*, 204(4):771–780.

Bénichou, O., Loverdo, C., Moreau, M., and Voituriez, R. (2011). Intermittent search strategies. *Reviews of Modern Physics*, 83(1):81.

Bentley, J. L., Weide, B. W., and Yao, A. C. (1980). Optimal expected-time algorithms for closest point problems. *ACM Transactions on Mathematical Software (TOMS)*, 6(4):563–580.

Beverly, B. D., McLendon, H., Nacu, S., Holmes, S., and Gordon, D. M. (2009). How site fidelity leads to individual differences in the foraging activity of harvester ants. *Behavioral Ecology*, 20(3):633–638.

Birk, A. and Carpin, S. (2006). Rescue robotics - a crucial milestone on the road to autonomous systems. *Advanced Robotics*, 20(5):595–605.

Bland, J. M. and Altman, D. G. (1995). Multiple significance tests: the Bonferroni method. *Bmj*, 310(6973):170.

Borenstein, J., Feng, L., and Everett, H. R. (1996). *Navigating mobile robots: Systems and techniques*. AK Peters, Ltd.

Bousso, P., Bhakta, N. R., Lewis, R. S., and Robey, E. (2002). Dynamics of thymocyte-stromal cell interactions visualized by two-photon microscopy. *Science (New York, N.Y.)*, 296(5574):1876–80.

Bousso, P. and Robey, E. (2003). Dynamics of CD8+ T cell priming by dendritic cells in intact lymph nodes. *Nature immunology*, 4(6):579–585.

Brambilla, M., Ferrante, E., Birattari, M., and Dorigo, M. (2013). Swarm robotics: a review from the swarm engineering perspective. *Swarm Intelligence*, 7(1):1–41.

Burlington, S. and Dudek, G. (1999). Spiral search as an efficient mobile robotic search technique. In *Proceedings of the 16th National Conf. on AI, Orlando Fl*.

Cannon, J. L., Asperti-Boursin, F., Letendre, K. A., Brown, I. K., Korzekwa, K. E., Blaine, K. M., Oruganti, S. R., Sperling, A. I., and Moses, M. E. (2013). PKCθ Regulates T Cell Motility via Ezrin-Radixin-Moesin Localization to the Uropod. *PloS one*, 8(11):e78940.

Celli, S., Day, M., Müller, A. J., Molina-Paris, C., Lythe, G., and Bousso, P. (2012). How many dendritic cells are required to initiate a T-cell response? *Blood*, 120(19):3945–3948.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical population biology*, 9(2):129–136.

Chechkin, A. V., Metzler, R., Klafter, J., and Gonchar, V. Y. (2008). Introduction to the theory of Lévy flights. *Anomalous Transport: Foundations and Applications*, pages 129–162.

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4):661–703.

Codling, E. A., Plank, M. J., and Benhamou, S. (2008). Random walk models in biology. *Journal of the Royal Society Interface*, 5(25):813–834.

Cox, I. J. (1991). Blanche-an experiment in guidance and navigation of an autonomous robot vehicle. *IEEE Transactions on robotics and automation*, 7(2):193–204.

De Boer, R. J., Oprea, M., Antia, R., Murali-Krishna, K., Ahmed, R., and Perelson, A. S. (2001). Recruitment times, proliferation, and apoptosis rates during the CD8+ T-cell response to lymphocytic choriomeningitis virus. *Journal of virology*, 75(22):10663–10669.

de Jager, M., Bartumeus, F., Kölzsch, A., Weissing, F. J., Hengeveld, G. M., Nolet, B. A., Herman, P. M. J., and van de Koppel, J. (2014). How superdiffusion gets arrested: ecological encounters explain shift from Lévy to Brownian movement. *Proceedings of the Royal Society B: Biological Sciences*, 281(1774):20132605.

Donovan, G. M. and Lythe, G. (2012). T-cell movement on the reticular network. *Journal of theoretical biology*, 295:59–67.

Donovan, G. M. and Lythe, G. (2016). T cell and reticular network co-dependence in HIV infection. *Journal of theoretical biology*.

Dunning, J. B., Danielson, B. J., and Pulliam, H. R. (1992). Ecological processes that affect populations in complex landscapes. *Oikos*, pages 169–175.

Dyson, F. (2004). A meeting with Enrico Fermi. *Nature*, 427(6972):297.

