A Geometrical Probability Approach to Location-Critical Network Performance Metrics

by

Yanyan Zhuang
B. Eng., Southeast University, 2005
M. Eng., Southeast University, 2008

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Computer Science

A Geometrical Probability Approach to Location-Critical Network Performance Metrics

by

Yanyan Zhuang
B. Eng., Southeast University, 2005
M. Eng., Southeast University, 2008

Supervisory Committee

Dr. Jianping Pan, Supervisor
(Department of Computer Science)

Dr. Sudhakar Ganti, Departmental Member
(Department of Computer Science)

Dr. Kui Wu, Departmental Member
(Department of Computer Science)

Dr. Aaron Gulliver, Outside Member
(Department of Electrical and Computer Engineering)

Supervisory Committee

Dr. Jianping Pan, Supervisor
(Department of Computer Science)

Dr. Sudhakar Ganti, Departmental Member
(Department of Computer Science)

Dr. Kui Wu, Departmental Member
(Department of Computer Science)

Dr. Aaron Gulliver, Outside Member
(Department of Electrical and Computer Engineering)

Abstract

The field of wireless communications has been experiencing tremendous growth with the ever-increasing dependence on wireless services. In the operation of a communication network, the network coverage and node placement are of profound importance. The network performance metrics can be modeled as nonlinear functions of inter-node distances. Therefore, a geometric abstraction of the distance between wireless devices becomes a prerequisite for accurate system modeling and analysis. A geometrical probability approach is presented in this dissertation to characterize the probabilistic distance properties, for analyzing the location-critical performance metrics through various spatial distance distributions.

Ideally, the research in geometrical probability shall give results for the distance distributions 1) over elementary geometries such as a straight line, squares and rectangles, and 2) over complex geometries such as rhombuses and hexagons. Both 1) and 2) are the representative topological shapes for communication networks. The current probability and statistics literature has explicit results for 1), whereas the results

for 2) are not in existence. In particular, the absence of the distance distributions for rhombuses and hexagons has posed challenges towards the analytical modeling of location-critical performance metrics in complex geometries. This dissertation is dedicated to the application of existing results in 1) elementary geometries to the networking area, and the development of a new approach to deriving the distance distributions for complex geometries in 2), bridging the gap between the geometrical probability and networking research.

The contribution of this dissertation is twofold. First, the one-dimensional Poisson point process in 1) is applied to the message dissemination in vehicular ad-hoc networks, where the network geometry is constrained by highways and city blocks. Second, a new approach is developed to derive the closed-form distributions of internode distances associated with rhombuses and hexagons in 2), which are obtained for the first time in the literature. Analytical models can be constructed for characterizing the location-critical network performance metrics, such as connectivity, nearest/farthest neighbor, transmission power, and path loss in wireless networks. Through both analytical and simulation results, this dissertation demonstrates that this geometrical probability approach provides accurate information essential to successful network protocol and system design, and goes beyond the approximations or Monte Carlo simulations by gracefully eliminating the empirical errors.

Contents

Supervisory Committee ii

Abstract iii

Table of Contents v

List of Tables ix

List of Figures x

Acknowledgments xiii

Dedication xv

1 Introduction 1

1.1 Wireless Communication Networks 1

1.1.1 Background 2

1.1.2 Motivation 3

1.2 Challenges 3

1.2.1 Theory 3

1.2.2 Application 4

1.3 Contributions of Dissertation 5

1.3.1 Application of One-Dimensional Random Distances 5

1.3.2 Random Distances Associated with Complex Geometries . . . 8

1.4 Outline of Dissertation 10

2 Background and Related Work 12

2.1 Geometrical Distributions 13

2.1.1 One-Dimensional Random Distances 13

2.1.2 Two-Dimensional Random Distances 13

2.1.3 Summary 16
2.2 Geometrical Probability Approach and Location-Critical Performance Metrics 16
2.2.1 Connectivity, Position-Based Routing and Hidden Terminals . 16
2.2.2 Path Loss, Fading and Shadowing 17
2.2.3 Interference, SINR and Channel Capacity 18
2.2.4 Stochastic Properties of Random Mobility Models 19
2.2.5 Summary 20

3 The Poisson Point Process and Vehicular Ad-Hoc Networks on Highways 21
3.1 Spatio-Temporal Vehicular Traffic Models 22
3.1.1 The Poisson Point Process 22
3.1.2 One-Dimensional Connectivity 23
3.2 Time and Location-Critical Framework for Vehicular Ad-Hoc Networks 25
3.2.1 Scalable Modulation and Coding (SMC) 25
3.2.2 TLC-based Emergency Message Dissemination 27
3.2.3 Forward Direction: Cluster Size Characterization 29
3.2.4 Backward Direction: Using the Reverse Traffic 34
3.3 Summary 40
3.4 Discussions and Future Work 42

4 Vehicular Message Dissemination in Two-Dimensional City Blocks 43
4.1 Percolation Theory and Two-Dimensional Connectivity 44
4.1.1 Percolation Theory 44
4.1.2 Connectivity in Two-Dimensional Vehicular Ad-Hoc Networks 45
4.2 Two-Dimensional Manhattan-Like City 46
4.2.1 Bond Probability 47
4.2.2 Ladder Connectivity 50
4.2.3 Lattice Connectivity 54
4.2.4 Network Connectivity 55
4.3 Summary 62
4.4 Discussions and Future Work 63
4.4.1 Connectivity, Collisions and Throughput 63
4.4.2 Vehicle Mobility 63

5 Random Distances Associated with Rhombuses: the Complex Geometry 65
5.1 State of the Art . . . 66
5.2 From Rectangles to Parallelograms . . . 67
5.2.1 The Quadratic Product Formulation . . . 67
5.2.2 Rectangles: a Simple Illustration . . . 68
5.2.3 Parallelograms: the Squeezed Rectangles . . . 69
5.3 Distance Distributions Associated with Rhombuses . . . 71
5.3.1 Distance Distribution within a Rhombus . . . 73
5.3.2 Distance Distribution between Parallel Rhombuses . . . 76
5.3.3 Distance Distribution between Diagonal Rhombuses . . . 80
5.3.4 Distance Verification . . . 82
5.4 Practical Results . . . 83
5.4.1 Statistical Moments of Random Distances . . . 83
5.4.2 Polynomial Fits of Random Distances . . . 84
5.5 Summary . . . 85
5.6 Discussions and Future Work . . . 85
5.6.1 Extension of Probability Function: From Rhombuses to Hexagons 86
5.6.2 Extension of Probability Density: Nonuniform Point Distribution 87

6 Random Distances Associated with Hexagons 88
6.1 Cell Shapes and Location-Critical Performance Metrics . . . 89
6.2 Distance Distributions Associated with Regular Hexagons . . . 90
6.2.1 Distance Distribution within a Regular Hexagon . . . 90
6.2.2 Distance Distribution between Adjacent Regular Hexagons . . 99
6.2.3 Distance Verification . . . 104
6.3 Practical Results . . . 106
6.3.1 Statistical Moments of Random Distances . . . 106
6.3.2 Polynomial Fits of Random Distances . . . 107
6.4 Performance Study Using Distance Distributions . . . 108
6.4.1 Sparse Network Scenario: the Nearest Neighbor . . . 109
6.4.2 Dense Network Scenario: the Farthest Neighbor . . . 111
6.4.3 Transmission Power Control . . . 114
6.5 Summary . . . 117
6.6 Discussions and Future Work . . . 117

7 Conclusions and Future Work **118**
7.1 Conclusions 118
7.1.1 Application of One-Dimensional Random Distances 118
7.1.2 Random Distances Associated with Complex Geometries . . . 119
7.2 Future Work 120
7.2.1 Vehicular Ad Hoc Networks with Real-World Traces 120
7.2.2 Wireless Channel Models 121
7.2.3 Inter-Disciplinary Research 121

Bibliography **123**

A Derivation of the Distance Distribution between Two Adjacent Rhombuses **134**
A.1 Long-Diagonal Case 134
A.2 Short-Diagonal Case 137

B Validation of Rhombus-Related Distributions by Recursion **140**

C Validation of Hexagon-Related Distributions by Recursion **145**

List of Tables

Table 5.1 Moments and Variances of Rhombus Distributions—Numerical vs Simulation Results 83
Table 5.2 Coefficients of the Polynomial Fit and the Norm of Residuals (NR) for Rhombus Distributions 85

Table 6.1 Moments and Variances of Hexagon Distributions—Numerical vs Simulation Results 107
Table 6.2 Coefficients of the Polynomial Fit and the Norm of Residuals (NR) for Hexagon Distributions 108

Table C.1 Distance Distributions and the Corresponding Equation Number 145

List of Figures

Figure 3.1 Scalable Modulation and Coding using the 64-QAM Constellation. . . . 26
Figure 3.2 Time-Location Critical Emergency Message Dissemination Scenarios. . . . 27
Figure 3.3 Comparison between $E[B]$ in [75] and $E[C]$ on the Expected Cluster Size. . . . 31
Figure 3.4 Gamma Approximation of the Cluster Size Distribution (Solid Curve—Analysis, Dashed Curve—Simulation). . . . 33
Figure 3.5 Three Cases to Extend the Forward Cluster by Reverse Traffic. 35
Figure 3.6 Travel Delay to Cross Forward Clusters and Comparison with [100]. . . . 38
(a) Travel Delay due to Reverse Traffic to Cross Forward Clusters (Solid Curve—Analysis, Dashed Curve—Simulation). . . . 38
(b) Travel Delay Comparison with [100], $R = 206\ m$. . . . 38
Figure 3.7 The Probability to Miss LDLD Deadlines (Solid Curve—Analysis, Dashed Curve—Simulation). . . . 41
(a) $R = 206\ m$, $D_2 = 5\ km$. . . . 41
(b) $R = 879\ m$, $D_2 = 5\ km$. . . . 41

Figure 4.1 Two-Dimensional Manhattan-Like City. . . . 47
Figure 4.2 Bond Probability Illustration. . . . 48
Figure 4.3 Bond Probability Validation. . . . 49
Figure 4.4 Ladder Connectivity Illustration. . . . 50
Figure 4.5 Ladder Connectivity Validation. . . . 52
Figure 4.6 Two Side Streets Form New Ladders (Solid Lines—Connected Street Segments, Dashed Line—Disconnected Main Street). . . . 54
Figure 4.7 Network Connectivity with Geo-Constrained Forwarding (GF). 57

Figure 4.8 Connectivity Probability: Geo-Constrained Forwarding (GF) vs. Unconstrained Forwarding (UF). 58
Figure 4.9 Network Connectivity with Unconstrained Forwarding (UF). . 59
Figure 4.10 Broadcast Overhead: Average Number of Transmissions. . . . 61
(a) Geo-Constrained Forwarding (GF). 61
(b) Unconstrained Forwarding (UF). 61

Figure 5.1 A Geometric Interpretation of Random Distances within a Rectangle. 68
(a) Random Points in a Rectangle. 68
(b) $Z = X^2 + Y^2$ when $a = 3$, $b = 2$. 68
Figure 5.2 A Rectangle Squeezed into a Parallelogram. 69
(a) A Rectangle Becomes a Parallelogram. 69
(b) $Z = X^2 + 2\cos\theta XY + Y^2$ when $a = 3$, $b = 2$ and $\theta = \frac{\pi}{3}$. 69
Figure 5.3 Random Points in Rhombuses. 72
Figure 5.4 Three Sub-cases for $Z = |A'B'|^2$. 73
Figure 5.5 Five Sub-cases for $Z = |R'S'|^2$. 76
Figure 5.6 Cumulative Distribution Functions and Simulation Results for Random Distances Associated with Rhombuses. 82
Figure 5.7 Polynomial Fit of the Distance Distribution Functions Associated with Rhombuses. 86
(a) Within a Single Rhombus 86
(b) Between two Parallel Adjacent Rhombuses 86
(c) Between two Long-Diag Adjacent Rhombuses 86
(d) Between two Short-Diag Adjacent Rhombuses 86

Figure 6.1 Hexagonal Cell Layout and Circular Approximations. 89
(a) Hexagonal Cells. 89
(b) Inscribed Circles. 89
(c) Enclosing Circles. 89
Figure 6.2 Relationship between a Hexagon and (a) Rhombuses (b) Triangles 91
(a) Rhombus Decomposition 91
(b) Triangle Decomposition 91
Figure 6.3 Geometric Interpretation of $Z = X^2+X(Y_1-Y_2)+Y_1^2+Y_1Y_2+Y_2^2$. 93
Figure 6.4 Sub-case when $0 \le z \le \frac{3}{4}$. 94
Figure 6.5 Sub-case when $\frac{3}{4} \le z \le 1$. 95

Figure 6.6 Sub-case when $1 \leq z \leq 3$. 96
Figure 6.7 Sub-case when $3 \leq z \leq 4$. 97
Figure 6.8 Cumulative Distribution Functions and Simulation Results for Random Distances within a Hexagon. 99
Figure 6.9 Random Points between Two Adjacent Hexagons: Different Cases with Rhombuses. 100
Figure 6.10 Cumulative Distribution Functions and Simulation Results for Random Distances within and between Hexagons. 105
Figure 6.11 Partial Recursion through Hexagons and Rhombuses. 105
Figure 6.12 Polynomial Fit. 108
(a) Within a Single Hexagon 108
(b) Between two Adjacent Hexagons 108
Figure 6.13 Nearest Neighbor Distribution. 110
Figure 6.14 Expected Distance to Nearest Neighbor. 112
(a) Distance vs. No. of Nodes. 112
(b) Distance vs. Network Size. 112
Figure 6.15 Farthest Neighbor Distribution. 113
Figure 6.16 Expected Distance to Farthest Neighbor. 115
(a) Distance vs. No. of Nodes. 115
(b) Distance vs. Network Size. 115
Figure 6.17 Expected Transmission Power to the Nearest Neighbor. 116
(a) Transmission Power vs. No. of Nodes. 116
(b) Transmission Power vs. Network Size. 116

Figure A.1 Five Sub-cases for $Z = |P'Q'|^2$. 134
Figure A.2 Four Sub-cases for $Z = |M'N'|^2$. 137

Acknowledgments

It doesn't matter if you try and try and try again, and fail.
It does matter if you try and fail, and fail to try again.

Charles Kettering

I would like to begin by expressing my appreciation to Dr. Jianping Pan, my thesis advisor and mentor for the past three and a half years. Dr. Pan has been an ideal supervisor, providing me with sparkling ideas, sage advice, insightful criticisms, as well as invaluable life advice. His knowledge, integrity, commitment and enthusiasm for research have always been an inspiration. Through numerous research discussion and meetings, not only had I acquired the various skills for conducting research, I also learned to be strong, resilient and independent. It was his unwavering patience and encouragement that kept me working when I wanted to give up. It was his strongest will that gave me the courage to challenge myself in order to achieve my full potential.

I would also like to thank Dr. Lin Cai from the Department of Electrical and Computer Engineering, who is not my supervisor or my committee member. Nevertheless, it was such a wonderful experience to work with Dr. Cai, who provides me her steadfast support and selfless help. Her advice and feedback on my research have significantly enhanced the quality of my work. I also admire her as a lady in engineering who is able to balance work and life, and having great achievement in the meantime.

My most heartfelt gratefulness goes to my late supervisor during my Master's program at Southeast University, Prof. Guanqun Gu. He was the first person who encouraged me to go outside of my home country and see the big world outside. It was my honor and privilege to have the guidance from such a great man. His strength and faith during his terminal illness gave me a new appreciation for the meaning and importance of life and personal achievement. Prof. Gu passed away half a year before my Master's thesis defense. However, his influence has always given me the great fortitude to overcome my personal barriers, chasing and accomplishing my dreams.

To complete a piece of work of this magnitude requires a network of support. Therefore, I am indeed indebted to many people. First, my thanks must go to the members of my dissertation committee, Dr. Sudhakar Ganti, Dr. Kui Wu and Dr. Aaron Gulliver, who have generously given their time and expertise to better my work.

I would also thank all the fellow graduate students, professors and colleagues in both the CS and EE departments for making it a friendly and lively working place, especially to Dr. Ulrike Stege and Mary Sanseverino for being such wonderful mentors. Victoria Li, Leandro Collares, Zhen Ling, Leon He, Yuanqian Luo, Zhe Yang, Le Chang, Tyler Cadigan, Rukhsana Ruby, David Cheperdak, Zhonghua Wei, Li Ji, Maryam Ahmadi, Sarshad Abubaker, Lei Zhang, Lei Zheng, Tianming Wei, Hoi-Ying Tsang, and in particular, Dr. Jun Tao—my previous mentor and current colleague. Thank you all for bearing with me for the past years. A special gratitude goes to Dr. Li Li from Communications Research Centre (CRC). Thank you so much for your kindest blessings, and for letting me believe that there are special moments in our lives that let us hear the call of eternity and see the real love and beauty. I also enjoy the company and support of my dearest friends from outside the engineering faculty, Sara Hockett, Mary Zihan Shi, Dr. Hua Lin, Mel Tremblay, etc. They have given their friendship and made my experience at UVic both more educational and more fun.

Finally, my gratitude goes to my parents for their love, support, encouragement, and everything that they have given to me, which I greatly needed and deeply appreciated.

Dedication

To My Dear Mentors, Family and Friends.

Scientific achievement does not stop.
It is a lifetime endeavor.
If you want to know the future, create it.
If you want to change the world, you can't sleep through.

Chapter 1

Introduction

1.1 Wireless Communication Networks

Recent years have seen the impressive penetration of wireless communication technologies, growing availability of wireless data access opportunities, and wide proliferation of wireless devices. The ever-increasing dependence on wireless services has also triggered a high demand for novel and exciting applications. With the wide deployment and convenient accessibility to communication network infrastructures, the emerging technologies are capable of providing ubiquitous network access to users, as well as high quality of services. Consequently, numerous research opportunities have appeared, as a result of the increased attention from the research community.

The wireless communication systems are primarily designed to provide cost-efficient wide-area coverage for users, with or without the assistance of an infrastructure. Infrastructure-based networks are the communication networks that have dedicated access points or base stations coordinating over their networking domains. Any communication is established between a fixed point and an arbitrary user, and thus the operation of the systems relies on centrally deployed devices. Typical examples are wireless local area networks (WLANs) and cellular systems. In contrast, the networks where wireless devices connect with one another either directly or via multi-hop are known as infrastructure-less networks, or ad-hoc networks. In an ad-hoc network, devices are usually limited in transmission and processing power. However, they can be autonomously deployed and relay information on behalf of other neighbors in the vicinity, acting as both terminals and relays simultaneously. Wireless sensor networks and vehicular ad-hoc networks are two important examples.

Both the infrastructure-based and infrastructure-less networks are major components of the current communication arena. As both of these communication technologies developing to their maturity, people's daily lives have become more connected than ever. Wireless communication networks have brought much convenience to today's world. But how to model and analyze these networks by combining the emerging wireless communication and networking technologies with a realistic mathematical model is still an open issue and a demanding task.

1.1.1 Background

Wireless transceivers use the radio channel as the medium for communications. Compared with the traditional wired Ethernet, wirelessly transmitted electromagnetic waves are not guided along any solid medium. Such an over-the-air transmission that is subject to noisy channel conditions is the primary reason for the severe attenuation or path loss of transmitted signals. The transmission between a pair of transceivers is also subject to unintended signals, or interference, transmitted simultaneously in the vicinity. The obstacles that cause reflection, diffraction or scattering of the signals make the wireless communication environment even more complicated. Therefore, the transmission of signal in the open air is significantly more complex than that in a guided medium. Such complexity poses challenges towards the modeling and analysis of wireless communication networks.

Among all the above physical phenomena, however, the locations and distances among transceivers, interferers and scatterers are the essential factors with the most significant impact on the transmitted wireless signals. A realistic mathematical model that captures the distance between randomly distributed wireless devices is therefore highly desired. As an example, the strength of a transmitted radio signal attenuates with the distance between the transmitter and receiver. The existence of interferers and scatterers, and the resultant inter-node distances affect the overall interference and multi-path fading. As another example, when wireless devices have a fixed transmission range, the inter-node distance determines the existence of a communication link, or the wireless connectivity. In this dissertation, the distance-related metrics, e.g., signal attenuation and connectivity, are referred to as location-critical performance metrics. They determine the ultimate performance of a wireless communication system.

1.1.2 Motivation

The above location-critical performance metrics are essentially nonlinear functions of inter-node distances. As transceivers are typically distributed over an area or volume, following a certain distribution, the distance between these nodes are determined by the network topology and node locations. In Geometrical Probability [67], such distance is captured by a probabilistic density function. Therefore, given a network coverage and the distribution of random nodes within the network, characterizing the distances among these nodes becomes a prerequisite for accurate system modeling and analysis.

In this respect, there are two major challenges to be tackled:

1. How to capture the distances and the spatial relationships among various wireless devices from a geometrical point of view; and

2. How to model and analyze the location-critical performance metrics, such as wireless connectivity, signal attenuation, etc., given such a characterization of spatial distances.

These two questions are particularly important for service providers and network operators, as they are critical to the proper planning and dimensioning of service infrastructure, and the provisioning of a consistent user experience.

1.2 Challenges

In a wireless communication network, the system performance metrics are highly dependent on the inter-node distances and network geometry. This dissertation aims at resolving the above two major challenges, from both the theory and application perspectives.

1.2.1 Theory

Relating to the first challenge, we refer to *theory* as the research in geometrical probability, from a mathematical and statistical aspect. Geometrical Probability [67] is an area of probability and statistics that studies the fundamental properties of geometrical objects, such as points, lines, planes and spheres. The distribution of distances between random points is an important aspect of this theory. With a certain node

distribution, the inter-node distances can be treated as random variables, which are best described by a statistical distribution. For instance, if nodes are uniformly distributed in the network coverage, the difference between the $\mathcal{X}$ and $\mathcal{Y}$-coordinates, denoted by X and Y, both follow a triangular distribution [34]. The Euclidean distance between these nodes is $D = \sqrt{X^2 + Y^2}$, whose distribution are typically complicated [67, 104–106, 108]. The study of the random distances associated with different geometric shapes, one of the fundamental probability measures in geometrical probability, has been a research topic with a rich mathematical history and background.

In this dissertation, straight lines, rectangles and squares are categorized as the elementary geometries, while rhombuses, parallelograms and hexagons are the complex geometries. In complex geometries, point coordinates are interdependent. Although in the existing literature, the closed-form results for elementary geometries have been obtained a long time ago, e.g., the distance distributions for rectangles [37, 38], the same problem for complex geometries remains unsolved.

1.2.2 Application

Aiming to address the second challenge, *application* is referred to as the probabilistic modeling and analysis for the location-critical performance metrics utilizing the geometrical probability theory. Given the random distance captured through a statistical distribution, probabilistic models can be constructed for analyzing the aforementioned performance metrics that are location-critical, e.g., connectivity, path loss, interference, etc. Such models are defined via a spatial distance density function. Since the distribution of random distances is an important aspect of the geometrical probability research, the approach used in this dissertation is thus a *geometrical probability approach*.

The traditional probabilistic methods providing statistical moments, particularly mean and variance, have been long existing in the literature. Other methods are based upon the empirical approximations and Monte Carlo simulations. Using moments or approximations, analytical models are able to provide a rough estimation of the system performance metrics, such as an upper or lower bound [100]. Using the statistical moments, empirical approximations or simulations, the complexity of analytical models has been maintained at a tractable level. However, a nonlinear relationship between the location-critical performance metrics and random distances

makes such performance evaluation much less accurate.

By the first example in Section 1.1.1, given D as the random variable denoting the distance between an arbitrary pair of transceivers, if the path loss exponent is α, then the signal attenuates as $D^{-\alpha}$. If the transmission power is P_t, then after being transmitted over distance D, the received signal strength at the receiver is proportional to $P_t D^{-\alpha}$. $P_t D^{-\alpha}$ is a nonlinear function of D, so are other location-critical performance metrics in wireless communication networks. From Jensen's inequality, $(\mathbb{E}[D])^{-\alpha} \leq \mathbb{E}\left[D^{-\alpha}\right]$, where $\mathbb{E}[\cdot]$ denotes the expectation and $\alpha \in [2, 6]$. As a result, the deviation increases drastically with the nonlinear path loss exponent α [107]. In the second example, for wireless devices that have a certain transmission range, the random variable D also determines the connectivity between devices. If the statistical distance distribution of D is given, then important insights into the transmitted signal and the properties of network connectivity can be obtained [110].

Henceforth, the terminology "point" is used in the context of theory, and "node" is used in the context of application. Similarly, "geometry" is used as a concept in mathematics and probability, and "topology" is used in the networking research.

1.3 Contributions of Dissertation

The contributions of this dissertation are twofold, addressing the above two challenges from both the theory and application perspectives. First, one-dimensional random distances are utilized for analyzing the connectivity properties in a vehicular ad-hoc network scenario, where the network topology is a highway or city blocks. Second, the closed-form distributions of inter-node distances associated with complex geometries are obtained for the first time in the literature. Previously, conducting analytical modeling in complex geometries was impossible.

1.3.1 Application of One-Dimensional Random Distances

The simplest network topology is a one-dimensional highway in vehicular ad-hoc networks (VANETs). VANETs are emerging paradigms in sensor networks, which use different sensing devices available in vehicles to gather environment information and provide intelligent traffic information services. Vehicles equipped with wireless transceivers communicate with each other through vehicle-to-vehicle communications,

where the location of each vehicle is constrained by the road structure. The communication among vehicles has a plethora of applications such as safety and emergency information dissemination for drivers, traffic monitoring, data collection and communication for road and traffic managers, advertisement for hotels and restaurants, entertainment and business services content distribution for passengers, etc.

Message dissemination is one of the most important applications in VANETs, which depends on the location-critical connectivity among vehicles. Given a transmission range, the underlying vehicle connectivity is determined by the inter-vehicle distance distribution. Such geometrical probability approach makes possible an in-depth study on the fundamental connectivity properties in a highway scenario, as well as other location-critical performance metrics. The high accuracy of this geometrical probabilistic approach is also demonstrated by the extension of this model to a two-dimensional Manhattan-like city. Compared with the state-of-the-art which uses mathematical simplification [75] and average analysis [100] in the analytical models, the research in this dissertation is a considerable further effort.

On the other hand, the opportunistic access to the roadside infrastructures from traveling vehicles, i.e., vehicle-to-infrastructure communications, also appeared as IEEE 802.11 access points and base stations opening up services to mobile clients [62, 63, 111, 112]. However, due to the random location and mobility of vehicles, and the limited roadside resources, the connectivity among vehicles and base stations are not naturally guaranteed. The deployment of base stations will also cost network operators and service providers more at up-front investments and maintenance. Hence, in this dissertation, the ad-hoc communication is proposed as the first-step solution for increasing the opportunity for connectivity and accessibility.

(1) One-Dimensional Highway

This is the application of geometrical probability along a line to a highway VANET scenario. By statistically analyzing different sets of empirical data, the authors of [13, 84, 86, 94, 100, 102] etc. found that the exponential distribution is a good match for inter-vehicle distances. Equivalently, the vehicle arrivals can be modeled as a Poisson point process. In probability theory and statistics, the Poisson distribution describes the probability of a given number of events occurring in a fixed interval of time or space [7]. Its application can be found in every field related to counting, such as the arrivals of phone calls in a telephony system, customers at a counter, and vehicles

passing a road observation point. Poisson point process is therefore widely used in traffic engineering and flow theory [41], and has been verified by statistical analysis and measurement [13, 84, 86, 94, 100, 102].

A time/location-critical (TLC) framework is proposed for emergency message dissemination, where vehicles at different distances to an accident site can receive information with different levels of details. Based on the memory-less property of Poisson distribution, in [110] we studied the message propagation in a highway scenario through the derivation of vehicle cluster size distribution. A vehicle cluster is defined as a finite number of vehicles that are connected sequentially with each other via multiple hops. Its size is the distance between the first and last vehicles in the same cluster. In contrast to the previous work in the literature [75, 100], no mathematical simplification or approximation is used in [110]. By observing a non-negligible probability that the message delivery cannot be guaranteed when propagating within the same vehicle cluster, reverse traffic is incorporated for further extending the vehicle connectivity to far-away clusters. The distribution of cluster size and the distance between clusters are critical to the characterization of network performance, such as the message propagation delay, and the likelihood of missing an emergency message.

(2) Two-Dimensional Manhattan-Like City

A more complicated topology is a square lattice, or the street blocks in a Manhattan-like city, where vehicles can disseminate messages to each other in perpendicular directions. Following the same Poisson distribution of vehicles in one dimension, and given the message propagation properties along a straight line, our work [109] extended the model in [110] to a two-dimensional city block scenario. Both the theoretical analysis in a ladder topology and the simulation in a lattice topology show that, the connectivity properties are significantly different from those in a highway scenario. Furthermore, the network connectivity is investigated by two different message forwarding schemes, with and without geographic constraints, respectively. The results are surprisingly similar to the percolation phenomenon [25], where there exists a critical threshold above which the entire network is connected with a high probability.

The problem of deriving the connectivity probabilities with an arbitrary location is a central problem of directed percolation [42] in Physics and Stochastic Processes. Although the authors in [25] derived the results in a very specific setting, the general

problem studied in [109] still remains a major challenge after many years of efforts.

1.3.2 Random Distances Associated with Complex Geometries

From highway to Manhattan grid, vehicles have the location constraints along straight lines. As the contribution from the theory perspective, the research in this dissertation extends the network topology to a two-dimensional space. Different from the previous vehicular network scenarios, the resultant geometrical probability models have much less constraint on the locations of wireless devices.

For the first time in literature, a geometrical probability approach is presented for obtaining the distance distributions associated with complex geometries: rhombuses and hexagons. The results provide a mathematical foundation for an accurate evaluation of the location-critical performance metrics. Based on the derived distance distributions, the geometrical probability model gives invaluable insights to the protocol design, as being a powerful, versatile tool that is built upon the elegant theory of geometric probability [93].

(1) Geometrical Probability in Two Dimensions

The distributions of random distances over elementary geometries such as squares, rectangles and circles, have well-established results in the geometrical probability literature [37,38,66–68]. In two-dimensional geometries, the derivation approaches in these works have one common assumption that, the coordinates of a point have to be independent. For instance, when using the Cartesian system, the distribution of $\mathcal{X}$-coordinate of a point in a rectangle is not affected by its $\mathcal{Y}$-coordinate; in a circle where the polar coordinate system is used, the distribution of radial coordinate and angular coordinate of a point are also independent. The approaches for these elementary geometries are not applicable to complex geometries, where the coordinates of a point are interdependent, such as in rhombuses and hexagons.

However, hexagons are one of the topological shapes most suitable for cellular systems [40], and rhombuses are for sectorized cells with directional antennas. Both of them have important applications in wireless communications, and other fields including city planning and transportation [28], forestry and chemistry [66], etc. Despite the recent deployment of picocells and femtocells [31] that are designed for indoor, small-scale cellular coverage, the hexagonal tessellation is the classic topology for out-

door, large-scale network coverage. From the perspective of network operators, such topology is critical to guaranteeing satisfactory and profitable service coverage while minimizing the deployment and maintenance costs. In the literature of geometrical probability, however, the problem for complex geometries remains unsolved. The development of new models for the random distances associated with complex geometries thus has become the critical factor between the theory of geometrical probability and the performance analysis for networks with complex topologies.

Our work [105,106] and [107] developed a new, unified approach through a quadratic product that tackled the above problem by presenting the explicit distance distributions for rhombuses and hexagons. Via an affine transformation in plane geometry [1], this product formulation not only handles the geometries where node coordinates are interdependent, its degenerated form also gives the exact same results for squares and rectangles as those in the classic geometrical probability research.

The novelty of this approach is twofold. First, from a mathematical point of view, there is no fixed reference point required, which makes the problem significantly more challenging and distinguishes the contribution of this dissertation from [15, 89, 103]. The results derived enable analytical models in a wider spectrum with less location constraint. Second, the coordinates of a node can be interdependent. The results are not only suitable for convex topologies, but also applicable to the networks with concave geometric shapes. Previously, conducting accurate analysis on network performance metrics has been intractable for complex network geometries. In this dissertation, this gap in the literature has been filled. The rigorousness and accuracy of the derived distributions have been verified through both mathematical validation and simulations. [105] and [106] also illustrate the use of our probabilistic distance models in a computation-effective manner with polynomial fitting.

(2) Application of Random Distances in Two Dimensions

As stated in Section 1.2.2, the major challenges in communication networks are the non-deterministic nature of wireless communication, and the nonlinear relationship between random distances and location-critical performance metrics. They make pertinent a rigorous characterization of network performance metrics by means of the closed-form distance distributions.

It is shown in our work [107] that, in both sparse and dense network scenarios, the state-of-the-art approximations are not accurate when analyzing the nonlinear,

location-critical network performance metrics. Examples are given for the analysis of the nearest neighbor distribution in a sparse network for improving energy efficiency, and the farthest neighbor distribution in a dense network for minimizing routing overhead. A further study on transmission power control shows that, while the current approximation methods contain errors and deviations that are inevitable, the geometrical probabilistic approach provides accurate results for network performance metrics. Therefore, the explicit distributions for rhombuses and hexagons in [105] and [106] not only solve an open problem in the geometrical literature, but also gracefully eliminate the errors in empirical and approximation methods. The probabilistic distance models hence bridge the gap between the network performance metrics, and the explicit distribution of random distances in complex geometries. The analytical results are critical to the fine tuning of protocol parameters, and the accurate modeling of performance metrics.

To sum up, the contributions of this dissertation have profound impact on the location-critical system performance metrics, from the perspective of an individual user, as well as the service providers and network operators. The aforementioned signal attenuation, connectivity probability, etc., are critical performance metrics and important service requirements that directly affect users experiences and satisfaction levels in data transmission. More importantly, the insightful results from the geometrical probability methods, which utilize the knowledge of network topology and user distribution, enable network engineers and service providers to effectively improve service quality, network planning, deployment and resource management.

1.4 Outline of Dissertation

The rest of this dissertation is organized as follows:

Chapter 2: Random distances associated with different geometric shapes, where node locations follow a certain distribution, have been research problems with a long mathematical history. In this chapter the classic work with the main focus in the field of mathematics and statistics, and the application of these results in communication networks are reviewed.

Chapter 3: Based on an exponential distribution of inter-vehicle distances, or Poisson point process, this chapter studies the fundamental limits of message propagation in a vehicular ad-hoc network scenario. On a one-dimensional highway,

the message propagation is always limited to a certain distance from the message source. Therefore, the reverse traffic is incorporated into the framework, further extending the vehicle connectivity among disconnected network components.

Chapter 4: This chapter extends the model in Chapter 3 from a one-dimensional highway to a two-dimensional Manhattan-like city, with the same Poisson vehicle distribution. Different from the highway scenario, there exists a critical threshold above which the entire network is connected with a high probability. This result is surprisingly similar to the percolation phenomenon.

Chapter 5: A unified approach is developed for the complex geometry, rhombuses, by the formulation through a quadratic product and an affine transformation in the plane geometry. The results are applicable to both convex and concave geometric shapes. The proposed approach is able to handle interdependent point coordinates, by separating the geometric shape and the characteristics of the random coordinates inside the shape.

Chapter 6: A regular hexagon can be divided into three congruent rhombuses. In this chapter, the derivation of distance distributions associated with hexagons is presented, through a probabilistic sum of various geometrical rhombus combinations. By analytical and simulation comparison, in both sparse and dense network scenarios, the state-of-the-art approximations of hexagon distributions are determined to be not accurate.

