

12-1-2013

# Abstract Models of Molecular Walkers

Oleg Semenov

Follow this and additional works at: [https://digitalrepository.unm.edu/cs\\_etds](https://digitalrepository.unm.edu/cs_etds)

---

## Recommended Citation

Semenov, Oleg. "Abstract Models of Molecular Walkers." (2013). [https://digitalrepository.unm.edu/cs\\_etds/35](https://digitalrepository.unm.edu/cs_etds/35)

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](mailto:disc@unm.edu).

Oleg Semenov

*Candidate*

Computer Science

*Department*

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

*Approved by the Dissertation Committee:*

Darko Stefanovic

, Chairperson

Cristopher Moore

Melanie Moses

Milan N. Stojanovic

Pavel L. Krapivsky

# **Abstract Models of Molecular Walkers**

by

**Oleg Semenov**

Specialist, Computer Science, Kazakh National Technical  
University, 2005

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

December, 2013

# Acknowledgments

I would like to thank my committee members for the help and advice I have received during the work on this dissertation. Paul Krapivsky introduced and explained to me the model of molecular spiders and its analysis. Cris Moore's advice formed my approach to the analysis of the model. Milan Stojanovic synthesized and tested the original molecular spiders and envisioned various experiments and applications of molecular spiders, many of which are discussed and studied in this dissertation. I would also like to acknowledge Melanie Moses for detailed discussions concerning swarm robotics. Finally, I would like to express my gratitude to my advisor, Darko Stefanovic, for his support and encouragement throughout my graduate studies. His advice helped me in all the time of research, programming, and writing of papers and this thesis.

Also, I would like to thank students within our Molecular Computing Group for their helpful advice, especially about presentations. I would like to thank my co-authors Mark J. Olah and David Mohr, who have greatly contributed to the work which is part of this dissertation.

This work was generously supported by the National Science Foundation under grants 0533065, 0829896, and 1028238. I would also like to thank NVIDIA Corporation for a hardware gift that made possible some of the simulations.

# **Abstract Models of Molecular Walkers**

by

**Oleg Semenov**

Specialist, Computer Science, Kazakh National Technical  
University, 2005

Ph. D., Computer Science, University of New Mexico, 2013

## **Abstract**

Recent advances in single-molecule chemistry have led to designs for artificial multi-pedal walkers that follow tracks of chemicals. The walkers, called molecular spiders, consist of a rigid chemically inert body and several flexible enzymatic legs. The legs can reversibly bind to chemical substrates on a surface, and through their enzymatic action convert them to products. We study abstract models of molecular spiders to evaluate how efficiently they can perform two tasks: molecular transport of cargo over tracks and search for targets on finite surfaces.

For the single-spider model our simulations show a transient behavior wherein certain spiders move superdiffusively over significant distances and times. This gives the spiders potential as a faster-than-diffusion transport mechanism. However, analysis shows that single-spider motion eventually decays into an ordinary diffusive motion, owing to the ever increasing size of the region of products. Inspired by cooperative behavior of natural molecular walkers, we propose a symmetric exclusion process (SEP) model for multiple walkers interacting as they

move over a one-dimensional lattice. We show that when walkers are sequentially released from the origin, the collective effect is to prevent the leading walkers from moving too far backwards. Hence, there is an effective outward pressure on the leading walkers that keeps them moving superdiffusively for longer times. Despite this improvement the leading spider eventually slows down and moves diffusively, similarly to a single spider. The slowdown happens because all spiders behind the leading spiders never encounter substrates, and thus they are never biased. They cannot keep up with leading spiders, and cannot put enough pressure on them.

Next, we investigate search properties of a single and multiple spiders moving over one- and two-dimensional surfaces with various absorbing and reflecting boundaries. For the single-spider model we evaluate by how much the slowdown on newly visited sites, owing to catalysis, can improve the mean first passage time of spiders and show that in one dimension, when both ends of the track are an absorbing boundary, the performance gain is lower than in two dimensions, when the absorbing boundary is a circle; this persists even when the absorbing boundary is a single site. Next, we study how multiple molecular spiders influence one another during the search. We show that when one spider reaches the trace of another spider it is more likely not to follow the trace and instead explore unvisited sites. This interaction between the spiders gives them an advantage over independent random walkers in a search for multiple targets. We also study how efficiently the spiders with various gaits are able to find specific targets. Spiders with gaits that allow more freedom of leg movement find their targets faster than spiders with more restrictive gaits. For every gait, there is an optimal detachment rate that minimizes the time to find all target sites.

# Contents

<b>Acknowledgments</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Molecular Spiders (Physical Level)</b>	<b>3</b>
<b>3 Computational Models</b>	<b>6</b>
<b>4 Single Spider in 1D</b>	<b>10</b>
4.1 The Antal-Krapivsky model . . . . .	11
4.2 Simulation Description . . . . .	14
4.3 Instantaneous Superdiffusion of Spiders . . . . .	15
4.4 Asymptotic Behavior and Distributions . . . . .	19
4.5 Mechanism of Transient Spider Super-diffusion . . . . .	23
4.5.1 The Boundary and Diffusive Metastates . . . . .	23
4.5.2 The Diffusive Metastate $D$ . . . . .	25
4.5.3 The Boundary Metastate $B$ . . . . .	26
4.5.4 How $B$ and $D$ States Explain Spider Motion . . . . .	31
4.6 First Passage Time . . . . .	33
4.6.1 Meaningfulness of MFPT . . . . .	33
4.6.2 MFPT of Spiders . . . . .	34
4.6.3 FPT Distribution . . . . .	39

*Contents*

<b>5</b>	<b>Single Spider in 2D</b>	<b>41</b>
5.1	Model . . . . .	41
5.1.1	Motion and Surface . . . . .	41
5.1.2	Boundaries and Starting Positions . . . . .	42
5.2	Simulation . . . . .	43
5.3	Mean Squared Displacement . . . . .	43
5.4	Spider on the Plane with a Circular Absorbing Boundary . . . . .	44
5.5	Spider on the Plane with a Circular Reflecting Boundary and an Absorbing Boundary in the Center . . . . .	47
5.5.1	2D Circle of Variable Radius With Target in the Middle and Spider Starting from the Boundary . . . . .	48
5.5.2	2D Circle of Fixed Radius With Target in the Middle and Spider Starting From Various Distances . . . . .	52
5.6	Comparison of Spider Performance in 1D and 2D Settings . . . . .	54
<b>6</b>	<b>Multiple Spiders with interactions on 2D surfaces</b>	<b>57</b>
6.1	T-junction experiment . . . . .	58
6.2	Search experiment . . . . .	61
6.2.1	Search . . . . .	62
6.2.2	Interaction . . . . .	62
6.2.3	Model and Simulation . . . . .	64
6.2.4	Results . . . . .	64
6.2.4.1	The effect of the gait . . . . .	66
6.2.4.2	The effect of the kinetics . . . . .	67
6.2.4.3	The effect of spider interactions . . . . .	68
6.2.4.4	The effect of exclusion alone . . . . .	68
6.2.5	Discussion . . . . .	69



<b>7</b>	<b>Multiple spider model with infinitely strong source</b>	<b>71</b>
7.1	The Multi-Spider Model . . . . .	72
7.1.1	Rebinding Gait . . . . .	73
7.2	Simulation Results for the Multi-Spider Model . . . . .	74
7.2.1	Experimental Setup . . . . .	75
7.2.1.1	Observed Superdiffusion of the Leading Spiders . . . . .	78
7.2.2	Comparison with Single Spiders . . . . .	82
7.3	Asymptotic Diffusion in the Multi-spider Model . . . . .	84
7.3.1	Effective size of the product sea . . . . .	85
7.3.2	Number of Released Spiders . . . . .	88
7.3.3	Spatial Distribution of Spiders . . . . .	89
7.3.3.1	Distribution of Spider Strides . . . . .	91
7.3.3.2	Density of leading spiders . . . . .	93
7.4	Importance of Multi-Pedal Gaits for Transport . . . . .	95
7.5	Analysis of Maximum Product Sea Size . . . . .	96
<b>8</b>	<b>Conclusion</b>	<b>101</b>
	<b>References</b>	<b>105</b>

# List of Figures

2.1	A molecular spider system. . . . .	3
2.2	The process of substrate cleavage by a deoxyribozyme (a molecular spider's leg). . . . .	4
4.1	Mean squared displacement, $\langle X^2 \rangle(t)$ . . . . .	16
4.2	Finite difference approximation of $\alpha(t)$ . . . . .	17
4.3	Displacement distribution for $r = 0.01$ at three characteristic times. . . . .	20
4.4	Comparison of displacement distributions for $r = 0.01$ and $r = 0.005$ at three characteristic times $\text{argmax}(\alpha(t))$ , $\hat{t}$ , and $t_{\text{max}}$ . . . . .	21
4.5	$D(t)$ as computed by Eq. 4.8. . . . .	22
4.6	Boundary and Diffusive metastates. . . . .	24
4.7	A realization $X = (P, F)$ of the spider Markov process. . . . .	25
4.8	Boundary state. . . . .	27
4.9	Initial configuration of the track. . . . .	34
4.10	Amplitude $A_1$ from Eq. 4.28 as a function of the cleavage and detachment rate $r$ . . . . .	35
4.11	Optimal $r$ values ( $r_{\text{opt}}(x)$ ) for various distances, and the corresponding $\langle \tau(x) \rangle$ . . . . .	37
4.12	Indicator $\chi$ shows how well the MFPT describes the first passage times of the individual spiders to any of the target sites. . . . .	38
4.13	Distribution of $\tau(x)$ for the $r = 1$ spider and the distance $x = 10000$ . . . . .	39
4.14	Distribution of $\tau(x)$ for the $r = 1$ and $r = 0.05$ spiders. . . . .	40
5.1	Mean squared displacement, $\langle X^2 \rangle(t)$ . . . . .	44
5.2	Finite difference approximation of $\alpha(t)$ . . . . .	45
5.3	Initial configuration of the 2D surface. All sites are initially substrates. . . . .	46

*List of Figures*

5.4	Amplitude $A_2$ from Eq. 5.1 as a function of the cleavage and detachment rate $r$ . . . . .	47
5.5	Optimal $r$ values ( $r_{\text{opt}}(x)$ ) for various radii. . . . .	48
5.6	Initial configuration of the 2D surface with target in the middle and circular reflecting boundary. . . . .	49
5.7	Amplitude $B_2$ from the Eq. 5.2 as a function of the cleavage and detachment rate $r$ . . . . .	50
5.8	Indicator $\chi$ for the 2D surface with target in the middle and circular reflecting boundary. . . . .	51
5.9	Initial configuration of the 2D surface with target in the middle and circular reflecting boundary of fixed radius. . . . .	52
5.10	Mean first passage time to a single point in the center of the circle with a fixed radius. . . . .	53
5.11	Indicator $\chi$ for the 2D surface with target in the middle and circular reflecting boundary of fixed radius. . . . .	54
5.12	Ratio between $\langle \tau(x) \rangle$ of the spiders with $r = 1.0$ and $r = r_{\text{opt}}$ . . . . .	55
5.13	Average density of products when spider with $r = 0.25$ and $k = 2$ is at distance 1000. . . . .	56
6.1	Neighborhoods of sites to which leg can move in a single step from its current position. . . . .	57
6.2	The initial state of the system and the spider gait for each simulated configuration. . . . .	65
6.3	Search time as a function of the kinetic parameter $r$ for the eight configurations simulated. . . . .	66
6.4	Enlarged plots from Figure 6.3. . . . .	70
7.1	A multiple molecular spider system with injection. . . . .	71
7.2	Estimate of $\langle X^2 \rangle(t)$ for spiders with the AK gait and the rebinding gait with $k = 2$ , $s = 2$ , and $r = 0.1$ . . . . .	74
7.3	The initial configuration used for the multi-spider model simulations. . . . .	76

List of Figures

7.4	Mean squared displacement for $R_s$ . . . . .	79
7.5	Smoothed finite difference approximation of $\alpha(t)$ for $R_s$ . . . . .	80
7.6	Effective diffusion rate ( $\tilde{D}(t)$ ) for $R_s$ . . . . .	82
7.7	Mean squared displacement for the leading spider $R_s$ in the multi-spider model versus a single AK spider. . . . .	83
7.8	Finite difference approximation of $\alpha(t)$ for the leading spider $R_s$ in the multi-spider model versus a single AK spider. . . . .	84
7.9	Comparison of the first passage time of the leading spider $R_s$ in the multi-spider model versus a single AK spider. . . . .	85
7.10	The effective size of the product sea ( $N_{\text{eff}}$ ). . . . .	86
7.11	Escape from an area of size $N_{\text{eff}}$ with one reflecting and one absorbing boundary. . . . .	87
7.12	The effective size of the product sea $\langle N_{\text{eff}}(t) \rangle$ grows with time, and hence leads to asymptotically diffusive motion for the multi-spider model. . . . .	88
7.13	Comparison of the effective size of the product sea ( $\langle N_{\text{eff}}(t) \rangle$ ) with the effective size of the product sea ( $\langle N(t) \rangle$ ). . . . .	89
7.14	Mean number of released spiders. . . . .	90
7.15	Mean spider density at $t_{\text{max}}$ , for $r = 1$ (a), and $r = 0.05$ (b). . . . .	90
7.16	Evolution of mean spider density through time, for $r = 1$ (a) and for $r = 0.05$ (b). . . . .	92
7.17	Mean density of spiders in each of the together (blue circles) and apart (red triangles) strides at $t_{\text{max}} = 10^7$ , for $r = 1$ (a) and for $r = 0.05$ (b). . . . .	92
7.18	Average density for spiders with legs together (blue circles) and apart (red triangles) plotted at several instants, for $r = 1$ (a) and for $r = 0.05$ (b). . . . .	93
7.19	Average density of the leading spider at $t_{\text{max}}$ , for $r = 1$ (a) and for $r = 0.05$ (b). . . . .	94
7.20	Average spider density of the leading spider plotted at several instants, for $r = 1$ (a) and for $r = 0.05$ (b). . . . .	95

*List of Figures*

7.21 Comparison of  $\langle X^2 \rangle(t)$  for the leading spider  $R_s$  in multi-spider simulations versus the  $k = 1$  multi-spider model (with corrected diffusion constant of  $D = 0.5$ .) . . . . . 97

7.22 Comparison of finite difference approximation of  $\alpha(t)$  for the leading spider  $R_s$  in multi-spider simulations versus the  $k = 1$  multi-spider model (with corrected diffusion constant of  $D = 0.5$ .) Horizontal lines show the threshold for ordinary diffusion at  $\alpha = 1$  and our defined threshold for superdiffusion at  $\alpha = 1.1$ . . . . . 98

7.23 Size of the effective product sea  $\langle N_{\text{eff}}(N) \rangle$  as a function of the number of visited sites  $N$ . . . . . 99

# List of Tables

4.1	Properties of the mean squared displacement and the superdiffusive regime defined by $\alpha(t) > 1.1$ . . . . .	18
4.2	The estimated diffusion coefficient $D$ for different $r$ values. . . . .	22
6.1	T-junction experiment results. . . . .	59
7.1	Parameters studied for the multi-spider model. . . . .	76
7.2	Number of KMC runs for each parameter value and minimum simulation time combination. . . . .	77
7.3	Properties of the MSD and the superdiffusive regime defined by $\alpha(t) > 1.1$ for $R_s$ . . . . .	79

# Chapter 1

## Introduction

Natural molecular motors play an important role in biological processes that are critical for the functioning of living organisms; they are the source of most forms of motion in living beings [37,46]. Molecular motors transport various cargoes inside cells, make possible for a cell itself to move, participate in the cellular division process, and when combined in large ensembles allow muscle contraction which leads to movements of macro-organisms. Molecular motors consume chemical energy and transform it into motion or mechanical work; they are more energy-efficient than the macroscopic heat engines. Beside their important role in biology, molecular motors can be used to perform various tasks in many synthetic nanoscale systems. This can lead to the creation of new materials and nano-scale devices that are applicable in medicine, electronics, energy production, optics, and many other fields. More specific examples of usage of molecular motors are [17] transport of molecules, supramolecular complexes, and nano- or microparticles and positioning them in a precise specific order; detection of chemical and biological entities (molecular motors can capture molecules and particles of interest and move them close to a sensor); assembly of nanoscale and microscale items in larger structures; self-propelled probes that can randomly sample the surface in order to produce an image of it; stretching and bending of molecular structures; microfluidic pumps that work in nanofluidic channels with submicron dimensions; molecular computation and communication; and control of optical and mechanical properties of materials. One of the widely studied molecular motors is kinesin, for example see [15,19,47]. Kinesin is a protein that moves directionally along microtubule filaments from one end to another, therefore it is also called a molecular walker. It has two identical "legs" (sometimes called heads or motor domains) that walk in a coordinated "hand-over-hand" fashion. The exact

mechanism of coordination between kinesin's legs is an active research area. Applications of kinesin and other biomolecular motors in synthetic systems have been extensively studied, see for example [2, 11, 18, 28]. In addition to that several synthetic molecular walkers have been designed [6, 22, 26, 31, 45, 48–51]. Natural motors like kinesin are very efficient, but they are protein-based and thus more complicated than synthetic ones. Synthetic motors on the other hand are simpler and can be used more easily in many applications listed above. As development of nano-scale devices and materials continues, modeling and efficient simulation are required to achieve a better quantitative understanding of their behavior, and to aid the design process.

In this dissertation, I introduce abstract theoretical models that are based on synthetic molecular motors called molecular spiders [24, 32]. Using the models I explore transport properties of a single spider moving along a one-dimensional track; first passage properties of a single spider in various one- and two-dimensional environments; the behavior of multiple spiders, in particular how efficient spiders can be for searching for target located on a bounded surface; and the release of spiders from a single localized source. I compare spiders' behavior with the behavior of simple random walkers, and show that for all studied scenarios spiders can exhibit more desirable properties than regular random walkers. Molecular transport properties at the nanoscale are important for assembly of nano-devices and similar applications where nano-particles need to be moved into specific positions. First passage and search properties are important for future applications, such as search for disease markers on cell surfaces. In this case bounded surfaces correspond to the surface of the cell, and target sites correspond to the disease markers.



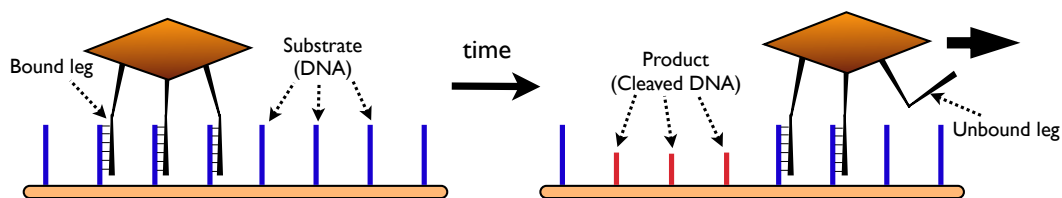
## Chapter 2

# Molecular Spiders (Physical Level)

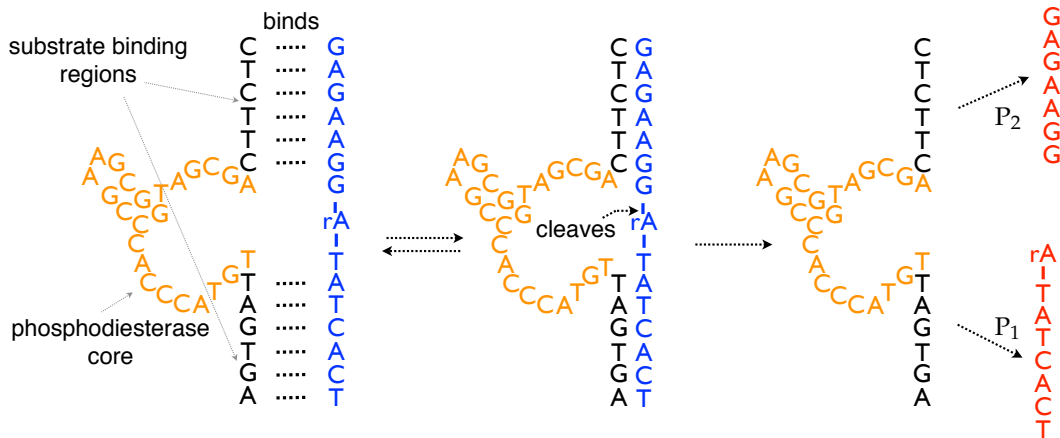
A molecular spider [24, 32] is a synthetic DNA-based molecular walker that is much simpler than the naturally occurring protein-based walkers such as kinesin. Its legs are independent and do not coordinate with each other. Unlike kinesin, spiders do not always move in one direction defined by the track. However, effective bias towards unvisited sites emerges from the properties of the legs and sites interactions when the spider is located on the boundary between visited and unvisited sites.

A molecular spider (Figure 2.1) consists of a rigid, chemically inert body (such as streptavidin—protein purified from the bacterium *Streptomyces avidinii*) to which are attached multiple, flexible enzymatic legs. The legs are deoxyribozymes—enzymatic sequences of single stranded DNA that can bind to and cleave complementary strands of a DNA substrate at the point of a designed ribonucleic base impurity. When a molecular spider is placed on a surface coated with the single-stranded DNA substrate, the legs bind to the substrate.

A bound leg can either detach from the substrate without modifying it, or it can catalyze the cleavage of the substrate, creating two product strands. The lower product remains bound to the surface, while the upper product is free to

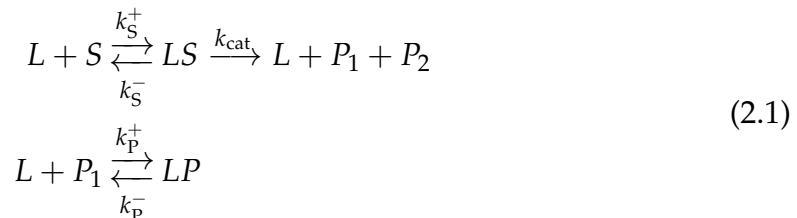


**Figure 2.1:** A molecular spider system. The spider moves over a surface of chemical sites as the legs attach and detach. A leg cleaves a substrate site, turning it into a product site when it detaches.



**Figure 2.2:** The process of substrate cleavage by a deoxyribozyme (a molecular spider's leg). First, the deoxyribozyme binds reversibly with a complementary substrate. The enzymatic core of the deoxyribozyme can then cleave the substrate. Finally, two products dissociate from the enzyme, while  $P_1$  remains bound to the surface,  $P_2$  becomes free to float away in solution. In the figure the "A" stands for Adenine and pairs with the "T", which stands for Thymine. The "C" stands for Cytosine and pairs with the "G", Guanine. The "rA" is a ribonucleic base, the point where cleavage of the substrate happens.

float away in solution. Because the lower product remains complementary to the lower part of the spider's leg, there is a residual binding to the product, although it is typically much weaker than the leg-substrate bond and thus much shorter lived. The leg kinetics are described by the five chemical reactions in Eq. 2.1 relating legs ( $L$ ), substrates ( $S$ ), and products ( $P_1$  and  $P_2$ ).



In these equations  $P_1$  refers to the lower part of the cleaved substrate that remains bound to the surface, and  $P_2$  refers to the upper part. Figure 2.2 illustrates the process of the substrate cleavage by the deoxyribozyme leg.

Two types of laboratory experiments to test molecular spiders have been performed, bulk experiments by Pei et al. [32] and single-molecule experiments by Lund et al. [24]. In bulk experiments multiple molecular spiders were released in a three-dimensional dextran-streptavidin matrix (a hydrogel made of dextran oligomers) where substrate sites were positioned at random locations. By measuring the loss of mass of the system it was shown that spiders move through the matrix and cleave substrates into products (as the upper parts of the cleaved substrates are washed away the system loses mass). In single-molecule experiments substrates were placed on a two dimensional surface (DNA origami) in the shape of a predetermined path. Then a spider was placed at the starting point at one end of the path, and thereafter the spider moved from the starting point to the other end of the path. The spider's progress was observed at several time points through atomic force microscopy (AFM).

# Chapter 3

## Computational Models

There are several categories of computational models that can be applied to the motion of molecular walkers. The most detailed models represent all elements of the system of interest as a set of interacting atoms. Those models can be simulated using the molecular dynamics method [16]. Molecular dynamics can give very detailed results, the closest possible to physical reality. It simulates physical movements and interactions of all atoms and individual molecules, but since those events happen at very small time-scales (less than a picosecond), typical simulations span only very short periods of real time and it would take enormous computational resources to simulate longer periods. Modern molecular dynamic simulations of proteins and DNA span nanoseconds to microseconds. For example, in molecular dynamic simulation of the satellite tobacco mosaic virus one million atoms were simulated for 50 ns [12], and in a study of the dynamics of the villin headpiece, 200,000 atoms were simulated for 500 ns [20]. From Ref. [32] we estimated that even the simplest molecular spider experiments we would like to simulate would involve at least 700,000 atoms and would span at least tens of minutes; also, we would like to study models where walkers move over surfaces of arbitrary sizes, and that increases the number of molecules by several orders of magnitude. Given the huge number of molecules involved in experiments with molecular spiders, and the long duration of those experiments, it is currently impossible to use molecular dynamics for our purposes.

In more abstract models the body and the legs of walkers are represented as indivisible geometric objects, the surface is represented as a set of several types of sites in a one-, two-, or three-dimensional space, and the legs randomly attach and detach from sites at constant rates. These rates correspond to actual chemical kinetic rates of the leg molecule reacting with a site molecule, and they depend

on the type of the site. Through its catalytic action, a leg can modify the type of the site it has attached to. When a leg attaches to a new site, it chooses it from all available sites according to some probability distribution  $f$ ; this distribution is influenced by the representation of the legs. If legs are considered independent from each other and the processes of attaching and detaching are of the Poisson type, then the Kinetic Monte Carlo method [8] can be used to simulate those models. Details of the representation of body, legs, and sites, distribution  $f$ , and structure of the set of sites define the abstraction level and the level of physical realism of the model. For example, in a detailed model of the real molecular spiders experiment, studied in [32], the surface can be represented as a set of sites randomly positioned in a two-dimensional space, and sites can be of two types, substrates and products. Initially all sites are substrates. The body is represented as a rigid tetrahedron with four flexible legs attached to its vertices. When a leg detaches from a site, it changes the site type into a product. In a simple and more abstract model the substrate and product sites can be placed on a regular infinite one-dimensional lattice, the body can be seen as a single point with two legs attached to it and  $f$  is a uniform distribution over the nearest neighbors of the site from which the leg just detached.

In the work of Olah [29, 30], a molecular spider model closer to physical reality was studied. The model has many parameters that correspond to the details of physical molecular spiders and surfaces. Surfaces in the model are two-dimensional and can have sites that are placed at random positions or placed on a regular grid, different shapes can be assigned to bodies of a spider, and legs can be attached to different parts of the body. Interactions of the legs with the surface sites are modeled in detail: the rates at which legs attach to the sites, detach from the sites, and modify the sites are all parameters of the model. The position of a spider body is also modeled, and its position affects the legs' movements, and movements of detached legs are taken into account as well. The goal of this more

realistic model is to understand how the motor properties of spiders depend on all those physical details. Those parameters cannot be studied in simpler and more abstract models. Particularly, one of the goals of that work is to verify that spiders can move against forces applied to them, and measure those forces. This analysis allows molecular spiders to be qualified as molecular motors.

In this work we choose to use more abstract models. We simplify many details of the model presented in [29]. We use only surfaces where sites form regular grids, the body of the spider is always represented as a single point to which all legs are attached, the position of the body is always defined as the mean of the leg positions, and movements of the body and the detached legs are not modeled. This approach has several advantages. It gives us more flexibility and makes it easier to adjust and combine the models to derive new ones. Some of those models can be used for explorations of particular applications of molecular walkers and others for enhancing certain properties of molecular walkers. It is also easier to write simulation software for more abstract models and make those simulations run faster. Furthermore, we stand a better chance of obtaining analytical results for simpler models. Finally, and probably most importantly, given some measured qualities of interest from real-life experiments and results of more realistic models, we can compare them with our results, and if they agree we can conclude that we captured the most important and relevant properties of the real-life system in our abstract representation while leaving out nonessential details.

Mathematical models of molecular spiders at the same levels of abstraction as our models have been proposed and studied. Antal and collaborators introduced the first abstract model of molecular spiders, and studied the motion of a single spider on a one-dimensional track. They investigated the movement of spiders with various numbers of legs and various gaits over products only [5] and showed that such spiders are equivalent to a regular diffusion; the diffusion con-

stants were computed for some gaits. Subsequently they introduced substrates and took into account that cleavage and detachment from substrates together take more time than the detachment from products [4], showing that this difference in residence time and the presence of multiple legs, together, bias a spider's motion towards fresh substrates when it is on a boundary between substrates and products. This important property was also observed experimentally [24]. Samii et al. investigated various gaits and numbers of legs [35,36], emphasizing the possibility of detachment from the track. Models of spiders in two dimensions have also been studied. In Ref. [3] Antal and Krapivsky evaluated the diffusion constant and the amplitude describing the asymptotic behavior of the number of visited sites for a single spider with various gaits placed on an infinite square lattice. Analytical results regarding the asymptotic behaviors (limit theorems, transience, recurrence, and rate of escape) of spiders have been derived in Refs. [7, 13]. In Refs. [14,21] the behavior of spiders in random environments was studied. Rank et al. showed that several spiders, each placed on a separate 1D track and connected to a single cargo particle move it faster and remain superdiffusive longer than a single spider on a single 1D track [34].

# Chapter 4

## Single Spider in 1D

Antal and Krapivsky introduced an abstract model of molecular spider motion [4, 5] over an infinite one-dimensional lattice (the AK model). All the models that we study are based on the AK model. The AK model simplifies the reaction rates in Eq. 2.1, setting the on-rates to be infinite, and the substrate dissociation rate to be 0, so that substrates are always cleaved to products before detachment. Essentially, the AK model defines

$$\begin{aligned}k_S^+ &= k_P^+ = \infty, \\k_{\text{cat}} &= r \leq 1, \\k_S^- &= 0, \text{ and} \\k_P^- &= 1.\end{aligned}\tag{4.1}$$

Under these conditions the spider motion can be studied as a function of the single rate parameter  $r \leq 1$ , which represents the ratio between the substrate cleavage and product detachment rates. Hence, a *residence-time bias* is established, where legs detach faster from previously visited sites than from unvisited sites. Antal and Krapivsky showed that the asymptotic behavior of this spider model is diffusive for all values of  $r$ . Thus, in the long time limit the AK spiders cannot be used as a faster-than-diffusion transport mechanism. However, our numerical simulations of the AK model reveal that when there is a residence-time difference between previously visited and unvisited sites, the spiders can move superdiffusively for time periods that span many orders of magnitude. However, our numerical simulations of the AK model reveal that when there is a residence-time difference between previously visited and unvisited sites, the spiders can move superdiffusively for a long period of time.

The results of this section are published in Refs. [38,39].



#### 4.1. The Antal-Krapivsky model

### 4.1 The Antal-Krapivsky model

The AK model effectively but not explicitly describes spider movement as a continuous-time Markov process. We reformulated the model to emphasize the states and transitions, and the Markovian nature of the transitions when the state is defined to include both the state of the spider and the state of the surface sites.