Edwards, A. M. (2011). Overturning conclusions of Lévy flight movement patterns by fishing boats and foraging animals. *Ecology*, 92(6):1247–1257.

ElHadidy, M. A. A. (2015). Optimal spiral search plan for a randomly located target in the plane. *International Journal of Operational Research*, 22(4):454–465.

Feinerman, O., Korman, A., Lotker, Z., and Sereni, J.-S. (2012). Collaborative search on the plane without communication. In *Proceedings of the 2012 ACM symposium on Principles of distributed computing*, pages 77–86. ACM.

Feuillet, V., Lucas, B., Di Santo, J. P., Bismuth, G., and Trautmann, A. (2005). Multiple survival signals are delivered by dendritic cells to naive CD4+ T cells. *European journal of immunology*, 35(9):2563–72.

Fink, W., Dohm, J. M., Tarbell, M. A., Hare, T. M., and Baker, V. R. (2005). Next-generation robotic planetary reconnaissance missions: a paradigm shift. *Planetary and Space Science*, 53(14):1419–1426.

Fisher, R. A. (1925). *Theory of statistical estimation*, volume 22. Cambridge Univ Press.

Flanagan, T. P., Letendre, K., Burnside, W. R., Fricke, G. M., and Moses, M. E. (2012). Quantifying the effect of colony size and food distribution on harvester ant foraging. *PloS one*, 7(7):e39427.

Fricke, G. M., Asperti-Boursin, F., Hecker, J., Cannon, J., and Moses, M. (2013). From Microbiology to Microcontrollers: Robot Search Patterns Inspired by T Cell Movement. In *Advances in Artificial Life (ECAL)*, volume 12, pages 1009–1016.

Fricke, G. M., Hecker, J. P., Black, S. R., Cannon, J. L., and Moses, M. E. (2015). Distinguishing Adaptive Search From Random Search in Robots and T cells. In *Proceedings of the Conference on Genetic and Evolutionary Computation*. ACM.

Fricke, G. M., Hecker, J. P., Cannon, J. L., and Moses, M. E. (2016a). Immune-inspired search strategies for robot swarms. *Robotica Journal Special Issue on Robotics Methods for Structural and Dynamics Modeling of Molecular Systems*, pages 1791–1810.

Fricke, G. M., Hecker, Joshua, P., Griego, Antonio, D., Tran, Linh, T., and Moses, Melanie, E. (2016b). A Distributed Deterministic Spiral Search Algorithm for Swarms. In *Proceedings of the 29th IEEE/RSJ 2016 International Conference on Intelligent Robots and Systems (IROS)*. Copyright 2016, IEEE reprinted with permission from the authors.

Fricke, G. M., Letendre, K. A., Moses, M. E., and Cannon, J. L. (2016c). Persistence and Adaptation in Immunity: T Cells Balance the Extent and Thoroughness of Search. *PLoS Computational Biology*, 12(3):e1004818.

Friedman, R. S., Beemiller, P., Sorensen, C. M., Jacobelli, J., and Krummel, M. F. (2010). Real-time analysis of T cell receptors in naive cells in vitro and in vivo reveals flexibility in synapse and signaling dynamics. *The Journal of experimental medicine*, 207(12):2733–2749.

Furusawa, C., Suzuki, T., Kashiwagi, A., Yomo, T., and Kaneko, K. (2005). Ubiquity of log-normal distributions in intra-cellular reaction dynamic. *arXiv preprint q-bio/0503040*.

Gabal, H. M. A. and El-Hadidy, M. A. A. (2015). Optimal searching for a randomly located target in a bounded known region. *International Journal of Computing Science and Mathematics*, 6(4):392–403.

Gérard, A., Patino-Lopez, G., Beemiller, P., Nambiar, R., Ben-Aissa, K., Liu, Y., Totah, F. J., Tyska, M. J., Shaw, S., and Krummel, M. F. (2014). Detection of Rare Antigen-Presenting Cells through T Cell-Intrinsic Meandering Motility, Mediated by Myo1g. *Cell*, 158(3):492–505.

Germain, R. N., Robey, E. a., and Cahalan, M. D. (2012). A decade of imaging cellular motility and interaction dynamics in the immune system. *Science (New York, N.Y.)*, 336(6089):1676–81.

Giraldeau, L.-A. and Caraco, T. (2000). *Social foraging theory*. Princeton University Press.

Glass, G. V., Peckham, P. D., and Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research*, 42(3):237–288.