Chapter 7: This chapter concludes the dissertation with further work. Future research plans beyond this dissertation, revolve around the probabilistic distance distributions under the impact of human and vehicle mobility, and the stochastic models of network performance metrics. The conditional distance distributions also have a profound impact on the wireless channel models and cooperative communications. It is anticipated that the applications of this research will be numerous and diverse.

Chapter 2

Background and Related Work

Random distances associated with different geometrical shapes, where node locations follow a certain distribution, have been research problems with a long mathematical history. Starting with a simple but important geometry, the distance distributions along a one-dimensional straight line allows an investigation on the location-critical performance metrics in a basic scenario. With a Poisson point process, the distance between the randomly and independently distributed points can be characterized by an exponential distribution.

In more practical scenarios, i.e., two-dimensional networks with finite sizes, the random distance distributions are more complicated. However, these network topologies are the most frequently encountered in practice. Looking back into the study on geometrical random distances over the past 60 years, considerable research efforts have been made to obtain the closed-form distributions in two dimensions. However, all the existing results for the random distances only have explicit forms in elementary geometries, such as rectangles [9, 68] and circles [69], or for the distance from a fixed reference point [66, 69].

This chapter presents the theoretical background of geometrical probability and the applications in the networking research. First, the classic works in geometrical probability are reviewed. These results include the closed-form distributions in simple geometries, and the empirical approximations in complex geometries. This chapter also introduces the application of these results in the existing literature. However, these applications in communication networks are still preliminary.

2.1 Geometrical Distributions

2.1.1 One-Dimensional Random Distances

In a one-dimensional network where nodes are distributed along a straight line, communication occurs if the distance between consecutive nodes is less than the transmission range. Albeit simple, the analytical results in the one-dimensional space are highly important, as they can provide bounds on the connectivity problem in higher dimensions by approximations [30, 32, 85].

One-dimensional networks have been studied in the context of cellular networks [54] and circuit-switched networks [35, 55, 113]. In a one-dimensional geometry, the most widely-used point distribution model is the Poisson point process. This can be seen as a special case for the birth-death process, or the branching process introduced by Gilbert [39]. Such a process is also the basic point distribution model in the field of continuum percolation [42].

As stated in Section 1.3.1, one-dimensional Poisson model is important since it represents a meaningful model for many applications. In particular, it is well-suited for vehicular ad-hoc networks, where the mobility of vehicles is constrained by the road structure. Through the statistical analysis of empirical data collected from real scenarios, the authors of [13, 84, 86, 94, 100, 102] etc., discovered that such a model is a good fit for highway vehicle traffic in terms of inter-vehicle distance and time distribution. Given a single parameter, the vehicle density λ, this model is able to describe the characteristics and variation of highway traffic.

The analysis provides insight into more complex or even two-dimensional networks. For example, the one-dimensional Poisson model on a highway can be extended to cover the cases of multiple lanes with different vehicle densities, where multiple lanes can be treated as a single lane with a higher traffic density. When the different lanes on the same road segment diverge, the traffic on a diverging lane can be modeled as a *thinning* Poisson process.

2.1.2 Two-Dimensional Random Distances

The study of the distribution of random distances in two-dimensional geometries dates back to the late 1940's [37, 38]. The problem of deriving the *expected* distance between random points was listed as problem number 75-12 of the Society for Industrial and Applied Mathematics Review [76]. While this problem has drawn considerable

attention from the literature (please refer to [36, 43, 92] and the references therein), obtaining the *distribution* of random distances, which leads to all statistical moments, turns out to be very challenging yet highly useful when analyzing the location-critical performance metrics in communication networks. Some research focused on the random distances when one of the endpoints of a link is fixed [66], whereas the problem becomes especially difficult when both the endpoints are random. Nevertheless, given the distribution of random distances in a two-dimensional space, more problems can be tackled in addition to the connectivity problems in one dimension.

(1) Two-Dimensional Geometrical Distributions Associated with Elementary Geometries

[9] is among the first efforts towards the derivation of distance distributions with random endpoints. In this work, the classical Crofton technique and its extensions were used for obtaining the geometrical distributions of random distances associated with circles and squares, both of which are elementary geometries. Later, [66, 68, 87] showed a few simple geometrical cases where their distance distributions can be derived analytically. [67] in particular, is a collection of methods for distance distributions in different elementary geometric shapes.

These methods either use local or global perturbations, differential equations or standard statistical techniques. By means of a geometrical probability approach, the resultant spatial distance distributions provide key insights into the understanding of the probabilistic nature of wireless communication networks, including our previous work [108]. However, many of these efforts either studied random distances from a fixed reference point, or the *chord length* where both endpoints are limited to the boundary of the geometry [9]. Another limitation of these works is that the employed technique only yields explicit distribution for a specific elementary network geometry [67]. Their inflexibility in handling interdependent point coordinates has limited these distance models to a certain number of geometries [37, 38, 108].

In elementary geometries, our previous work [108] proposed an energy consumption model in sensor networks, based on the probabilistic distance distribution when a sensing field is partitioned into a number of square grids. This model utilizes the probabilistic distribution functions of inter-node distances associated with squares, which are much more accurate than the work prior to [108]. In Chapter 5, a new model based on a quadratic product is presented to derive the distance distributions

in two-dimensional spaces. The new model is applicable to the scenarios when the coordinates of points are independent (elementary) and interdependent (complex).

(2) Two-Dimensional Complex Geometries and Empirical Approximations

Spatial reuse is a fundamental enabling tool to achieve a higher network capacity for the ever-increasing demand of wireless services. A wireless communication network is usually divided into congruent polygons, or cells. Examples exist for equilateral triangles, squares, and regular hexagons. Hexagons are typical in cellular systems as they provide the most economic coverage of the network, without leaving gaps or creating overlapping between cells. Moreover, any given point inside a hexagon is closer to the center of the hexagon than any point in an equal-area square or triangle [4]. However, the results of distance distributions associated with hexagons are lacking in the geometrical probability literature, which poses significant challenge to the networking research.

Hexagonal tessellation is the network topology applicable to both infrastructured and ad-hoc networks. In cellular systems, for example, the base station is usually located at the center or at one vertex of a hexagonal cell, whereas the subscribed users are located inside the cell with a certain distribution. If the base station uses directional antenna, then each of the sectorized cells is a rhombus. If the distance distributions associated with hexagons are given, analytical model can also be constructed for ad-hoc networks where no base station exists. In an ad-hoc network that is partitioned into a honeycomb, each hexagon is an autonomous entity in which nodes communicate directly with one another, e.g., the wireless sensor network partitioned into hexagonal clusters as in [29].

The hexagon topology is important for service providers for the proper planning and dimensioning of network coverage and service infrastructure. However, in all the existing works, no analytical models have been proposed for the applications of geometrical probability associated with hexagons in closed form. The state-of-the-art models are limited to the distribution with one fixed point in a hexagon [15, 17, 103], or similar circular approximations as those shown in Figure 6.1(b) and (c) [33, 60]. More recently, empirical approximations [14, 16] have been used to obtain the distance distributions in non-overlapping geometries. However, these methods do not provide much insight into the geometrical problem itself. Approximation errors are also inevitable, especially when the distance distributions are applied to the analysis

of complex, nonlinear performance metrics. The development of new models and approaches for the derivation of random distances associated with hexagons thus has become an imminent research problem.

2.1.3 Summary

Using the distance distribution functions derived for both elementary and complex geometries, further analysis can be conducted on the location-critical performance metrics in various network topologies. Such a geometrical probability approach has a wide range of applications in wireless communication networks, as will be shown in the following section. When compared with the models using explicit distributions, traditional methods based on moments and approximations are less accurate or even not applicable.

2.2 Geometrical Probability Approach and Location-Critical Performance Metrics

In a two-dimensional space, elementary or complex, the locations of different transceivers and interferers, and the Euclidean distances between them, have a profound impact on the operation of more general wireless communication networks. Initial efforts have appeared in the literature analyzing certain performance metrics that are distance-related, utilizing the random distances that are associated with geometries in two dimensions.

2.2.1 Connectivity, Position-Based Routing and Hidden Terminals

In the Boolean connectivity model [11, 12], two nodes are connected when their distance is smaller than the transmission range R. Based on this model, [73] derived the distribution of link distances, or hop distances, between two random radio transceivers in a wireless network covering a rectangular area. Given the transmission range R, [74] studied the joint distribution of link distances, i.e., two-hop connectivity, in a square area, which was based on the results in [73]. [18] investigated the discrete probability distribution of the minimum number of wireless hops, or hop count, between a random source and destination.

There are also ad-hoc routing protocols which make forwarding decisions based on the geographical location of a packet's destination [71]. For example, [96] derived the connection probability of the best path and the probability of having at least one alternate path. Furthermore, the k-th nearest neighbor distance is also crucial for relay and routing protocols [18]. [90] found that in a network where nodes are distributed according to a binomial point process (BPP), the distance from the source node to the k-th nearest neighbor follows a beta distribution.

Furthermore, hidden terminal occurs when two nodes that are more than R apart transmit simultaneously, while their transmissions collide at a third node that is less than R away from both. In contrast, if this third node is cooperating, it can serve as a relay, improving network throughput by exploiting spatial diversity. The joint distribution of the pairwise distances between any three random nodes, thus determines the impact of hidden terminals or the efficacy of a relay node.

The above metric, node connectivity, is critical to the reliability of message delivery, as well as the minimization of multi-hop energy consumption in our previous work [108]. Nodes are typically stationary, and a sequence of data forwarding results in different covered distances at each hop towards the destination. These performance statistics, determined by the pairwise distance between intermediate nodes, are crucial to the applications with energy constraints, yet requiring high message delivery ratio. However, the above works have tackled problems in elementary geometries, such as rectangular and circular networks. The analytical models in complex geometries are lacking.

2.2.2 Path Loss, Fading and Shadowing

In the Boolean connectivity model, the inter-node distance determines the existence of a communication link. However, when path loss, fading and shadowing are taken into account, the radio propagation model becomes more complicated than a disk with a fixed radius R.

In a wireless communication channel, the strength of transmitted signals falls off with the distance between transceivers at rate α, the path loss exponent. When there are no obstacles between a pair of transceivers, or when a line-of-sight (LOS) path exists, the attenuation of radio signals is proportional to the square of the distance, i.e., the free-space propagation when $\alpha = 2$. In a more realistic environment, the surface of the earth or obstacles cause a reflection, diffraction or scattering of the

signals, each of which leads to an extra loss of power over the distance between obstacles and transceivers. In [15], the analytical and approximated expressions for the path loss statistics are derived, although with the random distance distribution from the center of a hexagonal cell.

Power control schemes are needed to determine the required transmission power to overcome path loss. Moreover, if any of the transceivers are moving, the Doppler shift results in an even more complicated communication channel. After being reflected or diffracted, the area where the communication can happen successfully is no longer a disk with a fixed radius, but rather an irregular shape that is dependent on the node locations and network geometry. Meanwhile, the energy required to successfully deliver a packet increases nonlinearly with the distance between transceivers. The energy consumption (or received power) thus can be expressed as the (reciprocal) power-α of the node distance, such as in [15] and our previous work [108].

[46] presented a model for predicting the site-specific radio propagation characteristics, based on the geometrical probability of the layout in an indoor environment. The key parameters of the power/delay profile are directly related to the floor layout. However, the complicated ray-tracing in [46] has rather high computational cost. In [44], an analytical model of the frequency-selective indoor radio channel was presented, including the computation of the first and second moments of the received power, and a log-normal approximation of the received power distribution. [45] further derived the bounds for the moments of the received power, which depend on the volume and the surface area of a convex body. The above work, however, has only tackled the moments of the distances (means and variances) [44, 45], characterized the performance metrics for simple and very specific network topologies [46], or the distribution with a fixed point [15, 89, 90, 103].

2.2.3 Interference, SINR and Channel Capacity

The received signal at a receiver, albeit being attenuated or reflected, is superimposed with other unintended signals transmitted in the vicinity [89]. This leads to the mutual interference between wireless links that are active at the same time. It is possible that a set of nodes, even outside the interference range of the receiver, simultaneously transmit and their cumulative interference causes packet corruption at the receiver. Knowing the statistical distribution of node distances, the cumulative interference at the receiver can be modeled as an additive random variable. Co-channel

interference (CCI) in hexagonal cells is approximated in [17], where interference originates from neighboring cell centers. Similar distance model is used in the analysis of the interference factor in [21]. In [80], the outage probability, a measure that links the aggregate interference to the quality of service, is analyzed for Ultra-Wideband (UWB) systems. The network topology is only limited to the square-shaped cells, and the link distances are from a fixed point in a square. The outage probabilities in circular, square, and hexagonal cell geometries are compared in [79], where link distances are also from the cell center. Even the most recent works [10, 51, 78] model the locations of access points in a cellular system as a two-dimensional Poisson point process, due to the complexity of the problem. However, two-dimensional Poisson point process is only applicable to a network of an infinite size, and the cannot capture the hexagonal geometry of cells. In a general network where no reference point exists, approximations are used [89, 98].

The received signal strength and the interference by concurrent transmissions eventually determine the capacity of a system. A successful decoding of the received symbols is a random event with probability dependent on the ratio between the desired signal strength from an intended transmitter, and the unintended interference plus thermal noise, i.e., the signal-to-interference-and-noise ratio (SINR). Given an acceptable bit error rate (BER), the packet reception is successful if the SINR is greater than a certain threshold.

2.2.4 Stochastic Properties of Random Mobility Models

In a mobile network, if a device is allowed to move randomly along straight lines inside a certain region, its trajectory is formed by a set of polylines between random points [47]. By describing this mobility model as a discrete-time stochastic process, many of its fundamental stochastic properties can be investigated, with respect to the transition length of a mobile node between random points [19]. By also knowing the speed characteristics of the device, one can obtain the travel time and transition process. Similarly, the tour length of the traveling salesman problem (TSP) in a region with uniform demand density can be reduced to the same problem [81]. The resultant travel time is important to many time-sensitive applications where a minimal service latency is desired.

In the above work, approximation or theoretical bounds were given for certain network scenarios [14, 16, 44–46], or a reference point is needed in order to obtain an

explicit distance distribution, such as the work in [15,17,79,80,89,90,103]. In this dissertation, a new approach is proposed to derive the distance distributions when both endpoints of a link are random, and when the point coordinates are interdependent. As a result, more challenging problems can be solved. As mentioned above, only the interference and outage properties from base stations were analyzed in wireless communication systems in the current literature.

2.2.5 Summary

All the metrics listed in this chapter are related to the node location and distance, which are in close relation to the network deployment (i.e., the distribution of network devices) and geometry (network topology and size). They are particularly important at the network planning and dimensioning stage. Even though the knowledge of random distances is crucial in the networking research area, relatively little relevant work has been done to give a general, unified formulation, and no explicit results are available in the literature for complex topologies.

In this dissertation, the results of the existing geometrical probability give us the statistical distance distributions over elementary geometries, which are applied to the evaluation of location-critical performance metrics in one-dimensional network topologies. Meanwhile, the new models for deriving geometrical distributions for complex geometries are able to deal with both elementary and complex network topologies, and convex and concave communication regions, through a simple but elegant formulation.

In the next chapter, a time and location-critical message dissemination framework is proposed for vehicular communications in a highway scenario. Utilizing the Poisson distribution in one dimension, the location-critical performance metrics are analyzed in details.

Chapter 3

The Poisson Point Process and Vehicular Ad-Hoc Networks on Highways

Vehicular ad-hoc networks (VANETs) are emerging paradigms in sensor networks, which use different sensing devices available in vehicles, to gather environment information and provide intelligent traffic information services. VANET promises to enhance the road safety and travel comfort significantly in both highway and city scenarios. Message propagation, either for emergency or infotainment, constitutes a major category of VANET applications. It is particularly challenging in infrastructure-less vehicle-to-vehicle communication scenarios.

Early research on vehicle traffic theory has found that the inter-vehicle distances follow an exponential distribution [13, 84, 86, 94, 100, 102]. In this chapter, this traffic model is used for the study of network connectivity in a one-dimensional network topology. A time and location-critical framework is proposed for the emergency message dissemination among vehicles in a highway scenario, where vehicles at different distances to an accident site can receive information with different levels of details. This framework is achieved through the previously proposed scalable modulation and coding (SMC) scheme, which allows messages of different importance to be broadcast to different distances simultaneously. Such a unique feature fits well with the requirement of instant collision avoidance and advance travel planning in VANET.

The geometrical probability approach in this chapter gives accurate analysis of location-critical performance metrics that are crucial for message dissemination in

vehicular ad-hoc networks. Although the data dissemination in VANET and MANET has been studied extensively in the literature, to the best of our knowledge, this is the first time that multiple deadlines at different locations are taken into account at the same time. The performance metrics are analyzed through simpler approaches than the current literature [75, 100], yet with higher accuracy. In the next chapter, the extension of this model is applied to a two-dimensional Manhattan-like city. The term vehicles and nodes are used interchangeably, when it is obvious from the context.

3.1 Spatio-Temporal Vehicular Traffic Models

3.1.1 The Poisson Point Process

As a major component of the future intelligent transportation systems (ITS) [22], VANET brings huge economic and social impacts to the more connected lifestyles and activities, by enabling inter-vehicle communications with or without the assistance of roadside infrastructures. Vehicles equipped with different onboard sensing devices are used for gathering environmental information and providing intelligent traffic information services. With the allocation of a 75 MHz licensed band at 5.9 GHz for the Dedicated Short Range Communications (DSRC) [3] around the world, VANET has become increasingly popular and attracted considerable attention from both the academia and industry.

Many vehicle mobility models have been proposed in early research studies, including the car following model and other variants. More recently, through the statistical analysis of empirical data collected from real world scenarios, the authors of [13, 84, 86, 94, 100, 102] have arrived at a surprisingly similar conclusion that an exponential model is a good fit for highway vehicle traffic, in terms of inter-vehicle distance and inter-contact time distribution. Equivalently, the vehicle arrival process is modeled as a Poisson point process. As described in Section 1.3.1, Poisson distribution describes the probability of a given number of events occurring in a fixed interval. It is widely used in modeling the arrival process of vehicles, such as the models in traffic flow theory [41].

Message dissemination, either for emergency or infotainment, constitutes a major category of VANET applications. In a highway scenario, emergency message (EM) dissemination is the most important application for the safety of drivers and passengers. In VANET, message dissemination depends on the underlying location-

critical connectivity among vehicles, which is determined by the transmission range of equipped wireless devices and inter-vehicle distances. The transmission range of DSRC is typically 200 m up to 1 km, such as that in IEEE 802.11p [52]. On the other hand, given the vehicle arrivals as a Poisson distribution, the characteristics and variation of highway traffic can be captured with a single parameter: vehicle density λ, in the number of vehicles per meter.

3.1.2 One-Dimensional Connectivity

As in Section 3.1.1, the fundamental connectivity property of VANET depends on a given transmission range R, and inter-vehicle distances. Two consecutive vehicles on a road are directly connected when their distance is less than R. Therefore, the distance between two vehicles determines the existence of a communication link, i.e., the Boolean connectivity model [11, 12].

(1) Boolean Model and Network Connectivity

Existing studies show that the network connectivity in one-dimensional space is always limited. In practical situations, both R and $\lambda < \infty$. Therefore, for any two consecutive vehicles that are separated by distance d, the probability of a disconnection, $\Pr\{d \geq R\} = e^{-\lambda d}$, is strictly positive for all $d \geq 0$. As a result, the disconnection happens almost surely (a.s.) [1] [2]. Between these disconnections, a finite number of vehicles are connected sequentially with each other in a group via multiple hops. Such connected groups of vehicles are defined as *clusters*.

In contrast, in a two-dimensional space such as the city blocks in Manhattan, network connectivity can be guaranteed if the density among nearby vehicles is above a certain threshold. This is the so-called percolation phenomenon [25, 42, 57] where the entire network is connected almost surely. When a giant cluster with infinite size appears in a network, the network is said to be *percolating*. This case will be studied in the next chapter.

(2) Vehicle Clusters and One-Dimensional VANET

Consider a sequence of vehicles distributed along a one-dimensional highway or street segment. To analyze the network connectivity in such a one-dimensional space, the

[1] Let (Ω, F, P) be a probability space. An event E in F happens *almost surely* if $\Pr(E) = 1$. Equivalently, an event E happens almost surely if the probability of E not occurring is zero.

size of a vehicle cluster is particularly important, where the size is defined as the *distance* between the first and last vehicles in the same cluster. From the argument in above, the cluster size is finite with high probability (w.h.p.). Because all the vehicles in a cluster are connected to one another by single or multiple hops, the size of a cluster determines how far a message can be disseminated from one end of a cluster to the other end with a low delay.

When vehicles travel with relatively constant speed, messages cannot be propagated beyond a cluster, i.e., no communication happens between clusters in the same direction. However, the traffic in the reverse direction can be utilized to extend the connectivity between clusters, and carry-and-forward messages to vehicles that belong to different clusters. By taking the advantage of the reverse traffic, the one-dimensional connectivity becomes more dynamic and opportunistic, although the resulting message dissemination incurs a higher propagation delay. The carry-and-forward delay between clusters is dominated by the distance between forward and backward clusters in opposite directions, because the vehicle speed is much slower than the electromagnetic wave propagation within the same cluster.

(3) Related Work

Some of the existing work in the literature has studied the one-dimensional connectivity problem. In [30], nodes are assumed uniformly distributed in $[0, z]$ forming a one-dimensional ad-hoc, multi-hop radio network. By using Laplace Transforms, the authors derived the probability of network connectivity as a function of transmission range. However, an (n-1)-fold convolution is needed to obtain the probability of connectivity in a one-dimensional network with n nodes. A lower bound in a two-dimensional network is given by approximation.

In [26], a recursive formula giving the average number of hops between two connected nodes is derived, and the probability that a given number of nodes on a finite interval are all connected is computed in [85]. These performance metrics are evaluated with respect to the number of nodes. In vehicular ad-hoc networks, the location is critical to the network performance, especially for applications that are safety-related. Therefore we focus on the connectivity properties with respect to the distance between vehicles.

By using an equivalent $GI|D|\infty$ queuing model, the authors of [75] derived the connectivity probability in one-dimensional networks. In [75], node positions were

equivalent to customer arrivals, and the service process has a fixed duration R. The occurrence of a vacant space in a one-dimensional network is equivalent to a visit to the idle state of the queue. Because $\lambda < \infty$, the idle state will be visited almost surely within a finite time. The results were obtained by Laplace-Stieltjes Transforms (LST), and simplifications were made to obtain an approximated expression.

In this chapter, a time and location-critical (TLC) framework is proposed for the message dissemination among vehicles in a highway scenario. A new approach that is based on the memory-less exponential distribution is developed to analyze the fundamental connectivity problem, but without simplification or approximation. The derivation and simulation show that this new approach is simpler, and the results are much more accurate.

3.2 Time and Location-Critical Framework for Vehicular Ad-Hoc Networks

Among the vast array of potential ITS applications, emergency message (EM) dissemination is considered critical to all safety-related applications. When an accident occurs or a certain road condition is observed, referred to as a point-of-interest (POI), vehicles at different distances away need different levels of information in the propagated message. The vehicles close to the POI need detailed information immediately to react properly, e.g., slowing down or changing lanes, due to the short reaction distance and the resultant short reaction time. Meanwhile, the propagated message needs to be disseminated further to allow following vehicles to make location-aware decisions, e.g., detour at a highway exit or reroute whenever possible. It is crucial that the time and location criticality of EM dissemination should be taken into account at the same time.

Through an in-depth analysis on the one-dimensional VANET connectivity property, the location-critical performance metrics such as message propagation delay, and the likelihood of missing an emergency message, etc., are investigated.

3.2.1 Scalable Modulation and Coding (SMC)

This section first presents a TLC framework for EM dissemination, using the scalable modulation and coding (SMC) scheme proposed in [24]. The purpose is to allow

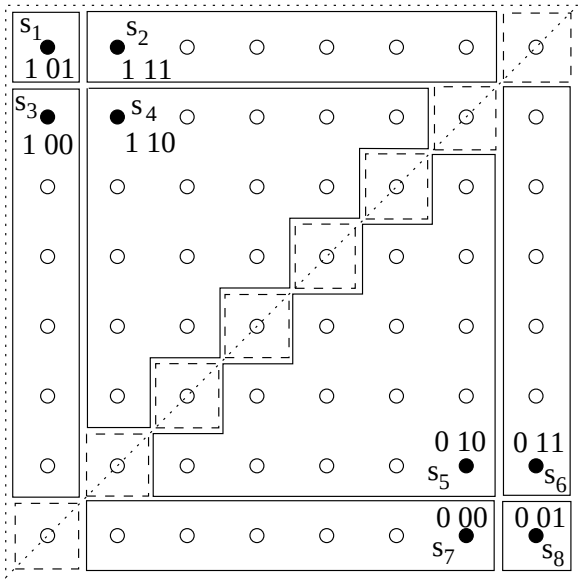


Figure 3.1: Scalable Modulation and Coding using the 64-QAM Constellation.

vehicles at different distances to the transmitter receive different levels of information in a *single* message transmission. By redefining the bit-to-symbol mapping in the modulation constellation, SMC [24] allows a transmitter to encode information of different importance simultaneously. In a nutshell, the receivers that are close to the transmitter and with a high SINR can decode more information bits of both high and low importance, i.e., the message is decoded with more detailed information. The receivers further away or with a lower SINR decode less information from the same broadcast transmission.

Rather than utilizing different modulation schemes to broadcast a single message at multiple times, allowing transceivers to encode/decode different information bits at the same time significantly reduces the non-deterministic medium access delay. Therefore, the SMC scheme fits well with the TLC framework for EM broadcast, where a low propagation delay is highly desired. The vehicles close to the POI need guaranteed, detailed information for quick maneuvering, e.g., cruise control or even autopilot, while further-away vehicles first receive an early warning and then obtain more detailed information as they approach the POI.

SDMD and LDLD Messages

With SMC, a message can be delivered over short distances with more detail (SDMD) and over long distances with less detail (LDLD) simultaneously. Refer to the bits of

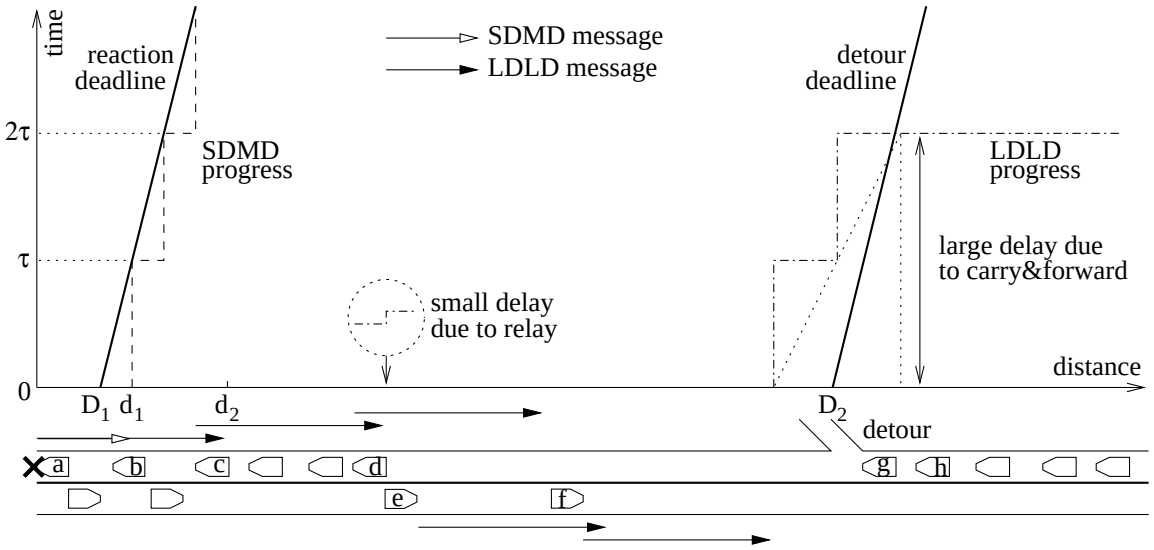


Figure 3.2: Time-Location Critical Emergency Message Dissemination Scenarios.

information to be transmitted over a longer transmission distance as the layer-1 bits (L1), i.e., the basic information in both SDMD and LDLD messages. Further, the bits of information over a short transmission distance, i.e., the extra information provided in SDMD only, are referred to as the layer-2 bits (L2). Figure 3.1 illustrates an SMC design using the 64-QAM constellation. Specifically, 8.5 dB and 25.5 dB are the SINR threshold required for L1 and L2 bits to achieve the 10^{-5} BER requirement. This SINR translates into an 879 m and 206 m transmission distance for L1 and L2 bits [110], which meets the standard requirement of IEEE 802.11p DSRC transmission range from 200 m to $1,000$ m [3]. More importantly, this is achieved by a single transmission to send one L1 bit and two L2 bits to different distances at the same time. For the details of the SMC scheme design, please refer to [24].

3.2.2 TLC-based Emergency Message Dissemination

Assuming that all vehicles are equipped with GPS devices. Figure 3.2 depicts the scenario for the TLC message dissemination framework. The x-axis shows the distance from the POI at location 0. For example, vehicle a has an accident at time 0 and broadcasts an emergency message with SMC, which allows the message to be delivered by both SDMD and LDLD at the same time. We denote the short and long transmission distance as d_1 and d_2, i.e., 206 m and 879 m [110].

(1) SDMD Message and Reaction Deadline

For the SDMD message dissemination, the transmission distance d_1 has to be greater than a distance threshold—the reaction deadline D_1. This distance threshold depends on the vehicle stopping distance, given the human/vehicle reaction time and travel speed. That is, any vehicles within D_1 from the POI must be notified immediately to avoid a pileup accident. The y-axis shows the time deadline at different locations, and the slope of "reaction deadline" is equal to $1/v_{\max}$, where $v_{\max}$ is the speed limit. To ensure that all vehicles receive the SDMD message before the time-location dependent deadline, the POI rebroadcasts the SDMD message periodically every τ seconds, with the constraint that $d_1 \geq D_1 + \tau v_{\max}$ in order to accommodate the rebroadcast interval.

(2) LDLD Message and Detour Deadline

On the other hand, the LDLD message should be rebroadcast further to reach as many vehicles as possible. In Figure 3.2, vehicle b receives both the SDMD and LDLD messages, and c receives only the LDLD message from the POI. Ideally, c will rebroadcast the LDLD message after a very small delay [56] compared with the POI's rebroadcast interval τ, assuming that all vehicles know their location and all messages contain location information. The same process repeats at vehicle d. However, when d rebroadcasts, the following vehicle g in the same direction is not close enough, i.e., the inter-vehicle distance is beyond d_2. But e in the opposite direction is in the transmission range of d. Therefore, e will rebroadcast the message to reach f. The opportunistic use of reverse traffic effectively extends the range of LDLD messages to reach more vehicles in the forward direction, when necessary and possible.

As demonstrated analytically later, even relaying by reverse traffic cannot guarantee to reach a vehicle arbitrarily far away from the POI. For example, in Figure 3.2, there is no vehicle in either forward or reverse direction which is close enough to f to further propagate the LDLD message. Thus the message has to be carried by f in the opposite direction while being periodically rebroadcast every τ seconds. Since the vehicle speed is much lower than the speed of electromagnetic waves, the LDLD message incurs a much higher propagation delay, which is dominated by the gap between forward and backward clusters and the vehicle speed. Relative to forward traffic, however, the message travels at the doubled vehicle speed, which is shown by the "LDLD progress" staircase curve in Figure 3.2. The purpose of the LDLD message dissemination is to reach another distance threshold D_2, where a detour exit

is available. Therefore, the warning message is received early enough to minimize the probability that vehicles miss the detour deadline. In Figure 3.2, once the LDLD staircase progress curve intersects with the detour deadline curve, any vehicles before the projection of this intersection on the x-axis and after D_2 will miss their detour deadline (e.g., g), since the LDLD message cannot reach these vehicles in time. In this example, the LDLD message reaches vehicle h before its detour deadline, so h will have the choice to take the detour exit, or travel along the same route and receive the SDMD message with more details later. In this chapter, we let $D_2 \in [1, 10]$ km, motivated by the real highway exit distribution [6] for the purpose of performance evaluation.

(3) Possible Extensions

There are other variants of the TLC framework. For example, in addition to LDLD, SDMD messages can also be relayed by both forward and reverse traffic when possible, conceivably with more intermediate vehicles. Further, more types of messages can be introduced and propagated to different distances away from the POI. For simplicity, we only use SDMD and LDLD in this section to illustrate the TLC framework.

In VANET, contention-based MAC protocols are utilized to resolve media access collisions, which result in random access delay. Since the MAC delay is much smaller than τ and the travel delay, it can be ignored within the same cluster, as shown in Figure 3.2. Many previous research efforts have addressed how to design effective MAC protocols for reliable EM dissemination in a cluster and how to quantify their performance, e.g., [20] etc. These efforts are orthogonal to the work in this chapter, and their delay bounds can be incorporated into our analysis as well.

3.2.3 Forward Direction: Cluster Size Characterization

A cluster is a connected group of vehicles on a one-dimensional highway, in which either SDMD or LDLD can be relayed directly or by multi-hop radio transmissions. The distance between the first and the last vehicles in the same cluster, or cluster size, is of great importance in SDMD and LDLD dissemination for collision avoidance and look-ahead decision making. It is also the foundation of other location-critical performance metrics that will be introduced shortly in this section. In the following, the expectation and distribution of cluster size are derived.

(1) Moments of Cluster Size

As in Section 3.1.2, the authors of [75] obtained the expected cluster size with an equivalent $GI|D|\infty$ queuing model, by means of the Laplace-Stieltjes Transforms (LST) and mathematical simplification, which are in turn based on the derivation in [61]:

$$E[B] = R + \frac{\int_0^R x f_X(x)\mathrm{d}x}{1 - F_X(R)}, \tag{3.1}$$

where $E[B]$ is the expected cluster size, R is the transmission range, and $F_X(x) = \int_{-\infty}^{x} f_X(x)\mathrm{d}x$ is the cumulative distribution function (CDF) of the inter-node distance. Assuming a Poisson point process with parameter λ, or an exponential inter-arrival distribution with mean $1/\lambda$, then $f_X(x) = \lambda x e^{-\lambda x}$ for $x > 0$. Therefore (3.1) gives

$$E[B] = R + \frac{\int_0^R \lambda x e^{-\lambda x}\mathrm{d}x}{e^{-\lambda R}} = \frac{1 - e^{-\lambda R}}{\lambda e^{-\lambda R}}. \tag{3.2}$$

Different from this approximated result, the approach in this section obtains the exact average cluster size without any simplification or approximation. By simply using the independent and identically distribution (i.i.d.) property of the inter-vehicle distance confirmed in [13] etc., this geometrical probability approach is also much more accurate[2]. Denote the random variable (RV) for cluster size by C, the following recursion provides the expectation of C:

$$\begin{aligned} E[C] &= E[C|X_1 < R] \times \Pr\{X_1 < R\} + 0 \times \Pr\{X_1 \geq R\} \\ &= (E[X_1|X_1 < R] + E[C']) \times \Pr\{X_1 < R\}, \end{aligned} \tag{3.3}$$

where X_1 is the distance between the first and second vehicles in the same connected cluster. For the second equality in (3.3), given the i.i.d. distribution of the inter-vehicle distance, the cluster size starting from any vehicle, recursively treated as the cluster head, is the same. This means that a sequence of k vehicles are connected to each other, only if the $k-1$ vehicles are also connected by taking away the first vehicle that is within distance R of the second vehicle in the sequence. Also, $E[C|X_1 < R] = E[X_1|X_1 < R] + E[C']$, where $E[X_1|X_1 < R]$ is the average distance between the first and the second vehicles in the cluster; and $E[C']$, which is equal to $E[C]$ asymptotically, is the average cluster size if the second vehicle is chosen as the cluster

[2]As the width of the street is negligible when compared with the street length, all vehicles in a two-way street are assumed along a line, no matter which lane they actually occupy.