We consider a system with a single  $k$ -legged spider. The legs step over sites on a regular lattice ( $\mathbb{Z}$ ). The states in the process are the combined state of the lattice sites and the state of the spider. Each lattice site is a substrate (uncleaved) or a product (cleaved). Initially all sites are substrates, so the state of the surface can be described by the set  $P \subset \mathbb{Z}$  of sites that have been cleaved. The state of the spider is described by the set  $F \subset \mathbb{Z}$  of foot locations—lattice sites with a leg attached. Together  $P$  and  $F$  completely define the state of the spider system, i.e., the state of the Markov process is  $X = (P, F)$ .

We call  $F$  a *configuration* of the legs. The *gait* of a spider is defined by what configurations and what transitions between configurations are allowed in the model. There are considerable possibilities for variations on the spider gait. Antal and Krapivsky describe the gait of a spider with the kinetics of Eq. 4.1. With  $k_S^+ = k_P^+ = \infty$ , a leg immediately reattaches after it detaches. Thus in any state  $X = (P, F)$  of the process, all  $k$  legs are attached. Together with the restriction that at most one leg may be attached to a site, this implies that

$$|F| = k. \tag{4.2}$$

Additionally, the legs are constrained by their attachment to a common body. If the spider has a point body with flexible, string-like legs of length  $s/2$ , then any two feet can be separated by at most distance  $s$ , thus

$$\max(F) - \min(F) \leq s. \tag{4.3}$$

This restriction is that of the “global spiders” of Antal and Krapivsky [5].

The transitions in the process correspond to individual legs unbinding and rebinding. When a spider is in configuration  $F$ , any foot  $i \in F$  can unbind and move to a nearest-neighbor site  $j \in \{i + 1, i - 1\}$  to form a new configuration  $F' = (F \setminus \{i\}) \cup \{j\}$  provided the new configuration does not violate one of the constraints of Eqs. 4.2 and 4.3. A transition  $i \rightarrow j$  is called *feasible* if it meets these constraints. The feasible transitions determine the gait of the spider. The nearest-neighbor hopping combined with the mutual exclusion of legs leads to a shuffling gait, wherein legs can slide left or right if there is a free site, but legs can never move over each other, and a leg with both neighboring sites occupied cannot move at all. If the legs of such a spider were distinguishable, they would always remain in the same left-to-right ordering.

The rate at which feasible transitions take place depends on the state of the site  $i$ . If  $i$  is a product the transition rate is 1, but if  $i$  is a substrate the transition occurs at a slower rate  $r < 1$ . This is meant to model the realistically slower dissociation rates from substrates. The effect of substrate cleavage is also captured in the transition rules. If for state  $X = (P, F)$  where  $i \in F \setminus P$ , the process makes the feasible transition  $i \rightarrow j$ , then the leg will cleave site  $i$  before leaving, and the new state will have  $P' = P \cup \{i\}$ .

In order to compactly represent the state of a spider process, Antal and Krapivsky introduced a graphical notation. The symbol  $\bullet$  represents a site occupied by a leg and  $\circ$  represents an unoccupied site. All sites initially have a hat  $\hat{\phantom{x}}$  indicating they are uncleaved (substrate) sites. A site is cleaved into a product when a leg detaches from it for the first time, denoted by removing the hat. For example, state  $(\{i - 2, i - 1, i, i + 1\}, \{i, i + 2\})$  can be illustrated thus:

$$\cdots \hat{i-4} \hat{i-3} \circ_{i-2} \circ_{i-1} \bullet_i \circ_{i+1} \hat{i+2} \hat{i+3} \cdots$$

#### 4.1. The Antal-Krapivsky model

Since the transition rates are translationally invariant on the lattice, we can generally omit the indexes on the sites.

Antal and Krapivsky have analytically studied the expectation of the random variable  $T(n)$ , which for the bipedal spider with  $s = 2$  is defined as the time when a leg steps onto an uncleaved site after  $n + 2$  sites have already been cleaved. When this event occurs the spider is always in the following position (or its reflection),

$$\dots \hat{\circ} \hat{\circ} \underbrace{\circ \circ \dots \circ \bullet \circ \hat{\circ}}_{n+2} \dots .$$

One can alternatively think of  $T(n)$  as the time at which the spider first visits  $n$  distinct sites not counting the three sites its legs span at that time. Since a spider always cleaves a substrate site it visits,  $T(n)$  is equivalent to the time for  $n + 2$  products to be formed. For the case  $s = 2, k = 2$ , when  $r = 1$ , it was found that

$$\langle T(n) \rangle = n^2 + n, \quad (4.4)$$

but more generally, when  $0 < r \leq 1$ , the leading coefficient is reduced to

$$\langle T(n) \rangle = \frac{3}{2} \frac{1+r}{2+r} n^2 + \frac{1}{r} n. \quad (4.5)$$

This implies that a large residence-time bias between unvisited and visited sites, corresponding to small values of  $r$ , leads to a faster mean time to visit  $n$  sites for large enough  $n$ . Antal and Krapivsky also showed that one-legged spiders do not exhibit this behavior. Thus, it is the combination of having more than one leg and the ability to irreversibly change the sites and hence rates that allows the

spider to move faster. While for small  $r$  values  $T(n)$  is smaller, it is still  $\mathcal{O}(n^2)$ , and hence not asymptotically faster than an ordinary diffusive process.

Antal and Krapivsky note that with  $r < 1$  there is an effective bias in the spider's motion when it has one leg on a substrate at the boundary between cleaved and uncleaved sites. In such a situation the spider moves with probability  $p_+$  in the direction away from previously visited sites and with probability  $p_-$  towards previously visited sites. Antal and Krapivsky calculated that the strongest bias is in the  $r \rightarrow 0$  limit when  $p_- = 3/8$  and  $p_+ = 5/8$ . In the next sections we show via simulation that spiders experience an initial period of superdiffusive behavior when  $r < 1$ , and in Section 4.5 we show how this behavior is caused by the effective bias, yet asymptotically dominated by diffusive motion over previously visited sites in the limit as  $t \rightarrow \infty$ .

## 4.2 Simulation Description

We use the Kinetic Monte Carlo method [8] to numerically sample traces of the spider Markov process. In our simulations, a single two-legged ( $k = 2$ ) spider with maximum leg separation constraint  $s = 2$  is placed on a one-dimensional infinite lattice of substrates and allowed to move according to the model. We vary the rate  $r$  to see how it influences the motion. The case  $r = 1$  corresponds to ordinary diffusion because there is no effective difference between substrates and products.

For Sections 4.3 and 4.4 for each value of  $r \in \{1, 0.5, 0.1, 0.05, 0.01, 0.005\}$  we simulate 5000 traces of the Markov process. We record samples of several random variables (e.g., mean squared displacement and first passage time) that are functions of time, distance, or the number of sites visited. To ensure that each simulation trace provides a sample for each measured value of the random variables,

### 4.3. Instantaneous Superdiffusion of Spiders

we run each simulation until all of the following conditions are met: (1) the time is greater than  $t_{\max} = 10^8$  time units; (2) the spider has visited at least  $c_{\max} = 10^4$  sites; and (3) the spider has moved at least a distance  $d_{\max} = 10^4$  sites away from the origin. We sample so that each of the plots with time on the  $x$ -axis that follow is obtained from 6000 measurement points equispaced for the independent variable axis of the plot (linear or logarithmic). For plots that have distance and number of cleaved sites on the  $x$ -axis we use  $10^4$  measurement points each.

For Section 4.6 we simulated 210 different  $r$  rates for up to a distance of 10000 using 21504 traces of the Markov process.

## 4.3 Instantaneous Superdiffusion of Spiders

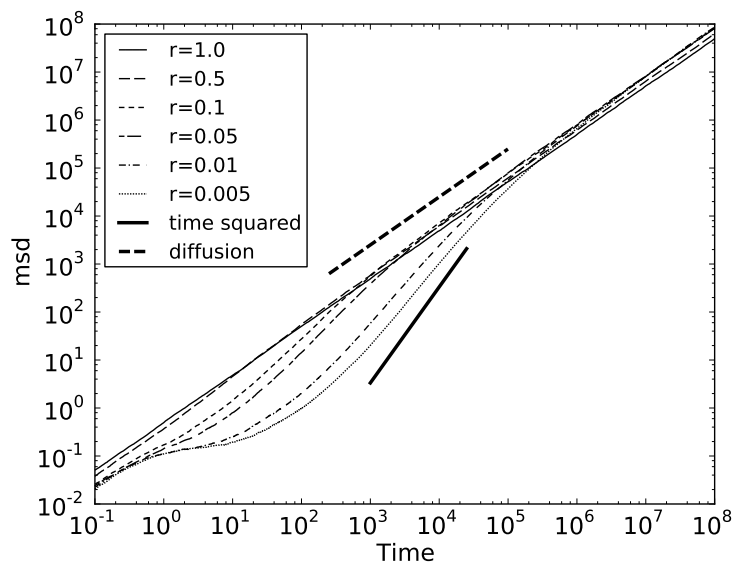
Superdiffusive motion can be quantified by analyzing the mean square displacement of a spider as a function of time. For diffusion in 1D space with diffusion constant  $D$ , the mean squared displacement is given by Eq. 4.6.

$$\langle X^2 \rangle(t) = 2Dt^\alpha \begin{cases} \alpha = 0 & \text{stationary} \\ 0 < \alpha < 1 & \text{subdiffusive} \\ \alpha = 1 & \text{diffusive} \\ 1 < \alpha < 2 & \text{superdiffusive} \\ \alpha = 2 & \text{ballistic or linear} \end{cases} \quad (4.6)$$

We shall say that the spider is moving *instantaneously superdiffusively* at a given time  $t$  if

$$\alpha(t) = \frac{d(\log_{10} \langle X^2 \rangle(t))}{d(\log_{10} t)} > 1. \quad (4.7)$$

This definition is similar to that used by Lacasta et al. [23] to describe transient superdiffusive behavior.

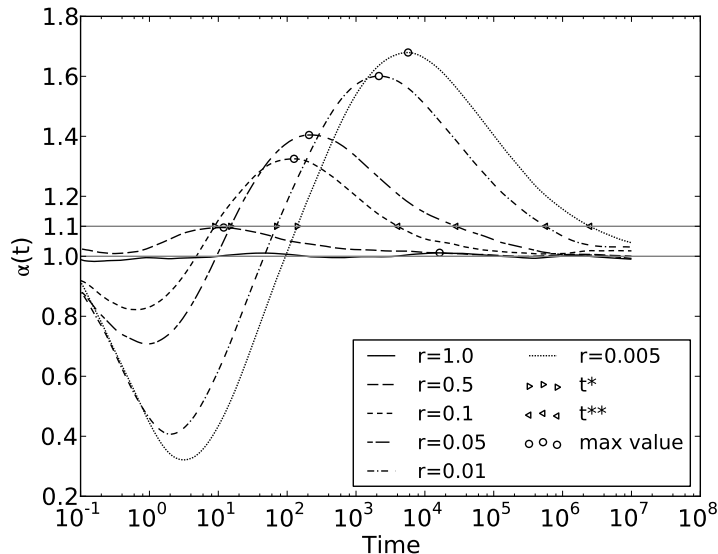


**Figure 4.1:** Mean squared displacement,  $\langle X^2 \rangle(t)$ .

Figure 4.1 shows  $\langle X^2 \rangle(t)$  for different  $r$  values. In this log-log plot, straight lines correspond to power laws, that is, to Eq. 4.6, and the parameter  $\alpha$  is given by the slope. A reference line for diffusion is shown to illustrate that the  $r = 1$  spider is ordinary diffusive, and all spiders eventually become ordinary diffusive asymptotically. A reference line proportional to  $t^2$  is also shown for comparison to ballistic motion, which shows that spiders with small  $r$  values experience significant periods of superdiffusive behavior.

We use finite difference methods to estimate  $\alpha(t)$  (Eq. 4.7). Figure 4.2 shows the result of using the Savitzky-Golay smoothing filter [33] on these estimates. The spiders with  $r = 1$  indeed move diffusively, with  $\alpha(t) \approx 1$  for all times. However, the spiders with  $r < 1$  show a pattern of three distinct diffusion regimes at different time scales. The first of these is an *initial regime* when the times are

### 4.3. Instantaneous Superdiffusion of Spiders



**Figure 4.2:** Finite difference approximation of  $\alpha(t)$ .

small enough that the mean number of cleavages is less than 1 and the spiders show significantly *subdiffusive* behavior. This can be explained by considering that the spider starts in the configuration defined by Antal and Krapivsky,

$$\cdots \hat{\circ} \bullet \circ \hat{\bullet} \hat{\circ} \cdots$$

From this state either the right leg moves at rate  $r$  or the left leg moves at rate 1, but if  $r \ll 1$  the mean time to move the right leg is large. Until the right leg has moved, the left leg is restricted to hopping between sites  $-1$  and  $0$ . Thus, the parameter  $r$  determines the time scale of this initial period as  $0 \leq t \leq 1/r$ .

When  $t > 1/r$ , the average number of cleavages is greater than one. After this time, the spider has taken several steps, and has cleaved out a small region (sea) of products which defines a boundary between regions of visited and unvisited sites. As Antal and Krapivsky noted, there is an effective outward bias for bipedal

$r$	$\max \alpha(t)$	$\operatorname{argmax} \alpha(t)$	$t^*$	$t^{**}$	$\hat{t}$
0.5	1.10	$1.20 \times 10^1$	-	-	$2.25 \times 10^1$
0.1	1.32	$1.26 \times 10^2$	$8.83 \times 10^1$	$4.06 \times 10^3$	$5.89 \times 10^2$
0.05	1.40	$2.08 \times 10^2$	$1.51 \times 10^2$	$2.83 \times 10^4$	$2.67 \times 10^3$
0.01	1.60	$2.15 \times 10^3$	$6.89 \times 10^2$	$5.68 \times 10^5$	$5.59 \times 10^4$
0.005	1.68	$5.70 \times 10^3$	$1.39 \times 10^3$	$2.49 \times 10^6$	$2.44 \times 10^5$

**Table 4.1:** Properties of the mean squared displacement and the superdiffusive regime defined by  $\alpha(t) > 1.1$ .

spiders near this boundary when  $r < 1$ . Figure 4.2 shows that spiders with small  $r$  values move superdiffusively in the period of time after the initial regime. Hence, we call this the *superdiffusive regime*. We quantify this regime as the period of time when  $\alpha(t) > 1.1$ . The choice of 1.1 is arbitrary, but is a sensible threshold that corresponds to a spider moving significantly superdiffusively. Using this threshold, we define  $t^*$  and  $t^{**}$  as the time when the spider enters and exits the regime of superdiffusive motion. Table 4.1 summarizes these values. We also compute the maximum value of  $\alpha(t)$ , and the time  $t$  at which the maximum is reached. These values show an increasingly significant superdiffusive regime for smaller values of  $r$ .

As predicted by Eq. 4.5, the spiders must eventually move diffusively. This leads to the third and final *diffusive regime*, in which all spiders asymptotically move with  $\langle X^2 \rangle(t) \propto t$ .

To quantify when a particular  $r$ -value spider is faster than the  $r = 1$  spider (in terms of  $\langle X^2 \rangle(t)$ ), we define  $\hat{t}(r)$  as the first time when  $\langle X^2 \rangle_r(t) > \langle X^2 \rangle_1(t)$ , and summarize the values in Table 4.1. Between the times  $\hat{t}$  and  $t^{**}$  the spider is farther on average than a diffusive spider and it is still moving faster by more than a constant factor. Thus during this interval, the spider is more efficient in every respect than an ordinary diffusive spider.

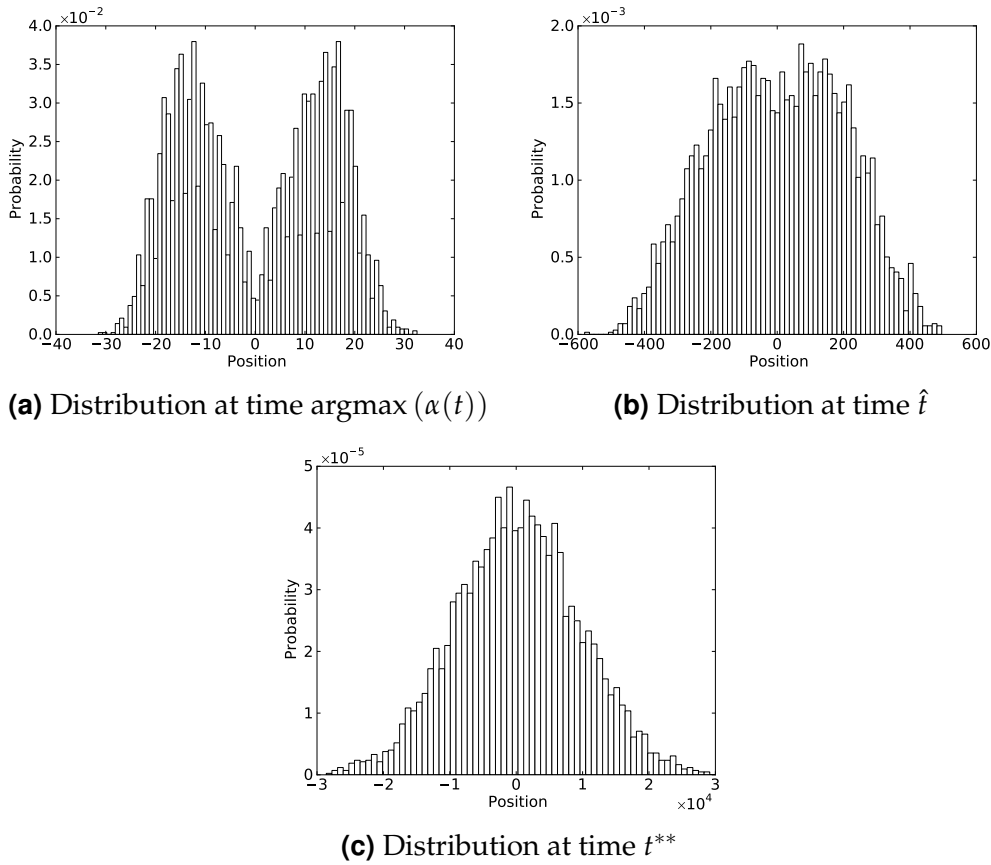


## 4.4 Asymptotic Behavior and Distributions

To describe a process as unbiased ordinary diffusive, one must show not just that the mean squared displacement increases linearly with time, but more specifically that the distribution of the displacement is Gaussian. Initially this is not true for spiders with  $r < 1$ . The bias at the boundary tends to keep spiders towards the outside of the region of cleaved products, leading to a bimodal distribution peaked around the average locations of the boundaries at that time. However, as time increases and the size of the sea of products grows, spiders spend increasingly more time moving in an unbiased, diffusive manner over these sites. This eventually leads to a more Gaussian-shaped distribution. Fig. 4.3 shows the displacement distributions for the  $r = 0.01$  spider at three times: at  $\text{argmax}(\alpha(t))$ , when the spider is moving most superdiffusively; at  $\hat{t}$ , when the spider mean squared displacement overtakes the  $r = 1$  spider; and at  $t_{\max}$ , when the spider is in the diffusive regime. Fig. 4.4 shows a comparison of the distributions at the same three times for the  $r = 0.01$  and the  $r = 0.005$  spiders. The spider with the smaller  $r$  value has a sharper peak near the boundary at time  $t = \text{argmax}(\alpha(t))$ , corresponding to the slower release from substrates.

At  $t_{\max} = 10^8$ , most of the spiders have  $\alpha(t) \approx 1$ . Table 4.2 shows the results of using the Shapiro-Wilk normality test [44] to test the hypothesis that the displacement distribution is Gaussian at time  $t_{\max}$ . The  $p$ -values are significant enough to support this hypothesis. However, note that the  $p$ -values are increasingly small for small  $r$  values. This likely indicates that the spider processes for small  $r$  values are still slowly moving towards ordinary diffusion, and hence are not quite normal, especially near the ends of the distribution due to the bias at the boundary.

Nevertheless, all the spiders are sufficiently close to normally distributed at



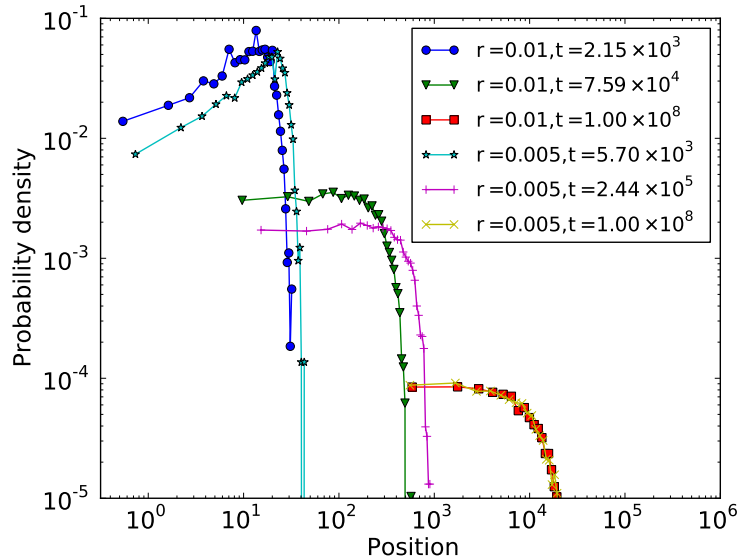
**Figure 4.3:** Displacement distribution for  $r = 0.01$  at three characteristic times.

time  $t_{\max}$  so that we can use

$$D(t) = \frac{\langle X^2 \rangle(t)}{2t} \quad (4.8)$$

as an approximation to the effective diffusion rate of the spiders. The value  $D(t_{\max})$  should be thought of as the diffusion constant an ordinary diffusive process would need in order to have the same mean squared displacement at time  $t_{\max}$  as the given spider process. In this way it can make sense to compute  $D(t)$  even at times when the spider processes are significantly subdiffusive or superdiffusive. For these times we interpret the  $D(t)$  value as a measure relating the msd of the spider process to that of an ordinary diffusive process with diffusion constant  $D$ . In Fig. 4.5, we use Eq. 4.8 to compute  $D(t)$  for all times. Finally, in Ta-

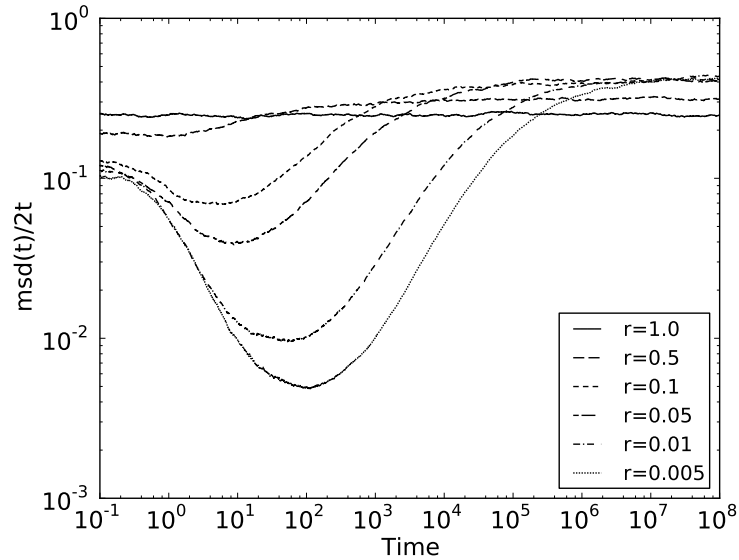
#### 4.4. Asymptotic Behavior and Distributions



**Figure 4.4:** Comparison of displacement distributions for  $r = 0.01$  and  $r = 0.005$  at three characteristic times  $\text{argmax}(\alpha(t))$ ,  $\hat{t}$ , and  $t_{\max}$ .

ble 4.2 we estimate  $D(t_{\max})$  with 95% confidence bounds for each value of  $r$ . The analytical value for  $r = 1$  is  $1/4$ , which is within the error bounds of our estimate. We should expect these values to be monotonically increasing with decreasing  $r$ , and this is true (within confidence intervals). However, the  $D(t_{\max})$  value for the  $r = 0.005$  spiders is not representative of their true long-term behavior, as these spiders still have not moved for long enough for their  $\langle T(n) \rangle$  value to surpass that of the  $r = 0.01$  spiders. The  $r = 0.005$  spiders are still moving superdiffusively enough at  $t_{\max}$  that the  $D(t_{\max})$  value is substantially smaller than its asymptotic value. The same would be true of the  $D(t_{\max})$  value for any spider with an even smaller  $r$ .

Of practical interest, from these diffusion rates we estimate that at time  $t_{\max}$  a spider with  $r = 0.005$  will be approximately 31% farther from the origin on average than an ordinary-diffusive spider with  $r = 1$  (or equivalently an ordinary random walker with  $D = 0.25$ ). Thus, given enough time, a spider with *slower*



**Figure 4.5:**  $D(t)$  as computed by Eq. 4.8.

enzymatic rate  $k_{\text{cat}}$  can transport objects significantly *farther*.

$r$	$D(t_{\text{max}} = 10^8)$	Shapiro-Wilk p-value (at $t_{\text{max}} = 10^8$ )
1.0	$0.247 \pm 0.010$	0.747
0.5	$0.313 \pm 0.012$	0.518
0.1	$0.413 \pm 0.016$	0.620
0.05	$0.407 \pm 0.016$	0.677
0.01	$0.435 \pm 0.017$	0.250
0.005	$0.417 \pm 0.016$	0.206

**Table 4.2:** The estimated diffusion coefficient  $D$  for different  $r$  values with 95% confidence bounds, and the  $p$ -value for Shapiro-Wilk normality test at time  $10^8$ , showing the distributions are reasonably normal at this time.

#### 4.5. Mechanism of Transient Spider Super-diffusion

### 4.5 Mechanism of Transient Spider Super-diffusion

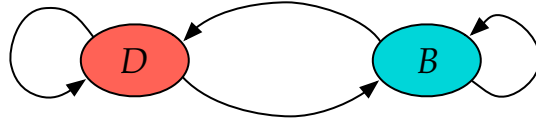
Our simulation results have shown that the spiders of the AK model for  $s = 2, k = 2$  move superdiffusively over a significant distance and time, and that this effect increases with decreasing values of  $r$ . Eventually, however, the motion decays to an ordinary diffusive walk.

In this section, we argue that there is a general principle underlying spider motion that can be understood by viewing spiders as existing in one of two metastates, a diffusive metastate  $D$  wherein a spider moves over visited sites, or a boundary metastate  $B$  wherein a spider moves ballistically away from the origin when it is on the boundary of uncleaved sites. Because the duration of a  $B$  period remains independent of the past, but the duration of a  $D$  period grows with time, eventually the spider will approach an ordinary diffusive motion. The  $s = 2, k = 2, r < 1$  AK spider model is the simplest model exhibiting the boundary/diffusive state decomposition and the resulting superdiffusive behavior.

#### 4.5.1 The Boundary and Diffusive Metastates

As explained in Section 4.1, in the AK model legs only hop to nearest-neighbor sites and cannot hop over one another. This leads to a shuffling gait. If the legs were distinguishable their ordering would not change. Thus for concreteness we can refer to a leftmost and a rightmost leg. Because the legs only move to nearest-neighbor sites, they cannot jump over any site; and because a leg always cleaves a substrate into a product, a spider cannot leave any substrates behind. Thus a spider with this shuffling gait will cleave out an interval of products so that for state  $X = (P, F)$ , we find

$$P = \{b_L(t) + 1, \dots, b_R(t) - 1\}, \quad (4.9)$$



**Figure 4.6:** A process moves between two metastates, a  $B$  state in which the spider is on the boundary between substrates and products, and a  $D$  state in which the spider is diffusing in the product sea.

where

$$b_L = \min(P) - 1, \quad \text{and} \quad b_R = \max(P) + 1. \quad (4.10)$$

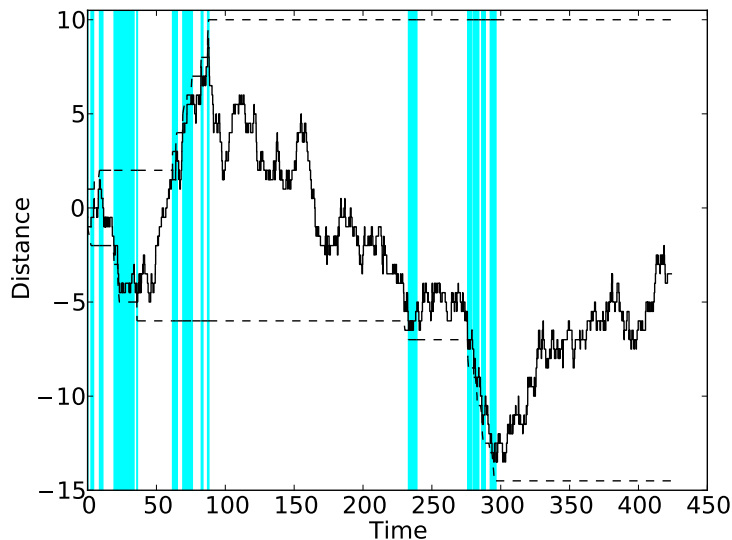
We call  $b_L$  and  $b_R$  the left and right boundaries, as they define the interval of products (Eq. 4.9) we call the *product sea*. This product sea includes the origin and contains no substrates within it. Thus, a spider in the product sea has all its legs on products so it must hop without bias at rate 1, and its motion is diffusive. Any state in which all the spider's legs are contained within the product sea belongs to the diffusive or  $D$  metastate. Formally, for state  $X = (P, F)$ ,  $X \in D$  if and only if  $F \subseteq P$ .

The only other possible state is for the spider to have a single leg on a substrate at one of the boundaries. This must be either the leftmost leg on  $b_L$  or the rightmost leg on  $b_R$ . No other situation is possible because of the shuffling gait of the legs enforced by nearest-neighbor hopping. In either of these cases, we say that the spider is in the boundary or  $B$  metastate, so that for  $X = (P, F)$ , we have  $X \in B$  if and only if  $b_L \in F$  or  $b_R \in F$ .

Together  $B$  and  $D$  form a partition of the state space for the spider Markov process. Thus, we can view a spider process as a (non-Markovian) stochastic process that moves between a  $B$  state and a  $D$  state (Fig. 4.6).

For a particular realization of the spider Markov process, we define a  $B$  *period* as an interval of time during which the spider is in the  $B$  metastate and a  $D$  *period*

#### 4.5. Mechanism of Transient Spider Super-diffusion



**Figure 4.7:** A realization  $X = (P, F)$  of the spider Markov process. We plot the mean body position as the mean of the feet locations,  $\sum_{i \in F} i / |F|$ . At each time the spider is in a  $B$  (shaded area) or  $D$  (white area) metastate. The top and bottom dashed lines show  $b_R$  and  $b_L$  respectively. Thus, at any time  $t$  the sites below the bottom dashed line and above the top dashed line have not yet been visited.

as an interval of time spent in the  $D$  state. Fig. 4.7 shows a particular simulated trace of the Markov process and the partitioning of time into  $B$  and  $D$  periods.

##### 4.5.2 The Diffusive Metastate $D$

The  $D$  metastate is the simpler state, as it corresponds to an unbiased diffusion over the product sea, and no sites can be cleaved while in the  $D$  state. Let  $\mathbf{E}[\tau_D(t)]$  be the mean duration of a  $D$  period that begins at time  $t$ . This quantity depends only on the size of the product sea  $P$ .