Goldstein, M. L., Morris, S. A., and Yen, G. G. (2004). Problems with fitting to the power-law distribution. *The European Physical Journal B-Condensed Matter and Complex Systems*, 41(2):255–258.

Goodrich, M. A., Morse, B. S., Gerhardt, D., Cooper, J. L., Quigley, M., Adams, J. A., and Humphrey, C. (2008). Supporting wilderness search and rescue using a cameraequipped mini UAV. *Journal of Field Robotics*, 25(12):89–110.

Gordon, D. M. (1994). How social insect colonies respond to variable environments. *Behavioral mechanisms in evolutionary ecology. The University of Chicago Press, Chicago*, pages 409–422.

Graw, F. and Regoes, R. R. (2012). Influence of the fibroblastic reticular network on cell-cell interactions in lymphoid organs. *PLoS Computational Biology*, 8(3):e1002436.

Groom, J. R., Richmond, J., Murooka, T. T., Sorensen, E. W., Sung, J. H., Bankert, K., von Andrian, U. H., Moon, J. J., Mempel, T. R., and Luster, A. D. (2012). CXCR3 chemokine receptor-ligand interactions in the lymph node optimize CD4+ T helper 1 cell differentiation. *Immunity*, 37(6):1091–1103.

Gunning, C. E. and Wearing, H. J. (2013). Probabilistic measures of persistence and extinction in measles (meta) populations. *Ecology letters*, 16(8):985–994.

Hao, Z. B., Sang, N., and Lei, H. (2008). Cooperative Coverage by Multiple Robots with Contact Sensors. In *Robotics, Automation and Mechatronics, 2008 IEEE Conference on*, pages 543–548. IEEE.

Harris, T. H., Banigan, E. J., Christian, D. A., Konradt, C., Wojno, E. D. T., Norose, K., Wilson, E. H., John, B., Weninger, W., Luster, A. D., and Others (2012). Generalized Lévy walks and the role of chemokines in migration of effector CD8+ T cells. *Nature*, 486(7404):545–548.

Hart, P. E., Nilsson, N. J., and Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *Systems Science and Cybernetics, IEEE Transactions on*, 4(2):100–107.

Hayes, A. T., Martinoli, A., and Goodman, R. M. (2001). Swarm robotic odor localization. In *Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on*, volume 2, pages 1073–1078. IEEE.

Hecker, J. P., Carmichael, J. C., and Moses, M. E. (2015). Exploiting clusters for complete resource collection in biologically-inspired robot swarms. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (in press)*.

Hecker, J. P. and Moses, M. E. (2015). Beyond pheromones: evolving error-tolerant, flexible, and scalable ant-inspired robot swarms. *Swarm Intelligence*, 9(1):43–70.

Hecker, J. P., Stolleis, K., Swenson, B., Letendre, K., and Moses, M. E. (2013). Evolving Error Tolerance in Biologically-Inspired iAnt Robots. In *Proceedings of the Twelfth European Conference on the Synthesis and Simulation of Living Systems (Advances in Artificial Life, ECAL 2013)*, pages 1025–1032.

Hochweller, K., Wabnitz, G. H., Samstag, Y., Suffner, J., Hämmerling, G. J., and Garbi, N. (2010). Dendritic cells control T cell tonic signaling required for responsiveness to foreign antigen. *Proceedings of the National Academy of Sciences of the United States of America*, 107(13):5931–6.

Hoff, N., Wood, R., and Nagpal, R. (2013). Distributed colony-level algorithm switching for robot swarm foraging. In *Distributed Autonomous Robotic Systems*, pages 417–430. Springer.

Hofmeyr, S. A. and Forrest, S. (2000). Architecture for an artificial immune system. *Evolutionary computation*, 8(4):443–473.

Hogg, R. V. and Ledolter, J. (1987). *Engineering statistics*. Macmillan Pub Co.

Holland, J. H. (1992). Complex adaptive systems. *Daedalus*, pages 17–30.

Hopkins, B. and Skellam, J. G. (1954). A new method for determining the type of distribution of plant individuals. *Annals of Botany*, 18(2):213–227.

Hu, H., Oyekan, J., and Gu, D. (2011). A school of robotic fish for pollution detection in port. *Biologically Inspired Robotics (Y. Liu and D. Sun, eds.)*, pages 85–104.

Hughes, B. D. (1996). *Random walks and random environments*. Clarendon Press Oxford.