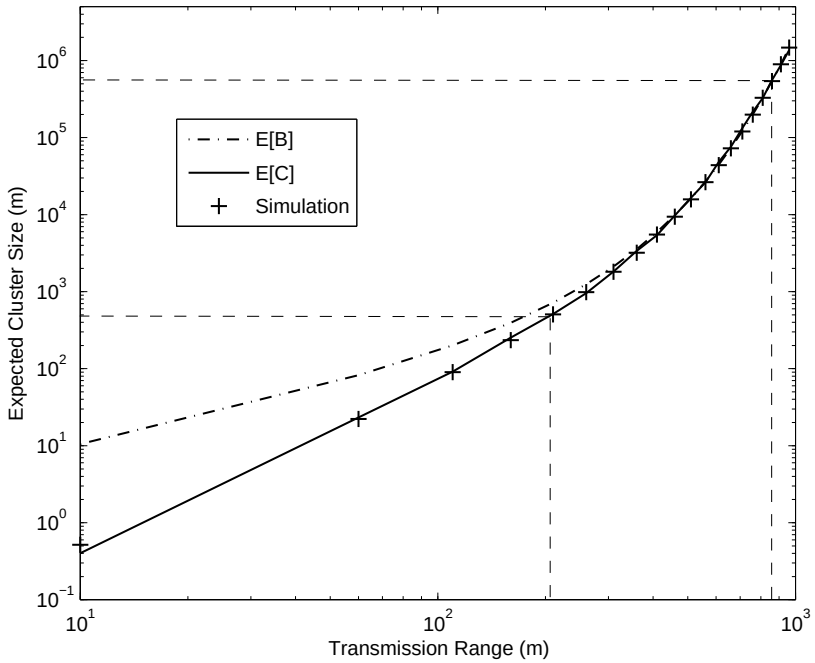


Figure 3.3: Comparison between $E[B]$ in [75] and $E[C]$ on the Expected Cluster Size.

head.

This above recursion is applicable to i.i.d. inter-vehicle distance distributions and is simpler than [75] and [100]. With an exponential distribution, $\Pr\{X_1 < R\} = 1 - e^{-\lambda R}$. Given the conditional expectation $\overline{X_1'} = E[X_1 | X_1 < R] = \int_0^R \frac{\lambda x e^{-\lambda x}}{1-e^{-\lambda R}} \mathrm{d}x$, $E[C]$ is

$$E[C] = \frac{1 - e^{-\lambda R}}{e^{-\lambda R}} \overline{X'}_1 = \frac{1 - e^{-\lambda R}(\lambda R + 1)}{\lambda e^{-\lambda R}}. \tag{3.4}$$

This simple approach gives the *exact* average cluster size without any simplification. Figure 3.3 compares the results from (3.2) with (3.4), and the simulation in Matlab, given the traffic density and transmission range. By distributing vehicles along a straight line according to the exponential inter-vehicle distances, whenever two consecutive vehicles are separated by a distance that is more than R apart, they are in different clusters. The average cluster size can thus be measured. In Figure 3.3, λ is 0.01 (number of vehicles per meter), and the y-axis is the expected cluster size in log scale. From the figure, when λR is sufficiently large, both approaches match the simulation well; but for small R and λ, the results given by (3.4) are much more accurate. For example, at $R = 206\ m$, $E[B] = 685\ m$ and $E[C] = 479\ m$, while the

simulation gives 478 m. The authors in [100] also obtained the average cluster size, or cluster length in their terminology. Their result was obtained via average analysis, which is based on the product of the average inter-vehicle distance and the average number of vehicles in a cluster. Although this approach gives correct result for the first moment, or expectation, it does not apply to higher-order moments.

In contrast, the approach in (3.3) is both simple and direct, and the higher-order moments of C can be obtained in a similar way. For instance, the second-order moment of C is

$$\begin{aligned} E[C^2] &= Pr\{X_1 < R\} \times E[(C + X_1)^2 | X_1 < R] \\ &= \frac{1 - e^{-\lambda R}}{e^{-\lambda R}} \times \left(2E[C]\overline{X'}_1 + \overline{X'}_1^2\right), \end{aligned} \tag{3.5}$$

where $\overline{X'}_1^2 = E[X_1^2 | X_1 < R] = \int_0^R \lambda x^2 e^{-\lambda x} \mathrm{d}x$. Higher-order moments, however, are not discussed in [100].

(2) Cluster Size Distribution

Let the sequence $\{X_i\}$ denote the random variables of the inter-vehicle distance between the i-th and the $(i+1)$-th vehicles, which are i.i.d. with $f_{X_i}(x) = \lambda e^{-\lambda x}$ for $x > 0$. $F_{X_i}(R) = \int_0^R f_{X_i}(x)\mathrm{d}x = 1 - e^{-\lambda R}$ is the probability that consecutive vehicles are within the transmission range of each other. Let $\{X_i'\}$ be the the sequence of random variables of the inter-vehicle distance, given that the i-th and $(i+1)$-th vehicles are in the same cluster, then

$$f_{X_i'}(x) = f_{X_i | 0 \le X_i \le R}(x | 0 \le x \le R) = \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda R}}, \tag{3.6}$$

for $0 \le x \le R$. Suppose that there are k vehicles in a cluster, the Laplace Transform of the cluster size distribution is

$$f^*_{C|k}(s) = f^*_{X_1' + X_2' + \ldots X_k'}(s) = \left[\frac{\lambda}{1 - e^{-\lambda R}} \times \frac{1 - e^{-(s+\lambda)R}}{s + \lambda}\right]^k. \tag{3.7}$$

If $f_{C|k}$ can be obtained by taking the inverse-Laplace Transform on (3.7), then the distribution function of cluster size C is

$$f_C(x) = \sum_{k=1}^{\infty} f_{C|k} \Pr\{k\}, \tag{3.8}$$

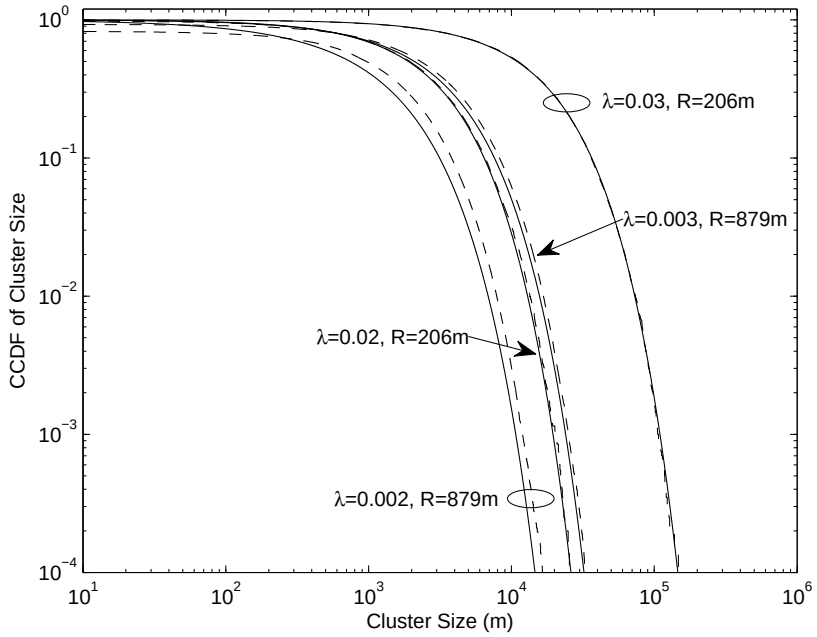


Figure 3.4: Gamma Approximation of the Cluster Size Distribution (Solid Curve—Analysis, Dashed Curve—Simulation).

where $\Pr\{k\} = (1 - e^{-\lambda R})^{k-1} e^{-\lambda R}$ is the probability that there are k vehicles in a cluster, a geometric distribution with $p = e^{-\lambda R}$. Unfortunately, (3.7) does not yield closed-form results by inverse-Laplace Transform. Thus it is very difficult to give the exact formula of the distribution function. In fact, C is the sum of k *truncated* exponential random variables given that they are smaller than the threshold R, and k itself follows a geometric distribution, which further complicates the derivation.

Although the closed-form distribution of the cluster size cannot be derived, we obtain its statistical moments without any simplification. Using these moments, the *Gamma* distribution is used for the approximation of cluster size distribution, since the Gamma distribution has been widely used to model the sum of a fixed number of exponentially distributed random variables but without the truncation. Similar to the average cluster size in Figure 3.3, the cluster size distribution, or the complementary cumulative distribution function (CCDF), can be obtained according to the first and second-order cluster size moments with different values of R and λ. In Figure 3.4, the dashed lines are from simulation results, and the solid lines are the CCDFs calculated

from the following Gamma approximation,

$$f_C(x) = x^{k-1}\frac{e^{-x/\theta}}{\theta^k \Gamma(k)}, \quad \text{for } x > 0, \tag{3.9}$$

where $k = (E[C^2]/E[C]^2 - 1)^{-1}$ and $\theta = E[C]/k$ ensure that the first and second-order moments of the Gamma approximated random variable are the same as $E[C]$ and $E[C^2]$ in (3.4) and (3.5). Gamma distribution clearly is a good fit for the cluster size distribution, according to the simulation results shown in the figure.

Once we have the cluster size distribution or approximation, the probability that a message can reach a location at distance d away from the source, is the probability that the cluster size is larger than d, i.e., $\int_d^\infty f_C(x)\mathrm{d}x$. From $f_C(x)$, it is clear that no matter how large the traffic density is, there is always a non-negligible probability that the gap between two adjacent vehicles is beyond the radio transmission range R. Once a vehicle is disconnected from a cluster, the remainder of the one-dimensional network is also disconnected. Moreover, the probability that a further-away location can be reached decays very quickly in the same direction as the message disseminates, suggested by the Gamma approximation of C. Thus, the message propagation in different directions should be explored.

3.2.4 Backward Direction: Using the Reverse Traffic

By observing the cluster size distribution, it can be found that the vehicle cluster is likely to be small and thus a large number of following vehicles are unable to receive the LDLD message. When the propagation of LDLD reaches the end of a cluster, it cannot be forwarded any further unless certain vehicles in the opposite direction pick up the message, and carry it towards the vehicles that are far behind so that the message is delivered between clusters. Although such message delivery incurs a large delay when compared with that inside a cluster (due to the much lower vehicle speed than that of electromagnetic waves), the reverse traffic can significantly extend the distance that can be covered by LDLD messages, e.g., [8, 100]. [100] gives the upper and lower bounds of the *average* delay caused by reverse traffic, or the "re-healing" time for messages to propagate cross clusters via average analysis and approximations. In this section, however, the "carry-and-forward" delay *distribution* is obtained. The vehicles in the reverse direction can also form clusters to relay the message, which further complicates the analysis.

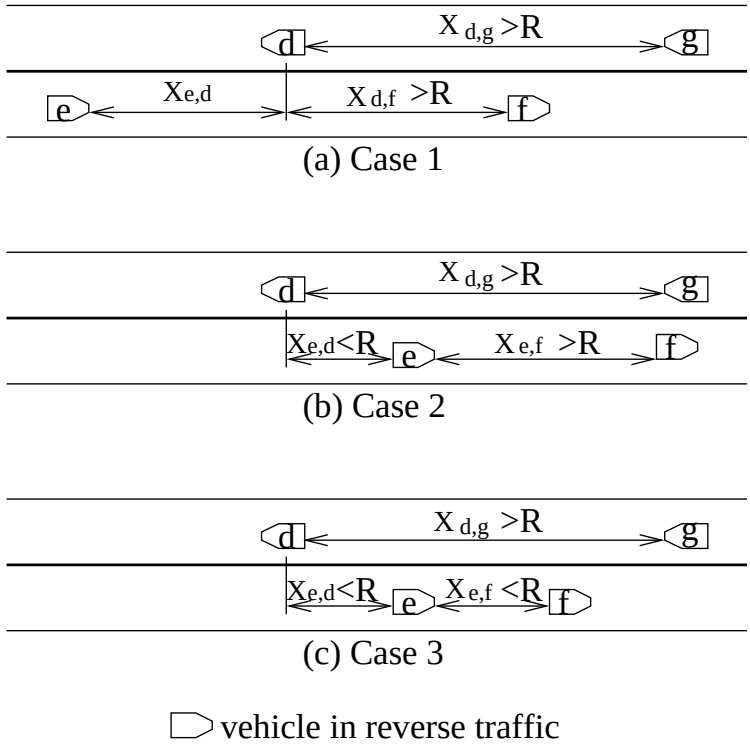


Figure 3.5: Three Cases to Extend the Forward Cluster by Reverse Traffic.

(1) Carry-and-Forward Delay

Denote $X_{i,j}$ as the distance between vehicle i and j at time 0, and $X_{i,0}$ as the distance between i and the POI. Without loss of generality, let vehicle d be the last vehicle in a cluster in the forward direction, g be the first vehicle of the following cluster in the same direction, and e be a vehicle in the reverse direction that can help relay or deliver the message to g, as shown in Figure 3.5. We consider the following three different cases to derive the distribution of the distance traveled by vehicles in the reverse direction to relay a message between consecutive clusters, which dominates the "carry-and-forward" delay when using the reverse traffic.

Case 1**:**

As shown in Figure 3.5(a), there is no vehicle located between d and g in the reverse direction and within the transmission range of d at the same time (i.e., $X_{d,f} > R$). The message needs to be relayed by e in the reverse direction, and $X_{e,0} < X_{d,0}$. The

probability for this case to occur is

$$P_1 = \Pr\{X_{d,f} > R\} = e^{-\lambda R}. \tag{3.10}$$

In this case, the message must be propagated from e to g when e travels close to g. The distance between e and d, $X_{e,d}$, follows the exponential distribution, and the distance distribution between d and g is $f_{X_{d,g}}(x) = \frac{\lambda e^{-\lambda x}}{e^{-\lambda R}}$ for $x \geq R$. Let Y_1 be this travel distance by vehicle e, then $Y_1 = X_{e,d} + X_{d,g}$ has the following distribution

$$f_{Y_1}(y) = \int f_{X_{e,d}}(x) f_{X_{d,g}}(y-x)\mathrm{d}x = \lambda^2 (y-R) e^{-\lambda(y-R)}, \tag{3.11}$$

for $y \geq R$.

Case 2:

As shown in Figure 3.5(b), e is within the communication range of the forward cluster, and it is closer to g than d is. The vehicle following e in the reverse direction, f, is outside both e and d's transmission ranges, therefore f cannot relay the message to g. Thus, $X_{e,0} \geq X_{d,0}$, $X_{e,d} \leq R$, and $X_{e,f} \geq R$. The probability of this case is

$$P_2 = \Pr\{X_{e,d} \leq R\} \Pr\{X_{e,f} \geq R\} = (1 - e^{-\lambda R}) e^{-\lambda R}. \tag{3.12}$$

In this case, the gap $Y_2 = X_{d,g} - X_{e,d}$. Since $f_{X_{e,d}}(x) = \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda R}}$ for $0 \leq x \leq R$, and $f_{X_{d,g}}(x) = \frac{\lambda e^{-\lambda x}}{e^{-\lambda R}}$ for $x \geq R$, the distribution of Y_2 is

$$\begin{aligned} f_{Y_2}(y) &= \int f_{X_{e,d}}(x) f_{X_{d,g}}(x+y)\mathrm{d}x \\ &= \frac{\lambda}{2} \begin{cases} (e^{\lambda y} - e^{-\lambda y})/(e^{\lambda R} - 1), & 0 \leq y \leq R, \\ (e^{\lambda R} + 1) e^{-\lambda y}, & y \geq R. \end{cases} \end{aligned} \tag{3.13}$$

Case 3:

As shown in Figure 3.5(c), both d and f are within the transmission range of e, although in different directions. Thus f, and possibly other vehicles in the same cluster as e and f in the reverse direction, can relay the message further, in order to

reach g faster. In this case, $X_{e,0} \geq X_{d,0}$, $X_{e,d} \leq R$, and $X_{e,f} \leq R$. Therefore,

$$P_3 = \Pr\{X_{e,d} \leq R\} \Pr\{X_{e,f} \leq R\} = (1 - e^{-\lambda R})^2. \tag{3.14}$$

Let $Y_3 = \max\{0, X_{e,g} - C'\}$, where C' is the distance that a message can be relayed within the cluster in the reverse direction, which has the same distribution as cluster size C. The density function of $X_{e,g}$ is derived in Case 2 as $f_{Y_2}(y)$, so $f_{Y_3}(y)$ can be obtained by $f_{Y_3}(y) = \int f_{X_{e,g}}(x) f_{C'}(x - y) \mathrm{d}x$. To simplify the calculation, C' is assumed to be a constant with value $E[C]$. Let δ be the Dirac delta function, and thus the distribution of Y_3 is approximated by

$$f_{Y_3}(y) \approx \begin{cases} f_{Y_2}(y + E[C]), & y \geq 0, \\ \delta \int_0^{E[C]} f_{Y_2}(x) \mathrm{d}x, & y = 0. \end{cases} \tag{3.15}$$

Based on the above three cases, the density function of the travel distance is $f_Y(y) = \sum_{i=1}^{3} P_i \times f_{Y_i}(y)$. Since messages can be directly transmitted between two vehicles that are within the transmission range of each other, the actual distance that will lead to the message propagation delay is $\max\{0, Y - R\}$, and the delay D for the reverse traffic to deliver the message to the next forward cluster has the CDF

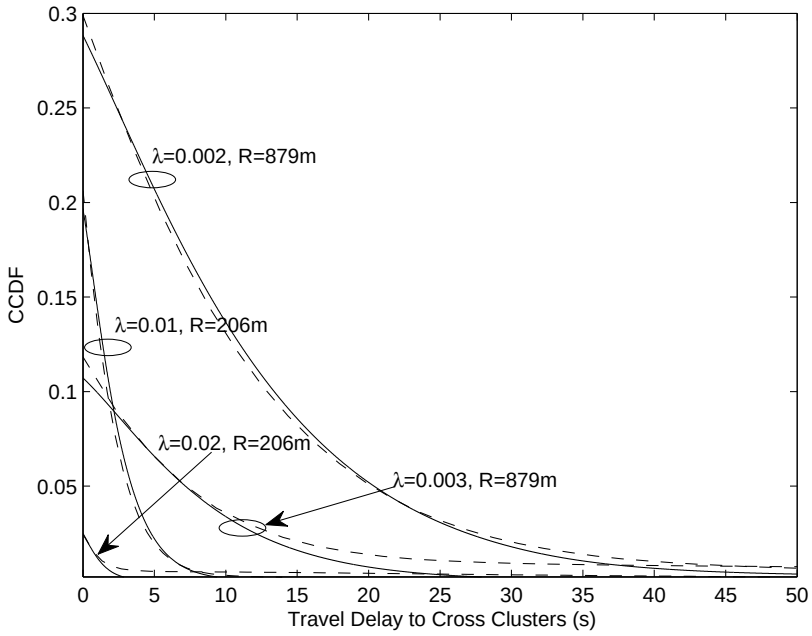
$$F_D(x) = F_Y(2xv + R) \qquad \text{for } x \geq 0, \tag{3.16}$$

where v is the speed of the vehicles in both directions. In [100], the upper bound of travel delay, or the best-case "re-healing" time, is essentially case 3 in Figure 3.5(c); and the lower bound (worst case) corresponds to case 1 and 2 in Figure 3.5(a) and (b). However, [100] only derived the first moment and the approach does not apply to higher-order moments.

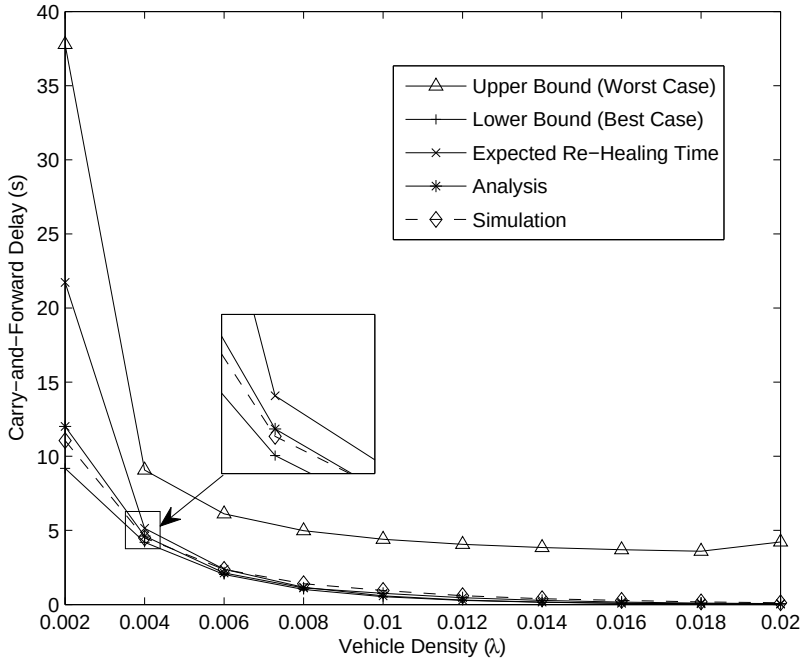
Note that although the derivation of (3.16) is based on constant vehicle speed, according to the invariance of the Poisson distribution under random shifts [72], if the density of the original Poisson point process is λ, after an i.i.d. shift, the resultant point process after the shift is still Poisson with rate λ.

Analytical and Simulation Results

Figure 3.6(a) shows the comparison between the analytical result given in (3.16), and the simulation results by CCDF $\overline{F}_D(x) = 1 - F_D(x)$. In the simulation, the travel time is calculated according to the distance between clusters and the given vehicle



(a) Travel Delay due to Reverse Traffic to Cross Forward Clusters (Solid Curve—Analysis, Dashed Curve—Simulation).



(b) Travel Delay Comparison with [100], $R = 206\ m$.

Figure 3.6: Travel Delay to Cross Forward Clusters and Comparison with [100].

speed, 120 km/h in this case. Traffic density λ is measured by the number of vehicles per meter. Although we simplify the cluster size by its expectation $E[C]$ in (3.15), the analysis and simulation match quite well. Compared with the results in [100] on the upper and lower bounds of the average "re-healing" time in Figure 3.6(b), the work in this section obtaining the "carry-and-forward" delay distribution is a considerable further effort. The analytical results, shown by the *Analysis* curve in the figure, are more accurate at low traffic density. Note that this delay, in the order of seconds, is much larger than the message propagation and MAC access delay within a cluster.

(2) Probability of Missing LDLD Deadlines

In addition to the carry-and-forward delay, the probability of missing an emergency message is another performance metric that is location-critical. We assume that the location of highway exits, i.e., D_2 in Figure 3.2, is known by using either digital map or GPS devices. For a tagged vehicle g in Figure 3.2, which is X_{g,D_2} away from the exit D_2 at time 0, the deadline to receive the LDLD message before g passes by the detour exit is $t_D = (X_{g,D_2})/v$. Let C_1 be the size of the cluster containing the POI, and C_k be the size of the k-th cluster in the forward direction. C_k's are the i.i.d. random variables with the distribution given in (3.8). Let G_k be the distance gap between C_k and C_{k+1}, the distribution of G_k is $f_{G_k}(x) = \lambda e^{-\lambda(x-R)}$ for $x \geq R$.

The probability that g is in the first cluster, i.e., $k = 1$, is $\Pr\{g \in C_1\} = 1 - F_C(X_{g,0})$. For $k > 1$, g is in the k-th cluster when $X_{g,0}$ is larger than the sum of the cluster sizes plus the inter-cluster gaps $\sum_{i=1}^{k-1}(C_i + G_i)$, but smaller than $\sum_{i=1}^{k-1}(C_i + G_i) + C_k$, i.e.,

$$\begin{aligned}
\Pr\{g \in C_k\} &= \Pr\left\{\sum_{i=1}^{k-1}(C_i + G_i) < X_{g,0} \;\&\; \sum_{i=1}^{k-1}(C_i + G_i) + C_k \geq X_{g,0}\right\} \\
&= \sum_{x=0}^{X_{g,0}} \Pr\left\{\sum_{i=1}^{k-1}(C_i + G_i) = x \;\&\; C_k \geq X_{g,0} - x\right\} \\
&= \sum_{x=0}^{X_{g,0}} \Pr\left\{\sum_{i=1}^{k-1}(C_i + G_i) = x\right\} \Pr\left\{C_k \geq X_{g,0} - x\right\}. \qquad (3.17)
\end{aligned}$$

Because x is a continuous random variable, the probability that g is in C_k is

$$\Pr\{g \in C_k\} = \int_0^{X_{g,0}} f_{\sum_{i=1}^{k-1}(C_i+G_i)}(x)[1 - F_{C_k}(X_{g,0} - x)]\mathrm{d}x. \qquad (3.18)$$

The delay for the reverse traffic to carry the message from C_k to C_{k+1} is D_k, which has the i.i.d. distribution given in (3.16). The probability that g misses the LDLD message before its deadline is

$$\Pr\{g \text{ miss } LDLD\} = \sum_{k=2}^{\infty} \Pr\{g \in C_k\} \left[1 - F_{\sum_{i=1}^{k-1} D_i}(t_D)\right]. \tag{3.19}$$

When k is large enough, we can use Gaussian random variables to approximate $\sum_{i=1}^{k-1}(C_i + G_i)$ and $\sum_{i=1}^{k-1} D_i$ in order to simplify the calculation.

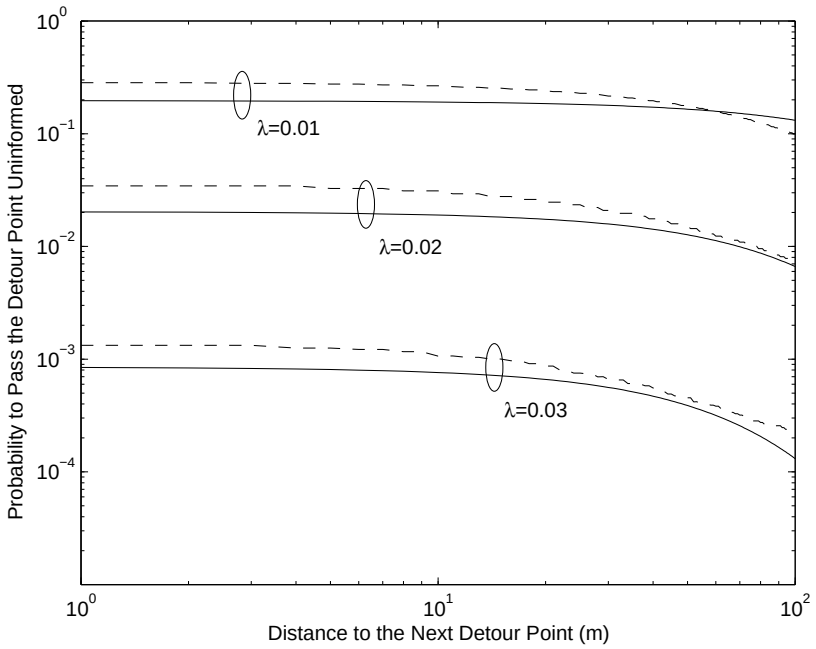
Analytical and Simulation Results

Figure 3.7(a) and (b) show the analytical and simulation results for the probability to miss LDLD deadlines, versus the distance between a vehicle and a highway exit. In both figures, the location of the exit is 5 km from the POI. When traffic density λ is higher, it is more likely that the reverse traffic can help disseminate LDLD messages and extend the cluster size in the forward direction, therefore the probability to miss LDLD deadline is much lower. Note that the y-axis is in log scale and the gap in the lower portion of the figure is actually smaller. Radio transmission range, on the other hand, also affects the cluster size. Thus using a larger R to cover longer distances can also reduce the risk of a vehicle missing LDLD messages before passing an exit, which substantiates the choice of using the SMC scheme.

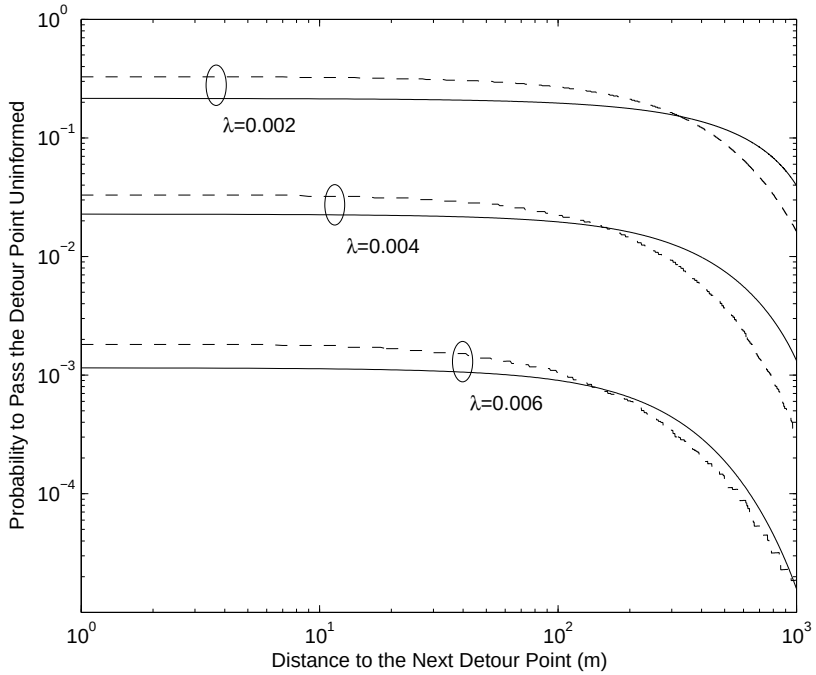
The simulation results match the analysis well. This demonstrates the high accuracy of the results given in (3.16), (3.18) and (3.19), which use the probabilistic distance distributions for the modeling and analysis of location-critical performance metrics in a one-dimensional ad-hoc network.

3.3 Summary

In this chapter, a time/location-critical (TLC) framework for emergency message (EM) dissemination is proposed in vehicular ad-hoc networks, which is achieved through the scalable modulation and coding (SMC) scheme. The exact average cluster size, as well as a close approximation to its distribution, is obtained through a simple approach than the current literature [75, 100]. The research in this chapter also conducted a rigorous analysis on the benefit of utilizing reverse traffic to extend the cluster size, by an investigation on the delay distribution. Extensive simulation



(a) $R = 206\ m$, $D_2 = 5\ km$.



(b) $R = 879\ m$, $D_2 = 5\ km$.

Figure 3.7: The Probability to Miss LDLD Deadlines (Solid Curve—Analysis, Dashed Curve—Simulation).

results have shown the efficacy of the TLC framework and the accurate analysis of the location-critical performance metrics.

In the next chapter, the existing model is extended from a one-dimensional highway to a two-dimensional Manhattan-like city, which is a more challenging problem.

3.4 Discussions and Future Work

Ad-hoc communication is proposed in this chapter as a solution for increasing the opportunity for network connectivity for emergency message dissemination. However, in certain extreme cases with very low vehicle density or short transmission range, the connectivity among vehicles cannot be guaranteed. Under such a circumstance, it is challenging for an arbitrary pair of vehicles to communicate, either directly or via multiple hops. Therefore, a hybrid structure supporting both vehicle-to-vehicle and vehicle-to-infrastructure communications needs to be introduced by service providers. Most existing work only studies the last-hop communication between vehicles and a single base station, i.e., vehicle-to-infrastructure communication as in [62,63,111,112]. In a hybrid scenario, the authors of recent work [77] studied two-hop connectivity and access probabilities between vehicles and different base stations. [101] extended the work in [77] to a multi-hop scenario. However, important performance metrics, such as message dissemination delay and the tradeoff between reduced delay and increased system cost, are not discussed in either [77] or [101]. The future work of this chapter is to further investigate such location-critical performance metrics in a hybrid structure, regarding the connectivity, delay, throughput for both individual vehicle and base stations.

Chapter 4

Vehicular Message Dissemination in Two-Dimensional City Blocks

The connectivity properties of vehicles randomly located in a large city area is more complicated than the problem along a highway. Each vehicle connects only locally to the vehicles within a fixed communication range. The connections among the vehicles across the city area rely on intermediate vehicles as relays, in both the current and perpendicular streets.

Compared with the one-dimensional networks in Chapter 3, however, the two-dimensional topologies such as city blocks, exhibit a very interesting property. The network connectivity can be guaranteed if the probability of connectivity among vehicles along the same street segment is above a certain threshold [88]. In other words, a giant cluster of connected vehicles appears, within which vehicles connect with each other via multiple hops. This phenomenon is known as percolation. The connectivity probability among vehicles is affected by both the vehicle density and the transmission range of onboard sensing devices. Compared with percolation which studies the macroscopic behavior of a network, this chapter studies the connectivity property at a more microscopic level, which differentiates the work in this chapter from existing work such as [23, 88].

In this chapter, the connectivity probability is derived for a ladder topology and formulated for a lattice topology. The latter has been an unsolved problem in directed percolation [42], where the most recent work [25] is only able to derive the results under a very specific setting. Therefore, extensive simulations are conducted using two message forwarding strategies, with and without geographic constraints. The

results also show the intrinsic tradeoff between different message forwarding schemes.

4.1 Percolation Theory and Two-Dimensional Connectivity

The vehicle connectivity in city blocks exhibits a close similarity to the percolation phenomenon. This section reviews the classic concepts and results in percolation theory. Although the problem investigated in this chapter is not identical to percolation, the difference between them is described in this section.

4.1.1 Percolation Theory

In classic percolation theory [42], a stochastic percolating process is modeled as the process of liquid seeping through a porous object. The object is usually modeled as a d-dimensional square lattice. Declare each edge in the lattice *open* with probability p, and closed otherwise, the percolation is related to the existence of an *infinitely connected component (or cluster)* of open edges. All components have finite sizes when $p < p_c$, but when $p \geq p_c$, the size of the connected component becomes infinite. Here p_c is the critical probability, or percolation threshold, above which the network percolates. The state corresponding to the existence (non-existence) of an infinite component is the super-critical (sub-critical) state. Compared with the one-dimensional network which is always disconnected, this is a fundamental difference in the network connectivity.

In contrast to the liquid penetration where edges can be open to all directions, another percolation process called directed percolation [25] has edges only open to certain directions. According to the prediction on directed percolation [57], if $p < p'_c$, the network connectivity probability will steadily decrease and eventually reach 0 with respect to the distance to the information source; if $p \geq p'_c$, the network connectivity probability will decrease gradually and eventually scatter around an asymptotic value. In directed percolation, the state corresponds to $p < p'_c$ ($p \geq p'_c$) is also called super-critical (sub-critical) state. These percolation processes are similar to the message propagation in two-dimensional VANET: we are interested in whether there exists a critical threshold that all the vehicles in a two-dimensional city are almost surely connected in a large component.