To derive  $\mathbf{E}[\tau_D(t)]$  we follow the analysis of Antal and Krapivsky [4], and consider that the spider always begins a  $D$  period in the state



From here it executes an unbiased random walk on the product sea in which each step corresponds to a  $\pm 1/2$  step in the mean of the leg locations. Thus, the process of exiting the  $D$  state is equivalent to that of a normal random walker exiting an interval of size  $M = 2N + 4$ , starting at position  $x = 2N + 1$ . For general  $M$  and  $x$  this time is

$$T(M, x) = \frac{x(M - x)}{2},$$

whence we obtain

$$T(2N + 4, 2N + 1) = \frac{3(2N + 1)}{2}. \quad (4.11)$$

Antal and Krapivsky [4] calculated that asymptotically

$$\mathbf{E}[N(t)] = \sqrt{t} \frac{\Gamma\left(\frac{3+3r}{4+2r}\right)}{\Gamma\left(\frac{5+4r}{4+2r}\right)}. \quad (4.12)$$

Now, combining Eqs. 4.11 and 4.12, allows us to show that asymptotically

$$\mathbf{E}[\tau_D(t)] = \frac{3}{2} \left( 2 \sqrt{t} \frac{\Gamma\left(\frac{3+3r}{4+2r}\right)}{\Gamma\left(\frac{5+4r}{4+2r}\right)} + 1 \right). \quad (4.13)$$

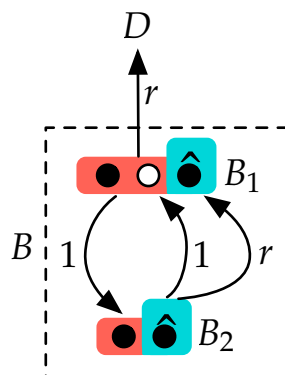
Notice that  $\mathbf{E}[\tau_D]$  grows with time, hence the  $D$  state is non-Markovian.

### 4.5.3 The Boundary Metastate $B$

In contrast to the  $D$  state, the  $B$  state is Markovian. For a  $B$  period we can compute the number of steps a spider takes ( $S_B$ ), the length of the  $B$  period ( $\tau_B$ ), and the number of cleavages the spider performs ( $C_B$ ). We find that each of these random



#### 4.5. Mechanism of Transient Spider Super-diffusion



**Figure 4.8:** To compute  $S_B$  and  $\tau_B$  we consider in detail the two states contained within the  $B$  state. A spider always enters the  $B$  state by moving to state  $B_1$ . It can leave the  $B$  state by moving its right leg, cleaving the site, and moving to the  $D$  state as there are no longer any legs on the boundary. If a spider in state  $B_1$  moves its left leg instead, it goes to state  $B_2$ . From  $B_2$  the spider can move either leg. It moves its right leg at rate  $r$  which cleaves the current boundary site, moving the boundary to the right and the spider back to state  $B_1$ . Also from state  $B_2$  the spider can move its left leg which moves the spider back to state  $B_1$  without changing the boundary.

variables is independent of time, independent of the size of product sea, and independent of the absolute position of the boundary. These conclusions show that the spider walking in a  $B$  period is essentially Markovian—independent of the past history of the spider, and translationally invariant. This means that as soon as the spider cleaves the boundary site and moves onto the new boundary site the process is renewed.

When  $s = 2$  and  $k = 2$ , legs can either be on adjacent sites or separated by a single unoccupied site. By definition, in the  $B$  state one of the legs is always on a substrate at the boundary. Without loss of generality, assume the spider is on the right boundary, so that the right leg is at  $b_R$ . Then the  $B$  metastate can be partitioned into two smaller metastates (Fig. 4.8), a state  $B_1$  in which the legs are separated by one site, and a state  $B_2$  in which the legs are adjacent. In either  $B_1$  or  $B_2$  each leg has exactly one transition it can make, and since one leg is on a product and one leg is on a substrate the total rate of transition out of either  $B_1$  or

$B_2$  is

$$R = 1 + r.$$

A spider can only leave the  $B$  metastate when it is in state  $B_1$  and the next action is to move the rightmost leg off the substrate. In the state  $B_2$ , either leg moving results in the spider moving to state  $B_1$ .

To derive the distribution for  $S_B$ , the number of steps the spider makes in the  $B$  state, we note that each  $B$  period begins with the spider moving into state  $B_1$ . From  $B_1$  the spider has a  $r/R$  probability of moving off the boundary into the  $D$  metastate. But with the remaining  $1/R$  probability, the spider moves to state  $B_2$  and subsequently back to  $B_1$ . Thus a  $B$  period can be thought of as  $Y \geq 0$  loops  $B_1 \rightarrow B_2 \rightarrow B_1$ , ending at state  $B_1$ , and a final move to state  $D$ , meaning that the number of steps taken in the  $B$  state will be

$$S_B = 2Y + 1. \quad (4.14)$$

Each time the spider is at  $B_1$  it has an independent  $1/R$  probability of making a loop through  $B_2$ , thus  $Y$  is geometrically distributed with mean  $1/R$ ,

$$\mathbf{P}[Y = y] = \left(\frac{1}{R}\right)^y \left(\frac{r}{R}\right) = \frac{r}{R^{y+1}}. \quad (4.15)$$

Combining Eqs. 4.14 and 4.15 gives

$$\mathbf{P}[S_b = s] = \begin{cases} 0, & s \text{ even} \\ \frac{r}{R^{\frac{s+1}{2}}}, & s \text{ odd} \end{cases} \quad (4.16)$$

Each of these  $S_B$  steps occurs with total rate  $R$ , hence the time for the  $i$ -th step is exponentially distributed with scale parameter  $1/R$ . Therefore, the duration of a  $B$  period, conditioned on the event that  $S_B = s$  steps are made in the period is gamma-distributed with probability distribution function

$$f_{\tau_B|S_B}(t|s) = \text{Gamma}(s, 1/R). \quad (4.17)$$

#### 4.5. Mechanism of Transient Spider Super-diffusion

Using the distribution of  $Y$ , we find the marginal probability distribution function as

$$\begin{aligned}
 f_{\tau_B}(t) &= \sum_{y=0}^{\infty} \text{Gamma}(2y+1, 1/R) (\mathbf{P}[Y=y]) \\
 &= \sum_{y=0}^{\infty} \left( \frac{t^{2y} e^{-Rt}}{R^{-(2y+1)} \Gamma(2y+1)} \right) \left( \frac{r}{R^{y+1}} \right) \\
 &= e^{-Rt} \sum_{y=0}^{\infty} \frac{t^{2y}}{(2y)!} \frac{r}{R^{-y}} \\
 &= r e^{-Rt} \sum_{y=0}^{\infty} \frac{(t \sqrt{R})^{2y}}{(2y)!} \\
 &= r e^{-Rt} \cosh(t \sqrt{R}).
 \end{aligned} \tag{4.18}$$

To compute  $C_B$ , the number of cleavages in a  $B$  period, we must pay closer attention to the transitions out of state  $B_2$  in Fig. 4.8. In state  $B_2$  either leg can move. If the leftmost leg moves, it is constrained to move left and the spider moves back to  $B_1$  without cleaving a site. If the rightmost leg moves, it cleaves the substrate, moves the boundary ( $b_R \rightarrow b_R + 1$ ), and the leg is constrained to move right onto the new boundary, leaving the spider in state  $B_1$  again but at a new absolute position.

A spider always enters a  $B$  period in state  $B_1$ . From this state there are two ways to cleave exactly one site. Either (1) the spider follows a sequence of non-cleaving moves ending in state  $B_2$  and then moves its right leg, cleaving that site and moving back to state  $B_1$ ; or (2) the spider follows a sequence of non-cleaving moves ending in state  $B_1$  and then moves its right leg, cleaving that site and exiting to the  $D$  metastate. Let  $Z_1$  and  $Z_2$  be the events (1) and (2) respectively. Then, for  $c \geq 1$  we can compute the distribution of  $C_B$  as

$$\mathbf{P}[C_B = c] = (\mathbf{P}[Z_1])^{c-1} \mathbf{P}[Z_2]. \quad (4.19)$$

Note that  $\mathbf{P}[C_B = 0] = 0$  since at least one substrate will be cleaved when the spider leaves the boundary.

To compute  $\mathbf{P}[Z_1]$  we must account for all the ways a spider can cleave exactly one substrate and return to  $B_1$ . The spider must first move to  $B_2$  with probability  $1/R$ , then it can move  $B_2 \rightarrow B_1 \rightarrow B_2$  an arbitrary number of times *without cleaving* by moving the left leg in state  $B_2$  with probability  $1/R$  and subsequently moving its left leg again when in state  $B_1$  with probability  $1/R$ . Finally, the spider will move its right leg with probability  $r/R$ , cleaving a site and returning to  $B_1$ . Thus,

$$\begin{aligned} \mathbf{P}[Z_1] &= \frac{1}{R} \times \sum_{i=0}^{\infty} \left(\frac{1}{R^2}\right)^i \times \frac{r}{R} \\ &= \frac{r}{R^2} \frac{R^2}{R^2 - 1} \\ &= \frac{r}{R^2 - 1} = \frac{1}{r + 2}. \end{aligned} \quad (4.20)$$

For event  $Z_2$ , the spider can leave the boundary by first moving  $B_1 \rightarrow B_2 \rightarrow B_1$  an arbitrary number of times *without cleaving* by moving the left leg with probability  $1/R$  when in state  $B_1$  and again moving the left leg with probability  $1/R$  when in state  $B_2$ , and finally in state  $B_1$  moving the right leg with probability  $r/R$  to exit to state  $D$ . Thus,

$$\begin{aligned} \mathbf{P}[Z_2] &= \sum_{i=0}^{\infty} \left(\frac{1}{R^2}\right)^i \times \frac{r}{R} \\ &= \frac{rR}{R^2 - 1} = \frac{r + 1}{r + 2}. \end{aligned} \quad (4.21)$$

#### 4.5. Mechanism of Transient Spider Super-diffusion

Therefore,

$$\begin{aligned} \mathbf{P}[C_B = c] &= (\mathbf{P}[Z_1])^{c-1} \mathbf{P}[Z_2] \\ &= \left(\frac{1}{r+2}\right)^{c-1} \left(\frac{r+1}{r+2}\right). \end{aligned} \quad (4.22)$$

Hence,  $C_B$  is geometrically distributed with mean

$$\mathbf{E}[C_B] = \frac{r+2}{r+1}. \quad (4.23)$$

Together these random variables characterize most of the important characteristics of the  $B$  periods. Each of  $\tau_B$ ,  $S_B$ , and  $C_B$  is independent of the state of the process when it enters the  $B$  period. For this reason we say that the  $B$  state is Markovian with respect to the  $B/D$  state decomposition of Fig. 4.6.

#### 4.5.4 How $B$ and $D$ States Explain Spider Motion

The random variable  $C_B$  is important because sites can only be cleaved during a  $B$  period. Also, because  $C_B$  is independent of the state of the system when it enters a  $B$  period, the only thing that affects the number of sites cleaved at time  $t$  is the number of  $B$  periods that have occurred. Let  $B(t)$  be the random variable giving the number of completed  $B$  periods at time  $t$ , and if the spider is in the middle of a  $B$  period, let  $K(t)$  be the number of sites it has cleaved up to time  $t$  in that period ( $K(t) = 0$  if the spider is in the  $D$  state). Recall Antal and Krapivsky's definition of  $N(t)$  as the number of sites cleaved at time  $t$ , to see that

$$N(t) = \sum_{i=1}^{B(t)} C_{B_i} + K(t). \quad (4.24)$$

Therefore,

$$\mathbf{E}[N(t)] = \mathbf{E}[B(t)] \mathbf{E}[C_B] + \mathbf{E}[K(t)]. \quad (4.25)$$

Eqs. 4.25 and 4.12 together allow us to show

$$\mathbf{E}[B(t)] = \left( \sqrt{t} \frac{\Gamma\left(\frac{3+3r}{4+2r}\right)}{\Gamma\left(\frac{5+4r}{4+2r}\right)} - \mathbf{E}[K(t)] \right) \frac{r+1}{r+2}. \quad (4.26)$$

Note that asymptotically  $\mathbf{E}[K(t)] \rightarrow 0$ , because, as we have shown,  $\mathbf{E}[\tau_D]$  (Eq. 4.13) increases with time while  $\mathbf{E}[\tau_B]$  (Eq. 4.18) and  $\mathbf{E}[C_B]$  (Eq. 4.23) are independent of time. As  $t \rightarrow \infty$ , the probability to be in a  $B$  period will tend to 0, and so also must  $\mathbf{E}[K(t)]$ . Thus for large  $t$ , Eq. 4.26 simplifies to

$$\mathbf{E}[B(t)] = \sqrt{t} \frac{r+1}{r+2} \frac{\Gamma\left(\frac{3+3r}{4+2r}\right)}{\Gamma\left(\frac{5+4r}{4+2r}\right)}. \quad (4.27)$$

The only way the spider can cleave substrates and increase its maximum distance from the origin is for it to be in a  $B$  state. In fact, if the spider never left the boundary (i.e., if  $\mathbf{P}[B \rightarrow D] = 0$ ), it would move ballistically away from the origin.

Thus, the  $B/D$  decomposition of Fig. 4.6 shows how the spider process is in essence a constant alternation between two types of motion: a ballistic motion away from the origin in the  $B$  state, and an ordinary diffusive motion over the contiguous sea of products. The spider repeatedly switches between these states, and the average amount of time spent in each state determines the average behavior of the spider (ballistic vs. diffusive). Because

$$\lim_{t \rightarrow \infty} \frac{d\mathbf{E}[B(t)]}{dt} = 0,$$

the spider initiates fewer and fewer  $B$  periods over time, and in the limit spends all of its time in the  $D$  state moving diffusively. *It is for this reason that the asymptotic behavior is diffusive.* However, because at least initially the spider spends a significant fraction of its time in the  $B$  period, there is a superdiffusive transient.

#### 4.6. First Passage Time

## 4.6 First Passage Time

In this section we study the mean first passage time (MFPT) and the distribution of the first passage time (FPT) of the spiders model. First passage properties of regular random walkers moving over discrete surfaces with various reflecting and absorbing boundaries have been extensively studied [9, 10]. Here we show that the difference between the time the legs spend on visited and on unvisited sites reduces the MFPT of two-legged spiders in various surface settings, and increases the MFPT of one-legged spiders (which behave like regular random walkers). We found that for the spider model the cleavage rate significantly affects the MFPT, and for any track length there exists an optimal cleavage rate.

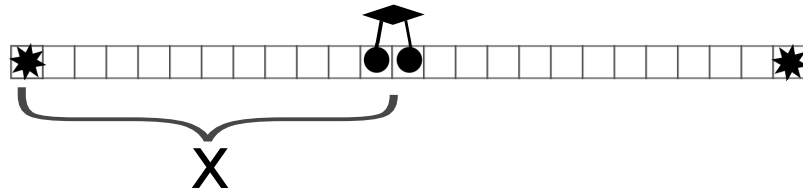
### 4.6.1 Meaningfulness of MFPT

A process to determine if MFPT is a valid characteristic of the first passage behavior was given in Ref. [25]. Similarly to that, we assess the meaningfulness of the mean first passage time in all studied settings. For every setting we estimate the distribution  $P(\omega)$  of the random variable  $\omega = \tau_1/(\tau_1 + \tau_2)$ , where  $\tau_1$  and  $\tau_2$  are first passage times of two independent spiders. Values of  $\omega$  close to 1/2 indicate that spiders act similarly in a particular setting. When  $\omega$  is close to 0 or 1, the process is not uniform, and MFPT is not a good measure of actual behavior. Distribution  $P(\omega)$  can have three distinct shapes: unimodal bell-shaped, bimodal M-shaped, and plateau-like, almost uniform behavior. Bell-shaped form with a maximum at  $\omega = 1/2$  indicates that MFPT can be considered as a valid measure of the first passage times of individual spiders. M-shaped form with two peaks close to 0 and 1, and local minimum at 1/2 indicates that MFPT is not a good measure of the first passage time of individual spiders. The plateau-like shape with zero second derivative at  $\omega = 1/2$  separates the two above cases. Just as in Ref. [25], to quantify the shape of  $P(\omega)$  we fit  $P(\omega)$  to the model  $\chi\omega^2 + c_1\omega + c_2$

for  $0.05 < \omega < 0.95$ . The sign of  $\chi$  indicates the shape of  $P(\omega)$ . In cases when  $\chi < 0$ , the distribution is bell-shaped;  $\chi > 0$  shows that the distribution is bimodal, M-shaped; and  $\chi = 0$  indicates that the distribution is almost uniform.

### 4.6.2 MFPT of Spiders

We measure the mean first passage time,  $\langle \tau(x) \rangle$ , where  $x$  is the absolute distance of the walker from the origin on a one-dimensional track. At that point the walker is absorbed by the boundary. Fig. 4.9 shows the initial configuration of the track where the walker is positioned in the middle, and the absorbing boundaries are shown as stars. The length of the track is  $2x$ .



**Figure 4.9:** Initial configuration of the track. All sites are initially substrates. Absorbing boundary is represented by stars.

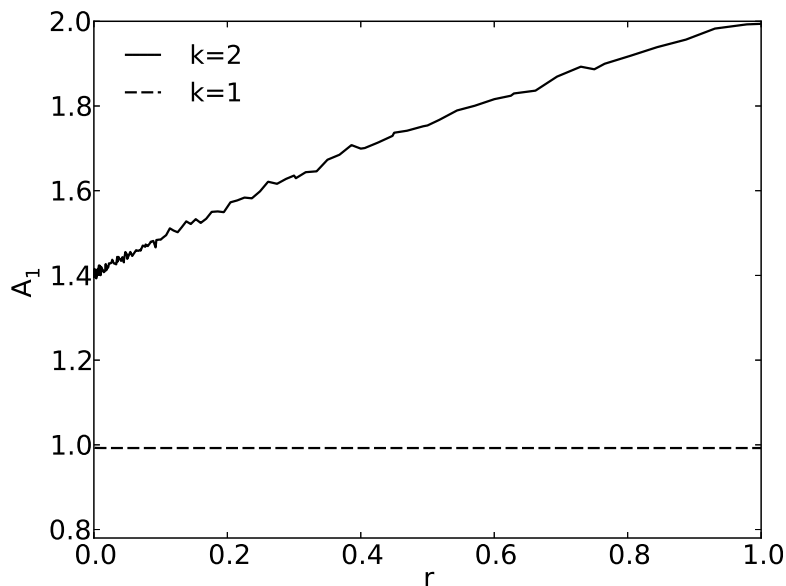
According to the results of our numerical simulations the MFPT of both one- and two-legged spiders is proportional to  $x^2$ . Eq. 4.28 shows the leading and sub-leading terms of  $\langle \tau(x) \rangle$ .

$$\langle \tau(x) \rangle \approx A_1(r)x^2 + a_1(r)x. \quad (4.28)$$

The amplitude  $A_1$  of the leading term describes the asymptotic behavior of the MFPT. For the one-legged spider we found that  $A_1$  does not depend on  $r$ . For the two-legged spider  $A_1$  increases with  $r$ , and approaches 2 when  $r = 1$ . As  $r$  approaches zero  $A_1 \approx 1.4$ . Fig. 4.10 shows the amplitude  $A_1$  for one- and two-legged spiders for 182  $r$  values less than 1.



#### 4.6. First Passage Time



**Figure 4.10:** Amplitude  $A_1$  from Eq. 4.28 as a function of the cleavage and detachment rate  $r$ . The data are for one-legged ( $k = 1$ ) and two-legged ( $k = 2$ ) spiders on a one-dimensional track.

For both one and two-legged spiders the amplitude  $a_1$  of the sub-leading term decreases monotonically and approaches 0 in the absence of memory ( $r = 1$ ). The amplitudes  $A_1$  and  $a_1$  show that the parameter  $r$  only increases the MFPT of one-legged spiders. For two-legged spiders varying  $r$  can decrease their MFPT.

In Section 4.3 we showed that when  $r < 1$ , the AK-model walkers go through three different regimes of motion—the initial, superdiffusive, and the diffusive stage. For lower  $r$  values the initial slow period is longer than for higher  $r$  values, but subsequently the superdiffusive period is longer and faster. For travel over shorter distances the initial period is more important and thus larger  $r$  values result in lower first passage times. For travel over longer distances the superdiffusive period is important and smaller  $r$  values give better results. Thus for every particular distance there is an optimal value of  $r$  ( $r_{\text{opt}}(x)$ ) that minimizes the MFPT. For example, for distance 2000 the spider with  $r = 0.05$  is faster than

the other (sampled)  $r$  values; but for distances 4000 and longer the spider with  $r = 0.01$  is faster.

We estimated  $r_{\text{opt}}(x)$  for tracks of various lengths (up to 10000) through simulations of 210  $r$  values. The results are shown in Fig. 4.11. The figure also shows the  $\langle \tau(x) \rangle$  that corresponds to  $r_{\text{opt}}(x)$  for each distance, i.e., the minimum  $\langle \tau(x) \rangle$  achievable by varying the parameter  $r$ . The optimum  $r$  monotonically decreases with distance. Smaller  $r$  values create a stronger bias towards unvisited sites, but they make individual steps slower when the spider is on the boundary. Fig. 4.11 shows that for better performance on shorter distances faster steps are more important than the bias towards unvisited sites, whereas for longer distances the bias dominates the MFPT.

We also estimate  $r_{\text{opt}}(x)$  by fitting. First, we assume that  $A_1(r)$  and  $a_1(r)$  have the functional form of Eq. 4.29, and find the constants  $c_1$  to  $c_8$  by fitting Eq. 4.29 to the estimates of  $A_1$  and  $a_1$ .

$$\begin{aligned} A_1(r) &= (c_1 r + c_2) / (c_3 r + c_4) \\ a_1(r) &= (c_5 r + c_6) / (c_7 r + c_8) \end{aligned} \tag{4.29}$$

Then, we substitute the results into Eq. 4.28 and find when  $\langle \tau(x) \rangle$  is minimized by extracting the derivative of  $\langle \tau(x) \rangle$  with respect to  $r$ . The predicted  $r_{\text{opt}}(x)$  is also drawn in Fig. 4.11. The derivation also shows that  $r_{\text{opt}}(x) \sim x^{-1/2}$ . To show this we, first, find the derivative of the assumed form of  $\langle \tau(x) \rangle$  (Eqs. 4.28 and 4.30) with respect to  $r$ .

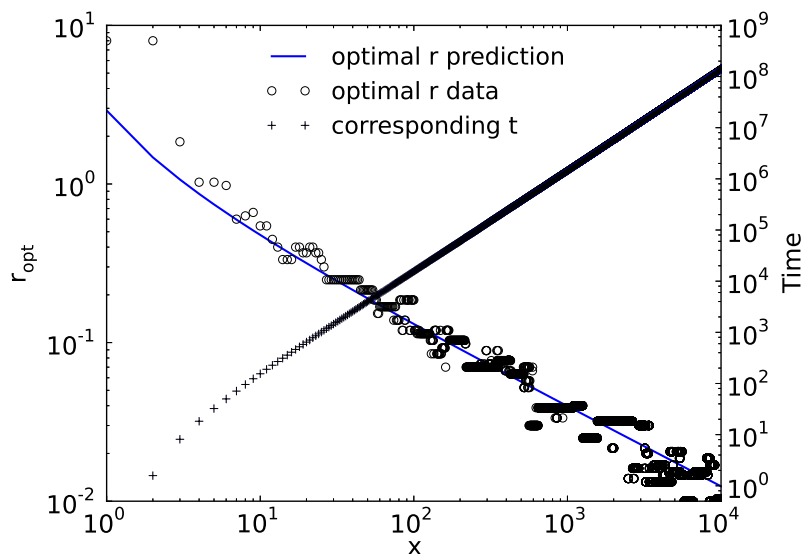
$$\begin{aligned} \frac{d\langle \tau(x) \rangle}{dr} &= \frac{c_1(c_3 r + c_4) - c_3(c_1 r + c_2)}{(c_3 r + c_4)^2} x^2 \\ &\quad + \frac{c_5(c_7 r + c_8) - c_7(c_5 r + c_6)}{(c_7 r + c_8)^2} x \\ &= \frac{c_1 c_4 - c_3 c_2}{(c_3 r + c_4)^2} x^2 + \frac{c_5 c_8 - c_7 c_6}{(c_7 r + c_8)^2} x \end{aligned} \tag{4.30}$$

#### 4.6. First Passage Time

The fitting of Eq. 4.29 to the estimates of  $A_1$  and  $a_1$  yields estimates of constants  $c_1$  to  $c_8$ . Next, we substitute the constants into Eq. 4.30 and find when the derivative is zero.

$$\frac{3.29r + 3.01}{r + 2.15}x + \frac{-4.35r + 1.34}{r} = 0$$

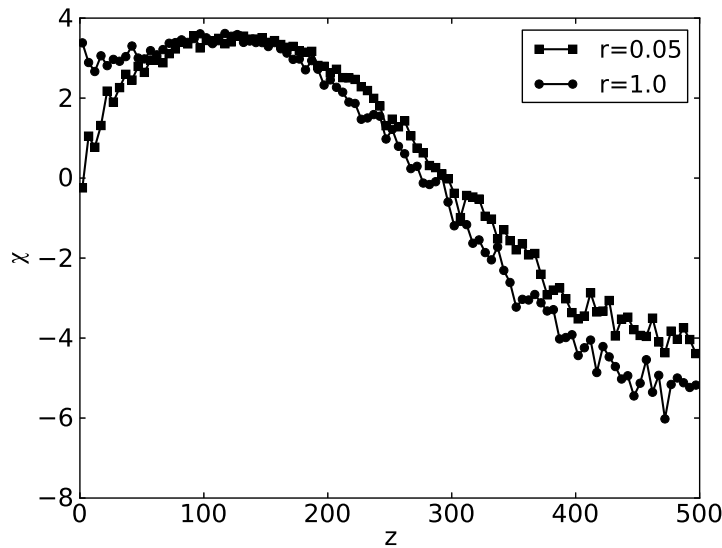
$$r = \frac{5.75 + \sqrt{4.05x - 1.34}}{8.09x - 2.68}$$



**Figure 4.11:** Optimal  $r$  values ( $r_{\text{opt}}(x)$ ) for various distances, and the corresponding  $\langle\tau(x)\rangle$ . Optimal  $r$  values are obtained in two ways; first, by fitting the data for  $\langle\tau(x)\rangle$  into the model of Eq. 4.28 and amplitudes  $A_1$  and  $a_1$  into the model of Eq. 4.29, and second, by simulating 210 values of  $r$  and choosing those that correspond to the lowest  $\langle\tau(x)\rangle$ . The estimation of  $r_{\text{opt}}(x)$  shows that  $r_{\text{opt}}(x) \sim x^{-1/2}$ .

Our estimates of the indicator  $\chi$  show that in the setting of the one-dimensional track and spider starting from the middle,  $\chi$  is always negative, independent of the values of the parameters  $x$  and  $r$ . Thus the MFPT is a meaningful, reliable measure of the first passage time of spiders in this setting. However, as the starting position is moved closer to one of the boundaries (and thus farther away from

the other)  $\chi$  grows, and eventually becomes positive. The positive values indicate that MFPT is not a good measure of the first passage time of spiders when they start close to the boundaries. Interestingly, when spiders start very close to the boundary, about 50 sites away or less, the parameter  $r$  starts to affect  $\chi$ . Fig. 4.12 shows dependence of  $\chi$  on the distance between the spider and one of the targets  $z$  for a track of size  $x = 1000$  and two  $r$  values: 0.05 and 1.0. When  $z \gtrsim 50$  the dependence of  $\chi$  on  $z$  is similar for both  $r$  values. However, when  $z \lesssim 50$ , the  $\chi$  value for the  $r = 0.05$  spider decreases while that of the  $r = 1.0$  spider remains steady.



**Figure 4.12:** Indicator  $\chi$  shows how well the MFPT describes the first passage times of the individual spiders to any of the target sites. The surface is one-dimensional with a circular reflecting boundary of radius  $x$ , the spider starts  $z$  sites away from the left absorbing boundary (site 0). For the  $r = 0.05$  spider  $\chi$  decreases when  $z \lesssim 50$ , while for the  $r = 1.0$  spider  $\chi$  stays about the same. For  $z \gtrsim 50$  the behavior of  $\chi$  is similar for both  $r$  values.

This behavior indicates that smaller  $r$  values make the first passage time of spiders more consistent when the starting position is located very close to one of the targets.

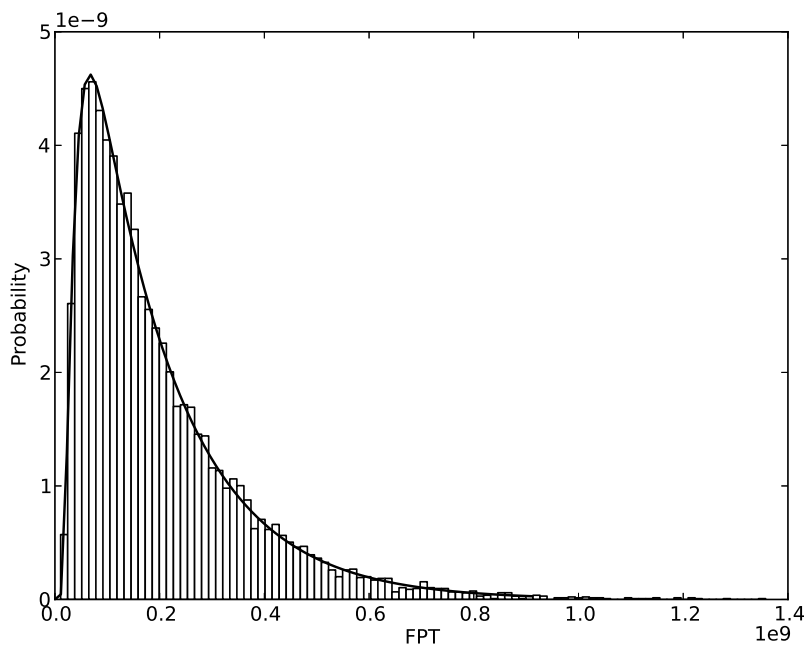
## 4.6. First Passage Time

### 4.6.3 FPT Distribution

For the same setting described in Section 4.6.2 we estimate the distribution of the spiders' first passage time,  $\tau(x)$ . For  $r = 1$ , when spiders are equivalent to regular random walkers, the distribution decays exponentially with time; in Ref. [27] the distribution is expressed in terms of the Jacobi theta function [1]:

$$P^*(t) = (1/x^2) \left( -\frac{1}{\pi^2} \frac{\partial \vartheta_3(z, e^{-\beta t/x^2})}{\partial z} \right) \Big|_{z=\pi/4}, \quad (4.31)$$

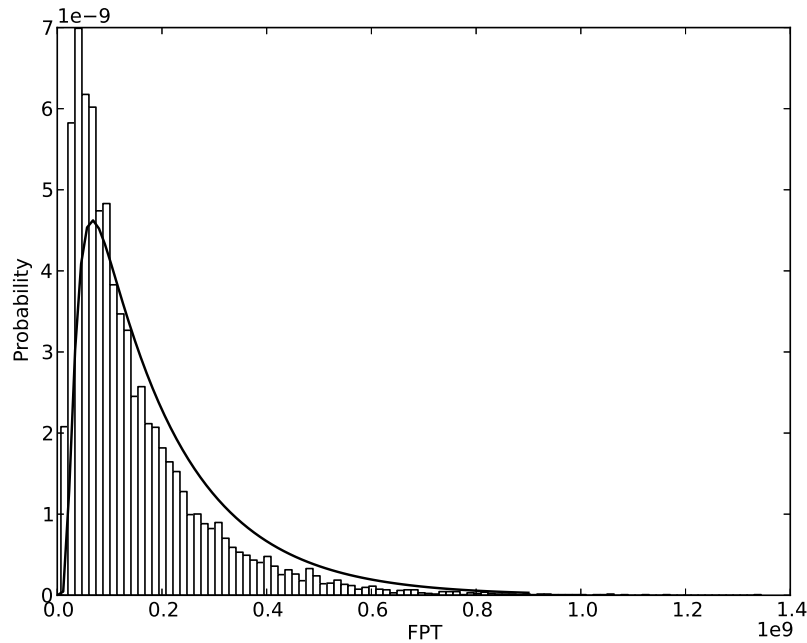
where  $\vartheta_3(z, q)$  is the Jacobi theta function. Fig. 4.13 shows agreement of the simulation results with Eq. 4.31. We also used Kolmogorov–Smirnov test against the



**Figure 4.13:** Distribution of  $\tau(x)$  for the  $r = 1$  spider and the distance  $x = 10000$ . The curve shows the theoretical prediction of Eq. 4.31, and the bars show simulation data.

hypothesis that when  $r = 1.0$   $P^*(t)$  is a distribution of  $\tau(x)$ ; for the distances  $x = 1000$ ,  $x = 5000$ , and  $x = 10000$  the  $p$ -values are 0.81, 0.81, and 0.42 respectively. The  $p$ -values are significant enough to support this hypothesis.