Humphreys, T. E., Ledvina, B. M., Psiaki, M. L., O'Hanlon, B. W., and Kintner Jr, P. M. (2008). Assessing the spoofing threat: Development of a portable GPS civilian spoofer. In *Proceedings of the ION GNSS international technical meeting of the satellite division*, volume 55, page 56.

Humphries, N. E., Weimerskirch, H., Queiroz, N., Southall, E. J., and Sims, D. W. (2012). Foraging success of biological Lévy flights recorded in situ. *Proceedings of the National Academy of Sciences*, 109(19):7169–7174.

Isbell, J. R. (1957). An optimal search pattern. *Naval Research Logistics Quarterly*, 4(4):357–359.

Jain, A. K. and Dubes, R. C. (1988). *Algorithms for clustering data*, volume 6. Prentice hall Englewood Cliffs.

James, A., Pitchford, J. W., and Plank, M. J. (2010). Efficient or inaccurate? Analytical and numerical modelling of random search strategies. *Bulletin of mathematical biology*, 72(4):896–913.

James, A., Plank, M. J., and Edwards, A. M. (2011). Assessing Lévy walks as models of animal foraging. *Journal of The Royal Society Interface*, 8(62):1233–1247.

Johnson, W. W. (1907). A treatise on the integral calculus founded on the method of rates, by William Woolsey Johnson...

Keeter, M., Moore, D., Muller, R., Nieters, E., Flenner, J., Martonosi, S. E., Bertozzi, A. L., Percus, A. G., and Levy, R. (2012). Cooperative search with autonomous vehicles in a 3d aquatic testbed. In *American Control Conference (ACC), 2012*, pages 3154–3160. IEEE.

Krapivsky, P. L., Redner, S., and Ben-Naim, E. (2010). *A kinetic view of statistical physics*. Cambridge University Press.

Krebs, J. R. (1978). Optimal foraging: decision rules for predators. *Behavioural ecology: an evolutionary approach*, pages 23–63.

Krummel, M. F., Friedman, R. S., and Jacobelli, J. (2014). Modes and mechanisms of T cell motility: roles for confinement and Myosin-IIA. *Current opinion in cell biology*, 30:9–16.

Langetepe, E. (2010). On the optimality of spiral search. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 1–12. Society for Industrial and Applied Mathematics.

Larralde, H., Trunfio, P., Havlin, S., Stanley, H. E., and Weiss, G. H. (1992). Territory covered by N diffusing particles. *Nature*, 355(6359):423–426.

Letendre, K., Asperti-Boursin, F., Donnadieu, E., Moses, M. E., and Cannon, J. L. (2015). Bringing Statistics Up To Speed With Data in Analysis of Lymphocyte Motility. *PloS one*.

Letendre, K. and Moses, M. E. (2013). Synergy in ant foraging strategies: memory and communication alone and in combination. In *Proceedings of the 15th annual conference on Genetic and evolutionary computation*, pages 41–48. ACM.

Levin, D. (2016). The Environment Constrains Successful Search Strategies in Natural Distributed Systems.

Lévy, P. and Borel, M. É. (1954). *Théorie de l'addition des variables aléatoires*, volume 1. Gauthier-Villars Paris.

Linderman, J. J., Riggs, T., Pande, M., Miller, M., Marino, S., and Kirschner, D. E. (2010). Characterizing the dynamics of CD4+ T cell priming within a lymph node. *The Journal of Immunology*, 184(6):2873–2885.

Lindquist, R. L., Shakhar, G., Dudziak, D., Wardemann, H., Eisenreich, T., Dustin, M. L., and Nussenzweig, M. C. (2004). Visualizing dendritic cell networks in vivo. *Nature immunology*, 5(12):1243–1250.

Liu, W., Winfield, A. F. T., and Sa, J. (2007). Modelling swarm robotic systems: A case study in collective foraging. *Towards autonomous robotic systems (TAROS 07)*, pages 25–32.

López-Ortiz, A. and Maftuleac, D. (2016). Optimal Distributed Searching in the Plane with and without Uncertainty. In *International Workshop on Algorithms and Computation*, pages 68–79. Springer.

Love, J., Amai, W., Blada, T., Little, C., Neely, J., and Buerger, S. (2015). The Sandia architecture for heterogeneous unmanned system control (SAHUC). In *Proc. SPIE 9464, Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR VI*. International Society for Optics and Photonics.