The problem of deriving the connectivity probabilities with arbitrary location is a central problem of directed percolation [42] in Physics and Stochastic Processes. In the most recent work, the authors in [25] derived the results for a directed percolation process on an $M \times N$ rectangular lattice whose vertical edges are open upward with a probability y, and horizontal edges are open toward the right with probabilities x and 1 in alternate rows. However, the general problem studied in [109] remains unsolved.

4.1.2 Connectivity in Two-Dimensional Vehicular Ad-Hoc Networks

In Chapter 3, the Poisson traffic model or the exponential inter-vehicle distance is applied to the network topology on a highway. In urban scenarios, vehicular traffic is much more complicated due to road grids, traffic lights and stop signs. However, the exponential distribution is widely used in the literature at a snapshot of the network. Although traffic lights make the vehicle arrival pattern more bursty, once the traffic becomes stable, a snapshot of the network flow still can be analyzed as a Poisson distribution. For example, [88] assumed the exponential distribution in a lattice-shaped road network, and studied the network connectivity as a percolation process, which is equivalent to the study at a snapshot of the network in this chapter. When the traffic is unstable, the model becomes complicated, which we leave as future work. Given a transmission range and the exponential inter-vehicle distance distribution, the connectivity properties and message dissemination in a two-dimensional VANET are investigated through a geometrical probability approach in this chapter.

In the existing work, [83] studied emergency message propagation with time constraints, and derived lower bounds on the probability that a vehicle at a certain distance can receive the messages on time. The derived lower bounds depend on channel reliability and message dissemination strategy. However, the inter-vehicle distance distribution is based on a group of n equally spaced vehicles, which is not a realistic distribution model in vehicular networks. In [53], the fraction of the vehicles that belong to the largest connected component in a two-dimensional graph is analyzed. Furthermore, it is proved in [88] that the Poisson distribution allows the existence of a finite critical vehicle density, and with non-Poisson distribution, there still exists a critical vehicle density, but at a larger value. [23] proved that the critical density exists for a hexagon model, and the connectivity probability increases sharply within a short interval around the density.

The above works in the literature either applied percolation theory in a scenario other than the traditional physics and biology systems, or proved the existence of a critical threshold. These existing works, which are based on the asymptotic properties of infinite networks, did not consider the likelihood that a message can propagate to a certain location in the network. Although the percolation process is closely related to the connectivity properties in a two-dimensional network, in practical VANET, the distance that a message can propagate throughout the network and the probability to reach such a distance, are crucial for both safety-related and infotainment applications.

The main focus in this chapter is the inter-vehicle connectivity, the location-critical performance metrics in a VANET. This is achieved through determining the probability that a message is delivered to a certain distance away from the source, through multi-hop connections between vehicles. Therefore, there is a difference with percolation. Percolation theory focuses on a coverage process, where an infinitely connected cluster does not guarantee messages are always reliably delivered: there may be a number of isolated nodes that are not connected to the giant cluster. The research in this chapter is more microscopic compared to the percolation process.

4.2 Two-Dimensional Manhattan-Like City

It has been proved that in a one-dimensional network, the network is almost surely disconnected [32]. Given the message propagation properties on a highway in Section 3.2, in a very similar scenario that is along a main street in city blocks, the probability that two nearby intersections are connected can be further obtained. This probability is the starting point for an investigation on the connectivity properties in two-dimensional city blocks, as shown in Figure 4.1. In this figure, point O is the source of information. The connectivity of message propagation for the two-dimensional *ladder* case is derived first for covering main-side streets, and the problem for the two-dimensional *lattice* case is formulated for the message propagation among city blocks, respectively.

As discussed in Section 4.1, the connectivity properties in a two-dimensional network are very similar to the percolation phenomenon. In this section, the connectivity properties of a Manhattan-like city are investigated, as the main performance metric that is location-critical. Based on traffic density λ, radio transmission range R, and street segment length d, the connectivity probability of two adjacent intersections

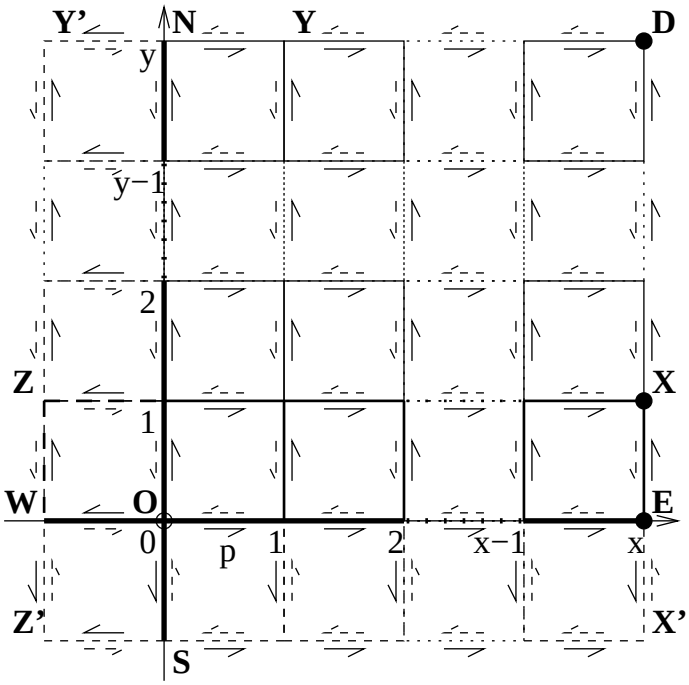


Figure 4.1: Two-Dimensional Manhattan-Like City.

p is derived. To derive p, this chapter assumes a homogeneous Poisson point process with arrival rate λ. However, regardless of how p is derived mathematically or obtained empirically, the approach to deriving the connectivity probability for the two-dimensional ladder case still applies.

4.2.1 Bond Probability

As defined above, p is equivalent to the *bond probability* in percolation theory. If wireless transmissions to other perpendicular streets are heavily shadowed, e.g., in cities with high-rise buildings along the street, p can be simplified as the probability that the one-dimensional cluster size is larger than the distance between two intersections, using the Gamma approximation of C in Section 3.2. In such a scenario, p can be calculated directly using the cluster size distribution in (3.9). This corresponds to the case when messages propagation cannot continue in a new street segment in a direction that is perpendicular to the current street. For more realistic situations when wireless transmissions can reach perpendicular streets, p can be derived as follows.

As shown in Figure 4.2, V_e, V_s, V_w and V_n represent the vehicles (if existing) closest to the right intersection on the east, south, west, and north streets, respectively. Their

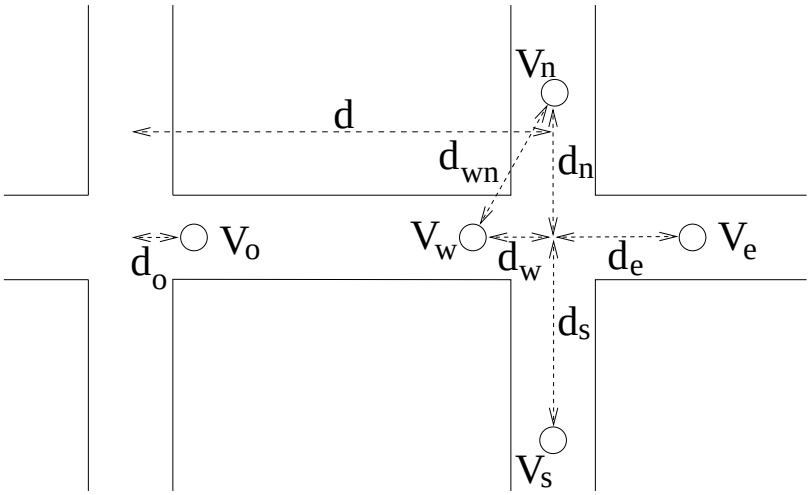


Figure 4.2: Bond Probability Illustration.

distance to the right intersection are d_e, d_s, d_w and d_n. Vehicle V_o is on the street between the two intersections and is closest to the left intersection. V_o is connected to the message source, and its distance to the left intersection is d_o. The distribution of d_o is a truncated exponential function $\frac{\lambda e^{-\lambda t}}{1-e^{-\lambda R}}$, for $0 \leq t \leq R$. We consider the following two disjoint cases where at least one of the vehicles V_e, V_s or V_n is connected to V_o via multi-hop. Therefore, the message propagation can continue in a new street segment, regardless of its direction, and the two intersections in Figure 4.2 are connected.

Case 1:

The cluster originating from V_o has a size larger than $d - d_o$, where d is the distance between two adjacent intersections. By considering the location of V_o and the cluster size distribution $f_C(x)$, the probability for this case is

$$p_{(1)} = \int_0^R \int_{d-t}^{\infty} f_C(x)\mathrm{d}x \frac{\lambda e^{-\lambda t}}{1 - e^{-\lambda R}}\mathrm{d}t. \tag{4.1}$$

In this case, V_e is connected to V_o in the same cluster, therefore the message is propagated along a straight line. V_s is also connected to the cluster, either because it is within the transmission range of V_w or V_e, or it is within the transmission range of V_n, and V_n is within the transmission range of V_w or V_e.

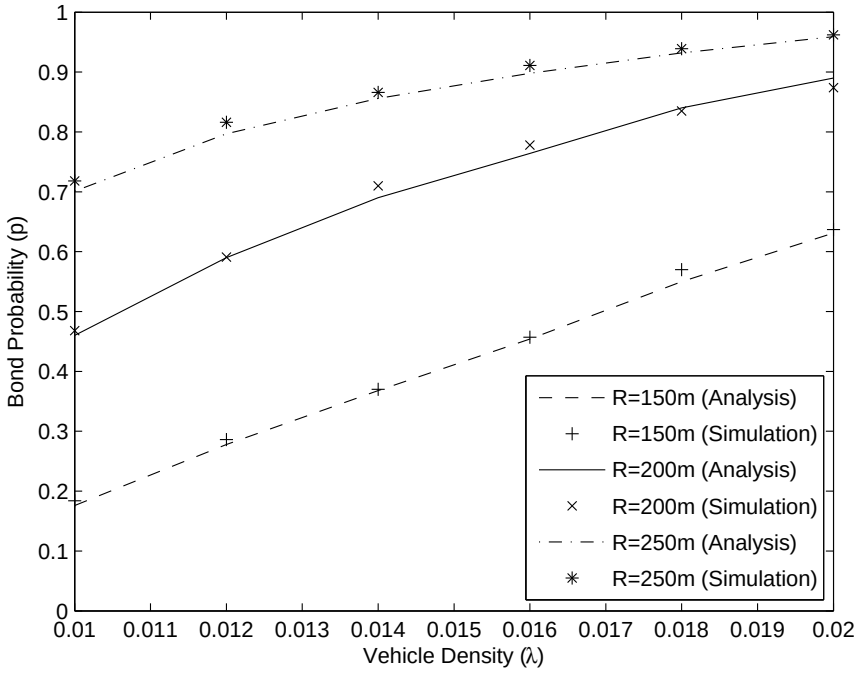


Figure 4.3: Bond Probability Validation.

Case 2:

In this case, the cluster size originating from V_o is smaller than $d-d_o$; the last vehicle connected to V_o is V_w, and $d_e + d_w > R$. In addition, one of V_n and V_s is within the transmission range of V_w. In this case, V_e is outside the direct transmission range of V_w, but at least one vehicle on the south or north street can receive the message and continue the propagation.

Denote the cluster size originating from V_o as x and let $d_0 = t$. To be connected to V_w, the minimum of d_s and d_n must be no greater than $\sqrt{(\eta R)^2 - (d-x-t)^2}$, where $\eta \in (0, 1]$ is the coefficient of the shadowing effect for signal transmissions to other perpendicular streets. Thus, given x and t, the conditional probability that at least one vehicle on the north or south street is within the transmission range of V_w is $1 - e^{-2\lambda\sqrt{(\eta R)^2-(d-x-t)^2}}$. The probability of this case is

$$p_{(2)} = \int_0^R \int_{d-t-\eta R}^{d-t} (1 - e^{-2\lambda\sqrt{(\eta R)^2-(d-x-t)^2}}) f_C(x)\mathrm{d}x \frac{\lambda e^{-\lambda t}}{1 - e^{-\lambda R}}\mathrm{d}t. \tag{4.2}$$

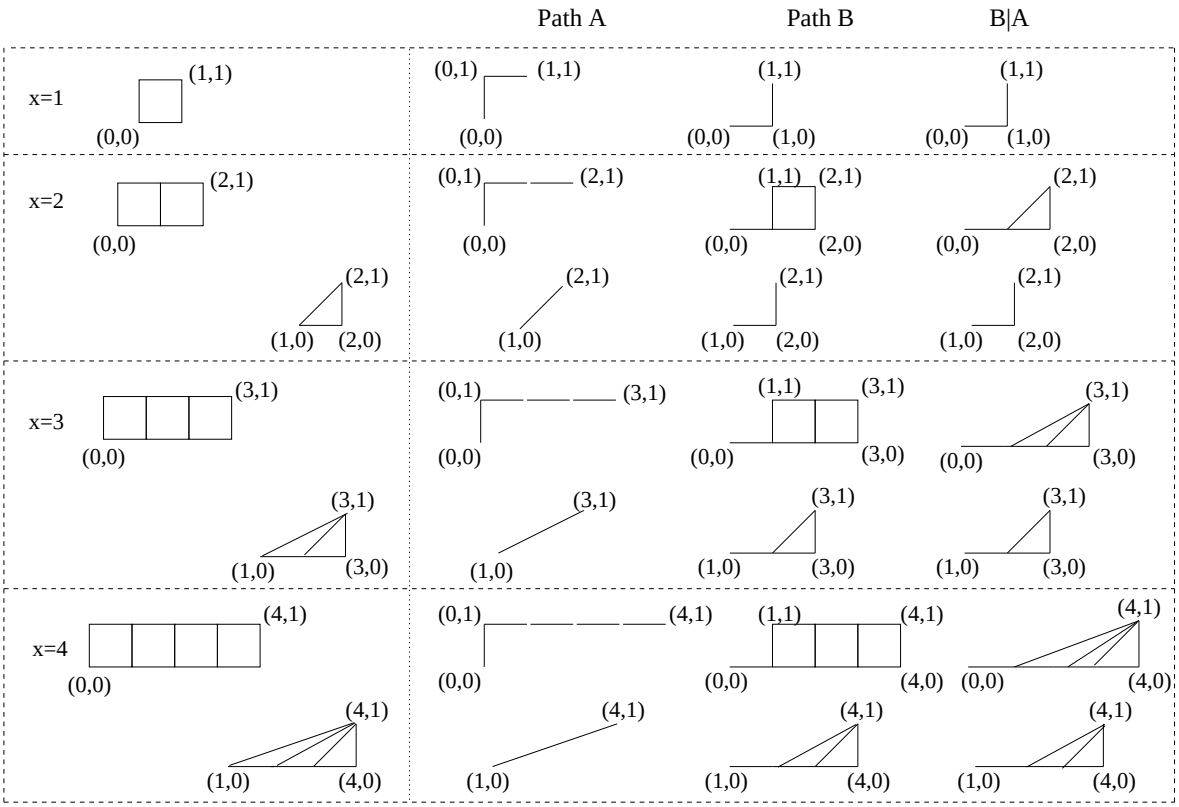


Figure 4.4: Ladder Connectivity Illustration.

Considering the above two disjoint cases, p is given by

$$p = p_{(1)} + p_{(2)}. \tag{4.3}$$

As shown in Figure 4.3, the analytical results given in (4.3) match well with the simulation, where the distance between two adjacent intersections is $d = 500\ m$ and $\eta = 1$. The simulation is done by distributing vehicles randomly on a two-dimensional square lattice, according to the exponential inter-vehicle distance distribution.

4.2.2 Ladder Connectivity

In this subsection the two-dimensional ladder connectivity is investigated through the bond probability derived in Section 4.2.1. The two-dimensional ladder refers to the two side streets, Z–X and Z′–X′, along the main street W–E in Figure 4.1. Events such as congestion detour messages are of interest to the main street and its immediate side streets. The connectivity probability between an arbitrary pair of intersections

in such a ladder topology is derived, with the constraint that the message can only be propagated further away from the source O, i.e., the geo-constrained forwarding which corresponds to the directed percolation.

In a ladder topology, as shown in Figure 4.4, the source is at $O(0,0)$ and the destination is at (x,y) where $y = \pm 1$. Here we use $y = 1$ for illustration. Looking at the first row in Figure 4.4, where $(1,1)$ is the destination, there are two paths to reach $(1,1)$: path A_1 where $(0,0)$ and $(1,1)$ are connected by the intermediate point $(0,1)$, and path B_1 where the intermediate point is $(1,0)$. Given that each edge is connected with probability p, by the principle of inclusion-exclusion (PIE), the following is the connectivity probability for $(1,1)$

$$\begin{aligned} P(1,1) &= P(A_1 + B_1) = P(A_1) + P(B_1) - P(A_1B_1) \\ &= P(A_1) + P(B_1) - P(B_1|A_1)P(A_1) \\ &= p^2 + p^2 - p^2 \times p^2 = 2p^2 - p^4. \end{aligned} \tag{4.4}$$

When $x \geq 1$, recursion is needed to derive the connectivity probability. For example, in the second row of Figure 4.4, where $(2,1)$ is the destination, it has paths A_2 and B_2, where B_2 is dependent on A_2. As shown in the $B|A$ column, event $(B_2|A_2)$ can be degenerated to a horizontal segment plus a triangle (i.e., $(1,1)$ and $(2,1)$ merge to a single point), given the segments in A_2 are already connected. That is, $P(B_2|A_2) = p * (p + p^2 - p^3)$, since the two paths from $(1,0)$ to $(2,1)$ in the degenerated triangle are independent, as illustrated in the second sub-row of the second row. On the other hand, $P(B_2)$ is simply $p * P(1,1)$, where $P(1,1)$ is given in (4.4). Using this recursive method, the connectivity probability for $(2,1)$ can be derived as

$$\begin{aligned} P(2,1) &= p^3 + p \times P(1,1) - P(B_2|A_2)P(A_2) \\ &= 3p^3 - 2p^5 - p^6 + p^7. \end{aligned} \tag{4.5}$$

Similar recursion can be done when $x = 3$. In the third row of Figure 4.4, we notice that $P(A_3)$ is simply p^4, $P(B_3)$ is related to $P(2,1)$, and the triangle in $B_3|A_3$ can be decomposed into two sub-cases, one is simply p and the other is related to $B_2|A_2$. Applying the same recursion for squares and degenerated triangles iteratively

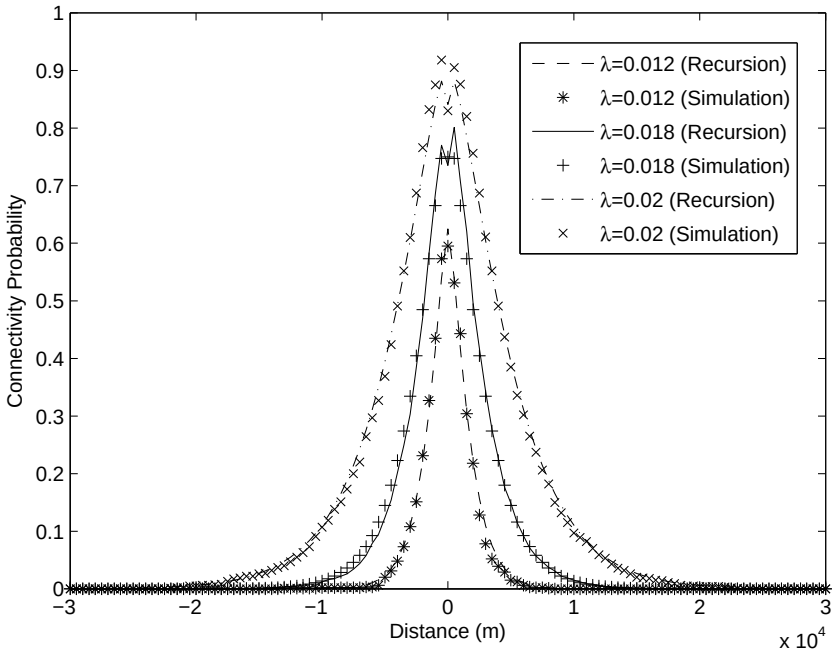


Figure 4.5: Ladder Connectivity Validation.

for all x's, the following for general $P(x,1)$ is obtained:

$$P(x,1) \quad = p\left[p^x + P(x-1,1) - p^x\theta(x)\right], \quad x \geq 1 \tag{4.6}$$

with $P(0,1) = p$, and

$$\theta(x) \quad = p\left[p + \theta(x-1) - p\theta(x-1)\right], \quad x \geq 1 \tag{4.7}$$

with $\theta(0) = 0$, and x can be any integer. The derivation is symmetric for the ladder connectivity in the other three quadrants for the main-side street scenarios.

Figure 4.5 shows the results from recursion (4.6) and (4.7), and compares them with simulation results. The vehicle transmission range is 200 m and the distance between two adjacent intersections is 500 m. The x-axis is the west-east coordinate of side street Z–X, which is 500 m away from the parallel main street. The message source O is on the main street with x-coordinate equal to 0. As shown in the figure, the recursive analysis and the simulation match well with each other. Similar results in the literature have not yet been found.

The results show that the connectivity probability decays fast with respect to the

distance, for all traffic densities. However, when compared with the results from the one-dimensional case, the connectivity probability at the same x-coordinate offset with the same traffic density is much higher. For example, with $\lambda = 0.02$, the connectivity probability at $x = 4\ km$ offset for the two-dimensional ladder case is above 0.5, whereas the connectivity for the one-dimensional case is only around 0.2 according to Figure 3.4. The reason is that there are more paths available from the source in the two-dimensional ladder case, whereas the one-dimensional case only has one path.

Extensions

Note that with a simple extension, (4.6) and (4.7) can be used to derive the connectivity probability where the bond probability of horizontal street segments is different from that of vertical street segments. For instance, if the intersections on the horizontal and vertical street segments are connected with different probabilities, p_1 and p_2, respectively, then we have $P'(1,1) = 2p_1p_2 - p_1^2p_2^2$ in (4.4), and (4.5) becomes

$$\begin{aligned} P'(2,1) &= p_1^2p_2 + p_1 \times P(1,1) - P(B_2|A_2)P(A_2) \\ &= p_1^2p_2 + p_1 \times \left(2p_1p_2 - p_1^2p_2^2\right) - p_1^2p_2 \times p_1 \times \left(p_2 + p_1p_2 - p_1p_2^2\right) \\ &= 3p_1^2p_2 - 2p_1^3p_2^2 - p_1^4p_2^2 + p_1^4p_2^3. \end{aligned} \tag{4.8}$$

Similarly, the extension of (4.6) and (4.7) are

$$P'(x,1) = p_1^x p_2 + p_1 P'(x-1,1) - p_1^x p_2 \theta'(x), \quad x \geq 1, \tag{4.9}$$

and

$$\theta'(x) = p_1 \left[p_2 + \theta'(x-1) - p_2\theta'(x-1)\right], \quad x \geq 1. \tag{4.10}$$

with $P'(0,1) = p_1$, and $\theta'(0) = 0$.

In addition to the above, this approach can be used to deal with dynamic traffic shifting. When the vehicles in the main street are no longer connected (e.g., due to detour), the two side streets can form new ladders themselves, and the above recursive method can be applied to recalculate the connectivity probability accordingly. Figure 4.6 shows an example when a main street is disconnected (dashed line) and two side streets become the new ladder. In this case, the horizontal intersections are

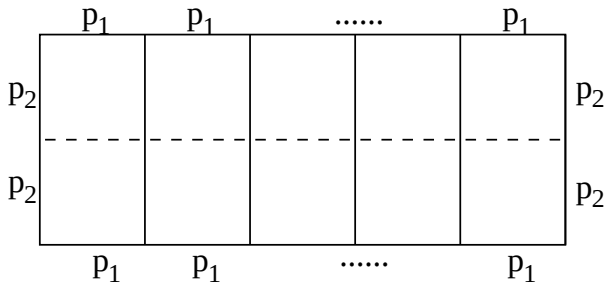


Figure 4.6: Two Side Streets Form New Ladders (Solid Lines—Connected Street Segments, Dashed Line—Disconnected Main Street).

still connected with probability p_1, but the vertical intersections are connected with probability p_2^2.

4.2.3 Lattice Connectivity

As mentioned in Section 4.1, the most recent work [25] in the literature derived the results for connectivity under a specific directed percolation process, where the horizontal edges in a lattice grid are open with alternating probabilities x and 1. The generalized connectivity problem is unsolved. For the general two-dimensional lattice case with each edge open with probability p, e.g., the origin-destination pair O–D in Figure 4.1, we have the following numerical formulation to derive the end-to-end connectivity probability. First, enumerate all the possible paths from $(0,0)$ to (x,y), and then still by the principle of inclusion-exclusion (PIE), $P(x,y)$ can be obtained by calculating the probabilities of different combinations of paths and crosschecking their overlapping street segments. As an example, we assume the destination point is $(2,1)$, and then all the possible paths starting from $(0,0)$ are

$$\begin{aligned} A: &\quad (0,0) \to (1,0) \to (2,0) \to (2,1) \\ B: &\quad (0,0) \to (1,0) \to (1,1) \to (2,1) \\ C: &\quad (0,0) \to (0,1) \to (1,1) \to (2,1) \end{aligned}$$

By PIE,

$$\begin{aligned} P(2,1) &= P(A+B+C) \\ &= P(A)+P(B)+P(C)-P(AB)-P(BC)-P(AC)+P(ABC) \\ &= p^3+p^3+p^3-p^5-p^5-p^6+p^7, \end{aligned}$$

which has the same result as (4.5). $P(AB) = P(BC) = p^5$ and $P(AC) = p^6$ since paths A and B together have 5 non-overlapping street segments, while paths A and C have 6 non-overlapping ones.

Extensions

When the bond probability of horizontal and vertical street segments are p_1 and p_2, respectively, then the above probability becomes

$$\begin{aligned} P'(2,1) &= P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC) \\ &= p_1^2 p_2 + p_1^2 p_2 + p_1^2 p_2 - p_1^3 p_2^2 - p_1^3 p_2^2 - p_1^4 p_2^2 + p_1^4 p_2^3, \end{aligned}$$

which has the same result as (4.8), according to the recursive formulation in (4.9) and (4.10).

Ideally, any $P(x, y)$ can be computed by this enumeration-combination method. Unfortunately, this approach suffers from the combinatorial explosion and does not scale. As a result, the computation becomes intractable when $x + y$ becomes larger. For instance, when $x = 5$, $y = 3$, the number of different paths is $\binom{x+y}{y} = 56$, and the number of different combinations of these 56 paths can be as many as $\binom{56}{28} = 7.6487 \times 10^{15}$, each of which has $|x| + |y| = 8$ street segments. If we store these street segments in a bit map, which requires 38 bits per path since there are $(x+1)y + x(y+1) = 38$ unique street segments, then the memory required will be $38 \times 7.6487 \times 10^{15}/8$ bytes $\approx 3.63 \times 10^7$ GB. The problem of deriving connectivity probabilities with arbitrary x and y is a central problem of directed percolation [25] in Physics and Stochastic Processes, which still remains unsolved after many years of efforts.

4.2.4 Network Connectivity

In Sections 4.2.2 and 4.2.3, we derived the connectivity probability for the two-dimensional ladder case, whereas the two-dimensional lattice case is not yet analytically solvable. In this section, we go further and evaluate the static connectivity [97] of these cases by simulation. This is an investigation of the location-critical performance metric at a snapshot of the network.

In vehicular ad-hoc networks, there are different ways for vehicles to forward the messages they have received. Geo-constrained forwarding (GF), the message forwarding scheme used in Section 4.2.2 and 4.2.3 for illustration, is commonly used to

avoid message redundancy, link contention, hidden terminals, and broadcast storm problems. Unconstrained forwarding (UF), on the other hand, is an extension of the former in the sense that vehicles can forward messages to neighbors in all directions, so messages may potentially go backwards, as shown by both the solid and dashed arrows in Figure 4.1. Unconstrained forwarding improves network connectivity, but may also increase the overhead, such as the broadcast cost in terms of the number of transmissions of each message. It is interesting to quantify and compare the performance and cost of these two message forwarding strategies, with or without geographic constraints. Unconstrained forwarding is equivalent to percolation process, whereas geo-constrained forwarding corresponds to *directed* percolation in [25].

In both message forwarding strategies, whenever a vehicle overhears the same message from the neighbors in the direction opposite to the source, it will not rebroadcast that message again, in order to reduce overhead and possible collisions; otherwise, the vehicle rebroadcasts the message periodically up to a retransmission limit. We use simulation to compare these two forwarding strategies, with and without geographic constraints, respectively, and reveal the tradeoff between them.

All simulation is done in Matlab. By using an exponential random variable generator and a square lattice map, vehicle locations are constrained to the vertical and horizontal streets. Streets are 500 m apart, and all vehicles have the same transmission range R, varying from 150 to 300 m with $\eta = 1$. Vehicles that are less than R apart are connected with one another, regardless they are located along the same street or perpendicular streets. All the results presented are averaged over $1,000$ simulation runs.

(1) Geo-Constrained Forwarding

In geo-constrained forwarding, similar to directed percolation [25], the probability of each street segment to be open is determined by p in (4.3). We first plot the network connectivity that measures the percentage of the vehicles connected to the message source in a 3×3 km^2 grid map. Figure 4.7 shows the change of network connectivity, from 0 to 1, with different vehicle densities. Such transitions are similar to the results in [23]. Combining Figures 4.3 and 4.7, we note that the critical threshold p_c, above which the network connectivity is close to 1, is much larger than 0.5, the critical threshold of the bond percolation in a lattice grid with no direction constraints [88].

In Figure 4.8, different sub-figures show the connectivity probability for the vehi-

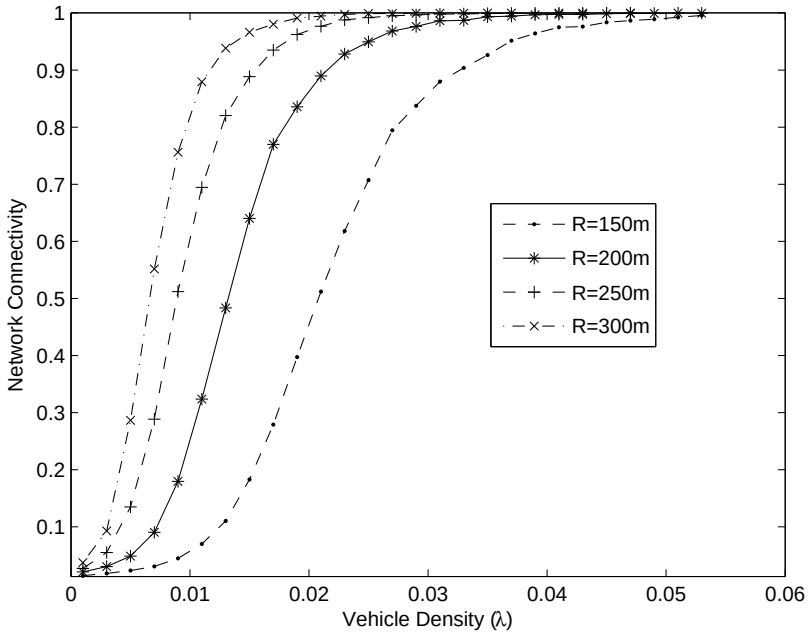
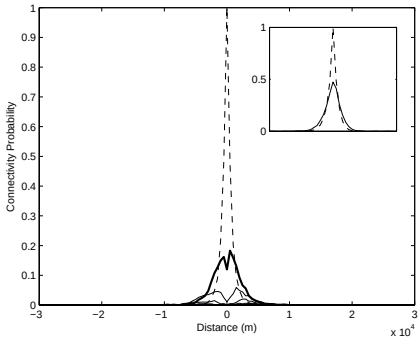


Figure 4.7: Network Connectivity with Geo-Constrained Forwarding (GF).

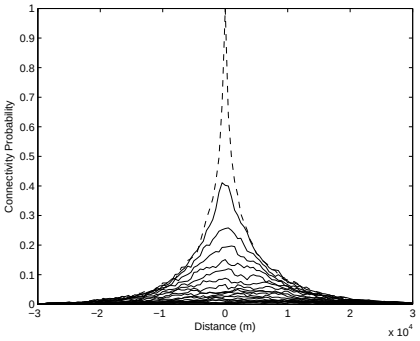
cles on different parallel streets in the lattice network, and the x-axis represents the coordinates along each street. The coordinate of the message source is $(0, 0)$, and the transmission range R is 200 m. Different sub-figures show the results with different vehicle densities λ, varying from 0.01 to 0.015 vehicles per meter. The values of λ are carefully chosen according to Figure 4.7, so that the bond probability p is below, close, or above the critical threshold for the directed bond percolation.

Typically, the farther away a street is from the message source, the lower the connectivity probability is for the vehicles on that street, and thus the message propagation within a few blocks away from the source (e.g., the bold curves in this figure) is of major concern. In Figure 4.8 we plot the connectivity probability up to 30 km, in order to give a big picture of the entire network, e.g., how the connectivity probability decays or converges. Figures 4.8(a)–(d) show the simulation results for geo-constrained forwarding: around the critical point with λ from 0.012 to 0.015, even with a slight change in traffic density, the connectivity probabilities change significantly; once the critical point is reached, the peak of the connectivity probability for each street converges to an asymptotic value, as shown in Figure 4.8(d), but the connectivity decays eventually over the distance.

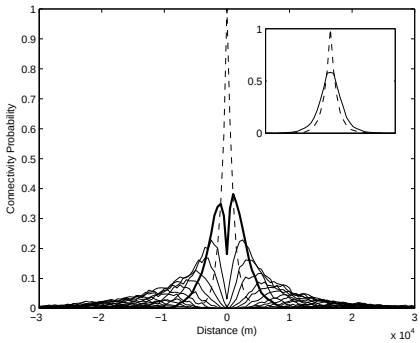
Another interesting observation is that, except the main street, the connectivity



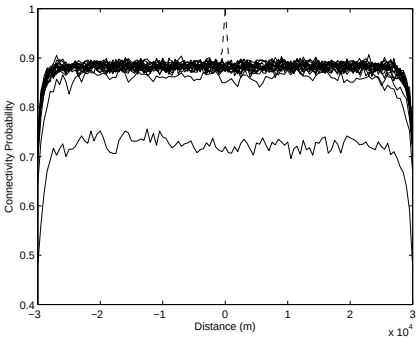
(a) GF, $\lambda = 0.01$



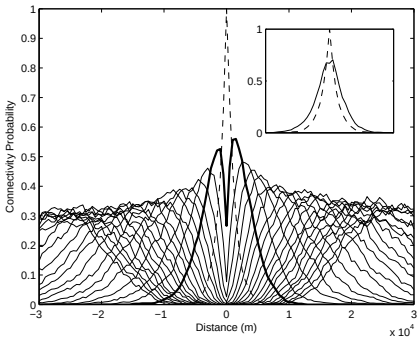
(e) UF, $\lambda = 0.01$



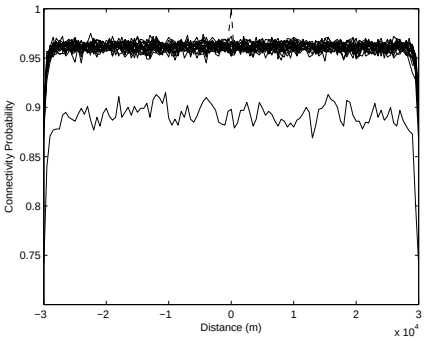
(b) GF, $\lambda = 0.012$



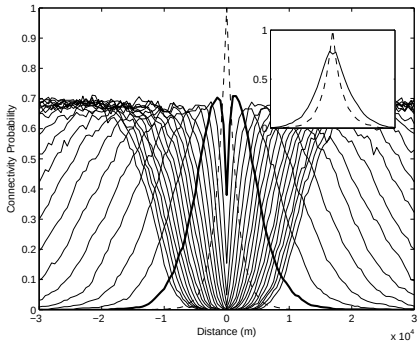
(f) UF, $\lambda = 0.012$



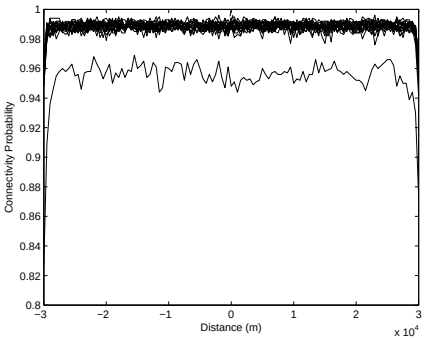
(c) GF, $\lambda = 0.0135$



(g) UF, $\lambda = 0.0135$



(d) GF, $\lambda = 0.015$



(h) UF, $\lambda = 0.015$

Figure 4.8: Connectivity Probability: Geo-Constrained Forwarding (GF) vs. Unconstrained Forwarding (UF).