Next, we compare FPT distributions of the spiders with  $r = 1$  and  $r < 1$ . Fig. 4.14 shows FPT distribution densities for the spiders with  $r = 1.0$  and  $r = 0.05$ . The curve shows the distribution density for the  $r = 1.0$  spider, as in Fig. 4.13, and the bars show the distribution density for the  $r = 0.05$  spider.



**Figure 4.14:** Distribution of  $\tau(x)$  for the  $r = 1$  and  $r = 0.05$  spiders. The distance is  $x = 10000$ . The curve shows the theoretical prediction of Eq. 4.31 for the  $r = 1.0$  spider, and the bars show simulation data for the  $r = 0.05$  spider.

The comparison shows that the distribution for the  $r = 0.05$  spider has more probability density concentrated closer to the origin. This shows that spiders with  $r = 0.05$  are more likely to reach boundary sites at  $x = 10000$  faster than  $r = 1.0$  spiders.

# Chapter 5

## Single Spider in 2D

In this chapter we describe properties of the spider model in two dimensions. Particularly, we study the mean square displacement (MSD) of the spider and investigate how the various boundary conditions affect the spider's mean first passage time (MFPT) when it moves over finite two-dimensional surfaces.

To study MFPT we start with an extension of the 1D model of Section 4.6 to 2D, i.e., the mean first passage time of a two-legged spider to a circle, where the spider starts in the center. For this surface we determined that the cleavage rate gives the spider an even greater advantage over a regular random walker. The advantage persists even when the target is a single site, and thus is much harder to find. In this second 2D model the circle is a reflecting boundary, its center is an absorbing boundary, and the spider starts from various distances from the center. As in 1D, every setting of the surface has a corresponding optimal cleavage rate that minimizes the first passage time to the boundary.

The results of this section are published in Ref. [38].

### 5.1 Model

The model is a direct extension of the model introduced in the section 4.1. It also can be seen as a direct extension of the AK model of the spiders on a plane [3] that takes various boundary conditions into account.

#### 5.1.1 Motion and Surface

In the model a single spider moves over finite two-dimensional regular lattices. Just as in 1D, a spider has  $k$  legs and it moves by detaching a leg from its site on

the lattice and reattaching it to a new site. Only one leg can detach at any given time, so a spider cannot detach all of its legs to leave the lattice. Each site can be occupied only by one leg at a time. There is a restriction on the maximum distance between any two legs  $S$  (the gait), and each leg can move to one of the nearest neighboring sites (4 sites in 2D; a diagonal step is not allowed) with equal probability as long as the move does not violate one of the constraints above. Here we use only two types of spiders. First, a spider with  $k = 1$ ; this spider is equivalent to a regular random walker. The parameter  $S$  does not affect this spider since it has only one leg. Second, a spider with  $k = 2$  and  $S = 2$ ; this spider is exactly the bipedal Euclidean spider with maximal separation 2 [3]. When such a spider is placed on a one-dimensional lattice, the model becomes equivalent to that of Section 4.1.

### 5.1.2 Boundaries and Starting Positions

We study the first passage time of spiders moving over two-dimensional regular lattices with various boundary conditions and initial configurations. In each case we are interested when the spider reaches the absorbing boundary.

The boundaries effectively make our surfaces finite. The surface is bounded by a circle of radius  $x$ . Here we study three types of boundaries and initial spider positions. First, the circle is an absorbing boundary, and the spider starts from the center of the circle. Second, the circle is a reflective boundary, while the center is an absorbing boundary, and the spider can start from any site on the circle. Third, we study the case when radius  $x$  is fixed and the spider starts  $y$  sites away from the target site at the center.

In all these settings  $x$  effectively defines the size of the surface. And in all those settings, except the last, we study the dependence of MFPT on  $x$ , i.e.,  $\langle \tau(x) \rangle$ . In the last case we study the dependence of MFPT on  $y$ , i.e.,  $\langle \tau(y) \rangle$ .



## 5.2. Simulation

## 5.2 Simulation

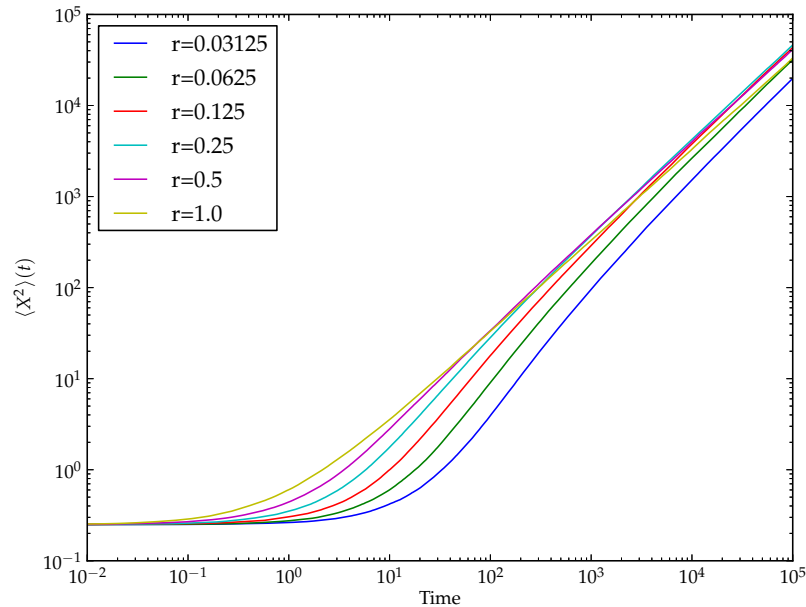
As in section 4.1, we use the Kinetic Monte Carlo method [8] to simulate many trajectories of the Markov process for every instance of the parameter set. In each case we record the first passage time to the absorbing boundary.

To measure  $\langle X^2 \rangle(t)$  in Section 5.3 we simulated 6 different  $r$  rates for up to a time  $10^5$  using 15000 traces. For Section 5.4 we simulated 100 different  $r$  rates for up to a distance of 1000 using 20000 traces. For Sections 5.5.1 we simulated 18 different  $r$  rates for up to a distance of 250 using 20000 traces. For Section 5.5.2 we simulated 15 different  $r$  rates for up to a distance of 100 using 20000 traces.

## 5.3 Mean Squared Displacement

As in Section 4.3, we measure  $\langle X^2 \rangle(t)$  and  $\alpha(t)$  of a two-legged spider moving over a two-dimensional plane. Figure 5.1 shows  $\langle X^2 \rangle(t)$  for different  $r$  values.

Figure 5.2 shows the result of using the Savitzky-Golay smoothing filter [33] on the estimates of  $\alpha(t)$ . The figure shows that, as in the 1D case, when  $r < 1$ , the spiders in 2D go through the three regimes of motion—the initial, superdiffusive, and the diffusive stage. However, in contrast to the 1D case, very small  $r$  values do not give spiders an advantage over spiders with  $r = 1$ . Spiders with smallest  $r$  values approach the diffusive stage before overtaking spiders with  $r = 1$ , but  $r$  values around 0.1 give spiders better advantage in 2D than in 1D. Thus, optimum  $r$  values are higher in 2D. In Section 5.6 we show that spiders with optimum  $r$  values have more advantage over spiders with  $r = 1$  in 2D than in 1D.



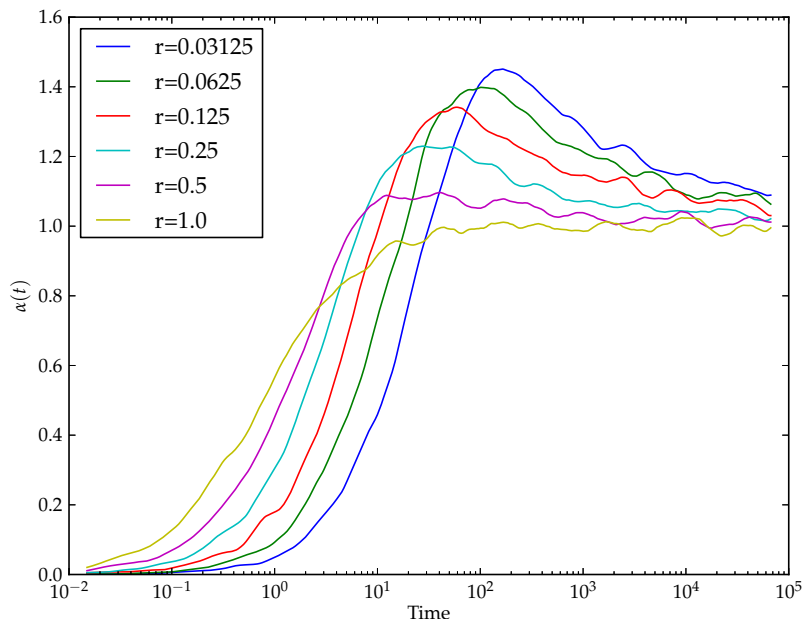
**Figure 5.1:** Mean squared displacement,  $\langle X^2 \rangle(t)$ .

## 5.4 Spider on the Plane with a Circular Absorbing Boundary

A direct extension of a one-dimensional track with length  $2x$  to two dimensions is a set of sites bounded by a circle of radius  $x$ . Fig. 5.3 shows the initial configuration of the surface with the spider positioned in the middle.

Using numerical simulations we found that, similarly to the spiders in one dimension, the MFPT of both one- and two-legged spiders is proportional to  $x^2$ . However, the sub-leading terms are much closer to the leading terms, and therefore are more important for estimating the MFPT. This implies that in two dimensions spiders (especially those with very small cleavage and detachment rate,  $r < 0.05$ ) approach the asymptotic behavior especially slowly. Eq. 5.1 shows the leading and sub-leading terms of  $\langle \tau(x) \rangle$ .

#### 5.4. Spider on the Plane with a Circular Absorbing Boundary

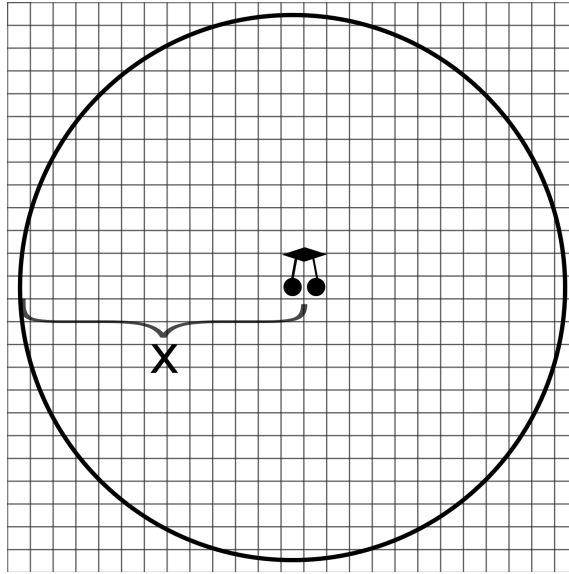


**Figure 5.2:** Finite difference approximation of  $\alpha(t)$ .

$$\langle \tau(x) \rangle \approx \begin{cases} A_2(r) x^2 + a_2(r) x^2 / \ln t & : k = 2 \\ A_2(r) x^2 + a_2(r) x^2 / (\ln t)^{0.88} & : k = 1 \end{cases} \quad (5.1)$$

The correlation  $(\ln t)^{-0.88}$  for the one-legged spider is unusual and slower than for the two-legged spider. Somewhat similar effects were observed in Ref. [3] in the estimation of the mean squared displacement. As in one dimension, the amplitude  $A_2$  of the leading term does not depend on  $r$  for the one-legged spider and is proportional to  $r$  for the two-legged spider. The sub-leading term's amplitude  $a_2$  monotonically decreases with  $r$  in both cases. Fig. 5.4 shows the amplitude  $A_2$  for the one- and two-legged spiders for 60  $r$  values that are less than 1.

For one-legged spiders, as in one dimension, lower  $r$  values only increase the MFPT. The comparison of Figs. 5.4 and 4.10 shows that for two-legged spiders  $A_2$  is affected more strongly by  $r$  than  $A_1$ . The  $A_2$  of the two-legged spider in two dimensions even intersects the  $A_2$  of the one-legged spider. The  $A_2$  starts at 2



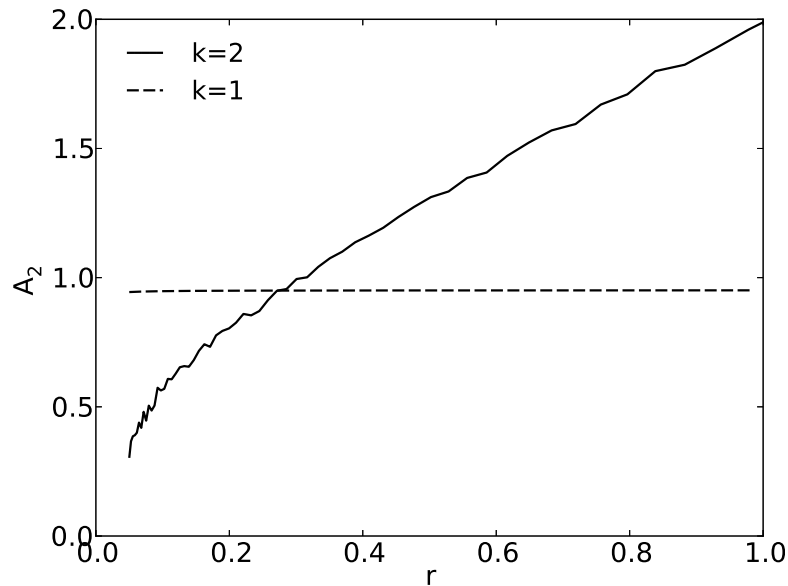
**Figure 5.3:** Initial configuration of the 2D surface. All sites are initially substrates. The spider starts at the center of the circle. The absorbing boundary is shown as a circle; as soon as either leg crosses the circle the target is considered to be found, and the experiment stops.

when  $r = 1$  and decreases towards  $\approx 0.5$  as  $r$  approaches zero.

Similar to the one-dimensional case, spiders with  $r < 1$  start more slowly than the no-memory spider with  $r = 1$ ; then they move faster, and finally they slow down and approach regular diffusion. But in contrast to spiders on a 1D track, the  $r$  values that correspond to fastest times are higher, and the MFPT of spiders with very small  $r$  values (less than 0.1) approaches the MFPT of spiders with  $r = 1$  very slowly. However, the transition towards the diffusive stage happens more slowly compared with 1D.

We estimated  $r_{\text{opt}}(x)$  for circles of various sizes (up to 1000) through simulations of 100  $r$  values, choosing the fastest ones for each distance. Comparison of the results is shown in Fig. 5.5. The figure also shows the fastest  $\langle \tau(x) \rangle$  that corresponds to  $r_{\text{opt}}(x)$  for each distance. Similarly to the 1D case, we also estimate

### 5.5. Spider on the Plane with a Circular Reflecting Boundary and an Absorbing Boundary in the



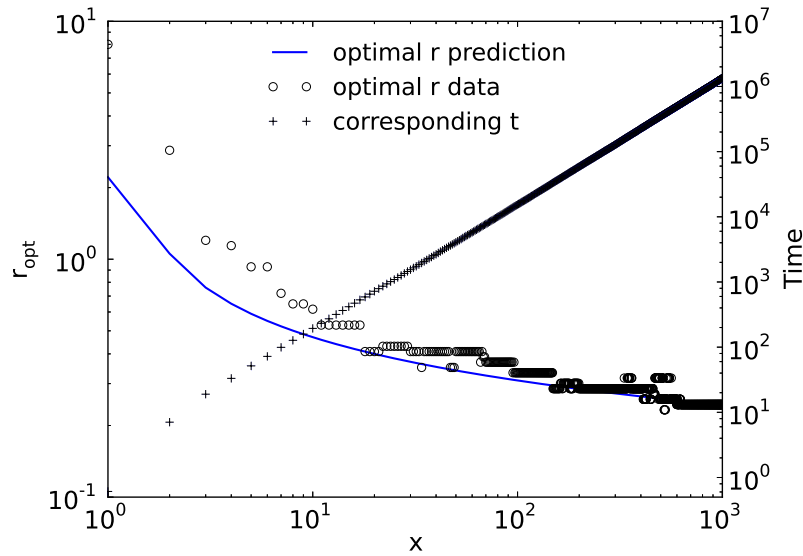
**Figure 5.4:** Amplitude  $A_2$  from Eq. 5.1 as a function of the cleavage and detachment rate  $r$ . The data is shown for the one-legged ( $k = 1$ ) and two-legged ( $k = 2$ ) spiders moving over two-dimensional surface with a circular absorbing boundary.

$r_{\text{opt}}(x)$  by fitting.

Our estimates of the indicator  $\chi$  for various  $r$  and  $x$  values show that in the setting of the two-dimensional circle with absorbing boundary at the perimeter  $\chi$  is always negative, and this does not change with the values of the parameters  $x$  and  $r$ . Just as in 1D, the MFPT is a meaningful measure of the first passage time of the individual spiders moving on a two-dimensional lattice.

## 5.5 Spider on the Plane with a Circular Reflecting Boundary and an Absorbing Boundary in the Center

When we change the contour of the 2D surface to be a circular reflecting boundary instead of an absorbing one, and place a single target site in the center, spiders



**Figure 5.5:** Optimal  $r$  values ( $r_{\text{opt}}(x)$ ) for various radii. Optimal  $r$  values are obtained in two ways; first, by fitting the data for  $\langle \tau(x) \rangle$  into the model of Eq. 5.1 and amplitudes  $A_2$  and  $a_2$  into the model of Eq. 4.29, and second, by simulating 100 values of  $r$  and choosing those that correspond to the lowest  $\langle \tau(x) \rangle$ .

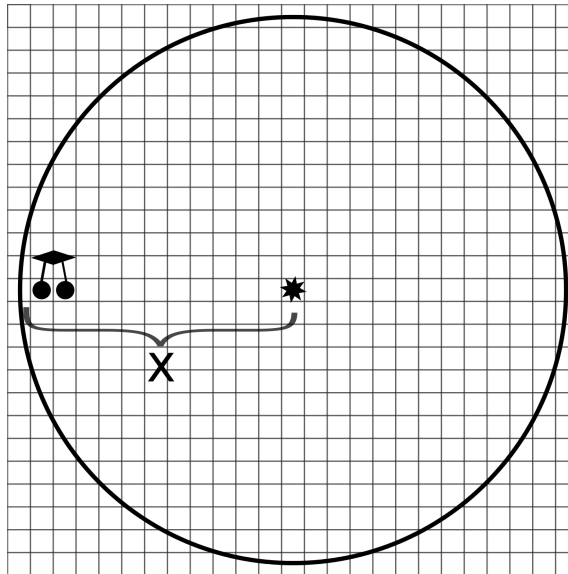
with memory ( $r < 1$ ) still have an advantage over spiders without memory ( $r = 1$ ). We consider two cases: (1) when the radius  $x$  of the circle is variable, and spider starts from any point on the contour, and (2) when the radius  $x$  is fixed, and the distance  $x$  between the starting position and the target is variable. Since both cases are circularly symmetric, all starting positions with the same distance from the center are equivalent. In both cases we study how  $\langle \tau(x) \rangle$  is affected by the parameter  $r$ , and how it is affected by  $x$  in (1) and  $y$  in (2).

### 5.5.1 2D Circle of Variable Radius With Target in the Middle and Spider Starting from the Boundary

Fig. 5.6 shows the initial configuration of the surface where the spider is positioned on the contour and the single target site (absorbing boundary) is shown as a star. Since the reflecting boundary is a circle and the absorbing boundary is a

### 5.5. Spider on the Plane with a Circular Reflecting Boundary and an Absorbing Boundary in the

single site in the center, all starting positions of the same distance from the target site are equivalent. In our simulations we pick a fixed position on the contour as a starting position for the spider. This position is shown as two small solid circles in Fig. 5.6.



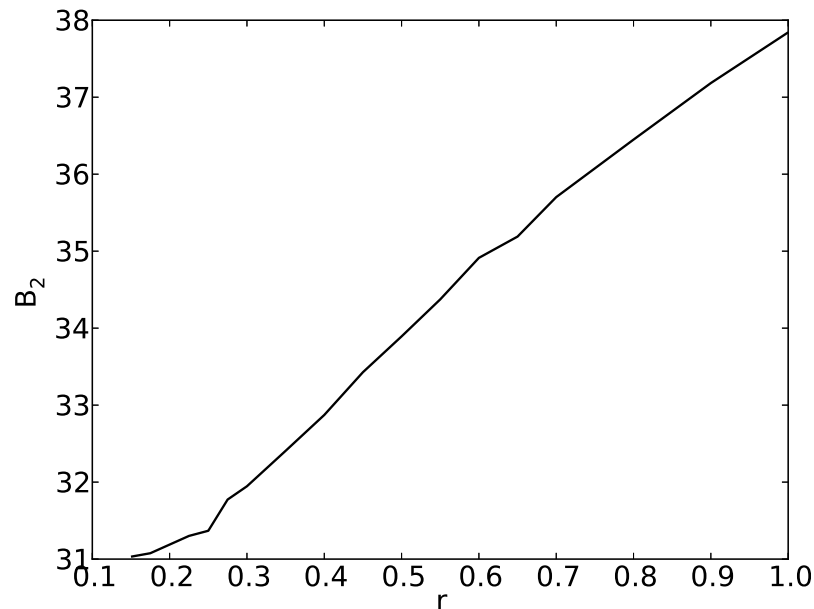
**Figure 5.6:** Initial configuration of the surface. All sites are initially substrates. The absorbing boundary is shown as a star. The spider starts at the periphery. The circle is a reflecting boundary and thus steps outside of the circle are not allowed.

We measure the first passage time of the walker from the contour of the surface to its center for various radii  $x$ .

Interestingly, the MFPT asymptotically grows slightly faster than  $x^2$ . The sub-leading term also grows faster than in the circular absorbing boundary case. Eq. 5.2 shows the leading and sub-leading terms of  $\langle \tau(x) \rangle$ .

$$\langle \tau(x) \rangle \sim B_2(r)x^2 + b_2(r)x^{1.5}. \quad (5.2)$$

Fig. 5.7 shows the amplitude  $B_2$  for the two-legged spiders for 20 different  $r$  values between 0.1 and 1.0.



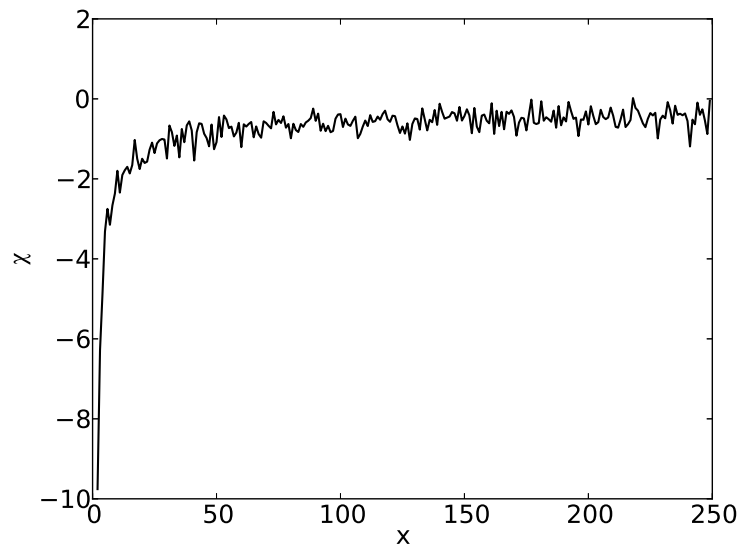
**Figure 5.7:** Amplitude  $B_2$  from the Eq. 5.2 as a function of the cleavage and detachment rate  $r$ . The data are shown for two-legged ( $k = 2$ ) spiders moving over a two-dimensional surface with a circular reflecting boundary. The single site in the middle of the circle is an absorbing boundary.

Values of  $r$  smaller than 1 give the spider an advantage, similar to the 2D configuration discussed above, but now even for shorter radii  $x$ . The typical time-line of this process can be characterized as follows. First, the spider starts moving, it eventually visits many sites without finding the target site, and leaves behind many smaller regions of substrates. Next, the spider starts to move only over visited sites more often, and eventually encounters regions of substrates of various shapes and sizes. At this stage, the spider with  $r = 1$  will not be affected by the substrate regions, and will move just as if they were visited sites. On the other hand, spiders with  $r < 1$  will become biased to stay on the substrates and explore those regions more thoroughly; this will increase their chances of finding the target site, since it must be in one of those unvisited areas. This scenario can explain how spiders with  $r < 1$  gain an advantage over spiders with  $r = 1$  on this surface.



### 5.5. Spider on the Plane with a Circular Reflecting Boundary and an Absorbing Boundary in the

The indicator  $\chi$ , in this setting, grows with the parameter  $x$ , and for  $x \gtrsim 17$   $\chi$  becomes close to 0. As in the previously described settings,  $\chi$  does not depend on the parameter  $r$ . Fig. 5.8 shows the dependence of  $\chi$  on  $x$  for  $r = 0.5$ . For small

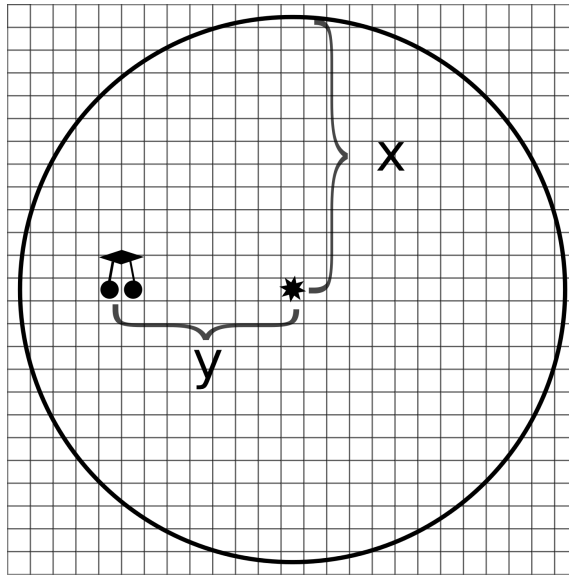


**Figure 5.8:** Indicator  $\chi$  shows how well the MFPT can describe the first passage times of the individual spiders to a single target. The surface is two-dimensional with a circular reflecting boundary of radius  $x$ , and spider starts from the contour.

surfaces (small values of  $x$ ), the  $\chi$  is negative, and thus in those cases MFPT is a good measure of the first passage time of individual spiders. However, for larger  $x$ , values of  $\chi$  close to 0 indicate that the shape of the distribution  $P(\omega)$  is close to uniform, and thus the MFPT is not as good a measure of individual behavior as it is in the settings when the spider searches for the contour. It also shows that the possible paths that the spider can take to locate a single target are more diverse than paths that lead to the contour.

### 5.5.2 2D Circle of Fixed Radius With Target in the Middle and Spider Starting From Various Distances

Fig. 5.9 shows the initial configuration of the surface where the walker is positioned at a distance  $y$  from the target site, and the radius  $x$  is set to 100 and remains constant.

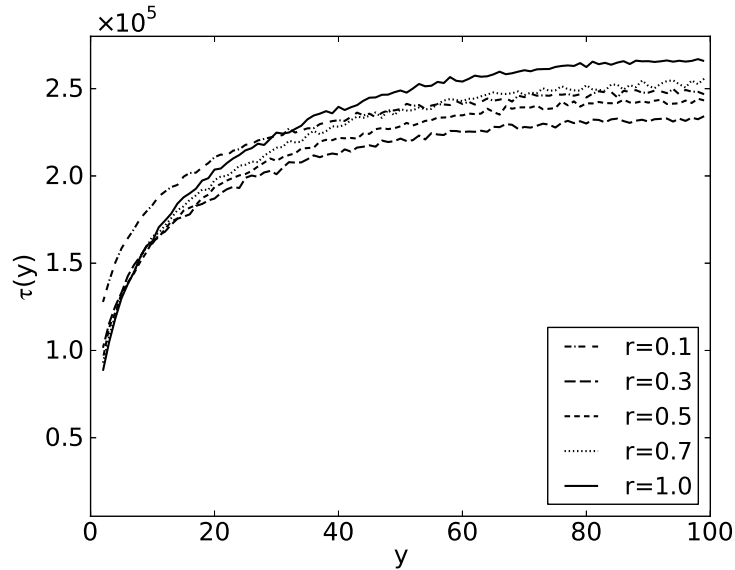


**Figure 5.9:** Initial configuration of the surface. All sites are initially substrates. The absorbing boundary is shown as a star. The spider starts  $y$  sites away from the center. A circle of constant radius  $x = 100$  is a reflecting boundary and thus steps outside of the circle are not allowed.

We measure the first passage time of the walker from the contour of the surface to its center for various distances  $y$ . Fig. 5.10 shows a plot of the MFPT.