Lu, Q., Hecker, J. P., and Moses, M. E. (2016). The MPFA: A Multiple-Place Foraging Algorithm for Biologically-Inspired Robot Swarms. *IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Mackay, I. R., Rosen, F. S., von Andrian, U. H., and Mackay, C. R. (2000). T-Cell Function and Migration - Two Sides of the Same Coin. *New England Journal of Medicine*, 343(14):1020–1034.

Maier, D. and Kleiner, A. (2010). Improved GPS sensor model for mobile robots in urban terrain. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 4385–4390. IEEE.

Mandelbrot, B. B. (1983). *The fractal geometry of nature*, volume 173. Macmillan.

Mårell, A., Ball, J. P., and Hofgaard, A. (2002). Foraging and movement paths of female reindeer: insights from fractal analysis, correlated random walks, and Lévy flights. *Canadian Journal of Zoology*, 80(5):854–865.

Matheu, M. P., Parker, I., and Cahalan, M. D. (2007). Dissection and 2-photon imaging of peripheral lymph nodes in mice. *Journal of Visualized Experiments: JoVE*, (7).

Maxwell, J. C. (1860). V. Illustrations of the dynamical theory of gases.Part I. On the motions and collisions of perfectly elastic spheres. *Philosophical Magazine Series 4*, 19(124):19–32.

McLachlan, G. and Peel, D. (2004). *Finite mixture models*. John Wiley & Sons.

Mempel, T. R., Henrickson, S. E., and Von Andrian, U. H. (2004). T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature*, 427(6970):154–159.

Méndez, V., Campos, D., and Bartumeus, F. (2013). *Stochastic foundations in movement ecology: anomalous diffusion, front propagation and random searches*. Springer Science & Business Media.

Michalet, X. (2010). Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. *Physical Review E*, 82(4):41914.

Michalet, X. (2011). Erratum: Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium [Phys. Rev. E 82, 041914 (2010)]. *Physical Review E*, 83(5):59904.

Michalski, R. S., Carbonell, J. G., and Mitchell, T. M. (1986). Machine learning: An artificial intelligence approach. Vol. II. page 251.

Miller, M. J., Hejazi, A. S., Wei, S. H., Cahalan, M. D., and Parker, I. (2004). 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(4):998–1003.

Miller, M. J., Wei, S. H., Cahalan, M. D., and Parker, I. (2003). Autonomous T cell trafficking examined in vivo with intravital two-photon microscopy. *Proceedings of the National Academy of Sciences*, 100(5):2604–2609.

Miller, M. J., Wei, S. H., Parker, I., and Cahalan, M. D. (2002). Two-photon imaging of lymphocyte motility and antigen response in intact lymph node. *Science Signaling*, 296(5574):1869.

Mirsky, H. P., Miller, M. J., Linderman, J. J., and Kirschner, D. E. (2011a). Systems biology approaches for understanding cellular mechanisms of immunity in lymph nodes during infection.

Mirsky, H. P., Miller, M. J., Linderman, J. J., and Kirschner, D. E. (2011b). Systems biology approaches for understanding cellular mechanisms of immunity in lymph nodes during infection. *Journal of theoretical biology*, 287:160–170.

Miyasaka, M. and Tanaka, T. (2004). Lymphocyte trafficking across high endothelial venules: dogmas and enigmas. *Nature Reviews Immunology*, 4(5):360–370.

Montgomery, D. C. (2012). *Design and analysis of experiments*. John Wiley & Sons, 8th edition.

Moore, T. and Stouch, D. (2016). A Generalized Extended Kalman Filter Implementation for the Robot Operating System. In *Intelligent Autonomous Systems 13*, pages 335–348. Springer.

Moses, M. E. (2005). Metabolic scaling from individuals to societies.

Müller, M. and Wehner, R. (1994). The hidden spiral: systematic search and path integration in desert ants, Cataglyphis fortis. *Journal of Comparative Physiology A*, 175(5):525–530.

Munoz, M. A., Biro, M., and Weninger, W. (2014a). T cell migration in intact lymph nodes in vivo. *Current opinion in cell biology*, 30:17–24.

Munoz, M. a., Biro, M., and Weninger, W. (2014b). T cell migration in intact lymph nodes in vivo. *Current opinion in cell biology*, 30C:17–24.