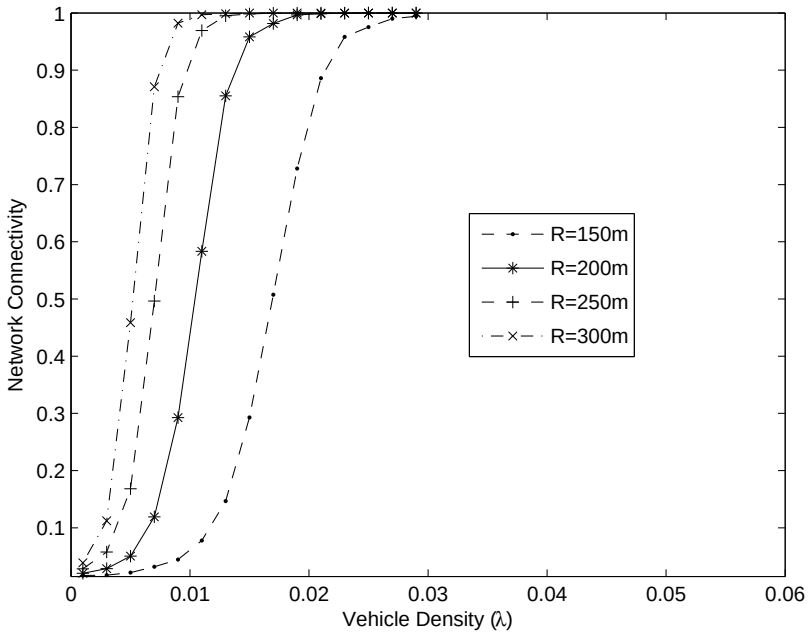


Figure 4.9: Network Connectivity with Unconstrained Forwarding (UF).

probability on all the other streets is not monotonic with respect to x. The dashed, Λ-shaped curves in Figures 4.8(a)–(d) are for the main street; the bold curves are for the side streets that are 3 blocks away; the solid curves in the zoom-in window show the results for the two-dimensional ladder case, i.e., one block away. All curves, except the ones for the main street, have an M-shape, which means being closer to the message source does not necessarily lead to a higher connectivity probability. This is because, the closer an intersection is to the source, the smaller number of paths exist between this intersection and the source, although each path is shorter. If an intersection is on either of the streets where the source is located, for example along W–E in Figure 4.1, vehicles can only forward data in one direction: as a result, the connectivity for them is the same as that in the one-dimensional case.

(2) Unconstrained Forwarding

Observing the cluster size distribution in the one-dimensional case and the limited connectivity probability in a two-dimensional lattice with geo-constrained forwarding, we find that clusters are likely to be small and a large number of vehicles are unable to receive the message. By allowing the message to go through other directions,

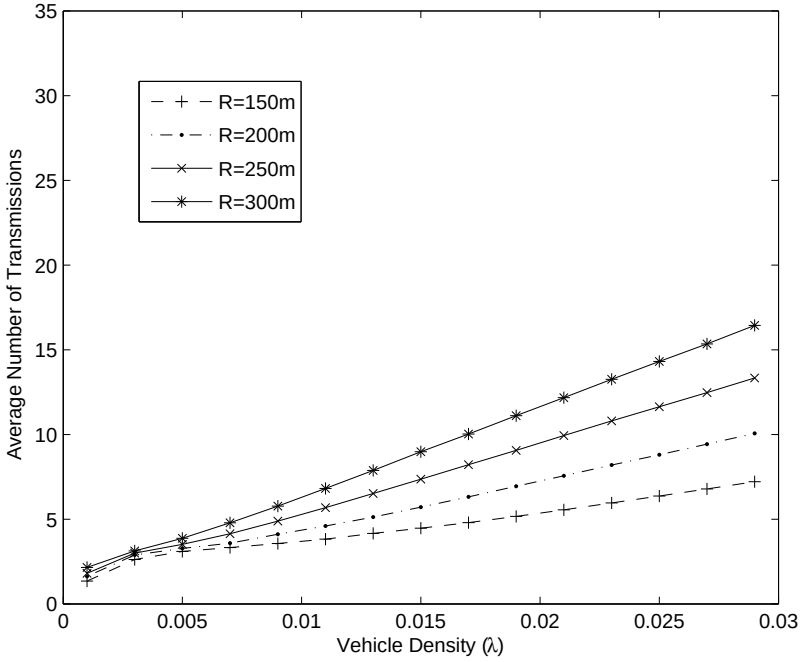
as indicated by the dashed arrows in Figure 4.1, clusters can be extended in these directions, i.e., including more neighbors for message delivery.

As shown in Figure 4.9, the transitional behavior in unconstrained forwarding is similar to that in geo-constrained forwarding, except that the critical threshold is much smaller, and thus the transitions are also sharper. p can be found in Figure 4.3, e.g., when $R = 200\ m$ and λ is slightly larger than 0.01, the corresponding p is exactly 0.5. From Figure 4.8, we can see another difference between these two forwarding schemes. At super-critical stage, the connectivity probability in geo-constrained forwarding converges to a value that is much smaller than 1, whereas in unconstrained forwarding, the connectivity probability is almost 1 as p approaches the critical value, given the same vehicle density. The Λ-shaped curves in Figures 4.8(e)–(h) still correspond to the main street, where the street farthest away from the source has a much lower connectivity probability due to the edge effect.

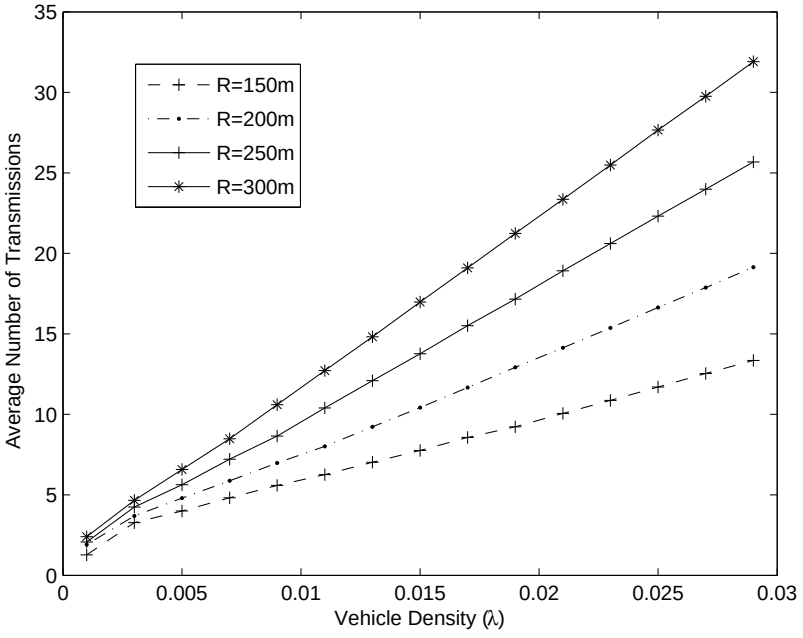
In contrast to geo-constrained forwarding, the connectivity probability here is either monotonically deceasing with distance or scattering around an asymptotic value. Taking a closer observation, the results in Figure 4.8(e) with $\lambda = 0.01$ for unconstrained forwarding is similar to Figure 4.8(b) with $\lambda = 0.012$ for geo-constrained forwarding, because the latter has relatively limited choices of path selection. As shown in Figures 4.8(e) and (f), when λ is only increased by 0.002, i.e., adding one vehicle every 500 m, the connectivity probability for vehicles increases drastically from below 20% to higher than 85%. When λ is further increased to 0.0135, as shown in Figure 4.8(g), the connectivity probability for most vehicles exceeds 95%. After this point, further increasing λ does not change the connectivity probability significantly, as shown in Figure 4.8(h). This tells us that when p is above the critical threshold 0.5, there is no more significant increase in the network connectivity. The above results can help the application to decide whether it is necessary to increase p, e.g., by using a larger transmission power, or use unconstrained forwarding to improve connectivity.

(3) Broadcast Cost

The high connectivity achieved by unconstrained forwarding is at the expense of a higher broadcast cost than geo-constrained forwarding. Higher broadcast cost will lead to severe collisions that negatively affect the throughput, as in our work [111,112]. Figures 4.10(a) and (b) show the average number of transmissions for each message



(a) Geo-Constrained Forwarding (GF).



(b) Unconstrained Forwarding (UF).

Figure 4.10: Broadcast Overhead: Average Number of Transmissions.

with these two forwarding schemes, using the same parameter settings for comparison. It is assumed that there is no transmission error for each broadcast. The vehicle within the transmission range and farthest away from the transmitter will rebroadcast the message. Geo-constrained forwarding apparently has a much lower broadcast cost, although the network connectivity in this case is limited. Therefore it reduces the amount of network resources required, such as the communication bandwidth and buffer space.

The decision whether to use geo-constrained or unconstrained forwarding, thus also depends on the application. If messages need to be propagated as far as possible, e.g., hotel, dining, or parking lot information, then using unconstrained forwarding, probably at a lower message frequency, can allow the information traverse the network without a high demand on vehicle density. Road hazard or traffic congestion, on the other hand, only has high impact on nearby streets. Therefore the messages need to be propagated within a few street blocks as soon as possible. In this case, geo-constrained forwarding is preferable due to its low broadcast cost and potentially fewer collisions.

4.3 Summary

The research in this chapter considers the fundamental limits of message propagation in VANET, and explored the vehicle and message dissemination characteristics in two-dimensional city scenarios. Starting from a one-dimensional street segment, the results are extended to the analysis of the network connectivity in the two-dimensional ladder and lattice cases. Through the simulation studies under two different message forwarding schemes with and without geographic constraints, the connectivity properties are surprisingly similar to the percolation phenomenon.

The tradeoff between vehicle density, transmission range, and the insightful knowledge obtained in this chapter thus enable vehicular network engineers and operators to effectively achieve high user satisfaction and good service coverage, with necessary implementation of different message forwarding schemes according to traffic density, user requirements and application categories.

4.4 Discussions and Future Work

4.4.1 Connectivity, Collisions and Throughput

Practical factors, such as packet loss due to wireless channel impairments or collisions and how the physical, link, and network layers react to them, can be further considered in deriving the bond probability, so the work can be extended to include these realistic wireless channel characteristics and the implementation details of the MAC and routing protocols. Results also can be used to study the tradeoff between connectivity and throughput: if each vehicle uses a lower data rate for transmission, then the transmission range can be larger, but each message transmission will occupy the channel longer, and more wireless resources will be consumed. Incorporating throughput, collision and scheduling analysis is critical in modeling vehicular networks, and it is open for further research. Nevertheless, the physical connectivity in this study is a necessary condition for the connectivity at the network and higher layers.

4.4.2 Vehicle Mobility

Thus far, the model in this chapter has been limited to analyzing a snapshot of the network, and has not considered vehicle mobility, or vehicles traveling at different speed on different road segments. As the snapshot may become time variant, one possible extension to the existing work is to use mobility trace in the existing simulation to estimate the distribution of the velocity of the vehicles, as well as the probability of making turns at each intersection, and study the characteristics of network connectivity under certain mobility models. Also, "carry-and-forward" messages between vehicles traveling in opposite directions can be applied for delay-tolerant applications. Different from unconstrained forwarding, carry-and-forward utilizes vehicle mobility, instead of forwarding directions, to opportunistically extend the cluster size. In two-dimensional networks, this problem is much more difficult since vehicles may turn to the other directions at different intersections. Furthermore, road grids, traffic lights and stop signs change the traffic arrival process, which we leave for our future work.

The contributions in Chapter 3 and Chapter 4 are from the application aspect. While the connectivity problem in these two chapters is nontrivial, it is more meaningful to extend these models to two dimensions with fewer location constraints than vehicular networks. Such two-dimensional networks occur most frequently in various

scenarios, such as wireless sensor networks, mobile ad-hoc networks, and cellular systems. Because the geometry is very different, there will be fundamental changes in both the mathematical models and applications. Next chapter will present a mathematical model from the geometrical probability theory aspect. For the first time in literature, the explicit forms of distance distributions for the complex geometries are obtained. These results enable the analytical models in a wider spectrum with fewer location constraints.

Chapter 5

Random Distances Associated with Rhombuses: the Complex Geometry

In this dissertation, straight lines, rectangles and squares are categorized as the elementary geometries, while rhombuses, parallelograms and hexagons are the complex geometries. In the literature of geometrical probability, the random distance associated with squares and rectangles have been studied with various approaches, and the results have been in existence for several decades [37, 38, 66–68]. From a geometric perspective, parallelograms and rhombuses are the *squeezed* rectangles and squares, respectively. Also, a regular hexagon can be divided into three congruent rhombuses. However, different from rectangles and squares, the coordinates of a random point in a complex geometry are no longer independent. We present in this chapter that through a quadratic product formulation, random distances associated with rhombuses can be obtained. The next chapter will extend these results and obtain the random distances associated with hexagons.

Challenging as they are to the geometrical probability research, rhombuses and hexagons are one of the basic building blocks in two-dimensional tiling. They have important applications in a wide variety of science and engineering fields, not only in wireless communication networks, but also in urban transportation, operations research, etc. In this chapter, the explicit probability density functions of the random Euclidean distances associated with rhombuses are derived, when both endpoints are randomly distributed in 1) the same rhombus, 2) two parallel rhombuses sharing a

side, and 3) two rhombuses having a common diagonal, respectively. By an affine transformation on the plane geometry, the product formulation presented in this chapter is able to handle the interdependency between point coordinates. The correctness of these distance distributions is validated by a recursion and a probabilistic sum, and the accuracy is validated through simulation.

5.1 State of the Art

In our previous work [108], we used a standard approach to derive the distribution of random distances between nodes within and between squares. Given two nodes on a Cartesian plane with coordinates (X_1, Y_1) and (X_2, Y_2), the Euclidean distance is $D = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$. Thus, the geometrical distribution of node distances, $P(D \leq d)$, is determined by the distribution of these coordinates. The standard approach was a four-step process, where the distributions of difference $V = X_1 - X_2$ (or $Y_1 - Y_2$), square $S = V^2$, sum $Z = S_X + S_Y$ and square-root $D = \sqrt{Z}$ were derived respectively. The drawback of this approach is, the condition that $Z = S_X + S_Y$ can be derived through convolution is based on the independence of S_X and S_Y, which requires the $\mathcal{X}$ and $\mathcal{Y}$ coordinates to be independent. Although this holds true for squares and rectangles, such an assumption makes this approach not applicable to other geometrical shapes.

[67, 68] used similar approaches and derived the distance distributions for rectangles. On the other hand, the derivation in [37, 38] are based on the joint distribution of $X_1 - X_2$ and $Y_1 - Y_2$, which leads to explicit results only when the corresponding point coordinates are independent. Successful as they were, the approaches in [37, 38, 67, 68, 108] cannot be extended to other geometrical shapes, such as parallelograms, rhombuses and hexagons, where the $\mathcal{X}$ and $\mathcal{Y}$ coordinates of each point are no longer independent.

In this chapter a general, unified approach is presented. We show that using a quadratic product formulation, the distributions of random distances associated with rhombuses (this chapter) and hexagons (next chapter) [105, 106] can be obtained in closed form through this approach which is much more simple and easy to visualize, with the corresponding geometric interpretation. They serve as a clear demonstration of the flexibility of the approach we are about to propose in this chapter. In the following, we use the terminology "node" and "point" interchangeably, when it is obvious from the context.

5.2 From Rectangles to Parallelograms

5.2.1 The Quadratic Product Formulation

The Euclidean distance between two points (X_1, Y_1) and (X_2, Y_2) on a Cartesian plane, $D = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$, is a function of their coordinate *differences*. Let random variable Z denote the squared Euclidean distance D^2, and $X = X_1 - X_2$ and $Y = Y_1 - Y_2$, then $Z = X^2 + Y^2$ is a function of X and Y. By going back to the definition of distance distribution, this section gives a unified approach by the quadratic product of the conditional probability function and the probability density of random variables. Using geometric integral, this approach decomposes the geometrical constraints and distributions of random points, i.e.,

The Probability Function $F_\omega(Z|X, Y)$, given in the form of a conditional probability that is a function of the coordinate differences X and Y with geometrical constraints Ω on (X_1, Y_1) and (X_2, Y_2);

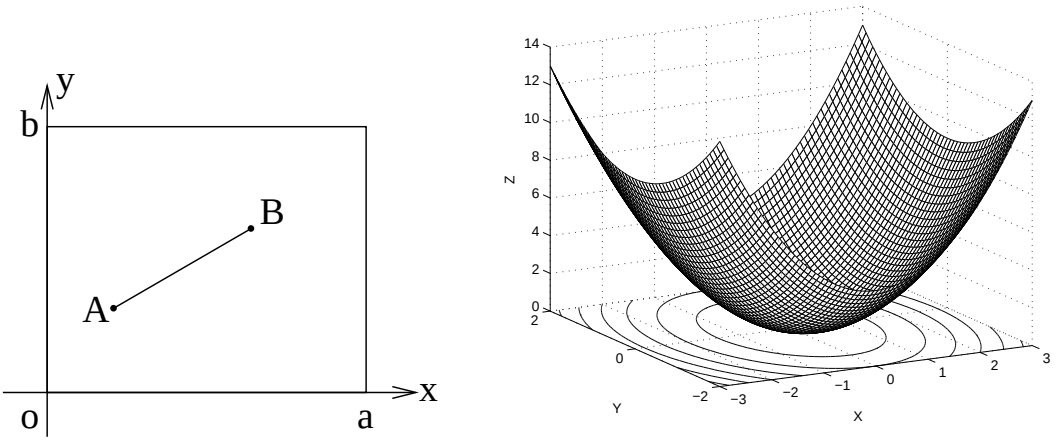
The Probability Density of the coordinate differences, $f_\omega(X, Y)$, where general distributions can be applied to (X_1, Y_1) and (X_2, Y_2).

By the definition of conditional probability, and Z as a function of X and Y as mentioned above, the corresponding formulation is given as follows

$$P_\Omega(Z \leq z) = \int\int_\omega P\left(Z(X, Y) \leq z | X = x, Y = y\right) \cdot f_{X,Y}(x, y) dx dy. \tag{5.1}$$

In (5.1), $F_\omega(Z|X, Y) = P\left(Z(X, Y) \leq z | X = x, Y = y\right)$ is the conditional probability function, and $f_\omega(X, Y) = f_{X,Y}(x, y)$ is the probability density function. Therefore, $P_\Omega(Z \leq z)$ can be obtained by the integration of Z projected onto the $\mathcal{X}$-$\mathcal{Y}$ plane according to ω, which is transformed from Ω. This integration thus gives the product formulation in (5.1).

The intuition behind this formulation is the correspondence between a geometric shape and the characteristics of the random coordinates inside the shape, which can be separated if independent. In complex geometries, both X and Y in (5.1) can be a function of other random variables, therefore, so is Z. In general, F_ω and f_ω both can be a function of multiple random variables, e.g., for hexagonal geometries. In such complex cases, the integration of Z will go to higher dimensions.



(a) Random Points in a Rectangle. (b) $Z = X^2 + Y^2$ when $a = 3$, $b = 2$.

Figure 5.1: A Geometric Interpretation of Random Distances within a Rectangle.

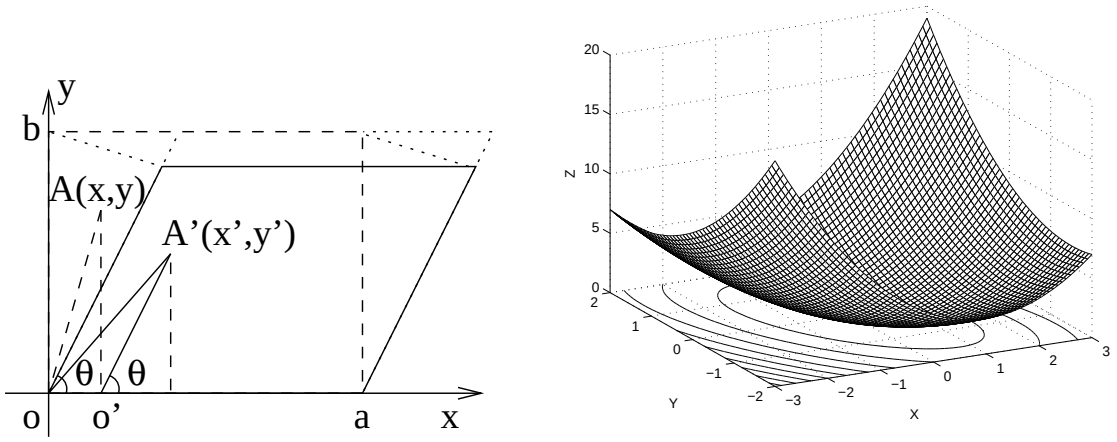
Regardless of the shape of a geometry and how the points are distributed, the product formulation is expressed in the same form as that in (5.1). Following this general formulation, many cases of different geometric shapes and distance distributions can be tackled as shown below.

5.2.2 Rectangles: a Simple Illustration

Figure 5.1(a) shows a rectangle of size $a \times b$, with two random points $A(X_1, Y_1)$ and $B(X_2, Y_2)$. Let $X = X_1 - X_2$ and $Y = Y_1 - Y_2$, then $Z = Z(X, Y) = X^2 + Y^2$. Because $X_1, X_2 \in [0, a]$ and $Y_1, Y_2 \in [0, b]$, we have $X \in [-a, a]$ and $Y \in [-b, b]$. Henceforth, uppercase letters are used to denote random variables, and lowercase letters for the sample values of the same random variable. Note that $X^2 + Y^2 = z$ is the equation of a circle with radius $\sqrt{z}$ for a given z, and $F_Z(z)$ is used to denote the cumulative distribution function of Z. By the product formulation in (5.1), the distribution of Z is

$$\begin{aligned} F_Z(z) &= P_\Omega(Z \le z) = P(X^2 + Y^2 \le z) \\ &= \iint P(x^2 + y^2 \le z | X = x, Y = y) f_{X,Y}(x, y) dx dy \end{aligned} \tag{5.2}$$

where $F_\omega = P(x^2 + y^2 \le z | X = x, Y = y)$ is the probability function, conditioned on the sample values $X = x$ and $Y = y$. In rectangles, we have $f_{X,Y}(x, y) = f_X(x) f_Y(y)$, where $f_X(x)$ and $f_Y(y)$ are the (marginal) probability density functions of X and Y,



(a) A Rectangle Becomes a Parallelogram. (b) $Z = X^2 + 2\cos\theta XY + Y^2$ when $a = 3$, $b = 2$ and $\theta = \frac{\pi}{3}$.

Figure 5.2: A Rectangle Squeezed into a Parallelogram.

since $X = X_1 - X_2$ and $Y = Y_1 - Y_2$ are stochastically independent in this case. The distributions of X and Y are

$$f_X(x) = \frac{1}{a^2}\begin{cases} a+x & -a \le x \le 0 \\ a-x & 0 \le x \le a \end{cases} \quad \text{and} \quad f_Y(y) = \frac{1}{b^2}\begin{cases} b+y & -b \le y \le 0 \\ b-y & 0 \le y \le b \end{cases}, \quad (5.3)$$

both following a symmetric triangular distribution.

The geometric interpretation of (5.2) in three-dimensional space is shown in Figure 5.1(b): a bowl cut off by the boundaries of X and Y at $\mathcal{X} = \pm a$ and $\mathcal{Y} = \pm b$. Figure 5.1(b) also shows the contour lines of Z projected onto the $\mathcal{X}$-$\mathcal{Y}$ plane, each of which corresponds to a specific value of Z and together forms a series of concentric circles centered at $(0, 0)$, as expected from the equation of the circle $X^2 + Y^2 = z$. From Figure 5.1(b), the geometrical constraint Ω is the bowl with cutoffs, and ω is the concentric circles transformed from Ω by projection. Moreover, ω is constrained by $[-a, a] \times [-b, b]$ on the $\mathcal{X}$-$\mathcal{Y}$ plane, therefore it determines the integral region in (5.2). The function $Z = X^2 + Y^2$ can be used for both rectangles and squares, whereas a different function will be used for parallelograms and rhombuses.

5.2.3 Parallelograms: the Squeezed Rectangles

Compare Figure 5.1(a) with Figure 5.2(a), a parallelogram is obtained by squeezing a rectangle. This is an affine transformation [1], which preserves straight lines (i.e.,

all points lying on a line initially still lie on a line after the transformation) and ratios of distances (e.g., the midpoint of a line segment remains the midpoint after the transformation).

(1) The Affine Transformation

Suppose that initially, a point $\mathrm{A}(x, y)$ lies in the rectangle shown in Figure 5.2(a). This point forms a right triangle OAO' with the $\mathcal{X}$-axis. Suppose $0 \leq \theta \leq \frac{\pi}{2}$, then squeezing the rectangle by $\frac{\pi}{2} - \theta$, the point $\mathrm{A}(x, y)$ becomes $\mathrm{A}'(x', y')$ in a parallelogram, which forms an obtuse triangle $\mathrm{OA}'\mathrm{O}'$ with the $\mathcal{X}$-axis. In Figure 5.2(a), from point $\mathrm{A}(x, y)$ to $\mathrm{A}'(x', y')$, there is the following affine transformation:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & \cos\theta \\ 0 & \sin\theta \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}, \quad \text{or} \quad \begin{cases} x' = x + y\cos\theta \\ y' = y\sin\theta \end{cases}. \tag{5.4}$$

From this relation, therefore, a rectangle is a degenerated case of a parallelogram when $\theta = \frac{\pi}{2}$. In the following, (X_1, Y_1) and (X_2, Y_2) are used as the random variables denoting the coordinates *before* the transformation in (5.4), while (X_1', Y_1') and (X_2', Y_2') as the random coordinates *after* this transformation. For the points within the same parallelogram, the squared Euclidean distance, $Z = D^2$, is

$$\begin{aligned} Z &= D^2 = (X')^2 + (Y')^2 = (X_1' - X_2')^2 + (Y_1' - Y_2')^2 \\ &= [(X_1 - X_2) + \cos\theta(Y_1 - Y_2)]^2 + [\sin\theta(Y_1 - Y_2)]^2 \\ &= (X_1 - X_2)^2 + 2\cos\theta(X_1 - X_2)(Y_1 - Y_2) + (Y_1 - Y_2)^2, \end{aligned} \tag{5.5}$$

where (X_1, Y_1) and (X_2, Y_2) are the corresponding coordinates in the original rectangle. As the value of θ changes, only the coefficient $2\cos\theta$ will change in (5.5). Still let $X = X_1 - X_2$ and $Y = Y_1 - Y_2$ be the *difference* of $\mathcal{X}$ and $\mathcal{Y}$-coordinates before the transformation, then

$$Z = X^2 + 2\cos\theta XY + Y^2.$$

where the distributions of X and Y are the same as in (5.3).

In analytic geometry, this equation satisfies the implicit equation of a non-degenerated real ellipse [58] when $\theta \neq \frac{\pi}{2}$. The probability function therefore becomes $F_\omega = P(x^2 + 2\cos\theta xy + y^2 \leq z | X = x, Y = y)$. When $\theta = \frac{\pi}{2}$, (5.5) degenerates to the

formulation for a rectangle in (5.2).

(2) The Geometric Interpretation

Without loss of generality, for a parallelogram of side length of a and b, and with uniformly distributed random points, then $X_1, X_2 \sim U[0, a]$ and $Y_1, Y_2 \sim U[0, b]$ in the original rectangle. Therefore, the distributions of X and Y are in (5.3). The geometric interpretation of $Z = X^2 + 2\cos\theta XY + Y^2$ in the three-dimensional space, assuming $a = 3$, $b = 2$ and $\theta = \frac{\pi}{3}$ as an example, is shown as the squeezed bowl in Figure 5.2(b). The projections of Z on the $\mathcal{X}$-$\mathcal{Y}$ plane are concentric ellipses centered at $(0, 0)$, with cutoffs at $\mathcal{X} = \pm a$ and $\mathcal{Y} = \pm b$. This is expected from the equation of the ellipse $X^2 + 2\cos\theta XY + Y^2 = z$. When compared with Figure 5.1(b), it is interesting to observe that, after a rectangle has been squeezed to a parallelogram, the shape of $Z = D^2$ has also been squeezed, and the projections on the $\mathcal{X}$-$\mathcal{Y}$ plane are squeezed from circles to ellipses. Moreover, the level of the squeeze from circles to ellipses in the projected domain is determined by the same squeeze from rectangles to parallelograms.

5.3 Distance Distributions Associated with Rhombuses

Rhombuses are the special case of parallelograms with $a = b$. Define a *unit rhombus* as the rhombus with an acute angle of $\theta = \frac{\pi}{3}$ and a side length of 1. The coordinates of a point in such a unit rhombus are $x' = x + \frac{y}{2}$ and $y' = \frac{\sqrt{3}}{2}y$ according to (5.4). Here x and y are the coordinates in the original unit square before being squeezed to a rhombus, and both follow $U[0, 1]$. Thus,

$$Z = Z(X, Y) = X^2 + XY + Y^2, \tag{5.6}$$

where the distributions of X and Y are (5.3) with $a = b = 1$. Note that this method can be applied to any general parallelograms and rhombuses where $\theta \in [0, \frac{\pi}{2}]$. By Z given in (5.6), the product formulation becomes

$$\begin{aligned} F_Z(z) &= P_\Omega(Z \le z) = P(X^2 + XY + Y^2 \le z) \\ &= \iint \left[P(x^2 + xy + y^2 \le z | X = x, Y = y) \cdot f_{X,Y}(x, y)\right] dxdy. \end{aligned} \tag{5.7}$$

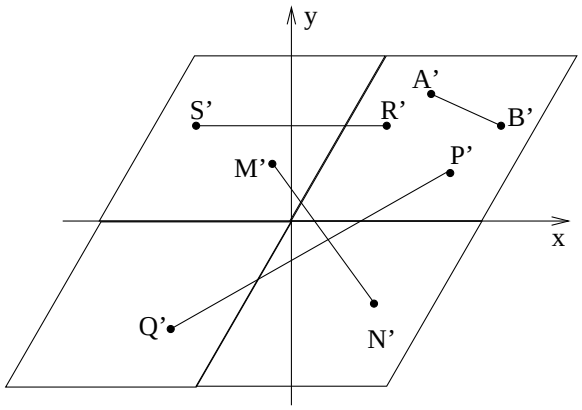


Figure 5.3: Random Points in Rhombuses.

Consequently, the elliptical projections of Z on the $\mathcal{X}$-$\mathcal{Y}$ plane, are cut off by $\mathcal{X} = \pm 1$ and $\mathcal{Y} = \pm 1$, the boundaries of f_ω that are defined by $f_X(x)$ and $f_Y(y)$. There are four cases of the geometric locations of two random points, when rhombuses are adjacent and similarly oriented, as shown in Figure 5.3: i.e., A′B′ that are within the same rhombus; R′S′ that are inside two parallel rhombuses sharing a side; P′Q′ and M′N′ that are inside two rhombuses sharing a common diagonal. Here P′Q′ and M′N′ are two different cases when a random node pair communicating across a concave geometry. In the following, we refer to them as long diagonal (long-diag) and short diagonal (short-diag), respectively.

Using the notation in (5.6), i.e., X and Y as the differences of $\mathcal{X}$ and $\mathcal{Y}$ coordinates *before* the affine transformation in (5.4), and assuming X_1, X_2, Y_1 and $Y_2 \sim U[0,1]$, we have the following four different cases of random distances as described above, by the displacement of rhombuses:

1. For A′ and B′, by the affine transformation, $X' = X_1' - X_2' = (X_1 - X_2) + \frac{1}{2}(Y_1 - Y_2)$, $Y' = Y_1' - Y_2' = \frac{\sqrt{3}}{2}(Y_1 - Y_2)$. According to (5.5), we have $X = X_1 - X_2$ and $Y = Y_1 - Y_2$ in (5.6);

2. For R′ and S′, we have $X = X_1 + X_2$ and $Y = Y_1 - Y_2$ in (5.6);

3. For P′ and Q′, we have $X = X_1 + X_2$ and $Y = Y_1 + Y_2$ in (5.6);

4. For M′ and N′, we have $X = -X_1 - X_2$ and $Y = Y_1 + Y_2$ in (5.6).

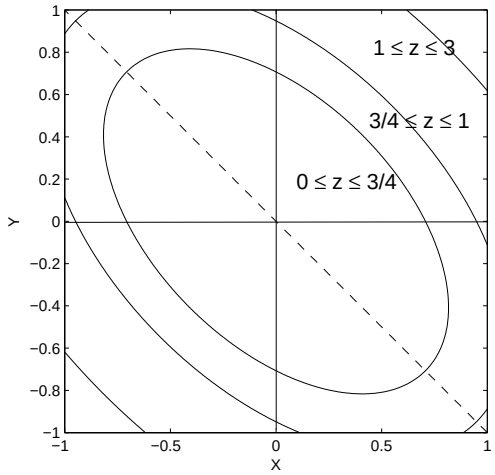


Figure 5.4: Three Sub-cases for $Z = |A'B'|^2$.

5.3.1 Distance Distribution within a Rhombus

Take $|A'B'|$ in Figure 5.3 as an example, i.e., the random distances within a unit rhombus. The projections of Z on the $\mathcal{X}$-$\mathcal{Y}$ plane are shown in Figure 5.4. Consider an *effective ellipse* as the projection of Z within the effective region of f_ω, which is determined by

$$f_X(x) = \begin{cases} 1+x & -1 \le x \le 0 \\ 1-x & 0 \le x \le 1 \end{cases} \quad \text{and} \quad f_Y(y) = \begin{cases} 1+y & -1 \le y \le 0 \\ 1-y & 0 \le y \le 1 \end{cases},$$

i.e., $\mathcal{X} \in [-1, 1]$ and $\mathcal{Y} \in [-1, 1]$. This region is divided into four compartments by $\mathcal{X} = 0$ and $\mathcal{Y} = 0$, the transitional values in $f_X(x)$ and $f_Y(y)$ above. From Figure 5.4 we can observe that, the effective ellipse has different overlapping areas with the four compartments, determined by a specific value of Z.