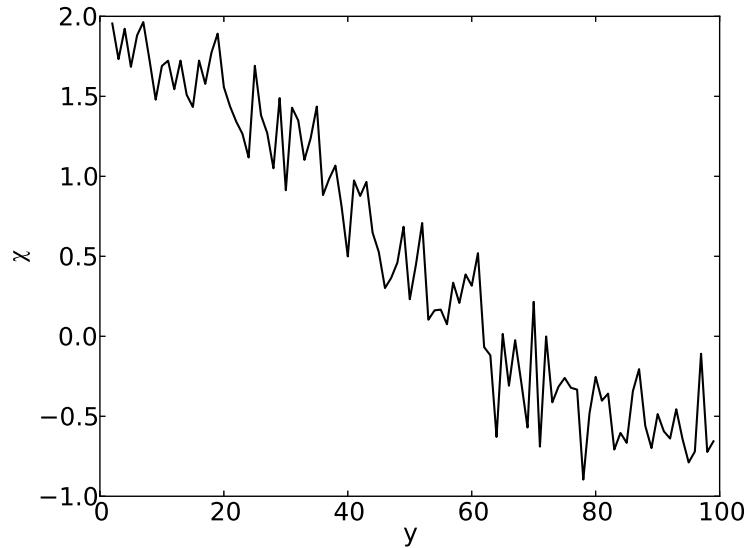
The shape of the curves is asymptotically logarithmic; this shows that the initial position of the spider does not significantly affect  $\langle \tau(y) \rangle$  when  $x$  is fixed. Even if the spider starts closer to the target, there are too many possible paths to the target in 2D, of which one is randomly chosen. Many of them are very long and initially lead the spider far away from the target.

5.5. Spider on the Plane with a Circular Reflecting Boundary and an Absorbing Boundary in the



**Figure 5.10:** Mean first passage time to a single point in the center of the circle with a fixed radius. The boundary is reflecting, and spider starts at various distances from the target.

The indicator  $\chi$ , in this setting, is positive for  $y \lesssim 80$  and decreases with the parameter  $y$ . For  $y \gtrsim 70$  as the surface configuration approaches the configuration discussed in the Section 5.5.1, and  $\chi$  becomes negative, but as in the case of the Section 5.5.1, it still remains close to 0. Fig. 5.11 shows the dependence of  $\chi$  on  $y$  for  $r = 0.5$ . The positive values of  $\chi$  for the smaller  $y$  values indicate that MFPT is not a good measure of the first passage times of the individual spiders, and paths that lead the spider to a target are very diverse and vary significantly in their lengths.



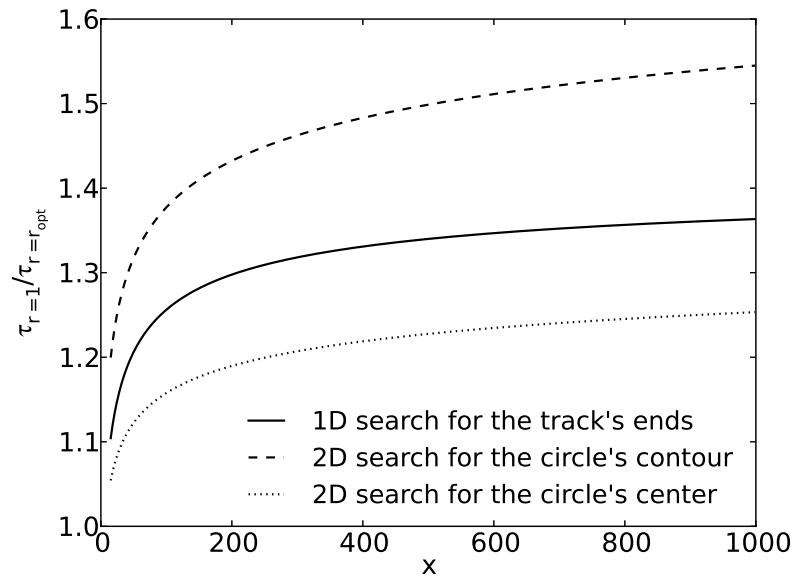
**Figure 5.11:** Indicator  $\chi$  shows how well the MFPT can describe the first passage times of the individual spiders to a single target. The surface is two-dimensional with a circular reflecting boundary of a fixed radius  $x = 100$ , and the spider starts at distance  $y$  from the target site.

## 5.6 Comparison of Spider Performance in 1D and 2D Settings

It is interesting to compare the advantage spiders with  $r < 1$  (i.e., with memory) enjoy over those with  $r = 1$  (i.e., without memory) in the described 1D and 2D settings. We compute the ratio of  $\langle \tau(x) \rangle$  for spiders with  $r = 1$  and  $\langle \tau(x) \rangle$  for spiders with  $r = r_{\text{opt}}(x)$ . This ratio is plotted in Fig. 5.12 against the surface size for 1D and 2D.

The plot shows that the cleavage rate  $r$  gives spiders more advantage on a 2D plane searching for a circle than on a 1D strip searching for its ends. This advantage can be attributed to the amount of substrates that spiders leave behind as they move away from the origin. In 1D, spiders do not leave any substrates behind, so there are no substrates between the left and the right ends of the sea

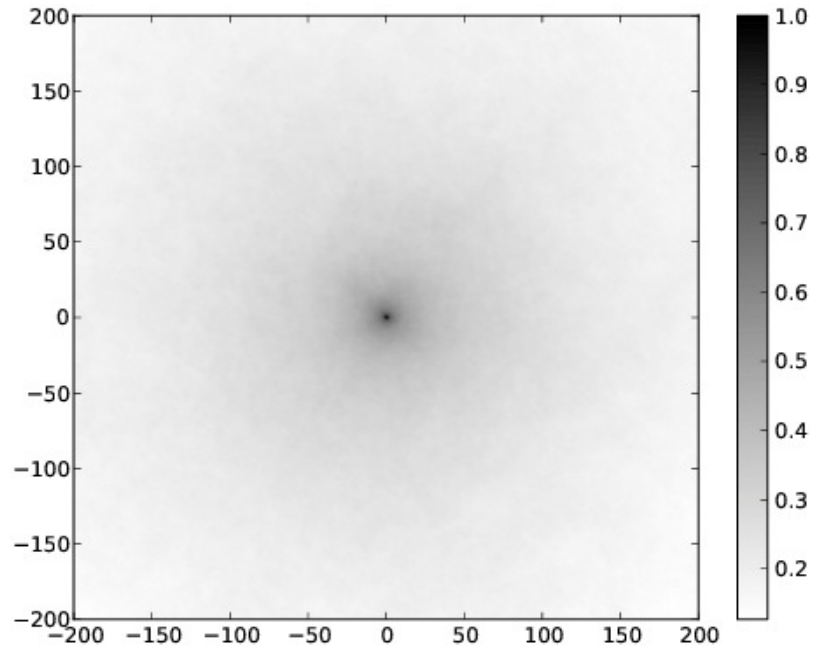
## 5.6. Comparison of Spider Performance in 1D and 2D Settings



**Figure 5.12:** Ratio between  $\langle \tau(x) \rangle$  of the spiders with  $r = 1.0$  and  $r = r_{\text{opt}}$ .

of products, and when a spider moves back towards the origin there are no substrates to bias it towards the boundary. In 2D, the shape of the product sea can be very complicated and there can be many substrates left behind. When a spider moves backward it has a high probability to still encounter substrates, which can bias it towards the boundary. The higher  $r_{\text{opt}}(x)$  values in 2D can be attributed to the direction of the emergent bias when the spider is on the border between visited and unvisited sites. In 1D, the border is simple, and its shape remains the same over time. It is defined by the two closest unvisited sites to the origin on the right and on the left side. As a result, when the spider with rate  $r < 1$  is on the border, it is always biased in the desired direction—away from the origin. In 2D, the shape of the border can be much more complex and is even not necessarily connected. This type of border leads to a much weaker bias towards the edge of the surface. In many cases the spider is not biased directly to the edge in the direction of the shortest path, and sometimes the spider even is biased back towards the circle's center. Despite that, the greater amount of substrates accessible by

the spider in 2D overcomes the weaker bias and makes spiders more efficient at finding the absorbing boundary. Fig. 5.13 shows the average density of products; the spider has a high probability of encountering substrates when it turns back towards the origin.



**Figure 5.13:** Average density of products when spider with  $r = 0.25$  and  $k = 2$  is at distance 1000.

## Chapter 6

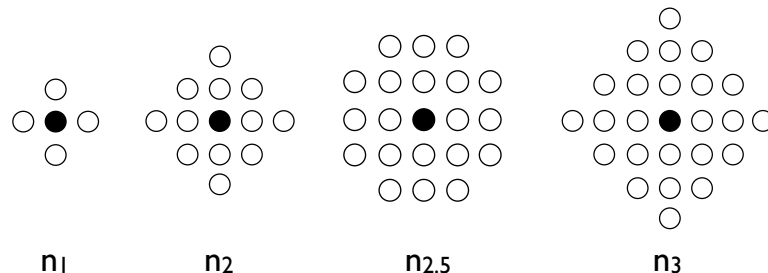
# Multiple Spiders with interactions on 2D surfaces

If we put several spiders of the AK-model on a two dimensional surface they will influence each other not only through the exclusion principle but also through their traces in a form of product sites. To test those influences two laboratory experiments have been proposed by Stojanovic: "T-junction" experiment and "Search" experiment <sup>1</sup>. We studied abstract models of both of those experiments and obtained some initial results. Both models use the same definition of spiders, which is a direct extension of spiders of the AK model. Spiders retain all the parameters and properties of spiders of the AK model in 1D, and we add a new parameter  $n$  that defines a neighborhood of sites to which leg can move in a single step from its current position. Figure 6.1 shows neighborhoods that we used in our models.

The results of this section are published in Refs. [42,43].

---

<sup>1</sup>Milan N. Stojanovic, personal communication, 2009



**Figure 6.1:** Neighborhoods studied. Black circle shows the current position of the leg. White circles show where a leg can potentially move in a single step. When a leg of a spider makes an actual step, that set of accessible sites is usually get reduced by the restriction that parameter  $s$  impose.

## 6.1 T-junction experiment

In T-junction experiments we calculate how a particular trajectory of the first spider affects the trajectory of the second spider; the two spiders are identical and are released sequentially to walk over a simple finite surface. The surface consists of a vertical and a horizontal track, where the top end of the vertical track connects to the middle of the horizontal track to form a T-junction. Both left and right ends of the horizontal track are traps, and all other sites are initially substrates. We consider the scenario that the first spider started to move from the bottom of the vertical track, then chose to move left at the intersection and ended up in the left trap, leaving a trace of products behind it. Given that the second spider is released at the intersection after the first spider has been trapped, we compute probability that the second spider will end up in the right trap. This probability can be seen as a measure of how strongly the first spider influences the second; if the first spider did not affect the second spider at all, that probability would be  $1/2$ , and if the second spider were not allowed to visit the sites visited by the first spider, that probability would be 1.

Because in this model the size of the surface is relatively small, we can compute the distribution of the spider final position exactly when it reaches one of the trap sites, no matter how many steps it takes to reach one of those sites. Also for small surfaces such as we used, we find that after about 60 steps the spider is almost certain to have reached the blue or the red trap, from which it cannot escape. Table 6.1 shows the outcomes of those experiments, as predicted by our analysis. The rightmost two columns show the probability that the spider will have reached the red and the blue states respectively. In the figures, we represent graphically the initial configurations of the tracks over which the spiders move: a white circle represents a product, a gray circle is an initial substrate, a black circles represents the initial position of the second spider, and finally the red and blue



### 6.1. T-junction experiment

circles represent the traps at the two opposite ends of the T-junction (a leg cannot detach from those sites).

The initial configurations in Table 6.1 should be viewed as plausible configurations of a T-junction track on an origami *after one spider has already completed its journey*. In other words, we are analyzing the movement of a *second* spider. Table 6.1 shows that for all studied gaits the trajectory of the first spider significantly influences that of the second spider. As  $r$  decreases, it becomes less likely that the second spider follows the first spider. The number of legs and the gait also influence this dependence: a second spider with more legs is less likely to follow the first spider. Also this bias is stronger for gaits that allow more freedom of legs movements.

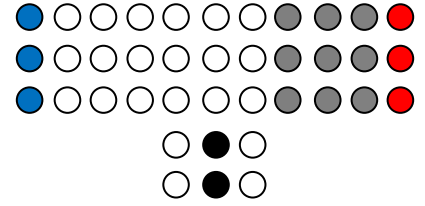
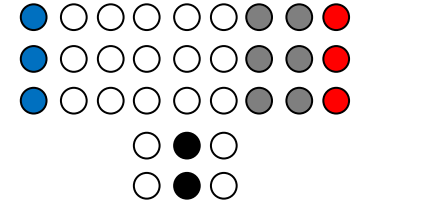
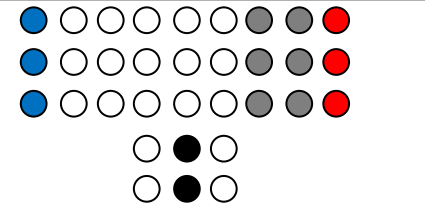
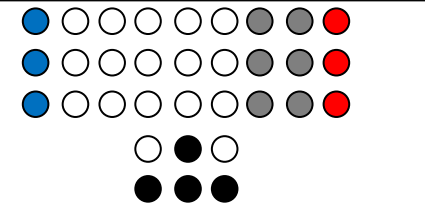
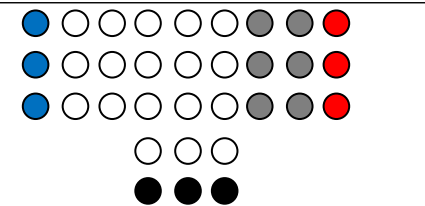
**Table 6.1:** T-junction experiment results.

Surface	Spider model	Probability finishing in red	Probability finishing in blue
	$k = 2, r = 0.1,$ $s = \sqrt{2}, n = n_1$	0.7464	0.2525
	$k = 2, r = 0.1,$ $s = \sqrt{2}, n = n_1$	0.6543	0.3451

Continued on next page

Chapter 6. Multiple Spiders with interactions on 2D surfaces

Table 6.1 – continued from previous page

Surface	Spider model	Probability finishing in red	Probability finishing in blue
	$k = 2, r = 0.1,$ $s = \sqrt{2}, n = n_1$	0.6678	0.3322
	$k = 2, r = 0.5,$ $s = \sqrt{2}, n = n_1$	0.5343	0.4642
	$k = 2, r = 1.0,$ $s = \sqrt{2}, n = n_1$	0.5	0.5
	$k = 4, r = 0.1,$ $s = 2, n = n_{2.5}$	0.7636	0.2364
	$k = 3, r = 0.1,$ $s = 2, n = n_{2.5}$	0.7449	0.2551
Continued on next page			

## 6.2. Search experiment

Table 6.1 – continued from previous page

Surface	Spider model	Probability finishing in red	Probability finishing in blue
	$k = 2, r = 0.1,$ $s = 2, n = n_{2.5}$	0.6885	0.3115

Table 6.1 – end.

Comparison between these analytical results from the model and potential experimental results (frequencies of red vs. blue outcome) can be used to validate the model. Alternatively, the model can be used to guide physical experiments by helping to choose those surface and spider geometries that yield the most bias.

## 6.2 Search experiment

In Search experiments we measure how fast three identical spiders can find three trap sites. The traps are placed in three corners of a rectangular lattice, and the spiders are released at the fourth corner at the same time. All sites except traps are initially substrates. Used Kinetic Monte-Carlo method, we simulated one- to four-legged spiders with several gaits and different values of  $r$ . We measured average time to find all traps and how it depends on the parameter  $r$  and the spiders' gait.

### 6.2.1 Search

In our search model, the lattice is of a finite fixed size, 22 by 32; the numbers may seem arbitrary, but they reasonably describe the DNA-origami tiles used in past molecular spider experiments [24]. We use three searching spiders, initially in one corner of the lattice. The search targets are the three special *trap* sites, in the three opposite corners. We assume that a leg that attaches to a trap remains forever bound to it.<sup>2</sup> Furthermore, when a spider's leg is thus trapped, *all* its legs cease moving. All remaining sites initially are ordinary cleavable substrates. Because there are as many targets as spiders, eventually each spider reaches a target. When all target sites have been found all motion stops (in this crude abstraction).

The chemical kinetics parameter  $r$  and the spiders' gait parameters  $n$ ,  $S$ , and  $k$  are the variable parameters of the model, and they will influence how fast the spiders move and search for targets. In the following section we begin our exploration of this parameter space.

### 6.2.2 Interaction

In the model spiders influence each other's motion in two ways: First, through the exclusion principle—if the legs of one spider occupy some set of sites, then another spider's legs are prevented from occupying any of those sites. Second, through stigmergy, or communication via modification of the environment—when one spider makes several moves it leaves behind a trail of products; another spider can subsequently encounter those products on its way and they will affect its future moves.

---

<sup>2</sup>In the laboratory, an uncleavable, pure-DNA substrate has been used [24] for the purpose. In envisaged applications, the targets presented on the cell surface will not necessarily be DNA strands. To bind to non-DNA targets, in addition to the legs the spider may carry an "arm", an aptamer molecule that specifically binds to the target.

## 6.2. Search experiment

Analysis of the t-junction model showed that spiders are biased towards unvisited sites when they are on the boundary between visited and unvisited sites. So when a spider encounters its own trace or other spiders' traces, it will more likely move towards fresh substrates. There is thus a repulsive stigmergic interaction: we expect spiders preferentially to avoid the traces of other spiders.

In consequence, in a target search application we expect the spiders to preferentially search for targets in new, unvisited locations. To evaluate the importance of exclusion and stigmergy for the search task we compare the performance of the model with two alternative modifications. In the first modification we make spiders move on separate and initially identical surfaces, i.e., each spider has its own surface and they cannot see each other. But, they still share the same trap sites, so if one spider finds a particular trap site, that site will appear as occupied for all other spiders. This modification to the original model removes both exclusion and stigmergy from the model, and allows us to evaluate how fast these *independent* spiders can find all traps. In the second modification, all spiders move on the same surface, but we each surface site contains three different chemical components, keyed to the three spiders. When a leg of a particular spider  $a$  leaves some site  $x$  it affects only the component  $x_a$  of site  $x$  corresponding to  $a$ : if  $x_a$  is a substrate it is converted into product. Thus in this modification a spider can "see" only its own traces and has no access to other spiders' traces. While a site  $x$  is occupied by spider  $a$ , the other two spiders are barred from it. Thus this modification of the model eliminates stigmergic communication between spiders but preserves the exclusion principle, allowing us to evaluate the importance of stigmergy alone.

### 6.2.3 Model and Simulation

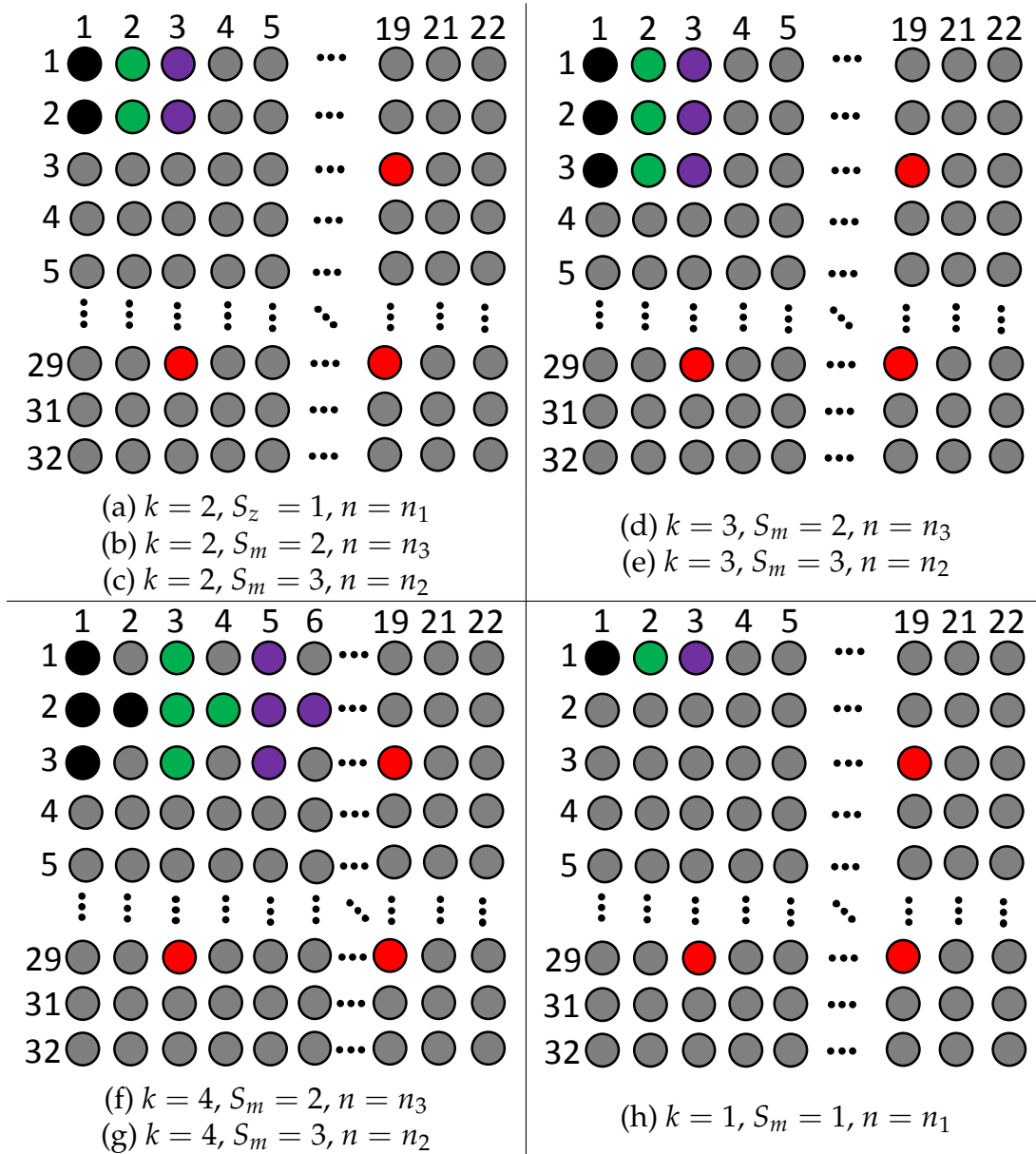
Combining the states of all the spiders and the surface gives a continuous-time Markov process for our model. We use the Kinetic Monte Carlo method [8] to simulate multiple trajectories of this Markov process. The simulation stops when all spiders are trapped, i.e., the search is complete, and then the simulated time is recorded. This time is an observation of a first-passage-time random variable. The results below will show the mean first passage time estimated from our traces.

### 6.2.4 Results

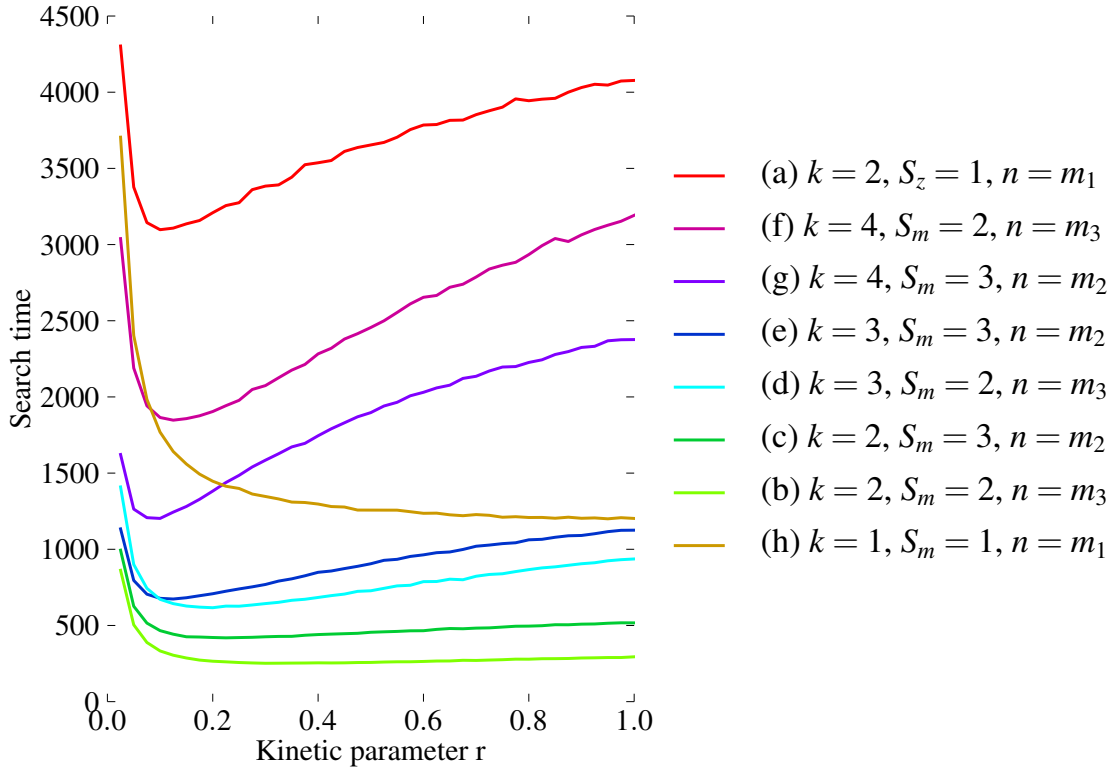
We carried out simulations using the following seven spider gaits: (a)  $k = 2$ ,  $S_z = 1$ ,  $n = n_1$ ; (b)  $k = 2$ ,  $S_m = 2$ ,  $n = n_3$ ; (c)  $k = 2$ ,  $S_m = 3$ ,  $n = n_2$ ; (d)  $k = 3$ ,  $S_m = 2$ ,  $n = n_3$ ; (e)  $k = 3$ ,  $S_m = 3$ ,  $n = n_2$ ; (f)  $k = 4$ ,  $S_m = 2$ ,  $n = n_3$ ; (g)  $k = 4$ ,  $S_m = 3$ ,  $n = n_2$ ; and a simple random walker, which can be viewed as a spider with parameters (h)  $k = 1$ ,  $S_m = 1$ ,  $n = n_1$ . Here  $S_m$  means that maximum distance between legs is given as the Manhattan ( $L_1$ ) distance and  $S_z$  means the Chebyshev ( $L_\infty$ ; maximum) distance. These gaits were chosen to correspond to different physically realistic molecular spiders, but it must be admitted that the space of possible plausible gaits is much larger, and we must defer its exploration to a future study.

Table 6.2 describes the chosen gaits, along with the initial spider positions and the target site positions, shown graphically. Initial positions for the gaits with equal number of legs  $k$  are equivalent, and thus Table 6.2 groups the gaits by  $k$ . Black, green, and violet circles represent the initial leg positions of the first, second, and third spider respectively; gray circles represent ordinary substrates; red circles represent the three target traps. The targets are non-specific, that is, any spider can be trapped by any target.

## 6.2. Search experiment



**Figure 6.2:** The initial state of the system and the spider gait for each simulated configuration. Configurations (a)–(g) are multi-legged spiders and configuration (h) is the control one-legged spider. Black, green, and violet circles represent the initial leg positions of the first, second, and third spider respectively; gray circles represent ordinary substrates; red circles represent the three target traps.



**Figure 6.3:** Search time as a function of the kinetic parameter  $r$  for the eight configurations simulated. (a)–(g) are multi-legged spiders; (h) is a one-legged spider.

Simulation results are shown comprehensively in Figure 6.3, and also individually for each gait in Figure 6.4 to reveal additional detail. Figure 6.4 also contains results for non-communicating spiders with and without exclusion.

### 6.2.4.1 The effect of the gait

In Figure 6.3 curves corresponding to the studied gaits have different shapes and are vertically separated. Thus, the gaits of the spiders greatly influence their performance, and spiders with particular gaits can be faster than regular random walkers. For the simulated surface, spiders with gaits that allow more freedom



## 6.2. Search experiment

for the legs to move (i.e., when a leg is moving to a new site it has more sites to choose from) achieve better performance. Two- and three-legged spiders with gaits (b), (c), (d), and (e) are the fastest among those simulated; they either have the larger neighborhood  $n = n_3$  or the longest possible distance between legs  $S_m = 3$ , which makes these gaits the least restrictive. Spiders with gait (a)  $k = 2$ ,  $S_z = 1$ ,  $n = n_1$  are the slowest, despite having the same number of legs  $k = 2$  as the best performing spiders, with gait (c): gait (a) is much more restrictive, and with its parameters  $S_z = 1$  and  $n = n_1$  it gives a small candidate set of new positions for the legs. Gaits (f)  $k = 4$ ,  $S_m = 2$ ,  $n = n_3$  and (g)  $k = 4$ ,  $S_m = 3$ ,  $n = n_2$  are also slower than gaits (b) through (e): they have the same parameters  $S_m$  and  $n$ , but the addition of an extra leg reduces the choice of sites for a moving leg (because the new position must be within a certain distance from each of the legs that remain attached), which leads to slower performance.

### 6.2.4.2 The effect of the kinetics

We now examine the influence of the kinetics (i.e., the difference between the substrates and the products displayed on the surface, which is amenable to adjustment in the laboratory) on the search performance of the spiders. Figure 6.4 shows the kinetic parameter  $r$  significantly affects the performance of the spiders. One-legged spiders always have better performance when  $r$  is bigger—this observation is similar to the results of the study of one-legged spiders in one dimension [4]. Thus, the presence of memory on the surface in the form of substrates and products does not improve the performance of a monovalent random walker. For multi-legged spiders each gait has an optimal  $r$  value that minimizes the search time.

### 6.2.4.3 The effect of spider interactions

Figure 6.4 shows that for all gaits except (h) (random walk) spiders with stigmergy perform better than non-communicating spiders when  $r$  is around its optimum value. As expected, for values of  $r$  closer to 1.0, the role of stigmergic communication is less significant, because the difference between visited and unvisited sites is small, and thus the traces of spiders influence their behavior only weakly. For the values of  $r$  between the optimum and 1.0, the difference between communicating and non-communicating spiders depends on the spiders' gaits. For the best performing gaits (b) and (c), the advantage of stigmergic communication grows as  $r$  decreases from 1.0 to its optimum value. For the gait (h) (random walk) the advantage of communication also grows as  $r$  decreases, while the search time of both communicating and non-communicating spiders grows. This happens because the presence of substrates only serves to slow the random spiders down, and non-communicating random spiders initially have more substrates in total (for non-communicating spiders there is initially a field of substrates per spider, while communicating spiders share just one field). Thus, non-communicating spiders use more time to clean their surfaces from substrates. For the gaits (d)–(g), as  $r$  decreases from 1.0 to its optimum value, the difference in search time between communicating and non-communicating spiders remains very small. For those gaits the advantage of stigmergic communication starts to appear only when  $r$  is close to its optimum value.