Nurzaman, S. G., Matsumoto, Y., Nakamura, Y., Koizumi, S., and Ishiguro, H. (2009). Yuragi-based adaptive searching behavior in mobile robot: From bacterial chemotaxis to Lévy walk. In *Robotics and Biomimetics, 2008. ROBIO 2008. IEEE International Conference on*, pages 806–811. IEEE.

Orians, G. H. and Pearson, N. E. (1979). On the theory of central place foraging. *Analysis of ecological systems*, pages 155–177.

Papi, F. (2012). *Animal homing*. Springer Science & Business Media.

Parker, L. E. (2009). Path planning and motion coordination in multiple mobile robot teams. *Encyclopedia of complexity and system science*, pages 5783–5800.

Pawley, J. B. and Masters, B. R. (2008). Handbook of biological confocal microscopy. *Journal of biomedical optics*, 13(2):9902.

Petrovskii, S., Mashanova, A., and Jansen, V. A. A. (2011). Variation in individual walking behavior creates the impression of a Lévy flight. *Proceedings of the National Academy of Sciences*, 108(21):8704–8707.

Pinciroli, C., Trianni, V., O'Grady, R., Pini, G., Brutschy, A., Brambilla, M., Mathews, N., Ferrante, E., Di Caro, G., and Ducatelle, F. (2012). ARGoS: a modular, parallel, multi-engine simulator for multi-robot systems. *Swarm intelligence*, 6(4):271–295.

Plank, M. J. and James, A. (2008). Optimal foraging: Lévy pattern or process? *Journal of The Royal Society Interface*, 5(26):1077–1086.

Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A. (2003). Mutual-information-based registration of medical images: a survey. *IEEE transactions on medical imaging*, 22(8):986–1004.

Potdar, A. A., Jeon, J., Weaver, A. M., and Cummings, P. T. (2008). Cell Migration Paths of Epithelial Cells Resemble Lévy Modulated Correlated Random Walk Pattern. In *Proceedings of the 2008 Annual Meeting of the American Institute of Chemical Engineers*.

Preston, S. P., Waters, S. L., Jensen, O. E., Heaton, P. R., and Pritchard, D. I. (2006). T-cell motility in the early stages of the immune response modeled as a random walk amongst targets. *Physical Review E*, 74(1):11910.

Przibram, K. (1913). Über die ungeordnete Bewegung niederer Tiere. *Pflügers Archiv European Journal of Physiology*, 153(8):401–405.

Pyke, G. H. (1984). Optimal foraging theory: a critical review. *Annual review of ecology and systematics*, pages 523–575.

Qian, H., Sheetz, M. P., and Elson, E. L. (1991). Single particle tracking. Analysis of diffusion and flow in two-dimensional systems. *Biophysical journal*, 60(4):910–921.

Raichlen, D. A., Wood, B. M., Gordon, A. D., Mabulla, A. Z. P., Marlowe, F. W., and Pontzer, H. (2014). Evidence of Lévy walk foraging patterns in human hunter-gatherers. *Proceedings of the National Academy of Sciences*, 111(2):728–733.

Ramsey, S. (2015). NASA Awards Grant to Manage Swarmathon Challenge (press release).

Raposo, E. P., Bartumeus, F., Da Luz, M. G. E., Ribeiro-Neto, P. J., Souza, T. A., and Viswanathan, G. M. (2011). How landscape heterogeneity frames optimal diffusivity in searching processes. *PLoS computational biology*, 7(11):e1002233.

Regner, M. (2001). Cross-reactivity in T-cell antigen recognition. *Immunology and cell biology*, 79(2):91–100.

Reynolds, A. M. (2010). Bridging the gulf between correlated random walks and Lévy walks: autocorrelation as a source of Lévy walk movement patterns. *Journal of the Royal Society Interface*, page rsif20100292.

Reynolds, A. M., Smith, A. D., Menzel, R., Greggers, U., Reynolds, D. R., and Riley, J. R. (2007). Displaced honey bees perform optimal scale-free search flights. *Ecology*, 88(8):1955–1961.

Ritchie, M. E. (2009). *Scale, heterogeneity, and the structure and diversity of ecological communities*. Princeton University Press.

Robbins, H. (1985). Some aspects of the sequential design of experiments. In *Herbert Robbins Selected Papers*, pages 169–177. Springer.

Ryan, A. and Hedrick, J. K. (2005). A mode-switching path planner for UAV-assisted search and rescue. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*, pages 1471–1476. IEEE.

Sahin, E. (2005). Swarm robotics: From sources of inspiration to domains of application. In *Swarm robotics*, pages 10–20. Springer.