It is then natural to sub-divide the ellipse at certain transitional values of Z. Although there are nine transitional points of f_ω at $\mathcal{X} = 0$, $\mathcal{X} = \pm 1$, $\mathcal{Y} = 0$ and $\mathcal{Y} = \pm 1$, because of symmetry, there are only four transitional values of Z as the ellipse expanding: z_1 where the ellipse is a point at the origin, z_2 where $\mathcal{X} = \pm 1$ or $\mathcal{Y} = \pm 1$ become the tangent lines of the ellipse, z_3 where the semimajor axis of the ellipse is equal to $\sqrt{a^2 + b^2} = \sqrt{2}$, and z_4 where the semiminor axis of the ellipse is equal to $\sqrt{2}$. The semimajor axis and semiminor axis of ellipse $X^2 + XY + Y^2 = z$ are $\sqrt{2z}$ and $\sqrt{\frac{2}{3}z}$ respectively, and thus we have $z_1 = 0$, $z_2 = \frac{3}{4}$, $z_3 = 1$ and $z_4 = 3$.

The three effective ranges of z are hence $\left[0, \frac{3}{4}\right]$, $\left[\frac{3}{4}, 1\right]$ and $[1, 3]$, yielding the three sub-cases as shown in Figure 5.4.

$0 \leq z \leq \frac{3}{4}$:

In this sub-case, the entire ellipse lies inside the effective region of f_ω. By looking at the first quadrant where $0 \leq x \leq 1$ and $0 \leq y \leq 1$, the density is $f_{X,Y}(x, y) = (1-x)(1-y)$. Therefore, $F_Z(z)$ is $\int_0^{\sqrt{z}} \int_0^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (1-x)dx(1-y)dy$. When $0 \leq x \leq 1$ and $-1 \leq y \leq 0$, $F_Z(z) = 2\int_{-\sqrt{z}}^{0} \int_{-y}^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (1-x)dx(1+y)dy$, by the symmetry with respect to $\mathcal{Y} = -\mathcal{X}$. Because the ellipse is also symmetric with regard to $\mathcal{Y} = \mathcal{X}$, we have the distribution of the squared distance $F_Z(z)$ as the following

$$\begin{aligned} F_Z(z) &= 2\left[\int_0^{\sqrt{z}} \int_0^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (1-x)dx(1-y)dy + 2\int_{-\sqrt{z}}^{0} \int_{-y}^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (1-x)dx(1+y)dy\right] \\ &= \left(\frac{2}{3} + \frac{\pi}{9\sqrt{3}}\right) z^2 - \frac{32}{9} z^{3/2} + \frac{2\pi}{\sqrt{3}} z. \end{aligned}$$

The probability density function (PDF) is the derivative of $F_Z(z)$, i.e.,

$$f_Z(z) = F_Z'(z) = \left(\frac{4}{3} + \frac{2\pi}{9\sqrt{3}}\right) z - \frac{16}{3}\sqrt{z} + \frac{2\pi}{\sqrt{3}}.$$

$\frac{3}{4} \leq z \leq 1$:

In this sub-case, the ellipse intersects with $\mathcal{X} = 1$, and the intersection points are $(1, -\frac{1}{2} \pm \sqrt{z - \frac{3}{4}})$. Therefore, by symmetry, $F_Z(z)$ is

$$\begin{aligned} F_Z(z) &= 4\left[\int_{-\sqrt{z}}^{0} \int_{-y}^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (1-x)dx(1+y)dy \right. \\ &\quad \left. - \int_{-\frac{1}{2}-\sqrt{z-\frac{3}{4}}}^{-\frac{1}{2}+\sqrt{z-\frac{3}{4}}} \int_{1}^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (1-x)dx(1+y)dy\right] \\ &\quad +2\int_0^{\sqrt{z}} \int_0^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (1-x)dx(1-y)dy \\ &= \frac{4}{\sqrt{3}}\left(2z + \frac{z^2}{3}\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{2}{3} - \frac{5\pi}{9\sqrt{3}}\right) z^2 - \frac{32}{9} z^{3/2} - \frac{2\pi}{\sqrt{3}} z + \frac{14z+3}{6}\sqrt{4z-3}, \end{aligned}$$

and the PDF is

$$f_Z(z) = \frac{8}{\sqrt{3}}\left(1+\frac{z}{3}\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{4}{3}-\frac{10\pi}{9\sqrt{3}}\right)z + \frac{10}{3}\sqrt{4z-3} - \frac{16}{3}\sqrt{z} - \frac{2\pi}{\sqrt{3}}.$$

$1 \le z \le 3$:

$F_Z(z)$ in this sub-case is

$$\begin{aligned} F_Z(z) &= 4\left[\int_0^{-\frac{1}{2}+\sqrt{z-\frac{3}{4}}}\int_y^1 (1-x)dx(1-y)dy + \int_{-\frac{1}{2}+\sqrt{z-\frac{3}{4}}}^{\sqrt{\frac{z}{3}}}\int_y^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (1-x)dx(1-y)dy\right] \\ &\quad +2\int_{-1}^0\int_0^1 (1-x)dx(1+y)dy \\ &= \frac{2}{\sqrt{3}}\left(\frac{z^2}{3}-2z\right)\sin^{-1}\frac{\sqrt{3}}{4}\frac{\sqrt{4z-3}-1}{\sqrt{z}} - \left(\frac{1}{3}+\frac{\pi}{9\sqrt{3}}\right)z^2 + \left(\frac{2\pi}{3\sqrt{3}}-1\right)z \\ &\quad +\frac{22z+15}{36}\sqrt{4z+3} + \frac{1}{4}, \end{aligned}$$

and the PDF is

$$f_Z(z) = \frac{4}{\sqrt{3}}\left(1-\frac{z}{3}\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} - \left(\frac{2}{3}-\frac{2\pi}{9\sqrt{3}}\right)z + \sqrt{4z-3} - \frac{2\pi}{3\sqrt{3}} - 1.$$

With $D = \sqrt{Z}$, the distance distribution $f_D(d)$ is

$$f_D(d) = F_Z'(d^2) = 2df_Z(d^2). \tag{5.8}$$

Combining the above three sub-cases and using (5.8), we have $f_{D_{\mathrm{I}}}(d)$ for the random distances inside a rhombus, as shown in (5.9).

Result 1 (Random distances within a unit rhombus)**.** The probability density function of the random distances between two uniformly distributed points that are both

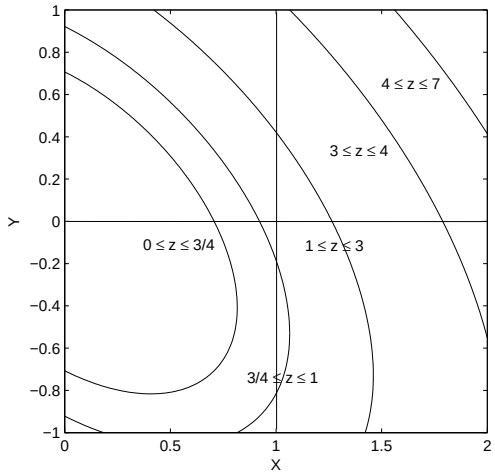


Figure 5.5: Five Sub-cases for $Z = |R'S'|^2$.

inside the same unit rhombus, is

$$f_{D_{\rm I}}(d) = 2d \begin{cases} \left(\frac{4}{3} + \frac{2\pi}{9\sqrt{3}}\right) d^2 - \frac{16}{3} d + \frac{2\pi}{\sqrt{3}} & 0 \le d \le \frac{\sqrt{3}}{2} \\ \frac{8}{\sqrt{3}} \left(1 + \frac{d^2}{3}\right) \sin^{-1} \frac{\sqrt{3}}{2d} + \left(\frac{4}{3} - \frac{10\pi}{9\sqrt{3}}\right) d^2 - \frac{16}{3} d + \frac{10}{3} \sqrt{4d^2 - 3} - \frac{2\pi}{\sqrt{3}} & \frac{\sqrt{3}}{2} \le d \le 1 \\ \frac{4}{\sqrt{3}} \left(1 - \frac{d^2}{3}\right) \sin^{-1} \frac{\sqrt{3}}{2d} - \left(\frac{2}{3} - \frac{2\pi}{9\sqrt{3}}\right) d^2 + \sqrt{4d^2 - 3} - \frac{2\pi}{3\sqrt{3}} - 1 & 1 \le d \le \sqrt{3} \\ 0 & \text{otherwise} \end{cases} . \tag{5.9}$$

It can be shown that at any transitional value of D between different sub-cases in (5.9), which is equal to $\sqrt{Z}$, $f_{D_{\rm I}}(d)$ has the same value in both adjacent sub-cases. This indicates that $f_{D_{\rm I}}(d)$ is a continuous function, as expected for the distribution of a continuous random variable.

5.3.2 Distance Distribution between Parallel Rhombuses

We then consider points R′ and S′ as shown in Figure 5.3, where $X = X_1 + X_2$ and $Y = Y_1 - Y_2$ are to be used in (5.6). For the two adjacent rhombuses that are parallel to each other and share a side, the distributions of X and Y are

$$f_X(x) = \begin{cases} x & 0 \le x \le 1 \\ 2 - x & 1 \le x \le 2 \end{cases}, \quad \text{and} \quad f_Y(y) = \begin{cases} 1 + y & -1 \le y \le 0 \\ 1 - y & 0 \le y \le 1 \end{cases} . \tag{5.10}$$

The elliptical contours of $Z = X^2 + XY + Y^2$ on the $\mathcal{X}$-$\mathcal{Y}$ plane now have to

be in $\mathcal{X} \in [0,2]$ and $\mathcal{Y} \in [-1,1]$, which is divided into four compartments by $\mathcal{X} = 1$ and $\mathcal{Y} = 0$, the transitional values in (5.10). As the value of Z increases, we have the effective ellipse expanding inside the region of f_ω, as shown in Figure 5.5, corresponding to the following five sub-cases.

$0 \le z \le \frac{3}{4}$:

$F_Z(z)$ in this sub-case is

$$\begin{aligned} F_Z(z) &= \int_0^{\sqrt{\frac{4z}{3}}} \int_{-\frac{x}{2}-\sqrt{z-\frac{3}{4}x^2}}^{-\sqrt{\frac{z}{3}}} (1+y)dyxdx + \int_{-\sqrt{\frac{z}{3}}}^{0} \int_0^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} xdx(1+y)dy \\ &\quad + \int_0^{\sqrt{z}} \int_0^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} xdx(1-y)dy = \frac{8}{9}z^{3/2} - \left(\frac{1}{3} + \frac{\pi}{18\sqrt{3}}\right) z^2, \end{aligned}$$

and the PDF is

$$f_Z(z) = \frac{4}{3}\sqrt{z} - \left(\frac{2}{3} + \frac{\pi}{9\sqrt{3}}\right) z.$$

$\frac{3}{4} \le z \le 1$:

The ellipse intersects with $\mathcal{X} = 1$ and $\mathcal{Y} = -1$, and the intersection points are $(1, -\frac{1}{2} \pm \sqrt{z-\frac{3}{4}})$ and $(\frac{1}{2} \pm \sqrt{z-\frac{3}{4}}, -1)$. $F_Z(z)$ in this sub-case is

$$\begin{aligned} F_Z(z) &= \left[\int_0^1 \int_{-\frac{x}{2}-\sqrt{z-\frac{3}{4}x^2}}^{-\sqrt{\frac{z}{3}}} (1+y)dyxdx - \int_{\frac{1}{2}-\sqrt{z-\frac{3}{4}}}^{\frac{1}{2}+\sqrt{z-\frac{3}{4}}} \int_{-\frac{x}{2}-\sqrt{z-\frac{3}{4}x^2}}^{-1} (1+y)dyxdx\right] \\ &\quad + \left[\int_1^{\sqrt{\frac{4z}{3}}} \int_{-\frac{x}{2}-\sqrt{z-\frac{3}{4}x^2}}^{-\sqrt{\frac{z}{3}}} (1+y)dy(2-x)dx + \int_1^{\sqrt{\frac{4z}{3}}} \int_{-\sqrt{\frac{z}{3}}}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} (1+y)dy(2-x)dx\right] \\ &\quad + \left[\int_{-\sqrt{\frac{z}{3}}}^{0} \int_0^{\sqrt{z}} xdx(1+y)dy + \int_{\sqrt{z}}^{1} \int_{-\sqrt{\frac{z}{3}}}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} (1+y)dyxdx\right] \\ &\quad + \int_0^{\sqrt{z}} \int_0^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} xdx(1-y)dy \\ &= -\frac{2}{\sqrt{3}}\left(2z + \frac{z^2}{2}\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{4\pi}{9\sqrt{3}} - \frac{1}{3}\right)z^2 + \frac{8}{9}z^{3/2} + \frac{2\pi}{\sqrt{3}}z - \frac{93z+15}{72}\sqrt{4z-3}, \end{aligned}$$

and the PDF is

$$f_Z(z) = -\frac{2}{\sqrt{3}}(z+2)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{8\pi}{9\sqrt{3}} - \frac{2}{3}\right)z + \frac{4}{3}\sqrt{z} - \frac{11}{6}\sqrt{4z-3} + \frac{2\pi}{\sqrt{3}}.$$

$1 \le z \le 3$:

In this sub-case, the ellipse intersects with $\mathcal{X} = 1$, with the intersection point $(1, -\frac{1}{2} + \sqrt{z - \frac{3}{4}})$. $F_Z(z)$ is

$$\begin{aligned} F_Z(z) &= \left[\int_0^1\int_0^1 xdx(1-y)dy - \int_{-\frac{1}{2}+\sqrt{z-\frac{3}{4}}}^1 \int_{-\frac{y}{2}+\sqrt{z-\frac{3}{4}x^2}}^1 xdx(1-y)dy\right] \\ &\quad + \int_{-1}^0\int_1^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (2-x)dx(1+y)dy + \int_{-1}^0\int_0^1 xdx(1+y)dy \\ &\quad + \int_0^{-\frac{1}{2}+\sqrt{z-\frac{3}{4}}}\int_1^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (2-x)dx(1-y)dy \\ &= \frac{2z^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{1}{3} - \frac{\pi}{9\sqrt{3}}\right)z^2 - \frac{16}{9}z^{3/2} + \left(\frac{2\pi}{3\sqrt{3}} + \frac{1}{2}\right)z \\ &\quad + \frac{10z-3}{36}\sqrt{4z-3} - \frac{11}{24}, \end{aligned}$$

and the PDF is

$$f_Z(z) = \frac{4z}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{2}{3} - \frac{2\pi}{9\sqrt{3}}\right)z - \frac{8}{3}\sqrt{z} + \frac{\sqrt{4z-3}}{3} + \frac{2\pi}{3\sqrt{3}} + \frac{1}{2}.$$

$3 \le z \le 4$:

The ellipse intersects with the $\mathcal{X}$-axis at $(\sqrt{z}, 0)$. $F_Z(z)$ in this sub-case is

$$
\begin{aligned}
F_Z(z) &= \frac{1}{2} + \left[\int_{-1}^{0} \int_{1}^{\sqrt{z}} (2-x)dx(1+y)dy + \int_{\sqrt{z}}^{2} \int_{-1}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} (1+y)dy(2-x)dx \right] \\
&\quad + \int_{0}^{1} \int_{1}^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (2-x)dx(1-y)dy \\
&= \left(\frac{2z}{\sqrt{3}} - \frac{z^2}{6\sqrt{3}} \right) \sin^{-1} \frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{2z}{\sqrt{3}} + \frac{z^2}{6\sqrt{3}} \right) \sin^{-1} \sqrt{\frac{3}{z}} + \left(\frac{1}{6} - \frac{\pi}{18\sqrt{3}} \right) z^2 \\
&\quad - \frac{16}{9} z^{3/2} + \left(\frac{3}{4} - \frac{2\pi}{3\sqrt{3}} \right) z + \frac{6z+3}{16} \sqrt{4z-3} + \frac{13z+16}{18} \sqrt{z-3} - \frac{55}{48},
\end{aligned}
$$

and the PDF is

$$
\begin{aligned}
f_Z(z) &= \left(\frac{2}{\sqrt{3}} - \frac{z}{3\sqrt{3}} \right) \sin^{-1} \frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{2}{\sqrt{3}} + \frac{z}{3\sqrt{3}} \right) \sin^{-1} \sqrt{\frac{3}{z}} + \left(\frac{1}{3} - \frac{\pi}{9\sqrt{3}} \right) z \\
&\quad - \frac{8}{3}\sqrt{z} + \frac{7}{12}\sqrt{4z-3} + \sqrt{z-3} + \frac{3}{4} - \frac{2\pi}{3\sqrt{3}}.
\end{aligned}
$$

$4 \le z \le 7$:

At $\mathcal{X} = 2$, the intersection point with the ellipse is $(2, -1+\sqrt{z-3})$. Then $F_Z(z)$ is

$$
\begin{aligned}
F_Z(z) &= 1 - \int_{-1+\sqrt{z-3}}^{1} \int_{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}}^{2} (2-x)dx(1-y)dy \\
&= \left(\frac{2z}{\sqrt{3}} - \frac{z^2}{6\sqrt{3}} \right) \left(\sin^{-1} \frac{\sqrt{3}}{2\sqrt{z}} + \sin^{-1} \sqrt{\frac{3}{z}} \right) + \left(\frac{\pi}{18\sqrt{3}} - \frac{1}{6} \right) z^2 - \left(\frac{2\pi}{3\sqrt{3}} + \frac{5}{4} \right) z \\
&\quad + \frac{6z+3}{16} \sqrt{4z-3} + \frac{z+6}{6} \sqrt{z-3} - \frac{23}{48},
\end{aligned}
$$

and the PDF is

$$
\begin{aligned}
f_Z(z) &= \left(\frac{2}{\sqrt{3}} - \frac{z}{3\sqrt{3}} \right) \left(\sin^{-1} \frac{\sqrt{3}}{2\sqrt{z}} + \sin^{-1} \sqrt{\frac{3}{z}} \right) + \frac{7}{12}\sqrt{4z-3} + \frac{\sqrt{z-3}}{3} \\
&\quad + \left(\frac{\pi}{9\sqrt{3}} - \frac{1}{3} \right) z - \frac{2\pi}{3\sqrt{3}} - \frac{5}{4}.
\end{aligned}
$$

Using (5.8), and combining the above five sub-cases, the result for the random distances of two parallel adjacent unit rhombuses is

Result 2 (Random distances between two parallel adjacent unit rhombuses sharing a side)**.** The probability density function of the random distances between two uniformly distributed points, one in each of the two adjacent unit rhombuses that are parallel to each other, is

$$f_{D_{\mathrm{P}}}(d) = 2d \begin{cases} \frac{4}{3}d - \left(\frac{2}{3} + \frac{\pi}{9\sqrt{3}}\right) d^2 & 0 \le d \le \frac{\sqrt{3}}{2} \\ -\frac{2}{\sqrt{3}}(d^2+2)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{8\pi}{9\sqrt{3}} - \frac{2}{3}\right) d^2 + \frac{4}{3}d - \frac{11}{6}\sqrt{4d^2-3} + \frac{2\pi}{\sqrt{3}} & \frac{\sqrt{3}}{2} \le d \le 1 \\ \frac{4d^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{2}{3} - \frac{2\pi}{9\sqrt{3}}\right) d^2 - \frac{8}{3}d + \frac{\sqrt{4d^2-3}}{3} + \frac{2\pi}{3\sqrt{3}} + \frac{1}{2} & 1 \le d \le \sqrt{3} \\ \left(\frac{2}{\sqrt{3}} - \frac{d^2}{3\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{2}{\sqrt{3}} + \frac{d^2}{3\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{1}{3} - \frac{\pi}{9\sqrt{3}}\right) d^2 & \\ \quad - \frac{8}{3}d + \frac{7}{12}\sqrt{4d^2-3} + \sqrt{d^2-3} + \frac{3}{4} - \frac{2\pi}{3\sqrt{3}} & \sqrt{3} \le d \le 2 \\ \left(\frac{2}{\sqrt{3}} - \frac{d^2}{3\sqrt{3}}\right)\left(\sin^{-1}\frac{\sqrt{3}}{2d} + \sin^{-1}\frac{\sqrt{3}}{d}\right) + \left(\frac{\pi}{9\sqrt{3}} - \frac{1}{3}\right) d^2 & \\ \quad + \frac{7}{12}\sqrt{4d^2-3} + \frac{\sqrt{d^2-3}}{3} - \frac{2\pi}{3\sqrt{3}} - \frac{5}{4} & 2 \le d \le \sqrt{7} \\ 0 & \text{otherwise} \end{cases} \tag{5.11}$$

5.3.3 Distance Distribution between Diagonal Rhombuses

(1) Long-Diag Adjacent Rhombuses

The distance distribution of two adjacent rhombuses when they have a common long diagonal is given as follows. The projection of Z is shown in Figure A.1, and the derivation is given in Appendix A.1.

Result 3 (Random distances between two diagonal adjacent unit rhombuses—the long-diag case)**.** The probability density function of the random distances between two uniformly distributed points, one in each of the two adjacent unit rhombuses that

have a common long diagonal, is

$$f_{D_{\mathrm{LD}}}(d) = 2d \begin{cases} \left(\frac{1}{3} - \frac{\pi}{9\sqrt{3}}\right) d^2 & 0 \le d \le 1 \\ -\frac{4d^2}{3\sqrt{3}} \sin^{-1} \frac{\sqrt{3}}{2d} + \left(\frac{\pi}{3\sqrt{3}} - 1\right) d^2 + \frac{8}{3} d - \frac{\sqrt{4d^2-3}}{3} - 1 & 1 \le d \le \sqrt{3} \\ \frac{4}{\sqrt{3}} \left(\frac{d^2}{3} - 2\right) \sin^{-1} \frac{\sqrt{3}}{2d} + \left(\frac{1}{3} - \frac{\pi}{9\sqrt{3}}\right) d^2 + \frac{8}{3} d - \frac{7}{3}\sqrt{4d^2-3} & \\ \quad + \frac{4\pi}{3\sqrt{3}} + 1 & \sqrt{3} \le d \le 2 \\ \frac{4}{\sqrt{3}} \left(\frac{d^2}{3} - 2\right) \sin^{-1} \frac{\sqrt{3}}{2d} + \frac{2d^2}{3\sqrt{3}} \sin^{-1} \frac{\sqrt{3}}{d} + \left(1 - \frac{\pi}{3\sqrt{3}}\right) d^2 & \\ \quad - \frac{7}{3}\sqrt{4d^2-3} + \frac{2}{3}\sqrt{d^2-3} + \frac{4\pi}{3\sqrt{3}} + 3 & 2 \le d \le \sqrt{7} \\ \frac{2}{\sqrt{3}} \left(4 - \frac{d^2}{3}\right) \sin^{-1} \frac{\sqrt{3}}{d} + \left(\frac{\pi}{9\sqrt{3}} - \frac{1}{3}\right) d^2 + 2\sqrt{d^2-3} & \\ \quad - \frac{4\pi}{3\sqrt{3}} - 2 & \sqrt{7} \le d \le 2\sqrt{3} \\ 0 & \text{otherwise} \end{cases} . \tag{5.12}$$

(1) Short-Diag Adjacent Rhombuses

The distance distribution of two adjacent rhombuses when they have a common short diagonal is given as follows. The projection of Z is shown in Figure A.2, and the derivation is given in Appendix A.2.

Result 4 (Random distances between two diagonal adjacent unit rhombuses—the short-diag case)**.** The probability density function of the random distances between two uniformly distributed points, one in each of the two adjacent unit rhombuses that have a common short diagonal, is

$$f_{D_{\mathrm{SD}}}(d) = 2d \begin{cases} \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right) d^2 & 0 \le d \le \frac{\sqrt{3}}{2} \\ \frac{8d^2}{3\sqrt{3}} \sin^{-1} \frac{\sqrt{3}}{2d} + \left(\frac{1}{3} - \frac{10\pi}{9\sqrt{3}}\right) d^2 + \frac{2}{3}\sqrt{4d^2-3} & \frac{\sqrt{3}}{2} \le d \le 1 \\ -\left(\frac{4d^2}{3\sqrt{3}} + \frac{8}{\sqrt{3}}\right) \sin^{-1} \frac{\sqrt{3}}{2d} + \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right) d^2 + \frac{8}{3} d - 3\sqrt{4d^2-3} & \\ \quad + \frac{8\pi}{3\sqrt{3}} + 1 & 1 \le d \le \sqrt{3} \\ \frac{8}{\sqrt{3}} \sin^{-1} \frac{\sqrt{3}}{d} - d^2 + \frac{8}{3} d + \frac{8}{3}\sqrt{d^2-3} - \frac{8\pi}{3\sqrt{3}} - 4 & \sqrt{3} \le d \le 2 \\ 0 & \text{otherwise} \end{cases} . \tag{5.13}$$

It can also be shown that at any transitional value of D between different sub-cases, $f_{D_{\mathrm{P}}}(d)$, as well as $f_{D_{\mathrm{LD}}}(d)$ and $f_{D_{\mathrm{SD}}}(d)$, always has the same value in both adjacent sub-cases, indicating a continuous distance distribution as expected.

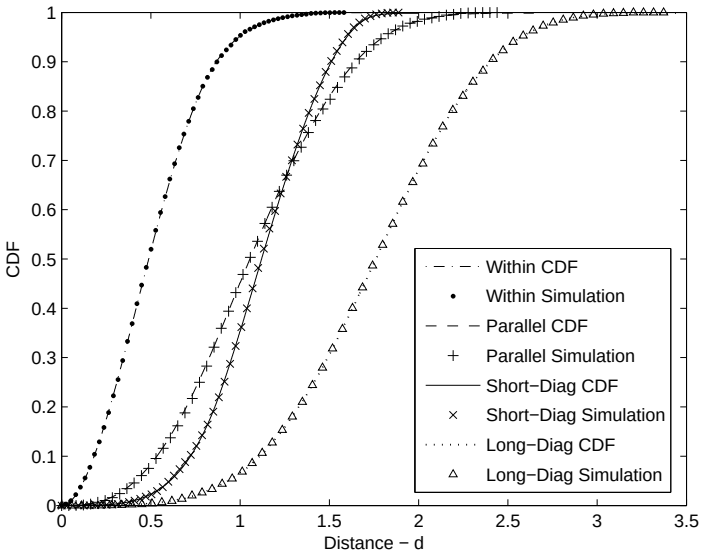


Figure 5.6: Cumulative Distribution Functions and Simulation Results for Random Distances Associated with Rhombuses.

Although unit rhombuses are assumed in (5.9)–(5.13), the distance distribution can be easily scaled by an arbitrary nonzero scalar. Let the scalar be $s > 0$, then

$$F_{sD}(d) = P(sD \leq d) = P(D \leq \frac{d}{s}) = F_D(\frac{d}{s}).$$

Therefore,

$$f_{sD}(d) = F'_D(\frac{d}{s}) = \frac{1}{s} f_D(\frac{d}{s}). \tag{5.14}$$

5.3.4 Distance Verification

(1) Verification by Simulation

Figure 5.6 shows the comparison between the cumulative distribution functions (CDFs) of the random distances, and the simulation results by generating $1,000$ pairs of random points with the corresponding geometric locations as shown in Figure 5.3. Figure 5.6 demonstrates that our distribution functions are very accurate when compared with the simulation.

Table 5.1: Moments and Variances of Rhombus Distributions—Numerical vs Simulation Results

Endpoint Geometry	PDF/Sim	$M_D^{(1)}$	$M_D^{(2)}$	Var_D
Within a Single Rhombus	$f_{D_{\mathrm{I}}}(d)$	$0.5123783359s$	$0.3333333333s$	$0.0708017742s^2$
	Sim	$0.5137344650s$	$0.3356749448s$	$0.0717518443s^2$
Between two Parallel Rhombuses	$f_{D_{\mathrm{P}}}(d)$	$1.0750863337s$	$1.3331823503s$	$0.1773717254s^2$
	Sim	$1.0749140141s$	$1.3318514164s$	$0.1764112787s^2$
Between two Long-Diag Adjacent Rhombuses	$f_{D_{\mathrm{LD}}}(d)$	$1.7570796617s$	$3.3330145688s$	$0.2456856311s^2$
	Sim	$1.7575618714s$	$3.3319543736s$	$0.2429306419s^2$
Between two Short-Diag Adjacent Rhombuses	$f_{D_{\mathrm{SD}}}(d)$	$1.1150961004s$	$1.3332064091s$	$0.0897670959s^2$
	Sim	$1.1156689048s$	$1.3341275970s$	$0.0894104918s^2$

(2) Validation by Recursion

By looking at Figure 5.3, the four adjacent rhombuses recursively form a larger rhombus, with side length of 2. According to (5.14), the distance distribution in this large rhombus is $f_{2D_{\mathrm{I}}}(d) = \frac{1}{2} f_{D_{\mathrm{I}}}(\frac{d}{2})$. On the other hand, if we look at any of the random endpoints of a link inside the large rhombus, they will fall into one of the four individual cases: both endpoints are inside the same small rhombus, with probability $\frac{1}{4}$; the two endpoints fall into two parallel rhombuses, with probability $\frac{1}{2}$; the two endpoints fall into two diagonal rhombuses (either long or short-diag), both with probability $\frac{1}{8}$. The resulting density function is given by a probabilistic sum, $f_{2D_{\mathrm{I}}}(d) = \frac{1}{4} f_{D_{\mathrm{I}}}(d) + \frac{1}{2} f_{D_{\mathrm{P}}}(d) + \frac{1}{8} f_{D_{\mathrm{LD}}}(d) + \frac{1}{8} f_{D_{\mathrm{SD}}}(d)$, where $f_{D_{\mathrm{I}}}(d)$, $f_{D_{\mathrm{P}}}(d)$, $f_{D_{\mathrm{LD}}}(d)$ and $f_{D_{\mathrm{SD}}}(d)$ are given in (5.9), (5.11), (5.12) and (5.13), respectively. To confirm that the above two equalities of $f_{2D_{\mathrm{I}}}(d)$ are equivalent, we have verified them mathematically in Appendix B.

5.4 Practical Results

5.4.1 Statistical Moments of Random Distances

The distance distribution functions given in Section 5.3.1 to 5.3.3 can conveniently lead to all the statistical moments of the random distances associated with rhombuses. Given $f_{D_{\mathrm{I}}}(d)$ in (5.9), for example, the first moment (mean) of d, i.e., the average

distance within a single rhombus, is

$$\begin{aligned} M_{D_{\rm I}}^{(1)} &= \int_0^{\sqrt{3}} x f_{D_{\rm I}}(x)dx = \frac{\sqrt{3}}{8} + \frac{3}{40} + \frac{1}{80}\left[7\ln\left(2\sqrt{3}+3\right) - 6\ln\left(2\sqrt{3}-3\right)\right] \\ &\approx 0.5123783359, \end{aligned}$$

and the second raw moment is

$$M_{D_{\rm I}}^{(2)} = \int_0^{\sqrt{3}} x^2 f_{D_{\rm I}}(x)dx = \frac{1}{3},$$

from which the variance (the second central moment) can be derived as

$$Var_{D_{\rm I}} = M_{D_{\rm I}}^{(2)} - \left[M_{D_{\rm I}}^{(1)}\right]^2 \approx 0.0708017742.$$

When the side length of a rhombus is scaled by s, the corresponding first two statistical moments given above then become

$$M_{D_{\rm I}}^{(1)} = 0.5123783359s, \quad M_{D_{\rm I}}^{(2)} = \frac{s}{3} \quad \text{and} \quad Var_{D_{\rm I}} = 0.0708017742s^2.$$

Table 5.1 lists the first two moments, and the variance of the random distances in the four cases given in Section 5.3.1 to 5.3.3, and the corresponding simulation results for verification purposes.

5.4.2 Polynomial Fits of Random Distances

Table 5.2 lists the coefficients of the degree-20 polynomial fits of the original PDFs derived in this section, from d^{20} to d^0, and the corresponding norm of residuals in the least squares fitting. A residual is defined as the difference between the actual value and the predicted value [5]. Figure 5.7 (a)–(d) plot the polynomials listed in Table 5.2 with the original PDFs. From the figure, it can be seen that all the polynomials match closely with the original PDFs. These high-order polynomials facilitate further manipulations of the distance distribution functions, with a high accuracy, such as the same polynomial fittings in [108].

Table 5.2: Coefficients of the Polynomial Fit and the Norm of Residuals (NR) for Rhombus Distributions

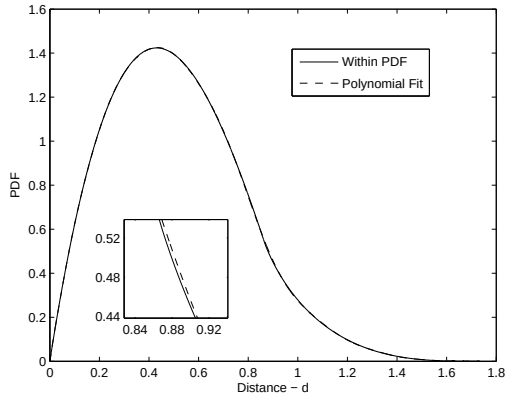
PDF	Polynomial Coefficients	NR
$f_{D_{\mathrm{I}}}(d)$	$10^8 \times [-0.000166 \;\; 0.002955 \;\; -0.024296 \;\; 0.122501$ $-0.423700 \;\; 1.065297 \;\; -2.013141 \;\; 2.916117 \;\; -3.273038$ $2.858823 \;\; -1.941184 \;\; 1.018567 \;\; -0.408430 \;\; 0.123035$ $-0.027166 \;\; 0.004245 \;\; -0.000446 \;\; 0.000029 \;\; -0.00000116$ $0.0000000892 \;\; -0.0000000000648]$	0.095901
$f_{D_{\mathrm{P}}}(d)$	$10^5 \times [0.000019 \;\; -0.000513 \;\; 0.006194 \;\; -0.045754$ $0.231329 \;\; -0.847940 \;\; 2.328313 \;\; -4.879792 \;\; 7.881444$ $-9.836044 \;\; 9.45481 \;\; -6.93816 \;\; 3.828928 \;\; -1.553959$ $0.4491195 \;\; -0.088213 \;\; 0.010963 \;\; -0.000782 \;\; 0.0000499$ $-0.0000000863 \;\; -0.0000000025988]$	0.059485
$f_{D_{\mathrm{LD}}}(d)$	$10^4 \times [0.0000002 \;\; -0.0000074 \;\; 0.0001242 \;\; -0.0012723$ $0.0089355 \;\; -0.045576 \;\; 0.174538 \;\; -0.511656 \;\; 1.160176$ $-2.042559 \;\; 2.78728 \;\; -2.927992 \;\; 2.339236 \;\; -1.395868$ $0.606494 \;\; -0.185173 \;\; 0.037762 \;\; -0.004740 \;\; 0.0003293$ $-0.0000098938 \;\; 0.0000000733372]$	0.011340
$f_{D_{\mathrm{SD}}}(d)$	$10^7 \times [-0.000022 \;\; 0.000493 \;\; -0.005116 \;\; 0.032591$ $-0.142436 \;\; 0.452238 \;\; -1.077816 \;\; 1.965337 \;\; -2.77016$ $3.029748 \;\; -2.567377 \;\; 1.674771 \;\; -0.831348 \;\; 0.308601$ $-0.083558 \;\; 0.015932 \;\; -0.002034 \;\; 0.000161 \;\; -0.00000698$ $0.00000013178 \;\; -0.0000000006098]$	0.147836

5.5 Summary

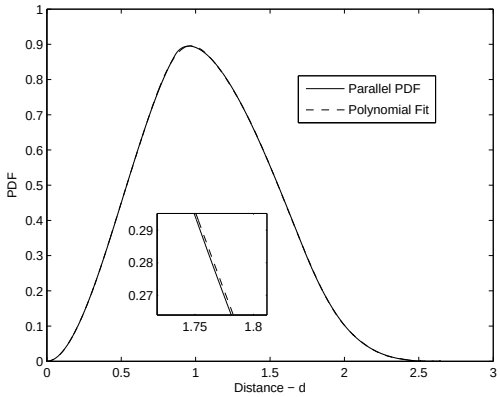
This chapter developed a new approach by a quadratic product, separating the constraint of geometric shapes and the geometric distribution of random points within or between these shapes. The closed-form probability density functions of random distances associated with rhombuses are obtained for the first time in the literature. The approach can be easily extended to parallelograms, either adjacent to or displaced from each other. Furthermore, by decomposing a hexagon into congruent rhombuses, this approach is promising in obtaining the distributions of random distances associated with hexagons.