### 6.2.4.4 The effect of exclusion alone

Comparison of non-communicating spiders with and without exclusion shows that the influence of exclusion on search time is less significant than the influence of communication. The effect of exclusion also depends on the number of legs. For gaits with one and two legs, the difference in search time between spiders

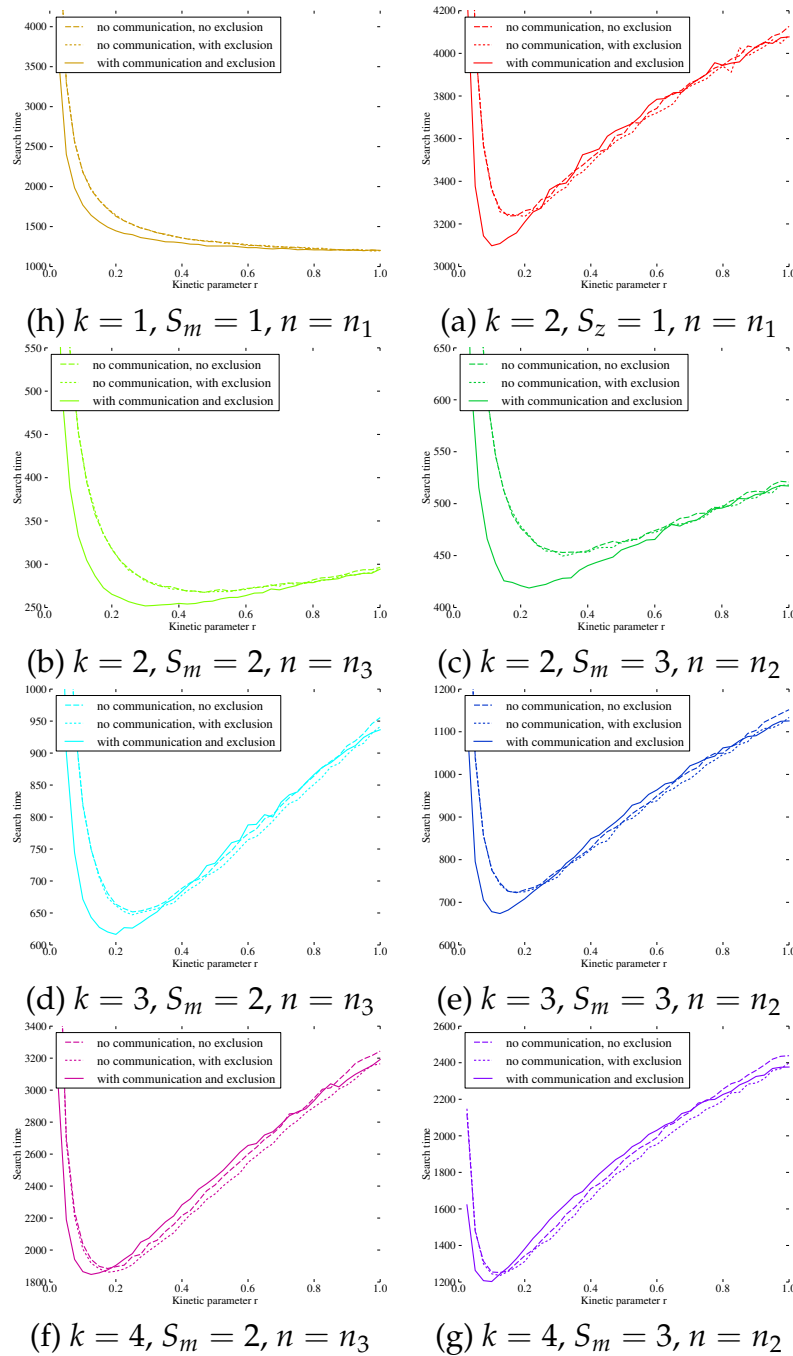
## 6.2. Search experiment

with and without exclusion is very small. For bigger spiders, with three and four legs, this difference grows, and spiders with exclusion are little bit faster than spiders without exclusion. In fact, for spiders with three and four legs and a range of  $r$  values of between the optimum and 1.0, spiders with exclusion and without communication perform the best, by a small margin. As with regular random walkers this happens because non-communicating spiders effectively have more substrates initially, but in contrast with regular random walkers the substrates make them faster, and this enhances search performance. But even for those gaits, communicating spiders with exclusion overtake non-communicating spiders as  $r$  approaches its optimum value.

### 6.2.5 Discussion

Our system can be viewed as a hierarchical multi-agent system: the system consists of multiple spiders, and each spider consists of multiple legs. The legs and the spiders interact through exclusion, and also stigmergically as they modify the surface; and the legs of one spider interact through kinematic constraints. But we are not free to design this multi-agent system as we please to achieve some system-level behavior; instead, we are severely restricted in the design of the agents: they are just molecules of a particular kind, and molecules are dumb. Thus, recalling the principle of complexity theory that simple local rules, iterated, may give rise to complex global behaviors, we are asking whether this also happens in a very primitive setting. We do not expect to be able to mimic the complexity of behaviors of even a single ant, which after all is millions times more structurally complex. However, we hope that our results will aid in the development of nanoscale molecular walkers.

Chapter 6. Multiple Spiders with interactions on 2D surfaces



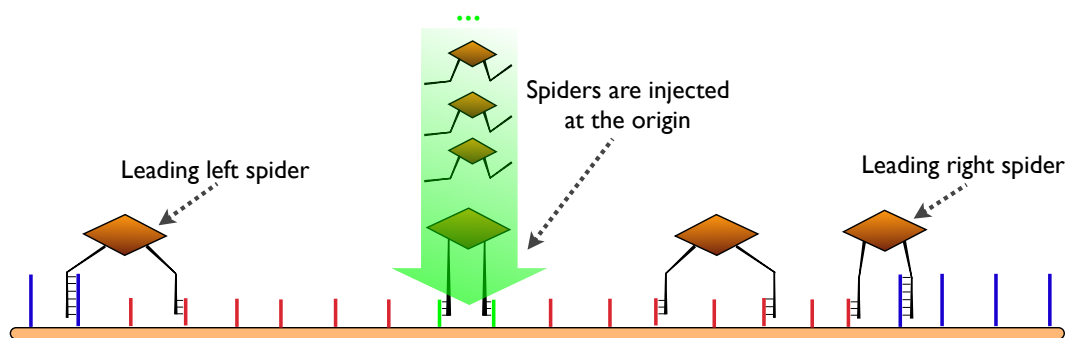
**Figure 6.4:** Enlarged plots from Figure 6.3, in which the existence of an optimum  $r$  value for the multi-legged spiders can be discerned better. And comparison with no communicating spiders with and without exclusion.

## Chapter 7

# Multiple spider model with infinitely strong source

One way to make spiders spend less time in the diffusive state in 1D is by reducing the size of the product sea that is accessible by the spider. We can do that by stochastically releasing multiple spiders at the origin. Spiders that were released later can act as moving barriers and prevent spiders that were released first from moving backward too far away from the boundary (Fig. 7.1). That works because, as in the AK model, every site can be occupied only by one leg of any spider and in combination with parameter  $s = 2$  this prevents the spiders from passing each other and changing their relative order. We call spiders that are furthest from the origin *leading* spiders. In 1D there are two leading spiders, a leftmost and a rightmost, and we are measuring the performance of the whole system by the performance of the leading spiders. Ideally spiders in the middle would make the size of the product sea that is visible by the leading spiders constant.

The results of this section are published in Ref [40,41].



**Figure 7.1:** A multiple molecular spider system with injection. A new spider is added at the origin whenever sites at the origin are unoccupied. Spiders can not pass each other.

## 7.1 The Multi-Spider Model

Cooperative and interactive behavior of spiders can be studied by extending the AK model to a *multi-spider* model that describes a system of  $S \geq 2$  spiders moving simultaneously. Every spider in the multi-spider model has identical values for parameters  $k$  (the number of legs) and  $s$  (the maximum leg spacing), and they all move over the same 1D surface of substrates. The (otherwise indistinguishable) spiders are enumerated as  $1, \dots, S$ , which allows the state of the system to be described as

$$X = (P, F_1, \dots, F_S). \quad (7.1)$$

Here,  $F_i \subset \mathbb{Z}$  gives the attached leg locations of spider  $i$ . In analogy to Eqs. 4.2 and 4.3, we maintain for all  $i$  that

$$|F_i| = k \quad (7.2)$$

and

$$\max(F_i) - \min(F_i) \leq s. \quad (7.3)$$

To extend the chemical exclusionary properties of spider legs to multi-spider systems, we add the restriction that any site on the surface can be occupied by only one leg of any spider, so that for all  $i, j$ :

$$F_i \cap F_j = \emptyset. \quad (7.4)$$

Finally, we define a mechanism to allow the addition of new spiders into the system, allowing  $S$  to grow with time, while maintaining the Markovian properties of the process. New spiders can be added at an injection site  $I = \{0, 1\}$  for any state in which the sites 0 and 1 are unoccupied. A new spider is added as a Poisson process with rate  $\lambda > 0$ , and the initial state of the new spider is  $F_{S+1} = \{0, 1\}$ . In the limit when  $\lambda = \infty$ , a new spider is added as soon as the injection site is unoccupied. Even in this limiting case, the presence of other spiders at the injection

## 7.1. The Multi-Spider Model

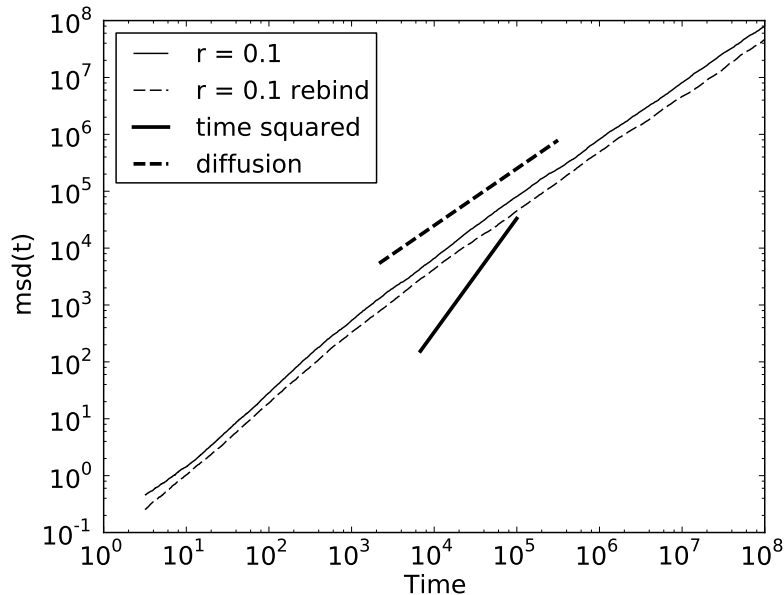
site and their finite rate of movement out of this site constrain the multi-spider system to a finite number of spiders at all times.

### 7.1.1 Rebinding Gait

With multiple spiders on a lattice, there are situations where a particular spider is completely blocked from movement. This happens when its legs are together (i.e., on adjacent sites) and other spiders occupy the sites to its immediate left and right. Thus, to simplify the Markov process description, we introduce a slight change to the gait of the walker. When a leg detaches from a site  $i$  it can move not only to sites  $i - 1$  and  $i + 1$  as in the AK model, but also back to site  $i$ . It chooses from any site in  $\{i - 1, i, i + 1\}$  with equal probability, provided none of the new configurations violates the constraints of Eqs. 7.2, 7.3, and 7.4. Thus, even if sites  $i - 1$  and  $i + 1$  are occupied, the leg has somewhere to go. We call this new spider gait the *rebinding gait*. It ensures that the probability that any particular leg will move is independent of the state of the rest of legs in the system, which simplifies the Monte Carlo simulation of the system.

Furthermore, the rebinding gait is more chemically realistic, as the enzymatic leg of a real molecular spider can always rebind to the site it just dissociated from, and its detachment should be independent of the state of the rest of the system.

From an analytical perspective, the effect of allowing rebinding is to slow the movement of walkers because of the potential for dissociations that do not move a spider leg. However, simulation results extended to very long times (Fig. 7.2) show that this change in effective rates does not qualitatively change the motion of a single spider. The rebinding gait leads to a constant-factor decrease in the mean squared displacement. In the remainder of the paper when comparisons of multi-spider systems with rebinding gait are made with single-spider systems, the single-spider systems have the normal AK gait, without rebinding. This gives



**Figure 7.2:** Estimate of  $\langle X^2 \rangle(t)$  for spiders with the AK gait and the rebinding gait with  $k = 2, s = 2$ , and  $r = 0.1$ .

them a constant factor advantage; however, as will be seen, even with this handicap the multi-spider systems are superior as transport devices.

## 7.2 Simulation Results for the Multi-Spider Model

The multi-spider system provides a simple model for cooperative transport using interacting walkers. In this application, walkers start at the origin of a 1D surface covered with energy-supplying substrate. The walkers move outwards in the plus and minus directions consuming substrate to bias their motion outwards, leaving an ever increasing sea of products ( $P \subset \mathbb{Z}$ ) in between the farthest sites visited in the plus and minus directions. Because spiders always cleave substrates when they detach from a site, and the spiders with  $s = 2, k = 2$  have a shuffling gait wherein they cannot hop over any substrates, the sea of products  $P$  is always contiguous and includes the origin. Thus, as explained in Sec. 4.5.1, there is a well



## 7.2. Simulation Results for the Multi-Spider Model

defined concept of a left ( $b_L$ ) and right ( $b_R$ ) boundary between the product sea and the unvisited substrates.

In the multi-spider model, the spiders fill this product sea, creating an exclusionary pressure that prevents the outermost spiders from moving past them. At any given time there will be one leftmost  $1 \leq L_s \leq S$  and one rightmost spider  $1 \leq R_s \leq S$ . We are interested in the location of these leading  $L_s$  and  $R_s$  spiders as the system evolves. When the Markov process is in state  $X$  as given by Eq. 7.1, the  $i$ -th spider's position is defined as the mean of  $F_i$ . We use a function  $\mu$  to describe the position of any spider  $1 \leq i \leq S$  as

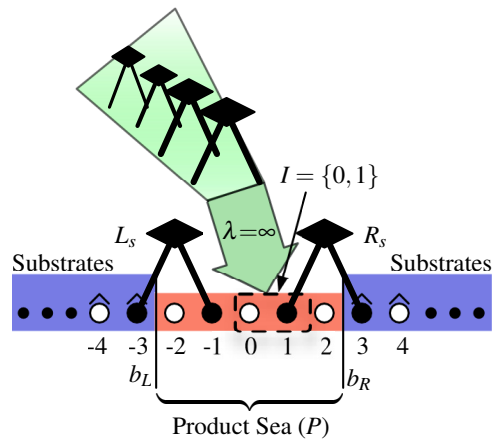
$$\mu(i) = \frac{\sum F_i}{k}. \quad (7.5)$$

Note that when  $k = 2$ ,  $s = 2$  as in the multi-spider model,  $\mu(i)$  only takes on half-integer values and the value of  $\mu(i)$  uniquely determines the value of  $F_i$ . This is in general not the case for larger values of  $k$  and  $s$ .

It is sometimes possible for the identities of the leftmost or rightmost spider to change. When  $\mu(L_s) > 2$ , all the spiders are to the right of the injection point  $I = \{0, 1\}$ , and when  $\mu(R_s) < 0.5$ , all the spiders are to the left of  $I$ . In these cases, when a new spider is injected at  $I$ , it becomes the new  $L_s$  or  $R_s$ , respectively. This situation becomes very unlikely as the number of spiders released increases; however, when reporting the MSD of  $L_s$  and  $R_s$ , these identity changes are important and are accounted for in our analysis.

### 7.2.1 Experimental Setup

To quantify the transport characteristics of the multi-spider injection system, we define a multi-spider system that begins with two spiders on either side of an injection site near the origin. The two initial spiders start on the boundary on



**Figure 7.3:** The initial configuration used for the multi-spider model simulations.

either side of an initial product sea (Fig. 7.3). The precise parameters studied are given in Table 7.1.

We use the Kinetic Monte Carlo (KMC) method [8] to numerically sample traces of the multi-spider Markov process. In Table 7.2 we show the number of runs of the Markov process sampled for each  $r$ -value and the minimum time each sample was run until.

**Table 7.1:** Parameters studied for the multi-spider model.

Parameter	Description
$k = 2$	number of legs
$s = 2$	maximum leg separation
$P = \{-2, -1, 0, 1, 2\}$	initial product sea
$S = 2$	initial number of spiders
$F_1 = \{1, 3\}$	initial location of rightmost spider
$F_2 = \{-3, -1\}$	initial location of leftmost spider
$I = \{0, 1\}$	injection site
$\lambda = \infty$	injection rate
$k_p^- = 1$	rate of detachment from products
$k_{\text{cat}} = r \leq 1$	rate of detachment from substrates

## 7.2. Simulation Results for the Multi-Spider Model

**Table 7.2:** Number of KMC runs for each parameter value and minimum simulation time combination.

$t_{\max}$	$r$ -value					
	1.0	0.5	0.1	0.05	0.01	0.005
$1.0 \times 10^7$	1800	1800	1800	1800	1800	1800
$1.0 \times 10^8$	200	-	200	-	-	-

The primary parameter of interest is the chemical kinetics parameter  $r$ . When  $r = 1$  there is no effective chemical difference between substrates and products, and hence no energy is available in the substrates to bias the outermost spiders' motion. When  $r < 1$  the chemical difference at the boundary acts to bias the motion of the leftmost spider ( $L_s$ ) and the rightmost spider ( $R_s$ ) away from the origin when they are at their respective boundaries. This creates effective superdiffusion for the leading spiders as long as they stay near the boundary.

Table 7.2 shows that we have many fewer simulation traces for  $t_{\max} = 10^8$  than for  $t_{\max} = 10^7$ . Indeed, these simulation counts are far fewer than the 5000 traces computed for the single spider model [39]. We were not able to compute more traces because simulation of the multi-spider model requires much more computational resources for large values of  $t_{\max}$  near  $10^8$ . Our KMC simulations of the spider systems consist of iterative computation of consecutive discrete events of the underlying Markov process. Every individual event takes a constant amount of computational (wall) time, but the simulation time intervals between events are exponentially distributed based on the total rate of all potential transitions from the current state. There are two possible transitions from every state of the single spider model: the left leg moves, or the right leg moves. Thus, the mean simulation time duration between events remains within  $1/2$  and  $1/(1+r)$  for every simulated step. Since  $r$  is a constant that does not depend on  $t_{\max}$ , the execution time of our single spider KMC algorithm is  $O(t_{\max})$ , i.e., it depends linearly on

the simulation time. However, in the multiple spider model the mean duration between events is not constant. When new spiders are injected, the number of different possible events in the system increases, and so the simulation time intervals between those events become smaller. Thus, to achieve desired maximum simulation time  $t_{\max}$  for the multi-spiders model, we need to simulate more discrete events than for the single spider model. In section 7.3.2 we estimate that the number of released spiders grows as  $O(\sqrt{t})$ . Hence, the execution time of the multi-spider KMC simulation algorithm is  $O(t_{\max}^{3/2})$ .

### 7.2.1.1 Observed Superdiffusion of the Leading Spiders

As discussed in Sec. 4.3, single spider systems with  $r < 1$  show transient superdiffusive behavior. Single spiders move faster than ordinary diffusion for a significant time and distance, but eventually slow down and move as an ordinary diffusion. Hence, the leading spiders ( $L_s$  and  $R_s$ ) of the multi-spider model also should move superdiffusively when they are near the boundary. This can be quantified by estimating the mean squared displacement of the leading spiders in the multi-spider model. Because the environment and the walker are symmetrical, we can, without loss of generality, represent the mean squared displacement of the outermost walkers by the position of the rightmost spider,

$$\langle X^2 \rangle(t) = \langle \mu(R_s)(t) \rangle. \quad (7.6)$$

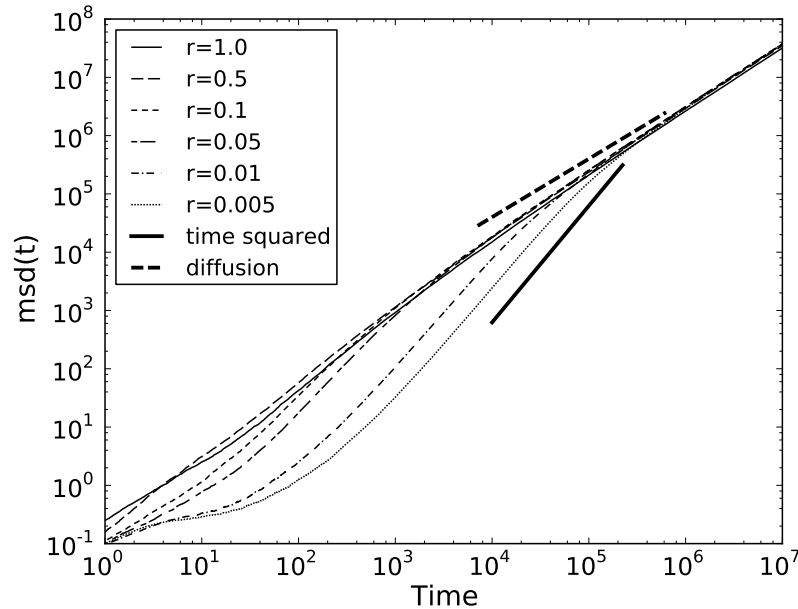
Figure 7.4 shows the KMC simulation estimate of  $\langle X^2 \rangle(t)$  for the multi-spider model on a log-log plot for each measured  $r$  parameter value. In this plot, straight lines correspond to power laws, that is, to Eq. 4.6, and the parameter  $\alpha$  is given by the slope. To show the instantaneous value of  $\alpha$ , we use finite difference methods to estimate  $\alpha(t)$  as in Eq. 4.7. Figure 7.5 shows the result of using the Savitzky-Golay smoothing filter [33] on these estimates of  $\alpha(t)$ .

## 7.2. Simulation Results for the Multi-Spider Model

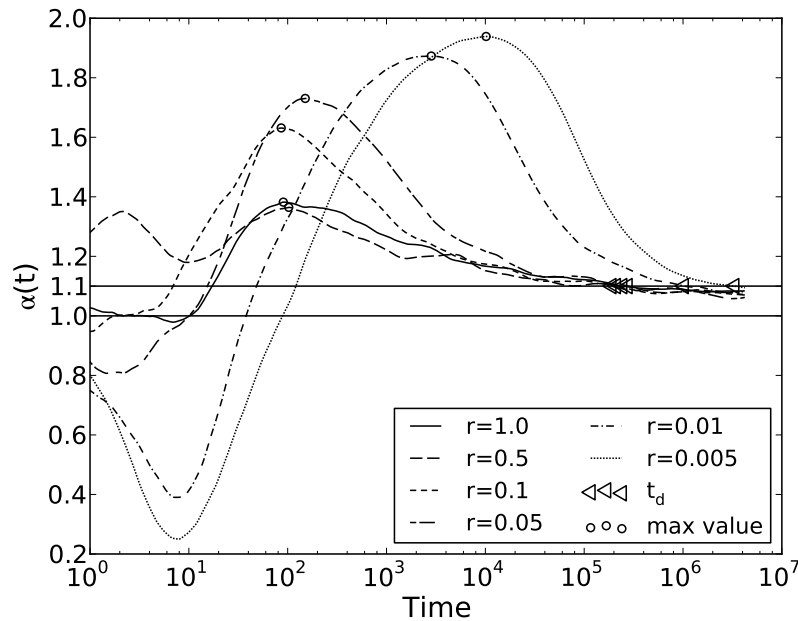
$r$	$k = 2$ spiders			$k = 2$ single-spider	$k = 1$ spiders
	1.0	0.05	0.005	0.05	1.0
$\alpha_{\max}$	1.38	1.73	1.93	1.43	1.41
$t_{\alpha_{\max}}$	$9.03 \times 10^1$	$1.50 \times 10^2$	$1.02 \times 10^4$	$2.17 \times 10^2$	$2.71 \times 10^1$
$t_d$	$2.43 \times 10^5$	$1.89 \times 10^5$	$3.31 \times 10^6$	$2.52 \times 10^4$	$1.14 \times 10^5$

**Table 7.3:** Properties of the MSD and the superdiffusive regime defined by  $\alpha(t) > 1.1$  for  $R_s$ . Results are compared for the multi-spider model as well as the single spider model (Sec. 7.2.2) and the  $k = 1$  spiders which are simple random walkers (Sec. 7.4).

These results show the same three-phase behavior as observed in single spider simulations (Sec. 4.3). The phases can be observed by noting the estimate of the value of  $\alpha(t)$  in relation to the horizontal line in Fig. 7.5 representing  $\alpha(t) = 1.0$ . Below this line, the leading spiders are moving subdiffusively and above this line the leading spiders are moving superdiffusively. For each value of  $r$ , the systems exhibit the following three sequential phases:



**Figure 7.4:** Mean squared displacement for  $R_s$ . Reference lines are shown for ordinary diffusion and ballistic motion.



**Figure 7.5:** Smoothed finite difference approximation of  $\alpha(t)$  for  $R_s$ . Horizontal lines define the threshold for ordinary diffusion at  $\alpha = 1.0$  and our defined threshold for superdiffusion at  $\alpha = 1.1$ .

- 1) The *initial* phase is defined when  $t < 1/r$ . At these times, very few steps have been made, and the value of  $\alpha(t)$  is largely dependent on peculiarities of the initial configuration, and not of relevance to transport phenomena.
- 2) The *superdiffusive* phase begins when  $t > 1/r$  and continues while the  $\alpha(t) \geq 1.1$ . During this phase the spiders move significantly faster than diffusion. Decreasing values of  $r$  lead to increasing maximum values of  $\alpha(t)$ , and a longer time until the spider returns to the  $\alpha(t) = 1.1$  threshold.
- 3) The *diffusive* phase begins when the instantaneous value of  $\alpha(t) = 1.1$ , and continues indefinitely, as even the leading spiders eventually spend almost all of their time diffusing over the products instead of cleaving new sites at the boundary.

Thus, even though a multi-spider system adds spiders at the origin as fast as

## 7.2. Simulation Results for the Multi-Spider Model

possible to prevent the leading spiders from moving too far backwards into the product sea, all multi-spider systems eventually decay to diffusion. This is the same qualitative behavior exhibited by single spider systems. However, multi-spider systems move superdiffusively over significantly longer times, and even when  $r = 1.0$ . They also reach a higher peak value of  $\alpha(t)$ . The superdiffusive properties of the multi-spider model can be quantified by examining the following:

- $\alpha_{\max} = \max_{t \geq 0} (\alpha(t))$ , the peak instantaneous value of  $\alpha(t)$ , which should satisfy  $1 \leq \alpha_{\max} \leq 2$ ;
- $t_{\alpha_{\max}} = \operatorname{argmax}_{t \geq 0} \alpha(t)$ , the time at which the peak of  $\alpha(t)$  is reached;
- and  $t_d$ , the time at which  $\alpha(t)$  drops below the threshold of 1.1, and enters the diffusive phase.

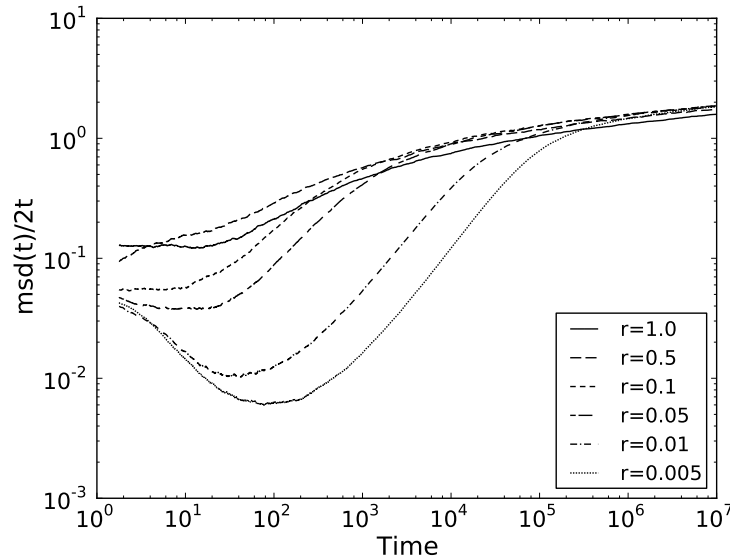
The estimates of these quantities are given in Table 7.3. The results show that  $\alpha_{\max}$  and  $t_d$  generally increase with decreasing  $r$ . Furthermore, the walkers with  $r = 0.005$  have peak  $\alpha(t)$  values above 1.9 and remain superdiffusive over times more than 6 orders of magnitude larger than the mean leg-product residence time.<sup>1</sup> Hence, for finite distances and times relevant to most transport processes the multi-spider model can achieve significant superdiffusive motion, and with small  $r$  values becomes nearly ballistic for significant spans of time and distance.

A further method of characterizing the transport behavior of multi-spider systems is to look at their effective instantaneous diffusion rate which (in 1D) is defined as

$$\tilde{D}(t) = \langle X^2 \rangle(t) / 2t. \quad (7.7)$$

---

<sup>1</sup>The value of  $k_p^-$  is a free parameter in the model. We choose time units so that rate  $k_p^-$  is normalized to 1, hence all time units are measured relative to  $1/k_p^- = 1$ .



**Figure 7.6:** Effective diffusion rate ( $\tilde{D}(t)$ ) for  $R_s$ .

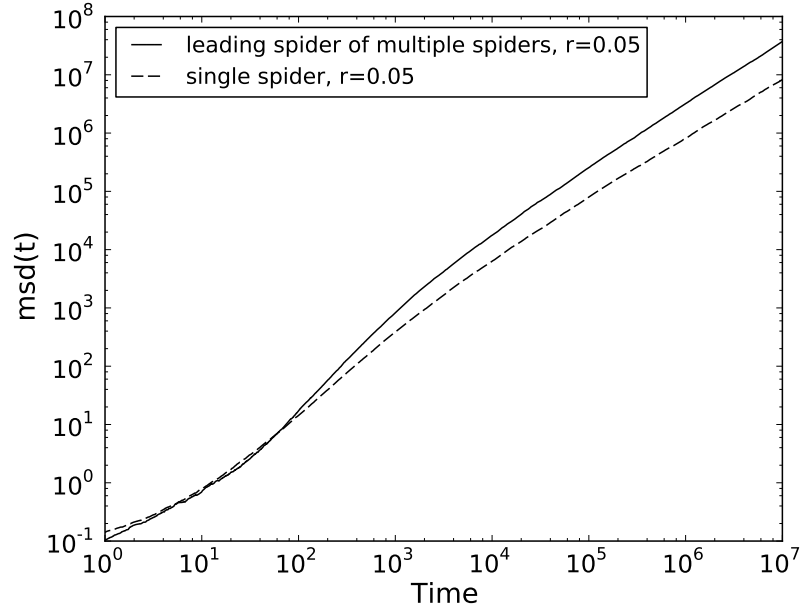
This is arrived at by setting  $\alpha = 1$  in Eq. 4.6. The value of  $\tilde{D}(t)$  can be thought of as the diffusion rate a simple diffusing particle would need to have the same mean squared displacement as the spider system at time  $t$ . Hence, greater values of  $\tilde{D}(t)$  correspond to faster transport systems. Figure 7.6 shows  $\tilde{D}(t)$  for the multi-spider systems. While initially the spiders with lower  $r$  have larger  $\tilde{D}$  values, eventually the spiders with the smallest  $r$  values are superior.