Sebbane, Y. B. (2011). *Lighter than air robots: guidance and control of autonomous airships*, volume 58. Springer Science & Business Media.

Seshadri, V. and West, B. J. (1982). Fractal dimensionality of L\'evy processes. *Proceedings of the National Academy of Sciences of the United States of America*, 79(14):4501.

Shalizi, C. R. (2006). Methods and techniques of complex systems science: An overview. In *Complex systems science in biomedicine*, pages 33–114. Springer.

Shlesinger, M. F. and Klafter, J. (1986). Lévy walks versus Lévy flights. In *On growth and form*, pages 279–283. Springer.

Shlesinger, M. F., Klafter, J., and Zumofen, G. (1999). Above, below and beyond Brownian motion. *American Journal of Physics*, 67:1253.

Skubch, H. (2012). *Modelling and Controlling of Behaviour for Autonomous Mobile Robots*. Springer Science & Business Media.

Stephens, D. W. and Krebs, J. R. (1986). *Foraging theory*. Princeton University Press.

Stoll, S., Delon, J., Brotz, T. M., and Germain, R. N. (2002). Dynamic imaging of T cell-dendritic cell interactions in lymph nodes. *Science*, 296(5574):1873–1876.

Stone, L. D. (1975). *Theory of optimal search*. Academic Press New York.

Stone, M. (1977). An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 44–47.

Sung, J. H., Zhang, H., Moseman, E. A., Alvarez, D., Iannacone, M., Henrickson, S. E., Juan, C., Groom, J. R., Luster, A. D., and von Andrian, U. H. (2012). Chemokine guidance of central memory T cells is critical for antiviral recall responses in lymph nodes. *Cell*, 150(6):1249–1263.

Sutantyo, D. K., Kernbach, S., Levi, P., and Nepomnyashchikh, V. A. (2010). Multi-Robot searching algorithm using Lévy flight and artificial potential field. In *Safety Security and Rescue Robotics (SSRR), 2010 IEEE International Workshop on*, pages 1–6. IEEE.

Tamura, K. and Naruse, K. (2014). Unsmooth field sweeping by Balistic random walk of multiple robots in unsmooth terrain. In *Soft Computing and Intelligent Systems (SCIS), 2014 Joint 7th International Conference on and Advanced Intelligent Systems (ISIS), 15th International Symposium on*, pages 585–589. IEEE.

Tarantino, N., Tinevez, J.-Y., Crowell, E. F., Boisson, B., Henriques, R., Mhlanga, M., Agou, F., Israël, A., and Laplantine, E. (2014). TNF and IL-1 exhibit distinct ubiquitin requirements for inducing NEMOIKK supramolecular structures. *The Journal of cell biology*, 204(2):231–245.

Taylor, L. R. (1961). Aggregation, variance and the mean. *Nature*, pages 732–735.

Taylor, S. J. (1953). The Hausdorff α-dimensional measure of Brownian paths in n-space. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 49, pages 31–39. Cambridge Univ Press.

Textor, J., Henrickson, S. E., Mandl, J. N., von Andrian, U. H., Westermann, J., de Boer, R. J., and Beltman, J. B. (2014). Random migration and signal integration promote rapid and robust T cell recruitment. *PLoS computational biology*, 10(8):e1003752.

Textor, J., Peixoto, A., Henrickson, S. E., Sinn, M., von Andrian, U. H., and Westermann, J. (2011). Defining the quantitative limits of intravital two-photon lymphocyte tracking. *Proceedings of the National Academy of Sciences*, 108(30):12401–12406.

Thoppian, M. R. and Prakash, R. (2006). A distributed protocol for dynamic address assignment in mobile ad hoc networks. *IEEE Transactions on Mobile Computing*, 5(1):4–19.

Tribus, M. and McIrvine, E. C. (1971). Energy and information. *Scientific American*, 225(3):179–188.

Trummel, K. E. and Weisinger, J. R. (1986). Technical NoteThe Complexity of the Optimal Searcher Path Problem. *Operations Research*, 34(2):324–327.

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, pages 99–114.

United States Department of Defense (2008). Global positioning system standard positioning service performance standard. *SPSGPS*, 4th Ed.:9–15.

U.S. Coast Guard (2002). U.S. Coast Guard Addendum to the United States National SAR Supplement (CGADD) COMDTINST M16130.2C.