5.6 Discussions and Future Work

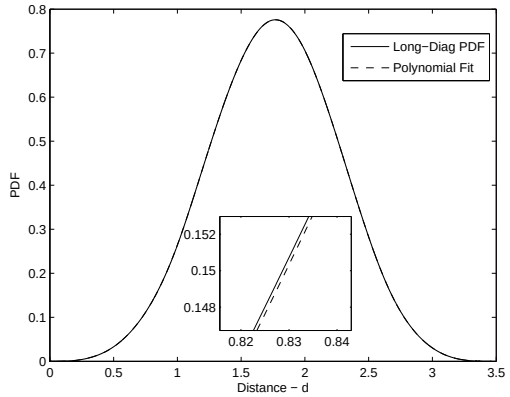
The quadratic product in (5.1) is the multiplication of the probability function and probability density, both of which can be extended to more general cases. The prob-



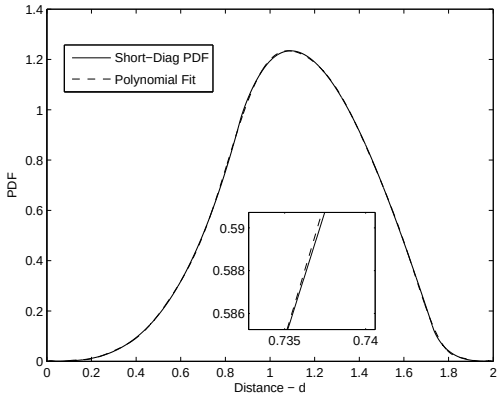
(a) Within a Single Rhombus



(b) Between two Parallel Adjacent Rhombuses



(c) Between two Long-Diag Adjacent Rhombuses



(d) Between two Short-Diag Adjacent Rhombuses

Figure 5.7: Polynomial Fit of the Distance Distribution Functions Associated with Rhombuses.

ability function determines a particular geometric shape, and the characteristics of the random coordinates inside the shape is dependent on the probability density. The product formulation in (5.1) is highly flexible by separating the geometrical constraints and the distributions of random points. Examples are shown in the following as the extensions to the current approach.

5.6.1 Extension of Probability Function: From Rhombuses to Hexagons

In a general case, random variable X or Y in (5.1) can be a function of other independent random variables, e.g., $X = f(X_1, X_2)$ and $Y = f(Y_1, Y_2)$. Therefore, the

probability function $F_\omega = P\left(Z(X,Y) \leq z | X = x, Y = y\right)$ and the probability density $f_\omega(X,Y) = f_{X,Y}(x,y)$ can both be complicated functions of multiple random variables. E.g., for hexagonal geometries, F_ω is a function of three random variables. The next chapter is dedicated to the derivation of random distances associated with regular hexagons, where the product formulation in (5.1) goes to higher dimensions.

5.6.2 Extension of Probability Density: Nonuniform Point Distribution

Throughout this chapter, it is assumed that points are uniformly distributed in a geometry. In (5.1), by separating the probability function F_ω with the probability density f_ω, the same approach can be extended to more general, nonuniform point distribution patterns. Consequently, (5.2) and the derivation thereafter can be applied. Depending on the point distribution, however, the final result may not have an explicit form.

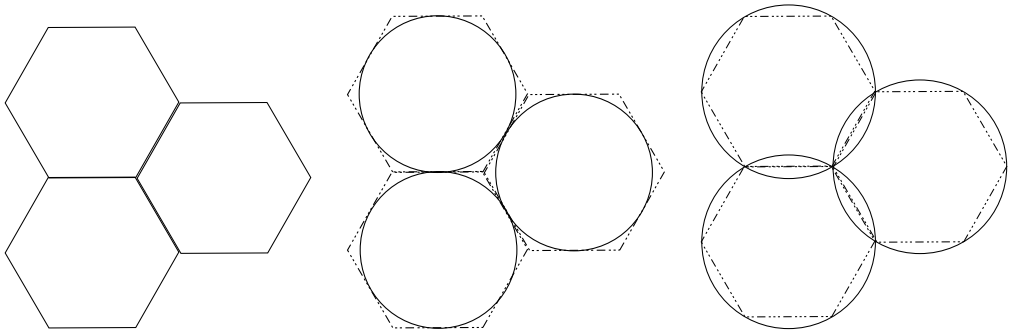
For example, [70] gives the solution to the distribution of the squared Euclidean distance between two bivariate normal vectors, where X and Y follow a bivariate normal distribution with mean μ and covariance matrix Σ. However, the results are in the form of infinite series expansions, which makes the application of these results difficult. Nevertheless, the case when point distribution follows bivariate normal or bivariate beta distribution has realistic applications in communication networks. Developing models for these nonuniform point distribution is challenging, especially in a network with a finite size. This problem deserves further investigation and thus is left as the future work of this dissertation.

Chapter 6

Random Distances Associated with Hexagons

Hexagons are magnificent geometries that commonly appear in nature, such as in the fields of biology, chemistry, astronomy, etc. The classic example is the honeycomb created by bees, who genetically build their cells in a hexagonal shape. The tessellated hexagons give the maximum amount of storage space per unit of wax that is around the cell. In field planting, [64] discovered that hexagons also yield the most balanced number of replications of plants under identical conditions of spatial competition between different plant species. These natural appearances inspired communication engineers the design of the honeycomb mesh networks, e.g., the one in [91] that is also based on the hexagonal plane tessellation. The most successful example is the currently deployed cellular communication networks.

In this chapter, a fundamental property of hexagons is investigated, i.e., the distributions of random distances between two endpoints, associated with a regular hexagon and two adjacent hexagons sharing a side. The results are derived for the challenging case when both endpoints are random. These results are presented for the first time in the literature, which are different from the existing results of elementary geometries that can be found in [67]. By dividing a hexagon into adjacent, congruent rhombuses, the distance distributions in the complex geometries are obtained by an extension of the approach in Chapter 5. Furthermore, the analytical models based on hexagon distributions are applied to the analysis of the nearest neighbor distribution in a sparse network for improving energy efficiency, and the farthest neighbor distribution in a dense network for minimizing routing overhead. Both the analytical and



(a) Hexagonal Cells. (b) Inscribed Circles. (c) Enclosing Circles.

Figure 6.1: Hexagonal Cell Layout and Circular Approximations.

simulation results demonstrate the high accuracy and promising potentials of this approach, whereas the current best approximations are not applicable in many scenarios. By utilizing the geometrical probability approach in the analysis of location-critical performance metrics, network protocols can benefit from the distance information prior to deployment.

6.1 Cell Shapes and Location-Critical Performance Metrics

From the geometrical probability point of view, a critical parameter of cellular networks is the shape of the cells. In practical networks, cells have irregular and complicated shapes because of the natural or man-made terrain, and user population. In analytical models, cells are represented as hexagons. In the early years of networking research, the honeycomb mesh networks is based on the hexagonal plane tessellation, inspired by the honeycomb created by bees. For bees, the tessellated hexagons give the maximum amount of storage space per unit of wax that is around the cell. In networking research, any given point inside a hexagon is closer to the center of the hexagon than any given point in an equal-area square or triangle [4]. The hexagonal representation, as shown in Figure 6.1(a), is frequently employed in the planning and analysis of wireless systems because of its flexibility and full network coverage.

In the current literature, however, the circular cells are often used as the approximations of hexagons, due to the reduced computational complexity. Examples are the inscribed and enclosing circles, as shown in Figure 6.1(b) and (c) respectively. However, the drawback of the circular cells is that, gaps must exist between cells

in order to avoid overlapping, such as Figure 6.1(b); or cells must partially overlap and include areas not belonging to the cell of interest, e.g., the enclosing circles in Figure 6.1(c).

In a wireless communication network, consider the transmission power, or energy consumption in cellular systems, the inscribed circles have a higher probability of assigning a low transmission power to the wireless devices. The initial network planning using inscribed circles thus cannot meet the requirement of covering all users in the cell area. In the other scenario where the enclosing circles are used, there is a higher probability of having a higher transmission power, i.e., using extra power to cover users in the overlapping area that does not belong to the cell of interest. Therefore, the wireless network resources are not efficiently utilized and transmission power is wasted, although with a lower computational complexity in the analytical model.

In all the existing works, no results exist for the random distances associated with hexagons in closed form. The development of new models and approaches for the derivation of random distances associated with hexagons thus has become an imminent research problem.

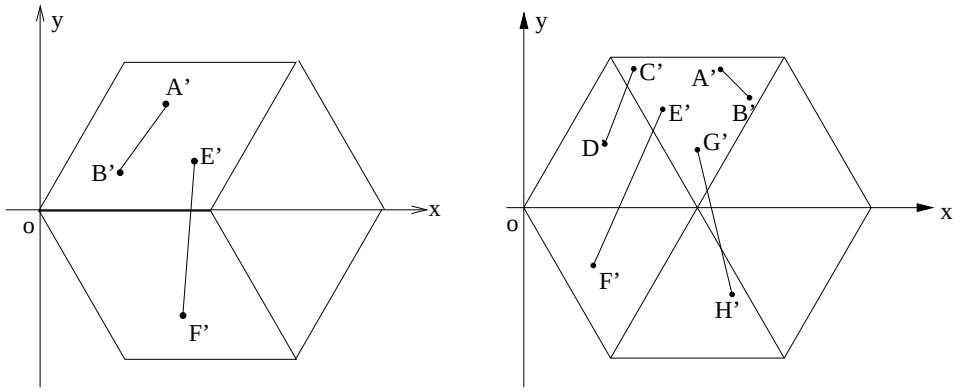
6.2 Distance Distributions Associated with Regular Hexagons

The distance distributions associated with rhombuses derived in Section 5.3 cover the basic cases when rhombuses are adjacent and similarly oriented. As mentioned at the beginning of this chapter, rhombuses are also the building blocks for hexagons. However, there are cases where the rhombuses in hexagons are having different orientations, or not adjacent. Such geometrical challenges are solved in this section, by another extension of the product formulation.

6.2.1 Distance Distribution within a Regular Hexagon

(1) Hexagon Decomposition

As shown in Figure 6.2(a), a regular hexagon can be decomposed into three congruent, adjacent rhombuses. Because of symmetry, random points are equally likely to fall in any one of these three rhombuses. Given the location of one endpoint of a random link, then the second endpoint falls in the same rhombus as the first point with



(a) Rhombus Decomposition (b) Triangle Decomposition

Figure 6.2: Relationship between a Hexagon and (a) Rhombuses (b) Triangles

probability $\frac{1}{3}$ (such as A′B′), and with probability $\frac{2}{3}$ falls in one of the two adjacent rhombuses (such as E′F′). Note that rhombus is not the only way to decompose a regular hexagon. As shown in Figure 6.2(b), a hexagon can be divided into six congruent equilateral triangles. However, to recover a hexagon from six triangles are more complicated than from three rhombuses. Four different cases need to be taken into consideration as in Figure 6.2(b): within a triangle (A′B′), between adjacent triangles sharing a side (C′D′), between two triangles 60 degrees apart (E′F′) and two diagonal triangles (G′H′). Using rhombus decomposition, only two cases are needed as shown in Figure 6.2(a).

All the three adjacent rhombuses inside a hexagon are equivalent, therefore we only consider two cases as shown in Figure 6.2, i.e., i) point A′ and B′ that are located in the same rhombus; ii) point E′ and F′ that are located in the two adjacent rhombuses (parallel but flipped with respect to $\mathcal{X}$-axis). The distribution of $|A'B'|$ has been derived in (5.9) in Section 5.3, by the observed relationship between the point coordinates before and after the affine transformation according to Figure 5.2(a). In this section, $|E'F'|$ and other variants associated with rhombuses will be derived.

(2) The Extended Product Formulation

We need to derive the distribution of $|E'F'|$ in order to get the result for the distance distribution within a regular hexagon. Following the same technique as that in (5.4), and denote the coordinates as $\mathrm{E}'(X_1', Y_1')$ and $\mathrm{F}'(X_2', Y_2')$, we have $X_1' = X_1 + \frac{Y_1}{2}$, $Y_1' = \frac{\sqrt{3}}{2}Y_1$, and $X_2' = X_2 + \frac{Y_2}{2}$, $Y_2' = -\frac{\sqrt{3}}{2}Y_2$, where X_1, Y_1, X_2 and $Y_2 \sim U[0,1]$, i.e., distributed uniformly in a unit square. Then we have the squared Euclidean distance

$Z = |E'F'|^2$ as

$$\begin{aligned} Z &= (X_1' - X_2')^2 + (Y_1' - Y_2')^2 \\ &= \left[(X_1 - X_2) + \frac{1}{2}(Y_1 - Y_2)\right]^2 + \left[\frac{\sqrt{3}}{2}(Y_1 + Y_2)\right]^2 \\ &= (X_1 - X_2)^2 + (X_1 - X_2)(Y_1 - Y_2) + Y_1^2 + Y_1Y_2 + Y_2^2. \end{aligned}$$

Let $X = X_1 - X_2$, we have

$$Z = Z(X, Y) = X^2 + X(Y_1 - Y_2) + Y_1^2 + Y_1Y_2 + Y_2^2. \tag{6.1}$$

Rewrite the RHS of (6.1), we get $Y_1^2+Y_2^2+X^2+Y_1Y_2+Y_1X-Y_2X$, which satisfies the general form of a non-degenerate quadratic surface [99]. However, (6.1) cannot be solved as easily as $|A'B'|^2$, because there are three independent random variables instead of two as in (5.5). By the product formulated in (5.1), the distribution of Z becomes a triple integral

$$\begin{aligned} F_Z(z) &= P_\Omega(Z \le z) = P\left(X^2 + X(Y_1 - Y_2) + Y_1^2 + Y_1Y_2 + Y_2^2 \le z\right) \\ &= \iiint P\left(x^2 + x(y_1 - y_2) + y_1^2 + y_1y_2 + y_2^2 \le z | X = x, Y = y\right) \\ &\quad f_{X,Y_1,Y_2}(x, y_1, y_2)dxdy_1dy_2, \end{aligned} \tag{6.2}$$

where the conditional probability function

$$F_\omega = P\left(x^2 + x(y_1 - y_2) + y_1^2 + y_1y_2 + y_2^2 \le z | X = x, Y = y\right)$$

determines a different constraint of the geometry. $f_\omega = f_{X,Y_1,Y_2}(x, y_1, y_2)$ is the joint density of random variables X, Y_1, Y_2. Therefore, the integral of the joint density function needs to go to the third dimension. Assuming that points are uniformly distributed, then

$$f_X(x) = \begin{cases} 1 + x & -1 \le x \le 0 \\ 1 - x & 0 \le x \le 1 \end{cases}, \quad \text{and} \quad f_{Y_1}(y) = f_{Y_2}(y) = \begin{cases} 1 & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}. \tag{6.3}$$

Because these three random variables, X, Y_1 and Y_2, are all independent, we have $f_\omega = f_{X,Y_1,Y_2}(x, y_1, y_2) = f_X(x)f_{Y_1}(y)f_{Y_2}(y)$.

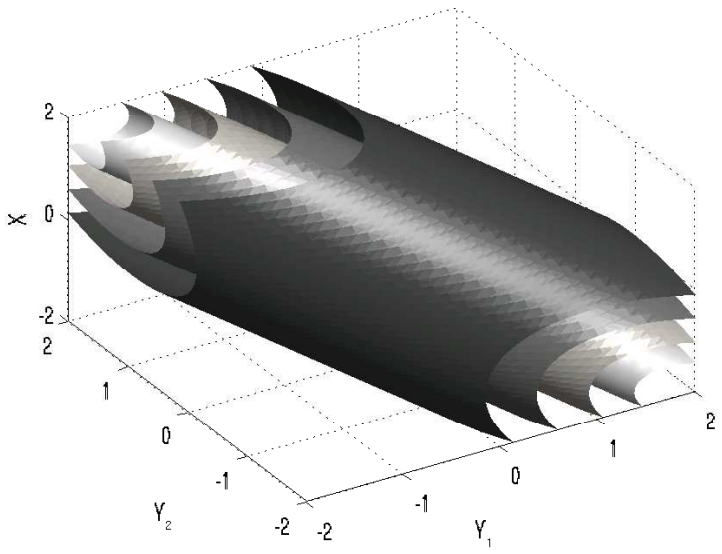


Figure 6.3: Geometric Interpretation of $Z = X^2 + X(Y_1 - Y_2) + Y_1^2 + Y_1Y_2 + Y_2^2$.

(3) The Geometric Interpretation

The geometric interpretation of $X^2 + X(Y_1 - Y_2) + Y_1^2 + Y_1Y_2 + Y_2^2$ with different values, or the *isosurfaces* of Z in (6.1), are concentric cylinders as shown in Figure 6.3, which is like a Swiss roll growing thicker as the value of Z increases. If the effective regions of f_ω are considered, then the Swiss roll is cutoff at the boundaries of $f_X(x)$, $f_{Y_1}(y)$ and $f_{Y_2}(y)$, because the density function f_ω is zero outside these boundaries. Consider $|E'F'|$ in Figure 6.2, we have $\mathcal{X} \in [-1, 1]$ and $\mathcal{Y}_1, \mathcal{Y}_2 \in [0, 1]$ according to (6.3). Therefore, in (6.2), the four-dimensional geometric shape of Z is projected onto the effective region of f_ω on $[0, 1] \times [0, 1] \times [-1, 1]$, in a similar way as the three-dimensional bowl projected to the two-dimensional plane.

Consider an *effective surface* as the part of the concentric cylinder that is inside the effective region of f_ω. As the value of Z increases, the isosurfaces of the Swiss roll are only effective inside the volume of $1 \times 1 \times 2$. These results in different cylindrical surfaces and cut off facets as shown in Figure 6.4–6.7.

$0 \leq z \leq \frac{3}{4}$:

In this sub-case, the effective surface of the cylinder is inside the boundary of f_ω, as shown in Figure 6.4. By symmetry, the distance distribution in this case is

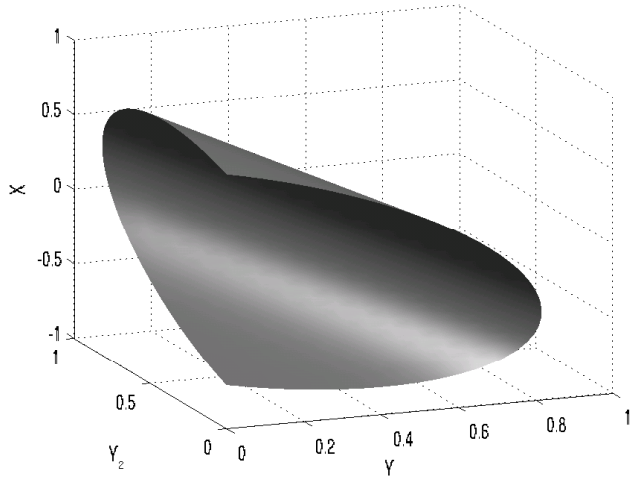


Figure 6.4: Sub-case when $0 \leq z \leq \frac{3}{4}$.

$$
\begin{aligned}
F(z) &= \int_{-2\sqrt{\frac{z}{3}}}^{0} (1+x)dx \iint\limits_{D(y_1,y_2)} dy_1 dy_2 + \int_{0}^{2\sqrt{\frac{z}{3}}} (1-x)dx \iint\limits_{D(y_1,y_2)} dy_1 dy_2 \\
&= 2\int_{0}^{2\sqrt{\frac{z}{3}}} (1-x)dx \iint\limits_{D(y_1,y_2)} dy_1 dy_2 \\
&= 2\left[\int_{0}^{\sqrt{z}} (1-x)dx \int_{0}^{\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} dy_2 \int_{0}^{-\frac{x+y_2}{2}+\sqrt{z-\frac{3}{4}(x-y_2)^2}} dy_1 \right. \\
&\quad \left. + \int_{\sqrt{z}}^{2\sqrt{\frac{z}{3}}} (1-x)dx \int_{\frac{x}{2}-\sqrt{z-\frac{3}{4}x^2}}^{\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} dy_2 \int_{0}^{-\frac{x+y_2}{2}+\sqrt{z-\frac{3}{4}(x-y_2)^2}} dy_1 \right] \\
&= \frac{8}{9}z^{3/2} - \left(\frac{1}{6} + \frac{\pi}{9\sqrt{3}}\right) z^2.
\end{aligned}
$$

Note that the greatest value of X, when either Y_1 or Y_2 equals 0, is $2\sqrt{\frac{z}{3}}$, rather than $\sqrt{z}$ when both Y_1 or Y_2 are 0. The corresponding PDF is

$$f(z) = \frac{4}{3}\sqrt{z} - \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right) z.$$

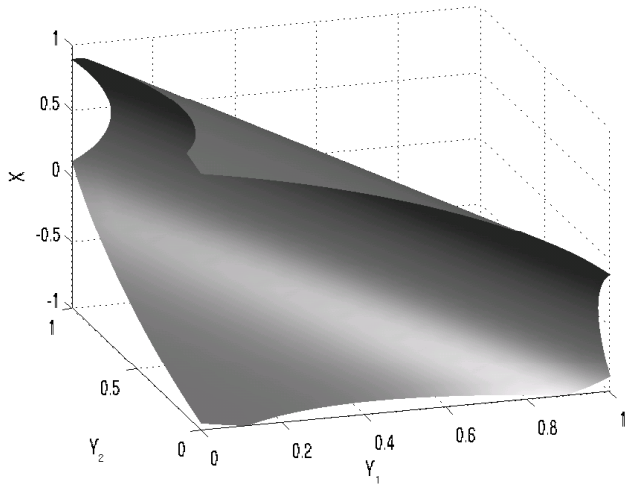


Figure 6.5: Sub-case when $\frac{3}{4} \leq z \leq 1$.

$\frac{3}{4} \leq z \leq 1$:

There are cutoffs on the cylinder at the boundaries, when $\mathcal{Y}_1 = 1$, $\mathcal{Y}_2 = 1$ and $\mathcal{X} = 1$, as shown in Figure 6.5. Because of symmetry, we look at the half plane when $0 \leq \mathcal{X} \leq 1$. When $\mathcal{Y}_1 = 0$ and $\mathcal{Y}_2 = 1$, the cutoffs on the $\mathcal{X}$-axis are $(0, 1, \frac{1}{2} \pm \sqrt{z - \frac{3}{4}})$. Meanwhile, when $\mathcal{Y}_1 = \mathcal{Y}_2 = 0$, there is also a cutoff on the $\mathcal{X}$-axis at $(0, 0, \sqrt{z})$. Given these cutoff points, the distribution function can be calculated as

$$
\begin{aligned}
F(z) &= 2\left[\int_0^{\frac{1}{2}-\sqrt{z-\frac{3}{4}}} (1-x)dx \int_0^{\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} dy_2 \int_0^{-\frac{x+y_2}{2}+\sqrt{z-\frac{3}{4}(x-y_2)^2}} dy_1 \right. \\
&\quad + \int_{\frac{1}{2}-\sqrt{z-\frac{3}{4}}}^{\frac{1}{2}+\sqrt{z-\frac{3}{4}}} (1-x)dx \int_0^{1} dy_2 \int_0^{-\frac{x+y_2}{2}+\sqrt{z-\frac{3}{4}(x-y_2)^2}} dy_1 \\
&\quad + \int_{\frac{1}{2}+\sqrt{z-\frac{3}{4}}}^{\sqrt{z}} (1-x)dx \int_0^{\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} dy_2 \int_0^{-\frac{x+y_2}{2}+\sqrt{z-\frac{3}{4}(x-y_2)^2}} dy_1 \\
&\quad \left. + \int_{\sqrt{z}}^{1} (1-x)dx \int_{\frac{x}{2}-\sqrt{z-\frac{3}{4}x^2}}^{\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} dy_2 \int_0^{-\frac{x+y_2}{2}+\sqrt{z-\frac{3}{4}(x-y_2)^2}} dy_1 \right] \\
&= -2\left(\frac{z^2}{3\sqrt{3}} + \frac{2z}{\sqrt{3}}\right) \sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{2\pi}{9\sqrt{3}} - \frac{1}{6}\right) z^2 + \frac{8}{9} z^{3/2} + \frac{2\pi}{\sqrt{3}} z - \frac{7}{6} z\sqrt{4z-3} \\
&\quad - \frac{\sqrt{4z-3}}{4},
\end{aligned}
$$

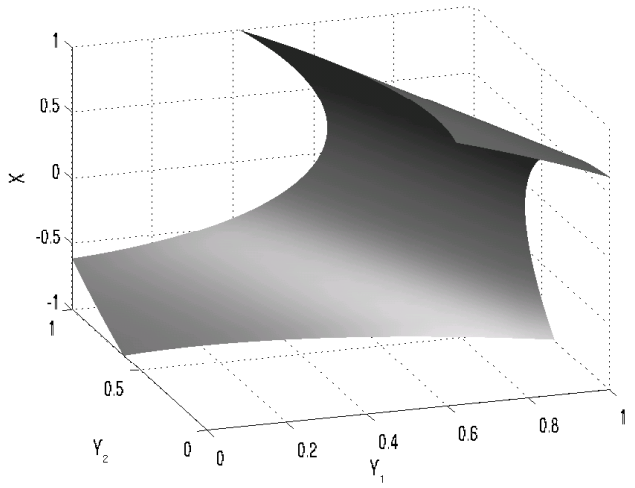


Figure 6.6: Sub-case when $1 \leq z \leq 3$.

and the PDF is

$$f(z) = -\frac{4}{\sqrt{3}}\left(\frac{z}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{4\pi}{9\sqrt{3}} - \frac{1}{3}\right)z + \frac{4}{3}\sqrt{z} - \frac{5}{3}\sqrt{4z-3} + \frac{2\pi}{\sqrt{3}}.$$

$1 \leq z \leq 3$:

When $\mathcal{Y}_1 = 1$ and $\mathcal{Y}_2 = 0$, the cutoff on the $\mathcal{X}$-axis is at $(1, 0, -\frac{1}{2} + \sqrt{z - \frac{3}{4}})$ when $0 \leq \mathcal{X} \leq 1$, as shown in Figure 6.6. Given this cutoff point, the distribution function can be calculated by subtracting the part outside the boundaries, i.e.,

$$\begin{aligned} F(z) &= 2\left[\int_0^1 (1-x)dx \int_0^1 dy_2 \int_0^{-\frac{x+y_2}{2}+\sqrt{z-\frac{3}{4}(x-y_2)^2}} dy_1 \right. \\ &\quad \left. - \int_0^{-\frac{1}{2}+\sqrt{z-\frac{3}{4}}} (1-x)dx \int_1^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} dy_1 \int_0^{\frac{x-y_1}{2}+\sqrt{z-\frac{3}{4}(x+y_1)^2}} dy_2 \right] \\ &= -\frac{z}{\sqrt{3}}\left(\frac{z}{3}+4\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{1}{6}+\frac{\pi}{9\sqrt{3}}\right)z^2 + \left(\frac{1}{2}+\frac{2\pi}{\sqrt{3}}\right)z - \frac{37}{36}z\sqrt{4z-3} \\ &\quad -\frac{7}{24}\sqrt{4z-3} - \frac{25}{24}. \end{aligned}$$

The PDF in this sub-case is

$$f(z) = -\frac{2}{\sqrt{3}}\left(\frac{z}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)z - \frac{3}{2}\sqrt{4z-3} + \frac{2\pi}{\sqrt{3}} + \frac{1}{2}.$$

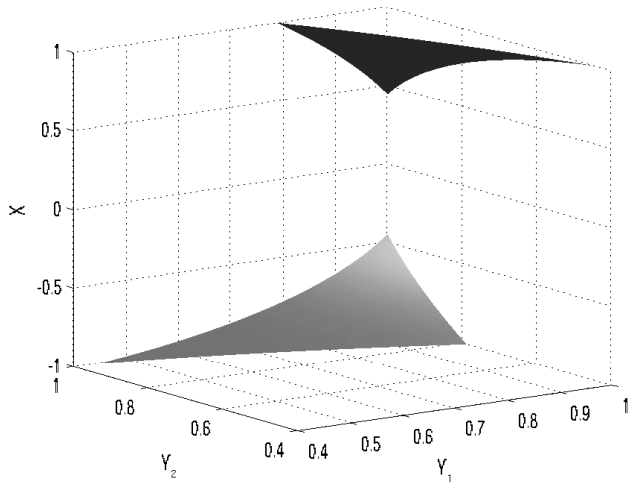


Figure 6.7: Sub-case when $3 \leq z \leq 4$.

$3 \leq z \leq 4$:

In this sub-case, it is easier to subtract the volume that is outside the cylinder but inside the boundaries, as shown in Figure 6.7. The cutoff on the $\mathcal{X}$-axis happens when $\mathcal{Y}_1 = \mathcal{Y}_2 = 1$, at $(1, 1, \pm\sqrt{z-3})$. The distribution function in this sub-case is thus

$$\begin{aligned} F(z) &= 1 - 2\int_{\sqrt{z-3}}^{1}(1-x)dx \int_{-\frac{x+1}{2}+\sqrt{z-\frac{3}{4}(x-1)^2}}^{1} dy_1 \int_{\frac{x-y_1}{2}+\sqrt{z-\frac{3}{4}(x+y_1)^2}}^{1} dy_2 \\ &= \left(\frac{z^2}{3\sqrt{3}} + \frac{8z}{\sqrt{3}}\right)\sin^{-1}\sqrt{\frac{3}{z}} - \left(\frac{1}{6} + \frac{\pi}{9\sqrt{3}}\right)z^2 - \left(2 + \frac{8\pi}{3\sqrt{3}}\right)z \\ &\quad + \frac{7}{3}z\sqrt{z-3} + 2\sqrt{z-3} + \frac{1}{3}. \end{aligned}$$

The corresponding PDF is

$$f(z) = \frac{2}{\sqrt{3}}\left(\frac{z}{3}+4\right)\sin^{-1}\sqrt{\frac{3}{z}} - \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right)z + \frac{10}{3}\sqrt{z-3} - \frac{8\pi}{3\sqrt{3}} - 2.$$

Combining the above four cases and using $f_D(d) = F'_Z(d^2) = 2df_Z(d^2)$, we have

Result 5 (Random distance between two parallel unit rhombuses—flipped with respect to $\mathcal{X}$-axis)**.** The probability density function of the random distances between uniformly distributed points, each inside one of the two parallel unit rhombuses flipped

with respect to $\mathcal{X}$-axis, is

$$f_{D_{\mathrm{PX}}}(d) = 2d \begin{cases} \frac{4}{3}d - \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right) d^2 & 0 \le d \le \frac{\sqrt{3}}{2} \\ -\frac{4}{\sqrt{3}}\left(\frac{d^2}{3} + 1\right) \sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{4\pi}{9\sqrt{3}} - \frac{1}{3}\right) d^2 + \frac{4}{3}d & \\ \quad - \frac{5}{3}\sqrt{4d^2 - 3} + \frac{2\pi}{\sqrt{3}} & \frac{\sqrt{3}}{2} \le d \le 1 \\ -\frac{2}{\sqrt{3}}\left(\frac{d^2}{3} + 2\right) \sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{2\pi}{9\sqrt{3}} + \frac{1}{3}\right) d^2 - \frac{3}{2}\sqrt{4d^2 - 3} & \\ \quad + \frac{2\pi}{\sqrt{3}} + \frac{1}{2} & 1 \le d \le \sqrt{3} \\ \frac{2}{\sqrt{3}}\left(\frac{d^2}{3} + 4\right) \sin^{-1}\frac{\sqrt{3}}{d} - \left(\frac{2\pi}{9\sqrt{3}} + \frac{1}{3}\right) d^2 + \frac{10}{3}\sqrt{d^2 - 3} & \\ \quad - \frac{8\pi}{3\sqrt{3}} - 2 & \sqrt{3} \le d \le 2 \\ 0 & \text{otherwise} \end{cases} \tag{6.4}$$

Distance Distribution within a Regular Hexagon

The decomposition of a hexagon into three adjacent, congruent rhombuses is shown in Figure 6.2 at the beginning of this section. The probability density function of the distance between two random points inside a hexagon, is $\frac{1}{3}$ of the density function of random distances within the same rhombus in (5.9), and $\frac{2}{3}$ of the density function just derived in (6.4) (with probability $\frac{1}{3}$, two random points are located in any of the two adjacent rhombuses. Because of symmetry, $\frac{2}{3}$ is the probability for (6.4) to occur). By defining a *unit hexagon* as the hexagon with the length of each side is 1, this probability density function is presented in (6.5).