## 7.2.2 Comparison with Single Spiders

The multi-spider systems can be directly compared with the single spiders to understand exactly how useful the additional interior spiders are for transport. Figure 7.7 shows the results of comparing the  $\langle X^2 \rangle(t)$  for a single spider, and the leading spider of the multi-spider model, both with  $r = 0.05$ . There is a significant transport advantage for the multi-spider system. Furthermore, Fig. 7.8 shows the estimate of  $\alpha(t)$  for these two systems, and the leading spider of the multi-spider

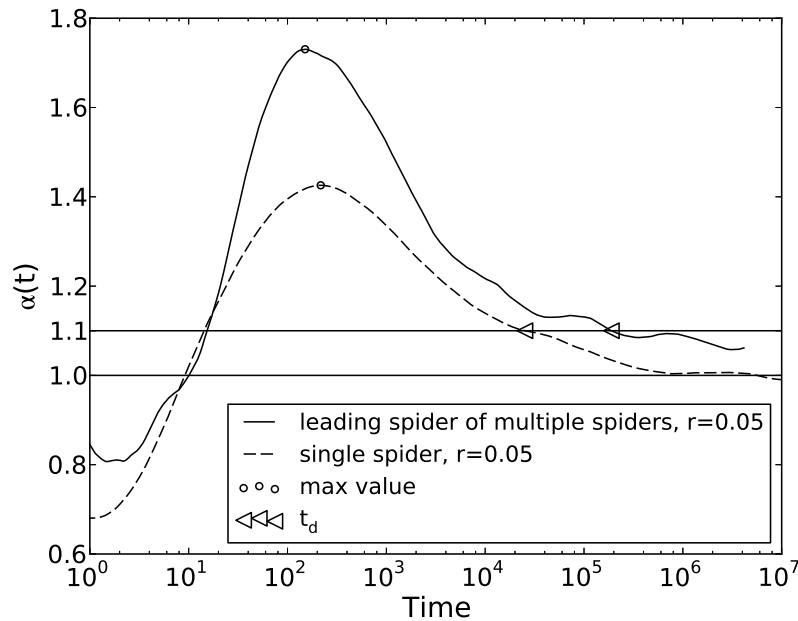


## 7.2. Simulation Results for the Multi-Spider Model



**Figure 7.7:** Mean squared displacement for the leading spider  $R_s$  in the multi-spider model versus a single AK spider.

system maintains a higher value of  $\alpha(t)$  at all times, and significantly longer superdiffusive period. The values of  $\alpha_{\max}$  and  $t_d$  are summarized in Table 7.3, and the multi-spider system is superior in both measurements. Finally, we compare the first passage times for the single versus multi-spider models in Fig. 7.9. The mean first passage time  $\langle \text{fpt}(d) \rangle$  is the average time for the leading spider to first visit a site at a distance  $d$  from the origin. At large times the multi-spider systems have a large advantage in this key transport statistic. Overall, the comparison with the single AK spider shows that for distances less than 3000 sites from the origin the leading spider of the multi-spiders model reaches unvisited sites up to 5.25 times faster than a single AK spider.

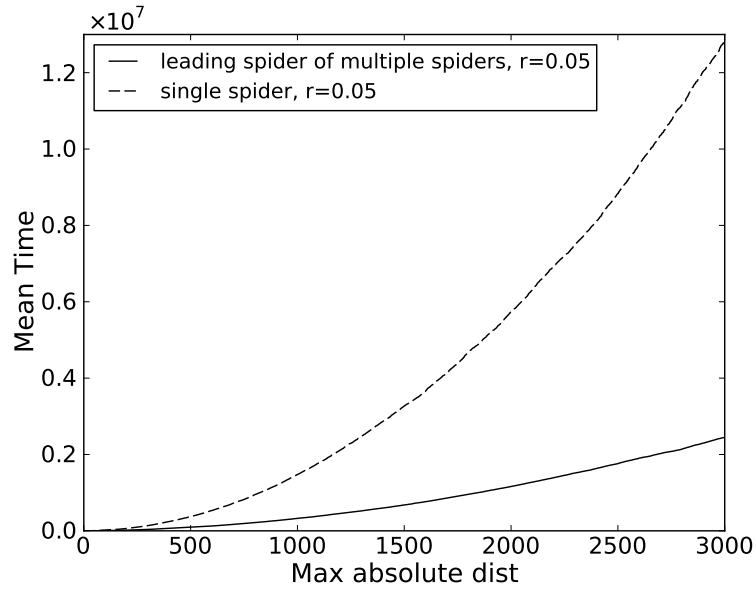


**Figure 7.8:** Finite difference approximation of  $\alpha(t)$  for the leading spider  $R_s$  in the multi-spider model versus a single AK spider.

### 7.3 Asymptotic Diffusion in the Multi-spider Model

In comparison with a single spider, the simulation results in Sec. 7.2 show that multi-spider systems exhibit larger values for  $\tilde{D}$ ,  $\alpha_{\max}$ , and  $t_d$ —all essential measures of transport efficiency. However, multi-spider systems still eventually decay to diffusion despite the unlimited supply of spiders injected at the origin. By analogy with the single spider model (Sec. 4.1), this implies that the effective size of the product sea as seen by the leading spiders is not constant. As the effective product sea grows, the leading spiders spend progressively less time at the boundary in the  $B$  state where they move ballistically, and more time in the diffusive  $D$  state where they move over the product sea. This leads to a value of  $\alpha(t) \rightarrow 1.0$ , as  $t \rightarrow \infty$ , and the leading walkers are effectively diffusing. The origins of this effect can be understood in detail by examining the effective size of the product sea (Sec. 7.3.1), the number of released spiders (Sec. 7.3.2), and their spatial distribu-

### 7.3. Asymptotic Diffusion in the Multi-spider Model



**Figure 7.9:** Comparison of the first passage time of the leading spider  $R_s$  in the multi-spider model versus a single AK spider.

tion (Sec. 7.3.3).

#### 7.3.1 Effective size of the product sea

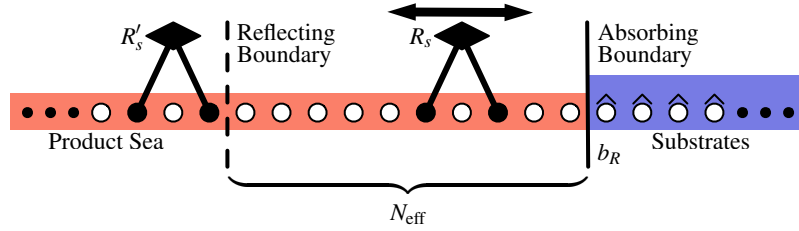
In 1D, spiders cannot move past each other, and thus the motion of the leading spiders  $R_s$  and  $L_s$  is constrained by the presence of their neighboring spiders. In particular when  $S \geq 3$  we can define the next-leading spiders,  $R'_s$  and  $L'_s$  as

$$R'_s = \max_{\substack{1 \leq i \leq S \\ i \neq R_s}} (\mu(F_i)) \quad (7.8)$$

and

$$L'_s = \min_{\substack{1 \leq i \leq S \\ i \neq L_s}} (\mu(F_i)). \quad (7.9)$$

When  $S \geq 4$ ,  $L'_s \neq R'_s$ , and the remaining  $S - 4$  spiders are called the interior spiders. The importance of the next-leading spiders,  $R'_s$  and  $L'_s$ , is that they define



**Figure 7.10:** The effective size of the product sea ( $N_{\text{eff}}$ ) as seen by leading spider  $R_s$  is the area between next-leading spider  $R'_s$  and the right boundary  $b_R$ .

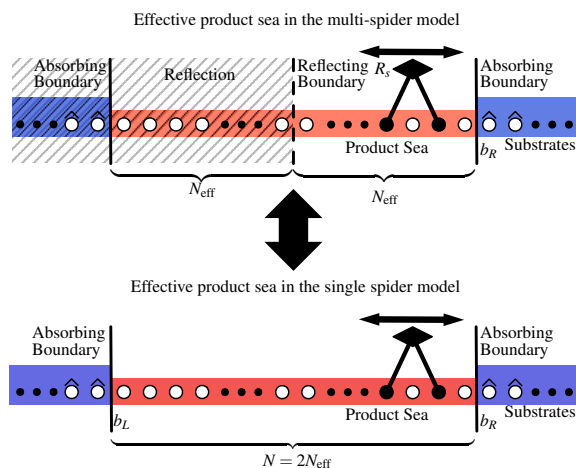
the effective size of the product sea as seen by the leading spiders. Without loss of generality we focus only on the rightmost spiders  $R_s$  and  $R'_s$  and define

$$\langle N_{\text{eff}}(t) \rangle = \langle \max(F_{R'_s}(t)) + 1 - b_R(t) \rangle, \quad (7.10)$$

where the system state at time  $t$  is given by  $X(t)$  from Eq. 7.1, and  $b_R(t)$  is the right boundary for that state as defined in Eq. 4.10. Figure 7.10 illustrates what  $N_{\text{eff}}$  is for a particular state  $X$ . The choice of name for  $\langle N_{\text{eff}}(t) \rangle$  is meant to correspond to Antal and Krapivsky's  $\langle N(t) \rangle$  [4], which is the mean number of sites cleaved by a single spider system at time  $t$ . It has been shown for single spiders that  $\langle N(t) \rangle = \Theta(\sqrt{t})$  [4], and that this implies that the time to leave the diffusive metastate  $\langle \tau_D(t) \rangle = \Theta(\sqrt{t})$  (Sec. 4.5.2). Since  $\langle \tau_D(t) \rangle$  grows with time, and  $\langle \tau_B(t) \rangle$  does not, single spiders eventually spend almost all their time diffusing in the  $D$  state and hence their motion is asymptotically diffusive.

The relation of  $\langle N(t) \rangle$  with  $\langle \tau_D(t) \rangle$  is based on the mathematics of the mean time for a random walker to escape an interval of size  $N$  with two absorbing boundaries. This gives the time for a walker to leave the product sea and return to the boundary. In the case of the multi-spider system,  $N_{\text{eff}}$  is also meant to represent the size of the region of products to escape from; however, it has one reflecting boundary at  $\max(F_{R'_s})$  and one absorbing boundary at  $b_R$ . The problem of escape from a region with one reflecting and one absorbing boundary is equivalent to escape from a region of twice the size with two absorbing boundaries

### 7.3. Asymptotic Diffusion in the Multi-spider Model



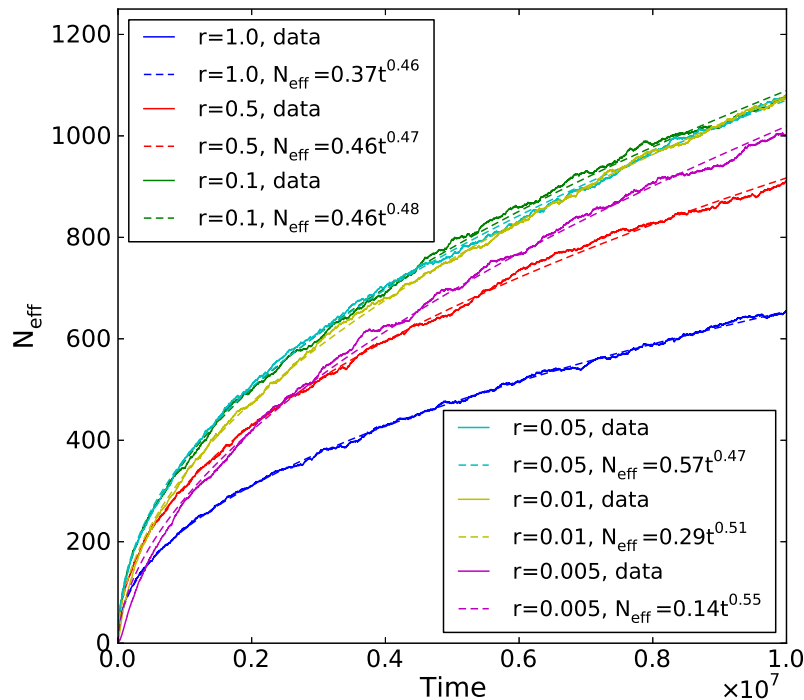
**Figure 7.11:** The problem of escape from an area of size  $N_{\text{eff}}$  with one reflecting and one absorbing boundary is equivalent to the problem of escape of a single spider from a region of size  $N = 2N_{\text{eff}}$  with two absorbing boundaries.

(Fig. 7.11). Hence, we note the relation:

$$\langle N_{\text{eff}}(t) \rangle = \langle N(t) \rangle / 2. \quad (7.11)$$

Further analysis of this relationship is deferred to Sec. 7.5. However, the importance of  $N_{\text{eff}}$  can be understood simply—if  $\langle N_{\text{eff}}(t) \rangle$  increases with time, then because  $\langle \tau_D(t) \rangle = \Theta(\langle N_{\text{eff}}(t) \rangle)$ , the leading spider  $R_s$  must eventually move diffusively. This is indeed the case for all values of  $r$ , as shown in Fig. 7.12. In this figure we used the Levenberg-Marquardt algorithm [33] to fit power laws to the estimates of  $\langle N_{\text{eff}}(t) \rangle$ . We see that, interestingly, the exponents are close to 0.5—exactly as with single spiders.

Thus,  $\langle \tau_D(t) \rangle$  for  $R_s$  also grows with time and therefore leads to asymptotic diffusion. However, in the multiple spider model the effective size of the product sea is much smaller than the total number of products ( $\langle N_{\text{eff}}(t) \rangle \ll |P(t)|$ ) because most of the product sea is filled with other spiders, whereas in the single spider model  $\langle N(t) \rangle = |P(t)|$ . Thus,  $R_s$  in the multi-spider model remains superdiffusive for much higher values of  $|P|$  than a single spider does. Fig. 7.13



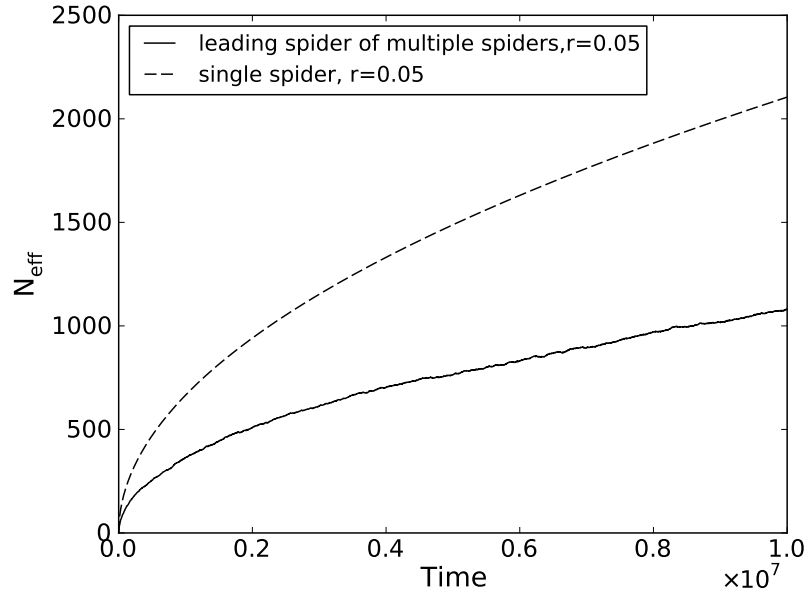
**Figure 7.12:** The effective size of the product sea  $\langle N_{\text{eff}}(t) \rangle$  grows with time, and hence leads to asymptotically diffusive motion for the multi-spider model.

compares the effective size of the product sea of the multi-spider model with the number of products of the AK model. While both  $\langle N(t) \rangle$  and  $\langle N_{\text{eff}}(t) \rangle$  grow with time, the single spider sees a much larger product sea, which explains the results of Sec. 7.2.2.

### 7.3.2 Number of Released Spiders

With injection rate  $\lambda = \infty$ , spiders in the multi-spider model are released at  $I = \{0, 1\}$  whenever possible, yet the presence of other spiders in the sites 0 and 1 prevents release, and so keeps the total number of spiders,  $S$ , finite. Hence,  $S(t)$  is a random variable giving the number of spiders released by time  $t$ . Figure 7.14 shows our estimates for  $\langle S(t) \rangle$ , for each studied value of  $r$ . Again we used the Levenberg-Marquardt algorithm to fit power laws, and observe approximate ex-

### 7.3. Asymptotic Diffusion in the Multi-spider Model

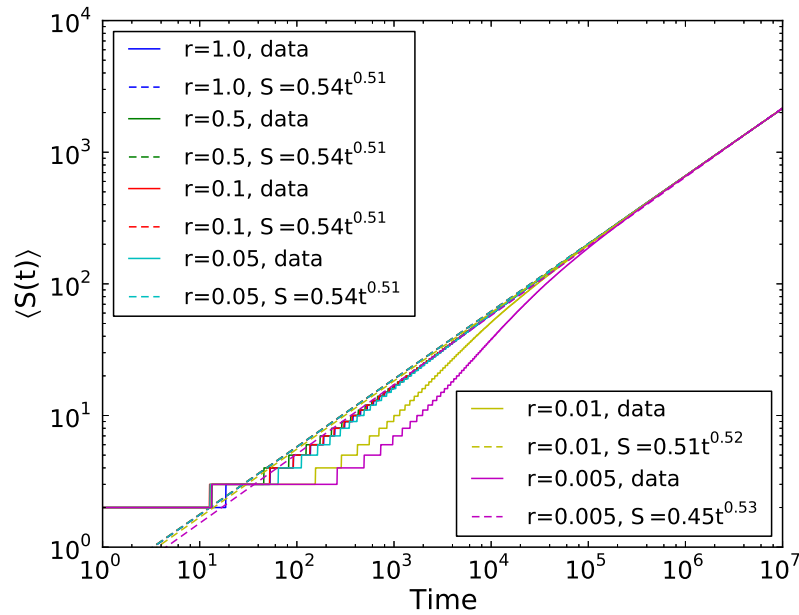


**Figure 7.13:** The effective size of the product sea ( $\langle N_{\text{eff}}(t) \rangle$ ) as seen by  $R_s$  in the multi-spider model versus the effective size of the product sea ( $\langle N(t) \rangle$ ) for a single walker in the AK model.

ponents of 0.5. Interestingly,  $\langle S(t) \rangle$  appears to be dependent on  $r$  only initially, whereas at later times the values of  $\langle S(t) \rangle$  for all  $r$  are nearly identical. Thus, the release of spiders, which occurs at the origin and away from the substrates at the boundaries, is independent of  $r$ , which limits the total number of released spiders regardless of how fast the leading spiders move.

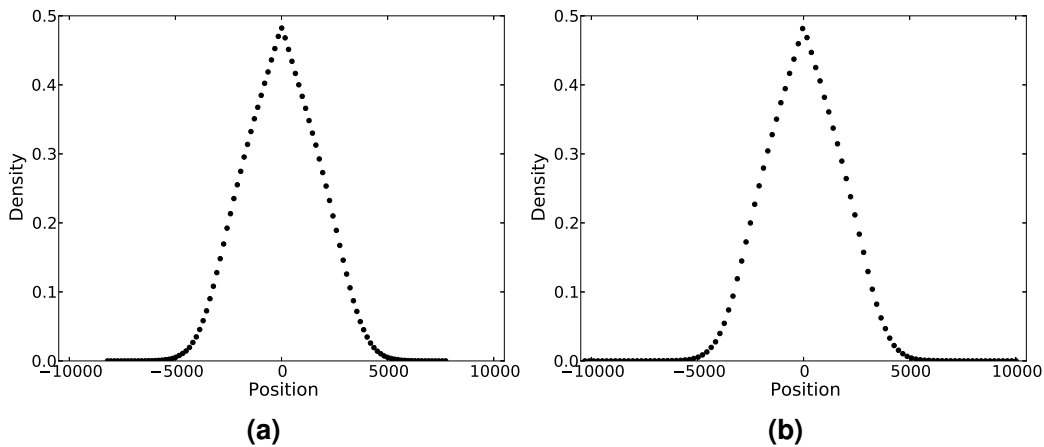
#### 7.3.3 Spatial Distribution of Spiders

Until now we have focused mainly on the position of the leading spiders  $R_s$  and  $L_s$ , but the behavior of the interior spiders controls the release of spiders at the origin and the effective product sea size near the boundary. At any time  $t$ , we measure the density of spiders using a histogram with 100 equally spaced bins over the maximum positions obtained by any spider in any simulation trace at that time. Each bin with  $n$  sites can contain at most  $n/2$  spiders, hence the maximum



**Figure 7.14:** Mean number of released spiders,  $\langle S(t) \rangle$ ; the log-log scale is necessary to exhibit the differences at early times.

density for each bin is 0.5. Figure 7.15 shows the density of spiders at  $t_{\max} = 10^7$  for  $r = 1$  and  $r = 0.05$ . The density of spiders near the origin is approximately the same in both cases and is nearly equal to the 0.5 maximum possible density. This



**Figure 7.15:** Mean spider density at  $t_{\max}$ , for  $r = 1$  (a), and  $r = 0.05$  (b).



### 7.3. Asymptotic Diffusion in the Multi-spider Model

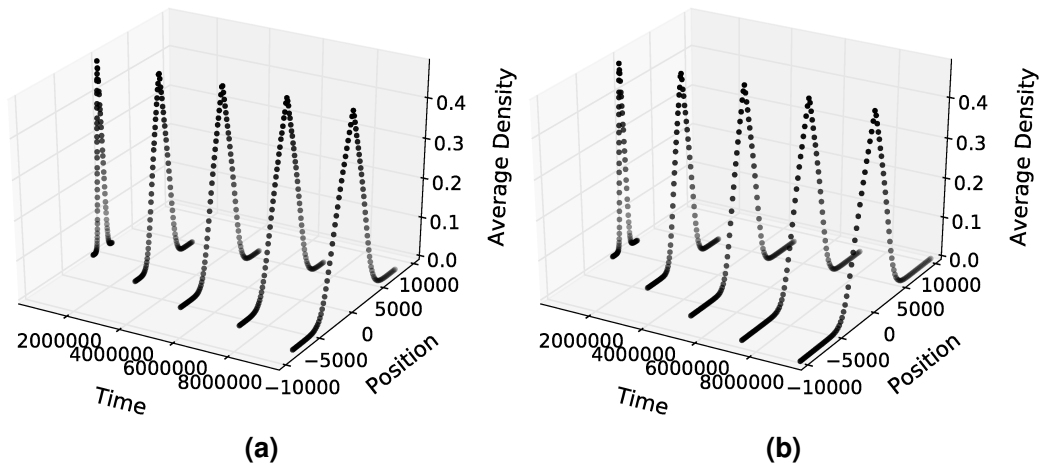
explains why  $\langle S(t) \rangle$  for large  $t$  is nearly the same for all  $r$  values. The sites near the origin are very densely packed for any  $r$ -value and the passive addition of spiders is not presented with many opportunities to add new spiders even when the injection rate is infinite.

The density of spiders falls off nearly linearly away from the origin, until approximately distance 4000 (at  $t_{\max} = 10^7$ ), where densities for both  $r$  values gradually transition to long tails with nearly 0 density. Essentially, the only difference between the densities for  $r = 1$  and  $r < 1$  is at the tails. The tails for  $r = 0.05$  are much longer, corresponding to the greater  $\langle X^2 \rangle(t_{\max})$  value observed for this  $r$ -value. This is to be expected as only the leading spiders  $R_s$  and  $L_s$  ever see a substrate, and all other spiders only move over products. Thus the rate  $r$  only affects the motion of the leading spiders and those spiders near to the leading spiders that have comparatively more space to diffuse in. The vast majority of the interior spiders are too far from the leading spiders to see the effects of the  $r$ -value. Hence, the distribution of spiders away from the boundaries is nearly identical for  $r = 1$  and  $r < 1$ .

This similarity in densities of interior spiders is present at all times, as can be seen in Fig. 7.16, which shows the evolution of these densities with time. For both  $r = 1$  and  $r = 0.05$ , the densities drop off linearly around the origin, until a critical point where they transition gradually to the near-zero density tails. The tails of the  $r = 0.05$  spider density remain longer as expected based on the MSD results of Fig. 7.4, but the evolution of the interior spider density is nearly independent of  $r$ .

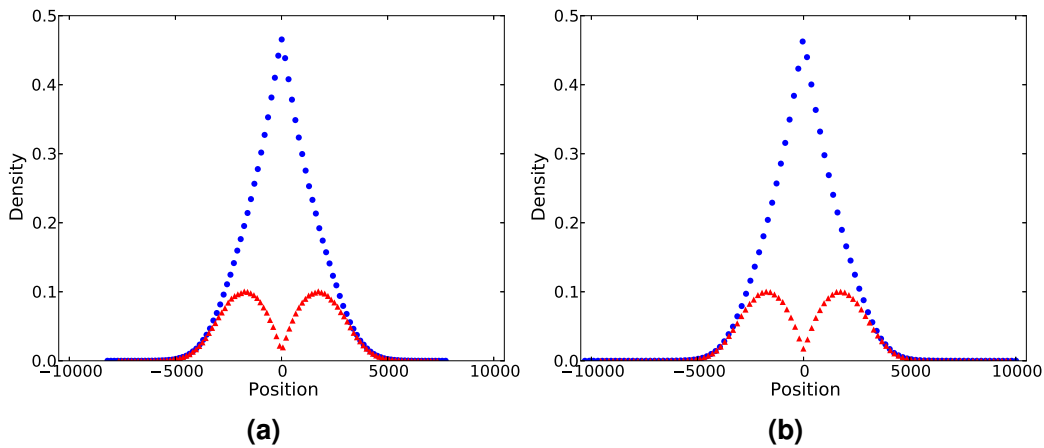
#### 7.3.3.1 Distribution of Spider Strides

Spiders with  $k = 2, s = 2$  only have two possible foot patterns: either both of their legs are together, or both are apart (Sec. 5.1). The together and apart patterns can



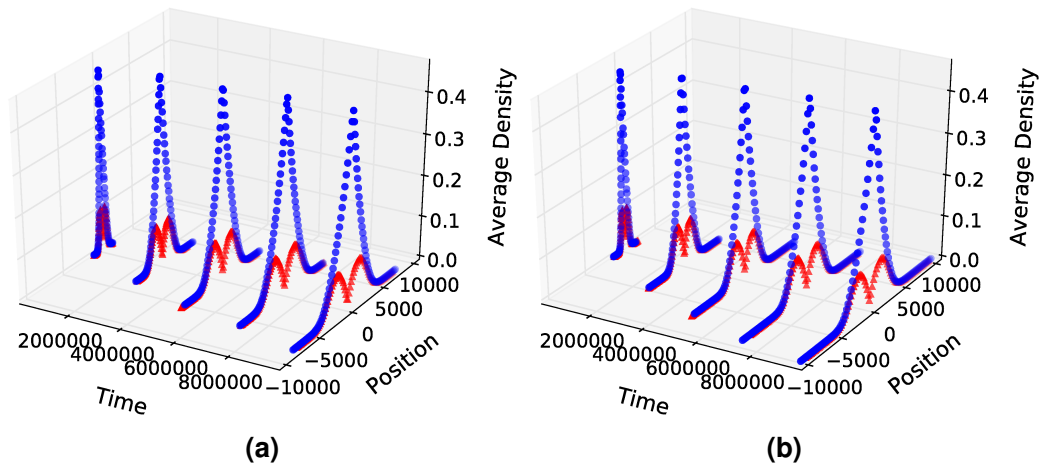
**Figure 7.16:** Evolution of mean spider density through time, for  $r = 1$  (a) and for  $r = 0.05$  (b).

be called the two possible *strides* of a spider. A single spider is equally likely to be in either of the two strides. However, the spiders in the multi-spider model show a curious distribution of leg patterns. Figure 7.17 shows the distribution of spiders in each of the two strides at time  $t_{\max}$ , and again the results are remarkably similar for both  $r = 1$  and  $r < 1$ . The high density of spiders near the origin leads to a



**Figure 7.17:** Mean density of spiders in each of the together (blue circles) and apart (red triangles) strides at  $t_{\max} = 10^7$ , for  $r = 1$  (a) and for  $r = 0.05$  (b).

### 7.3. Asymptotic Diffusion in the Multi-spider Model



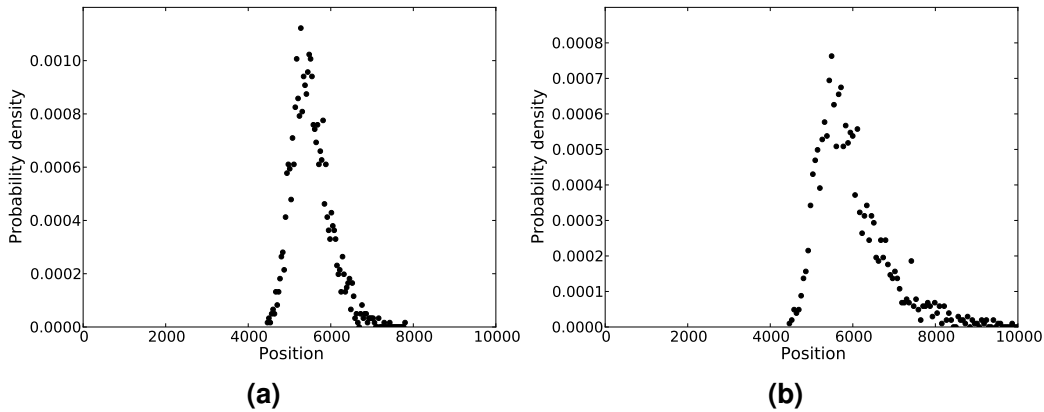
**Figure 7.18:** Average density for spiders with legs together (blue circles) and apart (red triangles) plotted at several instants, for  $r = 1$  (a) and for  $r = 0.05$  (b).

very strong bias towards the together stride.

This can be considered an emergent phenomenon that arises as a means to increase spider packing close to the injection source. Spiders in this high density region rarely get an opportunity to spread their legs into the apart stride because both neighboring sites are almost always occupied. Also of note, at the distances where the linear decrease in spider density is no longer apparent, the distributions of strides becomes equal again. This equality of stride distribution indicates that the spiders are no longer experiencing the extreme exclusionary pressure observed near the injection site. Instead, the spiders on the periphery are able to act more like single spiders, which have an equal distribution in strides. These effects are again apparent at all time scales, as seen in Figure 7.18.

#### 7.3.3.2 Density of leading spiders

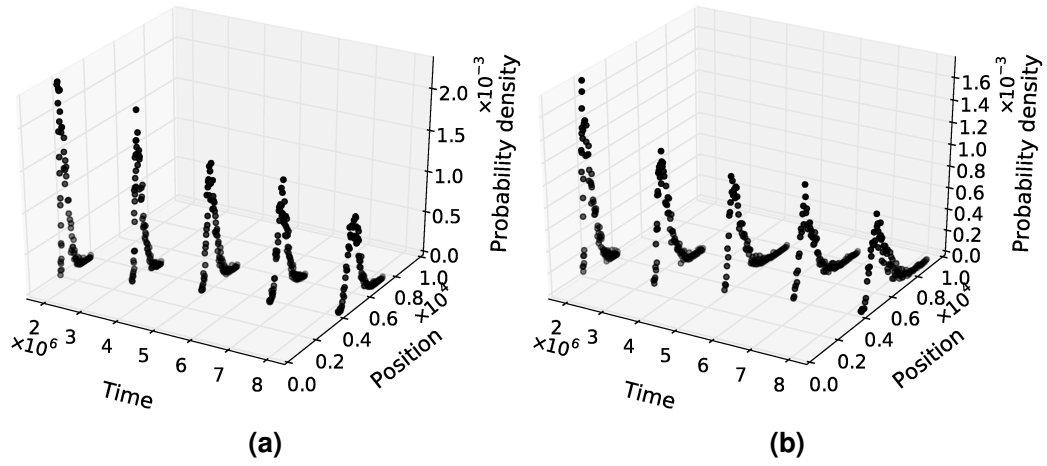
Unlike the interior spiders, which never see a substrate, the leading spiders  $L_s$  and  $R_s$  are strongly affected by the enzymatic rate  $r$ . When spiders are in the



**Figure 7.19:** Average density of the leading spider at  $t_{\max}$ , for  $r = 1$  (a) and for  $r = 0.05$  (b).

boundary metastate (Sec. 4.5.1), they move ballistically away from the origin, and the smaller the value of  $r$ , the less chance they have of exiting the boundary state and returning to the diffusive  $D$  state. Figures 7.19 and 7.20 show the probability distribution of  $\mu(R_s)$  (the leading spiders location) for both  $r = 1$  spiders and  $r = 0.05$  spiders. Particularly at the shorter times in Figure 7.20, there is a distinct difference in the distributions shape, with the  $r < 1$  distributions having much longer tails, and distinctly non-Gaussian shape. The mean of the distributions is the  $\langle X^2 \rangle(t)$ , reported in Sec. 7.2.1.1, which grows much faster for the  $r < 1$  multi-spider systems than for  $r = 1$ . However, the complete distributions shown in Figures 7.19 and 7.20 reveal more information, particularly that the tails of the distribution are much shorter on the left than the right. This arises from the exclusionary pressures exerted by the next leading spider  $R'_s$ . However, the results of Sec. 7.3.1 show that despite this exclusionary pressure, the distance between the two leading spiders  $\langle N_{\text{eff}}(t) \rangle$  continues to increase as  $\sqrt{t}$  regardless of the value of  $r$ . Hence, in the distributions at  $t_{\max}$  in Fig. 7.19, there is less distinction between the  $r = 1$  and  $r < 1$  walkers.