Van Dartel, M., Postma, E., van den Herik, J., and de Croon, G. (2004). Macroscopic analysis of robot foraging behaviour. *Connection Science*, 16(3):169–181.

Viitala, J., Korplmäki, E., Palokangas, P., and Koivula, M. (1995). Attraction of kestrels to vole scent marks visible in ultraviolet light.

Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Havlin, S., Da Luz, M. G. E., Raposo, E. P., and Stanley, H. E. (2000). Lévy flights in random searches. *Physica A: Statistical Mechanics and its Applications*, 282(1):1–12.

Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A., and Stanley, H. E. (1996). Lévy flight search patterns of wandering albatrosses. *Nature*, 381(6581):413–415.

Viswanathan, G. M., Bartumeus, F., V Buldyrev, S., Catalan, J., Fulco, U. L., Havlin, S., Da Luz, M. G. E., Lyra, M. L., Raposo, E. P., and Eugene Stanley, H. (2002). Lévy flight random searches in biological phenomena. *Physica A: Statistical Mechanics and Its Applications*, 314(1):208–213.

Viswanathan, G. M., Buldyrev, S. V., Havlin, S., Da Luz, M. G. E., Raposo, E. P., and Stanley, H. E. (1999). Optimizing the success of random searches. *Nature*, 401(6756):911–914.

Viswanathan, G. M., Da Luz, M. G. E., Raposo, E. P., and Stanley, H. E. (2011). *The physics of foraging: an introduction to random searches and biological encounters*. Cambridge University Press.

Viswanathan, G. M., Raposo, E. P., Bartumeus, F., Catalan, J., and da Luz, M. G. E. (2005). Necessary criterion for distinguishing true superdiffusion from correlated random walk processes. *Phys. Rev. E*, 72(1):11111.

Viswanathan, G. M., Raposo, E. P., and Da Luz, M. G. E. (2008). Lévy flights and superdiffusion in the context of biological encounters and random searches. *Physics of Life Reviews*, 5(3):133–150.

Von Mises, R. and Geiringer, H. (1964). *Mathematical theory of probability and statistics*, volume 75. Academic Press New York.

Von Neumann, J. (1951). The general and logical theory of automata. *Cerebral mechanisms in behavior*, 1:1–41.

Weber, T. R. (1995). An Analysis of Lemmings: A Swarming Approach to Mine Countermeasures in the VSW/SZ/BZ. Technical report, DTIC Document.

Weingarten, J. and Siegwart, R. (2005). EKF-based 3D SLAM for structured environment reconstruction. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3834–3839. IEEE.

Wiens, B. L. (1999). When log-normal and gamma models give different results: a case study. *The American Statistician*, 53(2):89–93.

Winfield, A. F. T. (2009). Foraging Robots. In Meyers, R. A., editor, *Encyclopedia of complexity and systems science*, pages 3682–3700. Springer, New York.

Winfield, A. F. T., Harper, C. J., and Nembrini, J. (2005). Towards dependable swarms and a new discipline of swarm engineering. In *Swarm robotics*, pages 126–142. Springer.

Worbs, T., Mempel, T. R., Bölter, J., von Andrian, U. H., and Förster, R. (2007). CCR7 ligands stimulate the intranodal motility of T lymphocytes in vivo. *The Journal of experimental medicine*, 204(3):489–495.

Zanette, D. H. (1999). Statistical-thermodynamical foundations of anomalous diffusion. *Brazilian journal of physics*, 29(1):108–124.

Zeng, M., Southern, P. J., Reilly, C. S., Beilman, G. J., Chipman, J. G., Schacker, T. W., and Haase, A. T. (2012). Lymphoid tissue damage in HIV-1 infection depletes naïve T cells and limits T cell reconstitution after antiretroviral therapy. *PLoS pathogens*, 8(1):e1002437.

Zhang, J., Leiderman, K., Pfeiffer, J. R., Wilson, B. S., Oliver, J. M., and Steinberg, S. L. (2006). Characterizing the topography of membrane receptors and signaling molecules from spatial patterns obtained using nanometer-scale electron-dense probes and electron microscopy. *Micron*, 37(1):14–34.

Zhao, K., Jurdak, R., Liu, J., Westcott, D., Kusy, B., Parry, H., Sommer, P., and McKeown, A. (2015). Optimal Lévy-flight foraging in a finite landscape. *Journal of The Royal Society Interface*, 12(104):20141158.