### 7.4. Importance of Multi-Pedal Gaits for Transport



**Figure 7.20:** Average spider density of the leading spider plotted at several instants, for  $r = 1$  (a) and for  $r = 0.05$  (b).

## 7.4 Importance of Multi-Pedal Gaits for Transport

Many molecular walkers, including the natural motors kinesin and dynein, are multivalent—they have two (or more) attachment sites. Interestingly, it has been shown that in the AK spider model, the superdiffusive effects are only present when the number of legs  $k \geq 2$ . Without the constraints imposed by multiple legs the residence time-bias at the boundary when  $r < 1$  does not lead to a bias in motion towards substrates.

The multi-spider model has two potential sources of bias to cause superdiffusive motion of the leading spiders: the residence time bias at the boundary when  $r < 1$ , and the effective bias caused by the exclusionary pressure of the interior spiders. We know that the multi-spider systems are transiently superdiffusive even when  $r = 1$  due to the exclusionary pressure, but what happens to their collective behavior when they have only a single leg? In fact when  $k = 1$  and  $r = 1$  an AK spider is equivalent to an ordinary random walker that moves left and right with rate 1. Thus, we measured  $\langle X^2 \rangle(t)$  for the leading walker of the multi-spider model with  $k = 1$  and  $r = 1$ . The values for  $\langle X^2 \rangle(t)$  are compared with the

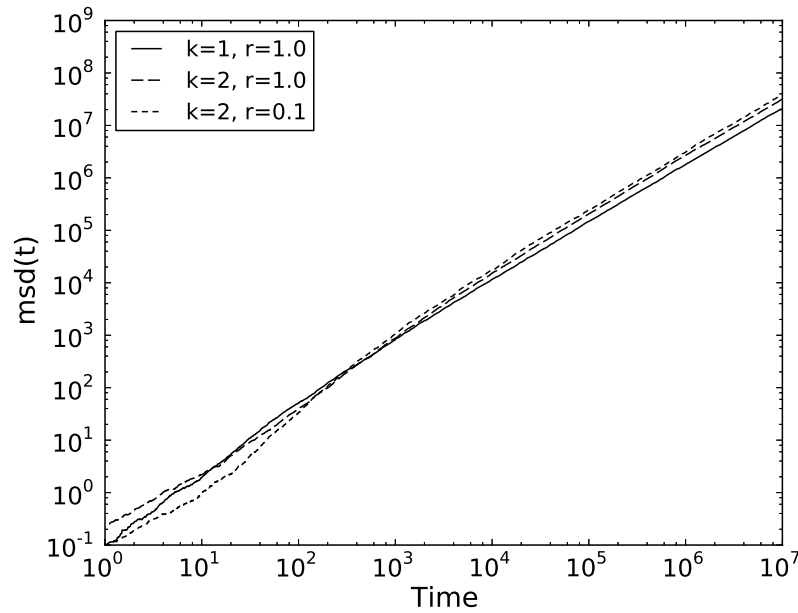
multi-spiders model with  $k = 2$  and  $r \in \{1.0, 0.1\}$ , and shown on a log-log scale in Fig. 7.21, and the corresponding values of  $\alpha(t)$  are shown in Fig. 7.22. The  $k = 1$  walkers do exhibit transient superdiffusive behavior, but their values of  $\alpha_{\max}$  and  $t_d$  are surpassed by  $k = 2$  spiders when  $r < 0.1$ , as summarized in Table 7.3. Spiders with  $k = 1$  achieve maximum  $\langle X^2 \rangle(t)$  when  $r = 1$  [4, 5], so by using  $r = 1$  in our comparison we are comparing with the most efficient single-legged spiders possible.

It is, however, necessary to make an adjustment of scales to correctly compare the  $\langle X^2 \rangle(t)$  values between  $k = 2$  and  $k = 1$  walkers. Since the position of a spider is defined as the mean of its attached leg positions, the  $k = 2$  spiders move by only distance 0.5 with each step, in contrast to the  $k = 1$  spiders which move by distance 1. Thus, in the analysis of  $k = 1$  spiders shown in Figures 7.21 and 7.22, the  $k = 1$  spiders move over a lattice with site spacing 0.5. In essence this correction can be thought of as adjusting the diffusion constant of the  $k = 1, r = 1$  walkers which have  $D = 1$  to that of the  $k = 2, r = 1$  walkers which have  $D = 0.5$ .

## 7.5 Analysis of Maximum Product Sea Size

Simulation results presented in Sec. 7.2.1.1 suggest that the leading spider  $R_s$  moves diffusively in the long time limit with the value of  $\alpha(t_{\max}) \approx 1$ . Section 7.3.1 showed that this happens because the mean effective size of the product sea as seen by  $R_s$ ,  $\langle N_{\text{eff}}(t) \rangle$ , grows with time approximately as  $\sqrt{t}$ . Hence, the duration of the  $D$  states  $\langle \tau_D(t) \rangle$  also grows with time, leading to asymptotically diffusive motion. Furthermore the effective product sea size  $N_{\text{eff}}$  can also be understood as being a function  $N$ , the number of sites cleaved. Figure 7.23 shows simulation estimates for  $\langle N_{\text{eff}}(N) \rangle$ , which at times close to  $t_{\max}$  is almost linear. Thus, while the leading spider is cleaving sites at the boundary, the interior spiders are not fol-

## 7.5. Analysis of Maximum Product Sea Size



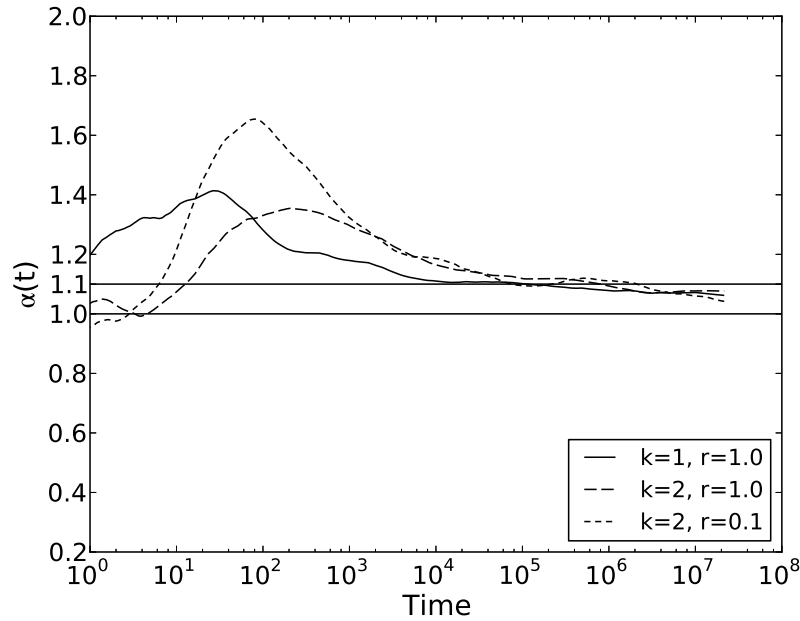
**Figure 7.21:** Comparison of  $\langle X^2 \rangle(t)$  for the leading spider  $R_s$  in multi-spider simulations versus the  $k = 1$  multi-spider model (with corrected diffusion constant of  $D = 0.5$ .)

lowing closely enough and the leading spider sees an increasingly large effective product sea. If a multi-spider system were to keep the leading spider superdiffusive as  $t \rightarrow \infty$ , it would have to ensure that  $\langle N_{\text{eff}}(N) \rangle$  does not grow too fast. The asymptotic bound that ensures this property can be found analytically.

The expected exit time for a random walker from an interval  $(0, M)$  with two absorbing boundaries at 0 and  $M$  is

$$\langle T_e(x) \rangle = \frac{x(M-x)}{2}, \quad (7.12)$$

where  $x$  is the starting position of the walker. As shown in Fig. 7.11 and explained in Sec. 7.3.1, the expected exit time from an interval with one absorbing and one reflecting boundary is the same as exit time from a interval with two absorbing boundaries of twice the size. Furthermore, when a spider moves over a region of product sites, its body position  $\mu(F)$  moves like a simple random walker with a step size of  $1/2$ .



**Figure 7.22:** Comparison of finite difference approximation of  $\alpha(t)$  for the leading spider  $R_s$  in multi-spider simulations versus the  $k = 1$  multi-spider model (with corrected diffusion constant of  $D = 0.5$ .) Horizontal lines show the threshold for ordinary diffusion at  $\alpha = 1$  and our defined threshold for superdiffusion at  $\alpha = 1.1$ .

When at time  $t$  the  $R_s$  spider moves off the boundary and into the  $D$  state, it enters a product sea of expected size  $\langle N_{\text{eff}}(N) \rangle$  with one absorbing and one reflecting boundary. The expected time to exit this interval is the expected duration of the  $D$ -state,  $\tau_D$ . This can be found by using Eq. 7.12 with  $M = 4\langle N_{\text{eff}}(N) \rangle - 5$  and  $x = 4\langle N_{\text{eff}}(N) \rangle - 8$ , which gives

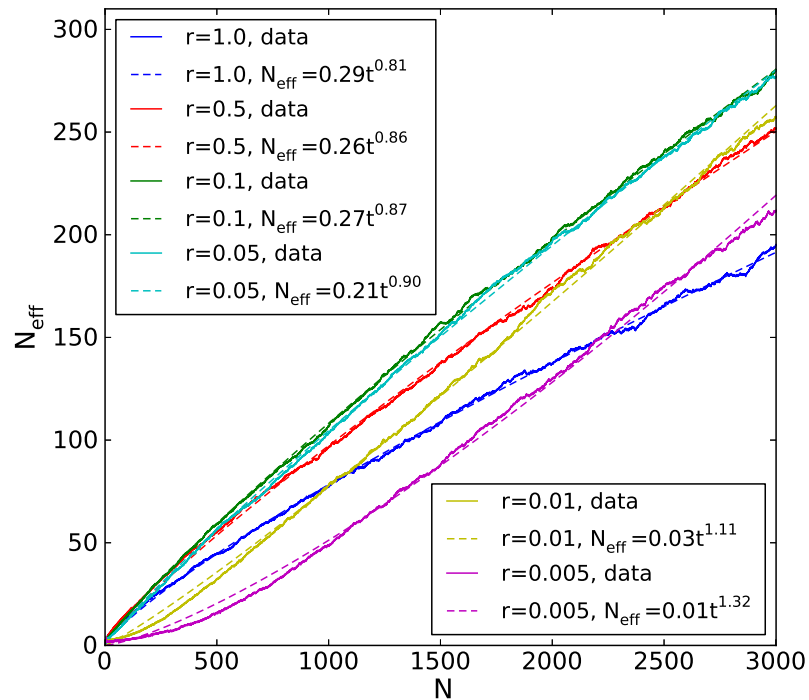
$$\langle \tau_D(N) \rangle = \langle T_e \rangle = \frac{3(4\langle N_{\text{eff}}(N) \rangle - 8)}{2}. \quad (7.13)$$

From Ref. [4], the average time interval during which the number of visited sites grows from  $N + 3$  to  $N + 4$  is

$$\langle \tau_N \rangle = \frac{1}{r} + \frac{1+r}{2+r} \mathbf{E}[T_e], \quad (7.14)$$



## 7.5. Analysis of Maximum Product Sea Size



**Figure 7.23:** Size of the effective product sea  $\langle N_{\text{eff}}(N) \rangle$  as a function of the number of visited sites  $N$ .

and the expected time to visit  $N + 3$  sites is

$$\langle T(N) \rangle = \sum_{i=0}^{N-1} \langle \tau_N \rangle. \quad (7.15)$$

Thus, we can write  $\langle N_{\text{eff}}(N) \rangle$  to be a function of  $N$ , and by substituting Eq. 7.13 into the sum in Eq. 7.15, we obtain

$$\langle T(N) \rangle = \frac{N}{r} + \frac{1+r}{2+r} \sum_{i=0}^{N-1} \frac{3(4\langle N_{\text{eff}}(i) \rangle - 8)}{2}. \quad (7.16)$$

Equation 7.16 shows that if  $\langle N_{\text{eff}}(N) \rangle = \Theta(N)$ , as Fig. 7.23 suggests, then  $\langle T(N) \rangle = \Theta(N^2)$ , which corresponds to diffusive motion. Hence, in order for the leading spider to be superdiffusive as  $t \rightarrow \infty$ , we require  $\langle T(N) \rangle = o(N^2)$ , which implied  $\langle N_{\text{eff}}(N) \rangle = o(N)$ . Unfortunately, as Fig. 7.11 shows, this is not

*Chapter 7. Multiple spider model with infinitely strong source*

the case for the multi-spider model—to maintain superdiffusive motion asymptotically, we need a mechanism stronger than passive injection at the origin.

# Chapter 8

## Conclusion

In this work we exhaustively studied the statistical properties of several molecular spider models and explored their potential for two applications, cargo transport and search. We demonstrated that for both tasks spiders with legs that spend more time on the unvisited sites ( $r < 1$ , spiders with memory) have an advantage over spiders without residence-time difference between previously visited and unvisited sites (spiders without memory, equivalent to regular diffusion).

Using Kinetic Monte Carlo simulations of the Antal and Krapivsky model we showed the unanticipated result that spiders move superdiffusively over a span of time before eventually moving diffusively as had been predicted analytically. This phenomenon can be explained by considering the natural decomposition of the process as switching between two metastates: a diffusive state  $D$  where a spider moves over the contiguous sea of product sites, and a boundary state  $B$  where the spider has a leg attached to a substrate at the boundary between visited and unvisited sites. This decomposition partitions the underlying continuous-time Markov process into  $B$  periods and  $D$  periods. The spider moves ballistically away from the origin during  $B$  periods, but moves diffusively over visited sites during  $D$  periods. The  $B$  state is Markovian in that the transitions from the  $B$  state are independent of the state of the system when it entered the  $B$  state. However, the transitions from the  $D$ -state depend on the size of the contiguous sea of products, and this size increases with time. This explains the apparent superdiffusion at short times when the spider spends more time in the  $B$  state, and the decay to ordinary diffusion at long times, as the spider spends nearly all of its time diffusing over previously visited sites in the  $D$  state. The AK model with  $k = 2$ ,  $s = 2$ ,  $r < 1$  is the simplest model of spider motion with this  $B/D$  state decomposition and the resulting superdiffusive effect. With  $k = 1$ , there is no bias at the bound-

aries, and without irreversible cleavage of sites and a rate  $r < 1$  there is no biasing effect at the boundaries. Thus, the superdiffusive effect depends on spiders having multiple legs and on the legs having the ability to modify sites so that future steps on those sites have different rates.

In order to reduce durations of  $D$  periods we proposed a model where spiders are sequentially released from the origin. The goal of multiple spiders in this model is to prevent the leading spiders from moving too far backwards. Hence, there is an effective outward pressure on the leading spiders that keeps them closer to the boundary and increases their chances to move with the bias towards unvisited sites. We showed that multi-spider systems move faster and farther than single spiders or systems with multiple simple random walkers. However, in the asymptotic limit as  $t \rightarrow \infty$ , we find that  $\alpha(t) \rightarrow 1$ , and the leading spiders of the multi-spider model move diffusively for all values of  $r$ . One way to understand this result is to note that the number of spiders released with time was  $\langle S(t) \rangle = O(\sqrt{t})$  and largely independent of  $r$ , even with the injection rate  $\lambda = \infty$  (Sec. 7.3.2). Under asymptotically superdiffusive motion of the leading spiders we would see the number of products cleaved  $N(t) = \omega(\sqrt{t})$ , and to fill this product sea would require  $S(t) = \Theta(N(t)) = \omega(\sqrt{t})$ , which is not achieved by the multi-spider model. Thus, we cannot seem to release spiders fast enough to support superdiffusion indefinitely. This failure can be understood by observing that the only spiders that actually get to attach to and cleave the energy bearing substrates are the leading  $R_s$  and  $L_s$  spiders. The other, interior spiders only ever walk on products. Thus, while there is some bias exerted on interior spiders by the exclusionary pressure of injected spiders near the origin, for the most part the motion of interior spiders is governed by diffusion. Hence, they do not move fast enough to get clear of the injection site at the origin to allow enough other spiders to be injected fast enough. Indeed, the density of spiders (Sec. 7.3.3) around the origin is nearly maximal, and also seemingly independent of  $r$ . Thus, it does

not really matter how small the value of  $r$  is (and hence how much biasing energy is contained within substrates), because the interior spiders see none of that energy and their diffusive motion is hence independent of  $r$ . Yet, the motion of these interior spiders remains the limiting factor for the injection rate of the new spiders needed to assist the leading spiders by reducing the effective size of the product sea. Hence, no matter how fast the leading spiders are able to move initially, they will inevitably be hindered by the insufficiently fast dispersal of the energy-deprived interior spiders.

For the search task, our studies show that the search time (MFPT) of single multi-legged spiders depends strongly on the kinetic parameter  $r$  in all studied cases. The MFPT is lower for  $r < 1$  (i.e., with memory) than for  $r = 1$  (i.e., without memory) in the one-dimensional case when the spider is searching for the ends of a one-dimensional track. For one-legged spiders the parameter  $r$  does not affect leading asymptotic behavior; however, it slows them down by increasing the constant of the sub-leading term. In the extension of this problem to two dimensions, when the spider is searching for the contour of a circle from its center, the advantage of having  $r < 1$  is even more significant, despite the less effective bias provided by the shape of the leftover substrates. Here the bias provided by the substrates does not always direct the spider towards the absorbing boundary. In contrast, on a 1D track, the substrates always bias the spider towards the closest end when one of the legs is attached to them. The disadvantage in 2D is overcome by the greater amount of substrates that are accessible to spiders. In 1D, the spider does not leave any substrates behind when it progresses towards the ends of the track. In 2D, the shape of the sea of products is complex, and many substrates are left behind. Those substrates can bias the spider towards the absorbing boundary when it starts to turn back towards the origin. When we reverse boundaries in 2D we make the absorbing boundary the single site in the center of the circle, and start the spider from the contour, the parameter  $r$  still improves the MFPT even

though it is difficult to find the single target site in this scenario.

When we introduce multiple spiders and multiple targets, the communication between spiders starts to influence their performance. Spiders communicate by seeing each other's traces, i.e. visited sites. Since multi-legged spiders are biased towards unvisited sites when they are in the border region, they are more likely to search for targets in unexplored locations. This property further reduces search time compared with spiders without memory, which are equivalent to regular random walkers.

In addition to the presented models, I studied several models with modifications that further enhance molecular spiders' properties, including substrates that can be cleaved several times and parallel tracks. Substrates that can be cleaved multiple times introduce more memory into the system and allow spiders to react differently with sites that have been visited more times than with sites that have been visited less frequently. Preliminary results show that this feature can improve spiders' performance. Also, when a spider moves over a small number of parallel tracks, in other words over a narrow but very long rectangle, its performance is also improved, compared with a single 1D track. Another approach to improve spiders' performance is to introduce multiple tracks of substrates into the model of multiple spiders with an infinitely strong source. Surprisingly, this modification made even one-legged spiders move superdiffusively during the whole observed time period for some  $r$  values; as it was shown in Section 4.1, a single one-legged spider is slowed down when  $r < 1$ . When  $r$  is less than some critical value the superdiffusive effect does not seem to fade away with time as it happened with previously studied models. A full exploration of these models is reserved for future work.

## References

- [1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, New York, NY, 1972.
- [2] Pier Lucio Anelli. A molecular shuttle. *Journal of the American Chemical Society*, 113:5131–5133, 1991.
- [3] Tibor Antal and P. L. Krapivsky. Molecular spiders on a plane. *Physical Review E*, 85:061927, Jun 2012.
- [4] Tibor Antal and Paul L. Krapivsky. Molecular spiders with memory. *Physical Review E*, 76(2):021121, 2007.
- [5] Tibor Antal, Paul L. Krapivsky, and Kirone Mallick. Molecular spiders in one dimension. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08027, 2007.
- [6] Jonathan Bath and Andrew J. Turberfield. DNA nanomachines. *Nature Nanotechnology*, 2(5):275–284, 2007.
- [7] Iddo Ben-Ari, Khalid Boushaba, Anastasios Matzavinos, and Alexander Roitershtein. Stochastic analysis of the motion of DNA nanomechanical bipeds. *Bulletin of Mathematical Biology*, 73:1932–1951, 2011.
- [8] Alfred B. Bortz, Malvin H. Kalos, and Joel L. Lebowitz. A new algorithm for Monte Carlo simulation of Ising spin systems. *Journal of Computational Physics*, 17(1):10–18, 1975.
- [9] C Chevalier, O Bénichou, B Meyer, and R Voituriez. First-passage quantities of Brownian motion in a bounded domain with multiple targets: a unified approach. *Journal of Physics A: Mathematical and Theoretical*, 44(2):025002, 2011.
- [10] S. Condamin, O. Bénichou, and M. Moreau. First-passage times for random walks in bounded domains. *Phys. Rev. Lett.*, 95:260601, Dec 2005.
- [11] Akihiro Enomoto, Michael Moore, Tadashi Nakano, Ryota Egashira, Tatsuya Suda, Atsushi Kayasuga, Hiroaki Kojima, Hitoshi Sakakibara, and Kazuhiro Oiwa. A molecular communication system using a network of cytoskeletal filaments. *NSTI-Nanotech*, 1:725–728, 2006.
- [12] Peter L. Freddolino, Anton S. Arkhipov, Steven B. Larson, Alexander McPherson, and Klaus Schulten. Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure*, 14(3):437 – 449, 2006.

- [13] Christophe Galletto, Sebastian Müller, and Serguei Popov. A note on spider walks. *ESAIM: Probability and Statistics*, 15:390–401, 0 2011.
- [14] Christophe Galletto, Sebastian Müller, Serguei Popov, and Marina Vachkovskaia. Spiders in random environment. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 8:129–147, 2011.
- [15] Arne Gennerich and Ronald D Vale. Walking the walk: how kinesin and dynein coordinate their steps. *Current Opinion in Cell Biology*, 21(1):59 – 67, 2009. Cell structure and dynamics.
- [16] J. M. Haile. *Molecular dynamics simulation: elementary methods*. Wiley professional paperback series. Wiley, 1997.
- [17] Henry Hess. Engineering applications of biomolecular motors. *Annu. Rev. Biomed. Eng.*, 13:429–450, 2011.
- [18] Henry Hess and Viola Vogel. Molecular shuttles based on motor proteins: active transport in synthetic environments. *Reviews in Molecular Biotechnology*, 82(1):67 – 85, 2001. Biomolecular nanotechnology.
- [19] Nobutaka Hirokawa, Yasuko Noda, Yosuke Tanaka, and Shinsuke Niwa. Kinesin superfamily motor proteins and intracellular transport. *Nature Reviews Molecular Cell Biology*, 10(10):682–696, 2009.
- [20] G Jayachandran, V Vishal, and VS Pande. Using massively parallel simulation and Markovian models to study protein folding: Examining the dynamics of the villin headpiece. *Journal of Chemical Physics*, 124(16), Apr 28 2006.
- [21] Róbert Juhász. Anomalous transport in disordered exclusion processes with coupled particles. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(11):P11015, 2007.
- [22] Euan R. Kay, David A. Leigh, and Francesco Zerbetto. Synthetic molecular motors and mechanical machines. *Angewandte Chemie International Edition*, 46(1-2):72–191, 2007.
- [23] A. M. Lacasta, J. M. Sancho, A. H. Romero, I. M. Sokolov, and K. Lindenberg. From subdiffusion to superdiffusion of particles on solid surfaces. *Phys. Rev. E*, 70(5):051104, Nov 2004.



## References

- [24] Kyle Lund, Anthony J. Manzo, Nadine Dabby, Nicole Michelotti, Alexander Johnson-Buck, Jeanette Nangreave, Steven Taylor, Renjun Pei, Milan N. Stojanovic, Nils G. Walter, Erik Winfree, and Hao Yan. Molecular robots guided by prescriptive landscapes. *Nature*, 465:206–210, May 2010.
- [25] Thiago G. Mattos, Carlos Mejia-Monasterio, Ralf Metzler, and Gleb Oshanin. First passages in bounded domains: When is the mean first passage time meaningful? *Physical Review E*, 86:031143, Sep 2012.
- [26] Richard A. Muscat, Jonathan Bath, and Andrew J. Turberfield. A programmable molecular robot. *Nano Letters*, 11(3):982–987, 2011.
- [27] Apoorva Nagar and Punyabrata Pradhan. First passage time distribution in random walks with absorbing boundaries. *Physica A: Statistical Mechanics and its Applications*, 320(0):141 – 148, 2003.
- [28] Dan V. Nicolau, Dan V. Nicolau, Jr., Gerardin Solana, Kristi L. Hanson, Luisa Filipponi, Lisen Wang, and Abraham P. Lee. Molecular motors-based micro- and nano-biocomputation devices. *Microelectronic Engineering*, 83(4-9):1582–1588, 2006.
- [29] Mark J. Olah and Darko Stefanovic. Multivalent random walkers — a model for deoxyribozyme walkers. In *DNA 17: Proceedings of The Seventeenth International Meeting on DNA Computing and Molecular Programming*, volume 6397 of *Lecture Notes in Computer Science*, pages 160–174. Springer, 2011.
- [30] Mark J. Olah and Darko Stefanovic. Superdiffusive transport by multivalent molecular walkers moving under load. *Phys. Rev. E*, 87:062713, Jun 2013.
- [31] Tosan Omabegho, Ruojie Sha, and Nadrian C. Seeman. A bipedal DNA brownian motor with coordinated legs. *Science*, 324(5923):67–71, 2009.
- [32] Renjun Pei, Steven K. Taylor, Darko Stefanovic, Sergei Rudchenko, Tiffany E. Mitchell, and Milan N. Stojanovic. Behavior of polycatalytic assemblies in a substrate-displaying matrix. *Journal of the American Chemical Society*, 128(39):12693–12699, 2006.
- [33] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical recipes in C++*. Cambridge University Press, New York, NY, 2002.
- [34] Matthias Rank, Louis Reese, and Erwin Frey. Cooperative effects enhance the transport properties of molecular spider teams. *Phys. Rev. E*, 87:032706, Mar 2013.

- [35] Laleh Samii, Gerhard A. Blab, Elizabeth H. C. Bromley, Heiner Linke, Paul M. G. Curmi, Martin J. Zuckermann, and Nancy R. Forde. Time-dependent motor properties of multipedal molecular spiders. *Phys. Rev. E*, 84:031111, Sep 2011.
- [36] Laleh Samii, Heiner Linke, Martin J. Zuckermann, and Nancy R. Forde. Biased motion and molecular motor properties of bipedal spiders. *Phys. Rev. E*, 81:021106, Feb 2010.
- [37] Manfred Schliwa and Gunther Woehlke. Molecular motors. *Nature*, 422(6933):759–765, 2003.
- [38] Oleg Semenov, David Mohr, and Darko Stefanovic. First passage properties of molecular spiders. *arXiv:1304.5572 [physics.bio-ph]*, April 2013.
- [39] Oleg Semenov, Mark J. Olah, and Darko Stefanovic. Mechanism of diffusive transport in molecular spider models. *Phys. Rev. E*, 83:021117, Feb 2011.
- [40] Oleg Semenov, Mark J. Olah, and Darko Stefanovic. Multiple molecular spiders with a single localized source—the one-dimensional case. In *DNA 17: Proceedings of The Seventeenth International Meeting on DNA Computing and Molecular Programming*, volume 6397 of *Lecture Notes in Computer Science*, pages 204–216. Springer, 2011.
- [41] Oleg Semenov, Mark J. Olah, and Darko Stefanovic. Cooperative linear cargo transport with molecular spiders. *Natural Computing*, 12(2):259–276, 2013.
- [42] Oleg Semenov, Darko Stefanovic, and Milan Stojanovic. The effects of multivalency and kinetics in nanoscale search by molecular spiders. In *Proc. of the Italian Workshop on Artificial Life and Evolutionary Computation*, pages 1–12, 2012. Published on CD, isbn 978-88-903581-2-8.
- [43] Oleg Semenov, Darko Stefanovic, and Milan Stojanovic. *Evolution, Complexity and Artificial Life*, chapter The Effects of Multivalency and Kinetics in Nanoscale Search by Molecular Spiders. Springer, 2013.
- [44] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52:591–611, 1965.
- [45] Jong-Shik Shin and Niles A. Pierce. A synthetic DNA walker for molecular transport. *Journal of the American Chemical Society*, 126(35):10834–10835, 2004.
- [46] Ronald D. Vale and Ronald A. Milligan. The way things move: Looking under the hood of molecular motor proteins. *Science*, 288(5463):88–95, 2000.

## References

- [47] Megan T Valentine and Susan P Gilbert. To step or not to step? how biochemistry and mechanics influence processivity in kinesin and Eg5. *Current Opinion in Cell Biology*, 19(1):75 – 81, 2007. Cell structure and dynamics.
- [48] Suvir Venkataraman, Robert M. Dirks, Paul W. K. Rothmund, Erik Winfree, and Niles A. Pierce. An autonomous polymerization motor powered by DNA hybridization. *Nature Nanotechnology*, 2:490–494, 2007.
- [49] Max von Delius, Edzard M. Geertsema, and David A. Leigh. A synthetic small molecule that can walk down a track. *Nature Chemistry*, 2(2):96–101, 2010.
- [50] Shelley F. J. Wickham, Jonathan Bath, Yousuke Katsuda, Masayuki Endo, Kumi Hidaka, Hiroshi Sugiyama, and Andrew J. Turberfield. A DNA-based molecular motor that can navigate a network of tracks. *Nature Nanotechnology*, 7(3):169–173, 2012.
- [51] Peng Yin, Hao Yan, Xiaojun G. Daniell, Andrew J. Turberfield, and John H. Reif. A unidirectional DNA walker that moves autonomously along a track. *Angew. Chem. Int. Ed.*, 43:4906–4911, 2004.