Result 6 (Random distance inside a regular unit hexagon)**.** The probability density function of the random distances between uniformly distributed points inside a regular unit hexagon is

$$f_{D_{\mathrm{H_I}}}(d) = \frac{2}{3}d \begin{cases} \left(\frac{2}{3} - \frac{2\pi}{9\sqrt{3}}\right) d^2 - \frac{8}{3}d + \frac{2\pi}{\sqrt{3}} & 0 \le d \le 1 \\ -\frac{4}{\sqrt{3}}\left(\frac{2d^2}{3} + 1\right) \sin^{-1}\frac{\sqrt{3}}{2d} + \frac{2\pi}{3\sqrt{3}}d^2 - 2\sqrt{4d^2 - 3} + \frac{10\pi}{3\sqrt{3}} & 1 \le d \le \sqrt{3} \\ \frac{4}{\sqrt{3}}\left(\frac{d^2}{3} + 4\right) \sin^{-1}\frac{\sqrt{3}}{d} - \left(\frac{4\pi}{9\sqrt{3}} + \frac{2}{3}\right) d^2 + \frac{20}{3}\sqrt{d^2 - 3} & \\ \quad - \frac{16\pi}{3\sqrt{3}} - 4 & \sqrt{3} \le d \le 2 \\ 0 & \text{otherwise} \end{cases} \tag{6.5}$$

It is interesting to see that the cases where $0 \le d \le \frac{\sqrt{3}}{2}$ and $\frac{\sqrt{3}}{2} \le d \le 1$ are exactly the same after calculating the probabilistic sum of the two density functions, therefore

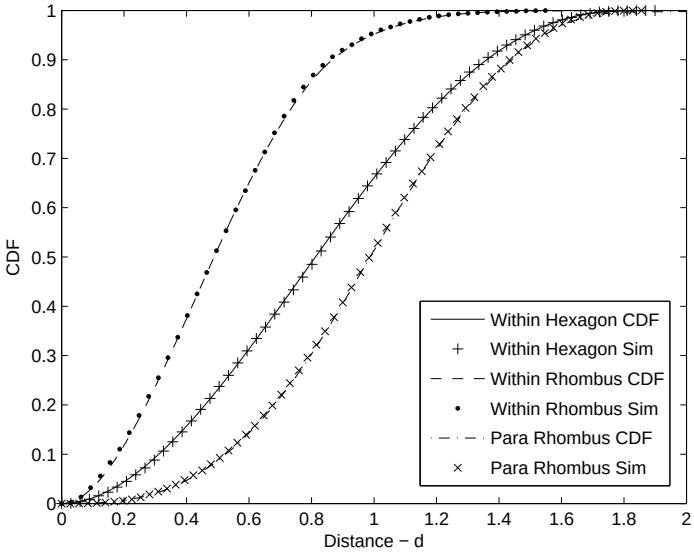


Figure 6.8: Cumulative Distribution Functions and Simulation Results for Random Distances within a Hexagon.

they are merged into one case. At any transitional value of D between different sub-cases, $f_{D_{\mathrm{H_I}}}(d)$ has the same values in both adjacent sub-cases. Figure 6.8 shows the CDFs of these three cases derived in (5.9), (6.4) and (6.5), and the verification through simulation, by generating $1,000$ pairs of random points with the geometric locations as shown in Figure 6.2. The results have demonstrated that our distribution functions are very accurate compared to the simulation.

6.2.2 Distance Distribution between Adjacent Regular Hexagons

Given two unit hexagons that are adjacent to each other, as shown in Figure 6.9, they can be decomposed into six congruent rhombuses in a similar way as that in Figure 6.2. If the two endpoints of a random link fall into each one of the adjacent hexagons, then each point is equally likely to fall in any one of the three adjacent rhombuses in a single hexagon, resulting in nine (3×3) possible combinations. However, by symmetry and distance invariance under rotation, many of these combinations are equivalent. We use letters A through J to denote different rhombuses in Figure 6.9.

With a slight abuse of notation, we also use $|AB|$, etc, to denote the distance between random points, one being located in rhombus A and the other inside B.

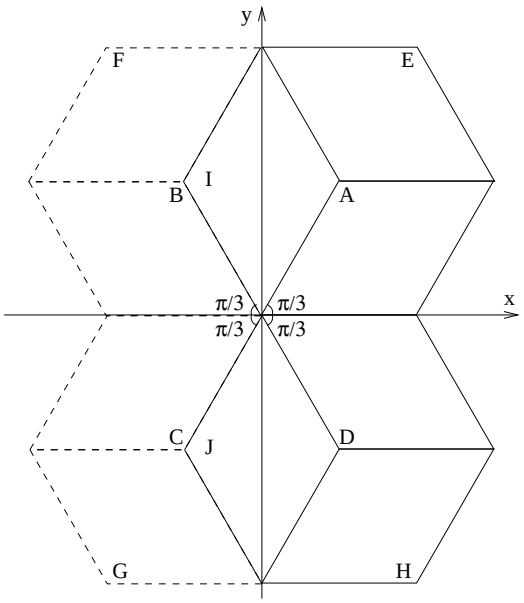


Figure 6.9: Random Points between Two Adjacent Hexagons: Different Cases with Rhombuses.

Thus we have the following six equivalent cases:

$$\begin{array}{lll} |AD| = |AI| & |AC| = |IJ| & |AB| = |DI| \\ |DE| = |AF| & |JE| = |DF| & |HE| \end{array} \tag{6.6}$$

The first two cases, $|AD| = |AI|$ and $|AC| = |IJ|$ have been derived in previous sections. Without loss of generality, we consider the upper right hexagon as the one where the first endpoint is located, i.e., in rhombus A, E or I, and the second endpoint can be in any one of rhombus D, H or J that is in the adjacent hexagon. Then the distance distribution between an arbitrary pair of endpoints, one in each of the adjacent hexagons, is $\frac{1}{9}(|AD| + |HE| + |IJ|) + \frac{2}{9}(|DE| + |DI| + |JE|)$, a statistical probabilistic sum among 3×3 different geometric locations.

Among the last four cases in (6.6), $|DE| = |AF|$ is the case where two rhombuses have the same orientation. Therefore by the same reasoning as in the rhombus geometry, the corresponding result should be obtained by a double integral of (5.5). In the rest of the three cases, $|AB| = |DI|$, $|JE| = |DF|$ and $|HE|$, the two rhombuses have different orientations, either by flipping or rotation. In the following, we give the derivation results of the four cases in (6.6) that have not yet been derived. At the end we combine them by a probabilistic sum, in order to obtain the distribution for the random distances between two adjacent hexagons sharing a side.

(1) $|DE| = |AF|$

In this case, the two rhombuses, e.g., D and E, are shifted by distance $\sqrt{3}$ apart.

$$f_{D_1}(d) = 2d \begin{cases} -\frac{2d^2}{3\sqrt{3}} \sin^{-1} \frac{\sqrt{3}}{2d} + \frac{\pi}{3\sqrt{3}} d^2 - \frac{\sqrt{4d^2-3}}{6} & \frac{\sqrt{3}}{2} \le d \le 1 \\ \frac{d^2+4}{\sqrt{3}} \sin^{-1} \frac{\sqrt{3}}{2d} - \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right) d^2 - \frac{2}{3}d + \frac{19}{12}\sqrt{4d^2-3} - \frac{4\pi}{3\sqrt{3}} - \frac{3}{4} & 1 \le d \le \sqrt{3} \\ -\left(\frac{d^2}{3\sqrt{3}} + \frac{10}{\sqrt{3}}\right) \sin^{-1} \frac{\sqrt{3}}{d} - \left(\frac{2d^2}{3\sqrt{3}} + 2\sqrt{3}\right) \sin^{-1} \frac{\sqrt{3}}{2d} - \frac{2}{3}d \\ \quad + \left(\frac{4}{3} + \frac{2\pi}{9\sqrt{3}}\right) d^2 - \frac{13}{6}\sqrt{4d^2-3} - \frac{11}{3}\sqrt{d^2-3} + \frac{16\pi}{3\sqrt{3}} + \frac{11}{2} & \sqrt{3} \le d \le 2 \\ \left(\frac{d^2}{3\sqrt{3}} + 2\sqrt{3}\right) \sin^{-1} \frac{\sqrt{3}}{d} - \left(\frac{2d^2}{3\sqrt{3}} + 2\sqrt{3}\right) \sin^{-1} \frac{\sqrt{3}}{2d} \\ \quad - \frac{13}{6}\sqrt{4d^2-3} + \frac{7}{3}\sqrt{d^2-3} + 2d - \frac{1}{2} & 2 \le d \le \frac{3\sqrt{3}}{2} \\ \left(\frac{2d^2}{3\sqrt{3}} + 4\sqrt{3}\right) \sin^{-1} \frac{3\sqrt{3}}{2d} + \left(\frac{d^2}{3\sqrt{3}} + 2\sqrt{3}\right) \sin^{-1} \frac{\sqrt{3}}{d} \\ \quad - \left(\frac{2d^2}{3\sqrt{3}} + 2\sqrt{3}\right) \sin^{-1} \frac{\sqrt{3}}{2d} - \frac{\pi}{3\sqrt{3}} d^2 - \frac{13}{6}\sqrt{4d^2-3} & \frac{3\sqrt{3}}{2} \le d \le \sqrt{7} \\ \quad + \frac{7}{3}\sqrt{d^2-3} + \frac{11}{6}\sqrt{4d^2-27} + 2d - 2\sqrt{3}\pi - \frac{1}{2} \\ -\frac{d^2}{3\sqrt{3}} \sin^{-1} \frac{3\sqrt{3}}{2d} + \left(\frac{\pi}{9\sqrt{3}} - \frac{1}{3}\right) d^2 + 2d - \frac{\sqrt{4d^2-27}}{4} - \frac{9}{4} & \sqrt{7} \le d \le 3 \\ 0 & \text{otherwise} \end{cases} \tag{6.7}$$

(2) $|AB| = |DI|$

The two rhombuses in this case, e.g., A and B, are flipped with respect to the $\mathcal{Y}$-axis.

$$f_{D_2}(d) = 2d \begin{cases} \left(\frac{2\pi}{9\sqrt{3}} - \frac{1}{6}\right) d^2 & 0 \le d \le \frac{\sqrt{3}}{2} \\ \frac{2}{\sqrt{3}} \left(\frac{2d^2}{3} + 1\right) \sin^{-1} \frac{\sqrt{3}}{2d} - \left(\frac{4\pi}{9\sqrt{3}} + \frac{1}{6}\right) d^2 + \sqrt{4d^2-3} - \frac{\pi}{\sqrt{3}} & \frac{\sqrt{3}}{2} \le d \le 1 \\ \frac{1}{\sqrt{3}} \left(\frac{4d^2}{3} + 1\right) \sin^{-1} \frac{\sqrt{3}}{2d} + \left(\frac{1}{3} - \frac{4\pi}{9\sqrt{3}}\right) d^2 + \frac{2}{3}\sqrt{4d^2-3} - \frac{2}{3}d \\ \quad - \frac{2\pi}{3\sqrt{3}} + \frac{1}{2} & 1 \le d \le \sqrt{3} \\ \frac{5}{\sqrt{3}} \sin^{-1} \frac{\sqrt{3}}{2d} - \frac{4}{\sqrt{3}} \left(\frac{d^2}{3} + 2\right) \sin^{-1} \frac{\sqrt{3}}{d} + \left(\frac{4\pi}{9\sqrt{3}} - \frac{1}{3}\right) d^2 \\ \quad + \frac{5}{3}\sqrt{4d^2-3} - 4\sqrt{d^2-3} - \frac{2}{3}d + \frac{8\pi}{3\sqrt{3}} - \frac{1}{2} & \sqrt{3} \le d \le 2 \\ \frac{5}{\sqrt{3}} \sin^{-1} \frac{\sqrt{3}}{2d} - \frac{2}{\sqrt{3}} \left(\frac{d^2}{3} + 2\right) \sin^{-1} \frac{\sqrt{3}}{d} + \left(\frac{2\pi}{9\sqrt{3}} - \frac{1}{6}\right) d^2 \\ \quad - + \frac{5}{3}\sqrt{4d^2-32}\sqrt{d^2-3} - 2d + \frac{4\pi}{3\sqrt{3}} - \frac{1}{2} & 2 \le d \le \sqrt{7} \\ \left(\frac{2d^2}{3\sqrt{3}} + 3\sqrt{3}\right) \sin^{-1} \frac{3\sqrt{3}}{2d} + \left(\frac{1}{6} - \frac{2\pi}{9\sqrt{3}}\right) d^2 + \frac{3}{2}\sqrt{4d^2-27} \\ \quad - 2d - \sqrt{3}\pi & \sqrt{7} \le d \le 3 \\ 0 & \text{otherwise} \end{cases} \tag{6.8}$$

(3) $|JE| = |DF|$

In this case, the two rhombuses, e.g., D and F, are both flipped with respect to the $\mathcal{Y}$-axis, and shifted by $\sqrt{3}$ apart.

$$f_{D_3}(d) = 2d \begin{cases} -\left(\frac{d^2}{3\sqrt{3}} + \frac{1}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{1}{6} + \frac{\pi}{9\sqrt{3}}\right)d^2 - \frac{5}{12}\sqrt{4d^2-3} \\ \quad + \frac{\pi}{3\sqrt{3}} + \frac{1}{4} & 1 \le d \le \sqrt{3} \\ \left(\frac{2d^2}{3\sqrt{3}} + \frac{4}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{2}{\sqrt{3}} - \frac{d^2}{3\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right)d^2 \\ \quad + \frac{7}{12}\sqrt{4d^2-3} + 2\sqrt{d^2-3} - \frac{13\pi}{6\sqrt{3}} - \frac{5}{4} & \sqrt{3} \le d \le 2 \\ \frac{d^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{2}{\sqrt{3}} - \frac{d^2}{3\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{1}{6} + \frac{\pi}{9\sqrt{3}}\right)d^2 \\ \quad + \frac{7}{12}\sqrt{4d^2-3} + \frac{\sqrt{d^2-3}}{3} - \frac{5\pi}{6\sqrt{3}} - \frac{1}{4} & 2 \le d \le \frac{3\sqrt{3}}{2} \\ \left(\frac{2}{\sqrt{3}} - \frac{d^2}{3\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{2d^2}{3\sqrt{3}} + 3\sqrt{3}\right)\sin^{-1}\frac{3\sqrt{3}}{2d} \\ \quad \frac{d^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{2\pi}{9\sqrt{3}} - \frac{1}{6}\right)d^2 + \frac{7}{12}\sqrt{4d^2-3} + \frac{\sqrt{d^2-3}}{3} & \frac{3\sqrt{3}}{2} \le d \le \sqrt{7} \\ \quad - \frac{3}{2}\sqrt{4d^2-27} + \frac{11\pi}{3\sqrt{3}} - \frac{1}{4} \\ \left(\frac{d^2}{3\sqrt{3}} - \frac{4}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{d} - \left(\frac{d^2}{3\sqrt{3}} + \frac{1}{2\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} \\ \quad - \left(\frac{2d^2}{3\sqrt{3}} + \frac{7\sqrt{3}}{2}\right)\sin^{-1}\frac{3\sqrt{3}}{2d} + \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{\sqrt{4d^2-3}}{4} & \sqrt{7} \le d \le 3 \\ \quad - \sqrt{d^2-3} - \frac{5}{3}\sqrt{4d^2-27} + \frac{11\pi}{2\sqrt{3}} + \frac{13}{4} \\ \left(\frac{d^2}{3\sqrt{3}} - \frac{4}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{d} - \left(\frac{d^2}{3\sqrt{3}} + \frac{1}{2\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} \\ \quad + \frac{5\sqrt{3}}{2}\sin^{-1}\frac{3\sqrt{3}}{2d} - \frac{\sqrt{4d^2-3}}{4} - \sqrt{d^2-3} + \frac{5}{6}\sqrt{4d^2-27} & 3 \le d \le 2\sqrt{3} \\ \quad - \frac{\pi}{2\sqrt{3}} - \frac{5}{4} \\ \left(\frac{d^2}{3\sqrt{3}} + \frac{8}{\sqrt{3}}\right)\sin^{-1}\frac{2\sqrt{3}}{d} - \left(\frac{d^2}{3\sqrt{3}} + \frac{1}{2\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} \\ \quad + \frac{5\sqrt{3}}{2}\sin^{-1}\frac{3\sqrt{3}}{2d} - \left(\frac{1}{6} + \frac{\pi}{9\sqrt{3}}\right)d^2 + \frac{5}{6}\sqrt{4d^2-27} & 2\sqrt{3} \le d \le \sqrt{13} \\ \quad + 2\sqrt{d^2-12} - \frac{\sqrt{4d^2-3}}{4} - \frac{31\pi}{6\sqrt{3}} - \frac{9}{4} \\ 0 & \text{otherwise} \end{cases} \tag{6.9}$$

(4) $|HE|$

The two rhombuses are both flipped with respect to the $\mathcal{X}$-axis, and shifted by $\sqrt{3}$ apart.

$$f_{D_4}(d) = 2d \begin{cases} \frac{2}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \frac{4}{\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{d} - \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 & \\ \quad + \sqrt{4d^2-3} + \frac{4}{3}\sqrt{d^2-3} - \frac{7\pi}{3\sqrt{3}} - 2 & \sqrt{3} \le d \le 2 \\ \frac{2}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} & \\ \quad + \sqrt{4d^2-3} - 2\sqrt{d^2-3} + \frac{\pi}{3\sqrt{3}} & 2 \le d \le \frac{3\sqrt{3}}{2} \\ \frac{2}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} & \\ \quad - 2\sqrt{3}\sin^{-1}\frac{3\sqrt{3}}{2d} + \sqrt{4d^2-3} - 2\sqrt{d^2-3} - \frac{2}{3}\sqrt{4d^2-27} & \frac{3\sqrt{3}}{2} \le d \le \sqrt{7} \\ \quad + \frac{10\pi}{3\sqrt{3}} & \\ \frac{1}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} & \\ \quad - 3\sqrt{3}\sin^{-1}\frac{3\sqrt{3}}{2d} + \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 + \frac{\sqrt{4d^2-3}}{2} - 4\sqrt{d^2-3} & \sqrt{7} \le d \le 3 \\ \quad - \sqrt{4d^2-27} + \frac{17\pi}{3\sqrt{3}} + \frac{9}{2} & \\ \frac{1}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} & \\ \quad + \left(\frac{2d^2}{3\sqrt{3}}+3\sqrt{3}\right)\sin^{-1}\frac{3\sqrt{3}}{2d} + \frac{\sqrt{4d^2-3}}{2} - 4\sqrt{d^2-3} & 3 \le d \le 2\sqrt{3} \\ \quad + \frac{3}{2}\sqrt{4d^2-27} - \frac{\pi}{3\sqrt{3}} & \\ \frac{1}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{2d^2}{3\sqrt{3}}+3\sqrt{3}\right)\sin^{-1}\frac{3\sqrt{3}}{2d} & \\ \quad + \frac{8}{\sqrt{3}}\sin^{-1}\frac{2\sqrt{3}}{d} - \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 + \frac{\sqrt{4d^2-3}}{2} + \frac{3}{2}\sqrt{4d^2-27} & 2\sqrt{3} \le d \le \sqrt{13} \\ \quad + \frac{4}{3}\sqrt{d^2-12} - \frac{17\pi}{3\sqrt{3}} - 8 & \\ 0 & \text{otherwise} \end{cases} \tag{6.10}$$

Distance Distribution between Adjacent Regular Hexagons

When the two endpoints of a given link fall into one of the two adjacent hexagons sharing a side, the probability density function of the random distances between these two endpoints is $\frac{1}{9}\left[f_{D_{\mathrm{PX}}}(d) + f_{D_{\mathrm{LD}}}(d) + f_{D_4}(d)\right] + \frac{2}{9}\left[f_{D_1}(d) + f_{D_2}(d) + f_{D_3}(d)\right]$, using the similar reasoning as that in Section 6.2.1. With this probabilistic sum, we thus have the following:

Result 7 (Random distances between two unit hexagons)**.** The probability density function of the random Euclidean distances between two uniformly distributed points,

one in each of the two adjacent unit hexagons sharing a side, is

$$f_{D_{\mathrm{H_A}}}(d) = \frac{2}{9}d \begin{cases} \left(\frac{\pi}{9\sqrt{3}} - \frac{1}{3}\right) d^2 + \frac{4}{3}d & 0 \le d \le 1 \\ \frac{2}{\sqrt{3}}(d^2+2)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{1}{3} + \frac{5\pi}{9\sqrt{3}}\right) d^2 + \frac{11}{6}\sqrt{4d^2-3} & \\ \quad - \frac{4\pi}{3\sqrt{3}} - \frac{1}{2} & 1 \le d \le \sqrt{3} \\ \frac{2}{\sqrt{3}}\left(\frac{d^2}{3} - 2\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+4\right)\sin^{-1}\frac{\sqrt{3}}{d} & \\ \quad + \left(1 + \frac{\pi}{3\sqrt{3}}\right) d^2 - \frac{7}{6}\sqrt{4d^2-3} - \frac{20}{3}\sqrt{d^2-3} + \frac{8\pi}{\sqrt{3}} + \frac{9}{2} & \sqrt{3} \le d \le 2 \\ \frac{2}{\sqrt{3}}\left(\frac{d^2}{3} - 2\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{1}{3} - \frac{\pi}{9\sqrt{3}}\right) d^2 - \frac{7}{6}\sqrt{4d^2-3} & \\ \quad + \frac{8\pi}{3\sqrt{3}} + \frac{1}{2} & 2 \le d \le \sqrt{7} \\ -\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+6\right)\sin^{-1}\frac{3\sqrt{3}}{2d} - \frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} & \\ \quad + \left(\frac{1}{3} + \frac{5\pi}{9\sqrt{3}}\right) d^2 - 4\sqrt{d^2-3} - \frac{11}{6}\sqrt{4d^2-27} + \frac{28\pi}{3\sqrt{3}} + \frac{9}{2} & \sqrt{7} \le d \le 3 \\ \frac{2}{\sqrt{3}}\left(\frac{d^2}{3} + 12\right)\sin^{-1}\frac{3\sqrt{3}}{2d} - \frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} & \\ \quad + \left(\frac{\pi}{9\sqrt{3}} - \frac{1}{3}\right) d^2 - 4\sqrt{d^2-3} + \frac{19}{6}\sqrt{4d^2-27} - \frac{8\pi}{3\sqrt{3}} - \frac{9}{2} & 3 \le d \le 2\sqrt{3} \\ \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+12\right)\left(\sin^{-1}\frac{3\sqrt{3}}{2d} + \sin^{-1}\frac{2\sqrt{3}}{d}\right) - \left(\frac{2}{3} + \frac{4\pi}{9\sqrt{3}}\right) d^2 & \\ \quad + \frac{19}{6}\sqrt{4d^2-27} + \frac{16}{3}\sqrt{d^2-12} - \frac{16\pi}{\sqrt{3}} - \frac{25}{2} & 2\sqrt{3} \le d \le \sqrt{13} \\ 0 & \text{otherwise} \end{cases} \tag{6.11}$$

6.2.3 Distance Verification

(1) Verification by Simulation

Figure 6.10 shows the comparison between the CDFs of the random distances both within the same hexagon and between adjacent hexagons, computed from (6.5) and (6.11), and the simulation results by generating $2,000$ pairs of random points with each of the corresponding geometric locations. Figure 6.10 demonstrates that our distance distribution functions are very accurate when compared with the simulation results.

(2) Validation by Recursion

As shown in Figure 6.11, a hexagon with a side length of 2 can be decomposed into three unit hexagons, and three unit rhombuses A, J and K. With the scaling in (5.14), the distance distribution in the large hexagon is $f_{2D_{\mathrm{H_I}}}(d) = \frac{1}{2} f_{D_{\mathrm{H_I}}}(\frac{d}{2})$. On the other hand, if we look at the two random endpoints of a given link inside the large hexagon,

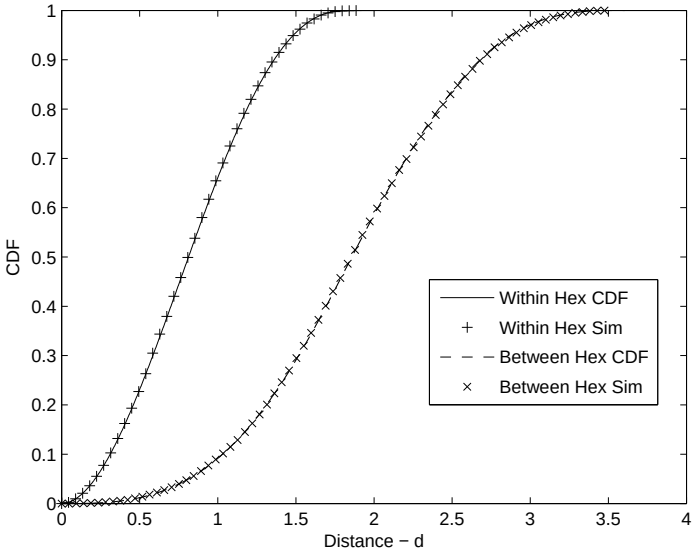


Figure 6.10: Cumulative Distribution Functions and Simulation Results for Random Distances within and between Hexagons.

they will fall into one of the three following cases: i) both endpoints fall inside one of the three small hexagons (BDF, ECG or HIL), with probability $\frac{3}{4} \times \frac{3}{4}$; ii) one of the endpoints falls into one of the small hexagons, and the other endpoint falls into one of the three rhombuses A, J or K, with probability $2 \times \frac{3}{4} \times \frac{1}{4}$; iii) both endpoints fall into one of the rhombuses, with probability $\frac{1}{4} \times \frac{1}{4}$.

Each of these three cases includes several more detailed sub-cases as follows:

Case i) Given the location of the first endpoint of a particular link, the second endpoint

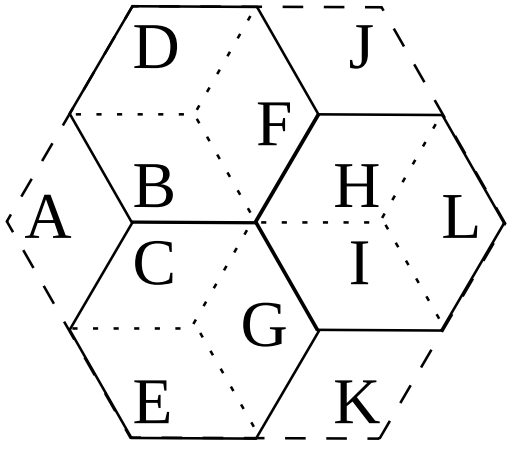


Figure 6.11: Partial Recursion through Hexagons and Rhombuses.

will fall in the same hexagon as the first one with probability $\frac{1}{3}$, and in one of the two adjacent hexagons with probability $\frac{2}{3}$. The unconditional probabilities of these two sub-cases are $\frac{9}{16} \times \frac{1}{3} = \frac{3}{16}$ and $\frac{9}{16} \times \frac{2}{3} = \frac{3}{8}$, respectively.

Case ii) Without loss of generality, suppose the first endpoint is located in A. By symmetry, the second endpoint falls into any one of the four rhombuses B, D, F and H with probability $\frac{2}{9}$, and into L with probability $\frac{1}{9}$. Thus the unconditional probabilities are $\frac{3}{8} \times \frac{2}{9} = \frac{1}{12}$ each for $|AB|$, $|AD|$, $|AF|$ or $|AH|$, and $\frac{3}{8} \times \frac{1}{9} = \frac{1}{24}$ for $|AL|$.

Case iii) If the first endpoint is in A, then by symmetry, the second endpoint is still located in A with probability $\frac{1}{3}$, and in either one of J or K with probability $\frac{2}{3}$. The unconditional probabilities of these two sub-cases are $\frac{1}{16} \times \frac{1}{3} = \frac{1}{48}$ and $\frac{1}{16} \times \frac{2}{3} = \frac{1}{24}$, respectively.

In short, we have the probabilistic sum as

$$\begin{aligned} f_{2D_{\mathrm{H_I}}}(d) &= \frac{3}{16} f_{D_{\mathrm{H_I}}}(d) + \frac{3}{8} f_{D_{\mathrm{H_A}}}(d) + \frac{1}{12} \left[f_{D_{\mathrm{PX}}}(d) + f_{D_2}(d) + f_{D_1}(d) + f_{D_3}(d) \right] \\ &\quad + \frac{1}{24} \left[f_{D_4}(d) + f_{D_5}(d) \right] + \frac{1}{48} f_{D_{\mathrm{I}}}(d), \end{aligned} \tag{6.12}$$

where $f_{D_5}(d)$ is the density function of $|AL|$ in Figure 6.11, which is the only distance distribution function that has not been given yet.

The result of $f_{D_5}(d)$ and the mathematical verification of all the distribution functions derived in this section, through the above probabilistic sum in (6.12) and a scaling in (5.14), is give in Appendix C.

6.3 Practical Results

6.3.1 Statistical Moments of Random Distances

The distance distribution functions given in Sections 6.2.1 and 6.2.2 can conveniently lead to all the statistical moments of the random distances associated with hexagons. Given $f_{D_{\mathrm{H_I}}}(d)$ in (6.5), for example, the first moment (mean) of d, or the average

Table 6.1: Moments and Variances of Hexagon Distributions—Numerical vs Simulation Results

Geometry	PDF/Sim	$M_D^{(1)}$	$M_D^{(2)}$	Var_D
Within a Single Hexagon	$f_{D_{\mathrm{H_I}}}(d)$	$0.8262542775s$	$0.8333333333s$	$0.1506291100s^2$
	Sim	$0.8263306317s$	$0.8335924725s$	$0.1507701596s^2$
Between two Adjacent Hexagons	$f_{D_{\mathrm{H_A}}}(d)$	$1.8564318344s$	$3.832947195s$	$0.3866080394s^2$
	Sim	$1.8583366966s$	$3.8326819696s$	$0.3792666917s^2$

distance within a regular hexagon, is

$$\begin{aligned} M_{D_{\mathrm{H_I}}}^{(1)} &= \int_0^2 x f_{D_{\mathrm{H_I}}}(x)dx = \frac{7\sqrt{3}}{30} - \frac{7}{90} + \frac{1}{60}\left[28\ln\left(2\sqrt{3}+3\right) + 29\ln\left(2\sqrt{3}-3\right)\right] \\ &\approx 0.8262542775, \end{aligned}$$

and the second raw moment is

$$M_{D_{\mathrm{H_I}}}^{(2)} = \int_0^2 x^2 f_{D_{\mathrm{H_I}}}(x)dx = \frac{5}{6},$$

from which the variance, or the second central moment, can be derived as

$$Var_{D_{\mathrm{H_I}}} = M_{D_{\mathrm{H_I}}}^{(2)} - \left[M_{D_{\mathrm{H_I}}}^{(1)}\right]^2 \approx 0.1506291100.$$

When the side length of a hexagon is scaled by s, the corresponding first two statistical moments given above then become

$$M_{D_{\mathrm{H_I}}}^{(1)} = 0.8262542775s, \quad M_{D_{\mathrm{H_I}}}^{(2)} = \frac{5s}{6} \quad \text{and} \quad Var_{D_{\mathrm{H_I}}} = 0.1506291100s^2.$$

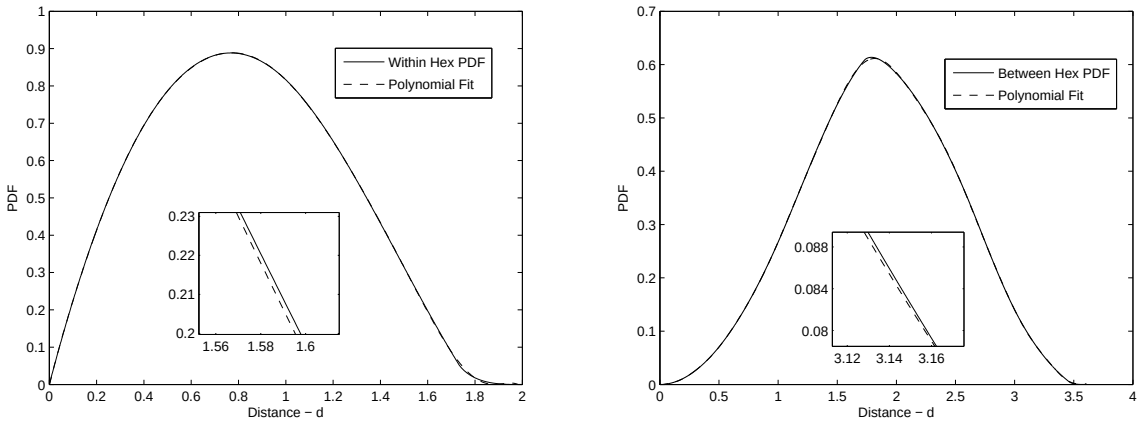
Table 6.1 lists the first two moments, and the variance of the random distances in the two cases given in Sections 6.2.1 and 6.2.2, and the corresponding simulation results for verification purposes.

6.3.2 Polynomial Fits of Random Distances

Table 6.2 lists the coefficients of the high-order polynomial fits of the original PDFs given in this section, from the highest degree (degree-10 for (6.5) and degree-20 for (6.11)) to d^0, together with the corresponding norm of residuals. Figure 6.12 (a)–(b) plot the polynomials listed in Table 6.2 with the original PDFs. From the figure,

Table 6.2: Coefficients of the Polynomial Fit and the Norm of Residuals (NR) for Hexagon Distributions

PDF	Degree	Polynomial Coefficients	NR
$f_{D_{\mathrm{H_I}}}(d)$	10	$10^2 \times [-0.0146710 \;\; 0.136604 \;\; -0.538052 \;\; 1.167903$ $-1.525478 \;\; 1.230615 \;\; -0.605940 \;\; 0.175147$ $-0.043772 \;\; 0.025830 \;\; -0.000025]$	0.075608
$f_{D_{\mathrm{H_A}}}(d)$	20	$10^4 \times [0.00000035 \;\; -0.000013 \;\; 0.000207 \;\; -0.002094$ $0.014469 \;\; -0.072522 \;\; 0.272508 \;\; -0.782682 \;\; 1.736254$ $-2.986406 \;\; 3.976655 \;\; -4.072372 \;\; 3.169347 \;\; -1.841066$ $0.778001 \;\; -0.230634 \;\; 0.045522 - 0.005534 \;\; 0.000394$ $-0.0000103 \;\; 0.00000007092]$	0.191157



(a) Within a Single Hexagon

(b) Between two Adjacent Hexagons

Figure 6.12: Polynomial Fit.

it can be seen that both polynomials match closely with the original PDFs. These high-order polynomials facilitate further manipulations of the distance distribution functions, with a high accuracy.

6.4 Performance Study Using Distance Distributions

Using the distance distribution functions derived in this chapter, we are able to further analyze the location-critical performance metrics in wireless communication networks. Define neighbors as the nodes that are within the one-hop transmission range of a

target node. The underlying routing schemes have the option to choose a particular neighbor to forward a message. Starting from a challenging scenario where nodes are sparsely deployed, we analyze the nearest neighbor of a node in a hexagonal network. In a dense network, we analyze the farthest neighbor that a node can reach through a single transmission. By choosing appropriate neighbors for packet relay, the underlying routing protocols can benefit from the minimization of energy consumption and routing overhead in these two scenarios, respectively.

As discussed in Section 2.1.2, the distance distributions of circles have been widely used in the existing literature to approximate hexagonal topologies: the inscribed or enclosing circles, and the circles with the same area as the hexagon [33, 60], due to the reduced computational complexity. If the side length of a hexagon is s, the radius of the circle having the same size as the hexagon is $r = \sqrt{\frac{3\sqrt{3}}{2\pi}}s$. For the inscribed and enclosing circle, r is $\frac{\sqrt{3}}{2}s$ and s, respectively. When the communication involves adjacent hexagons, however, circular geometries must leave gaps, or partially overlap and include areas not belonging to the cell of interest. Otherwise, gaps must exist between cells in order to avoid overlappings. Therefore, we also show in this section that when compared with the explicit hexagon distribution models, traditional circular approximations are less accurate or even not applicable, in analyzing both the statistical distribution and expected average of the system performance metrics.

6.4.1 Sparse Network Scenario: the Nearest Neighbor

In a sparse network where the network size is much larger than the communication range of a node, an energy-efficient routing protocol will choose the nearest neighbor as the packet forwarder in order to conserve energy at each node. This is particularly important when the network device has very limited energy supply. Suppose there are n nodes located in the same area, for a random node i, the minimum distance from $n-1$ nodes to i is

$$\delta = \min\{d_1, d_2, \cdots, d_{n-1}\}. \tag{6.13}$$

The distribution function of δ is

$$F_\Delta(\delta) = 1 - P(d_i \geq \delta)^{n-1} = 1 - [1 - F_D(\delta)]^{n-1},$$

where $F_D(\cdot)$ is any of the distance distribution CDFs derived in the last section. The

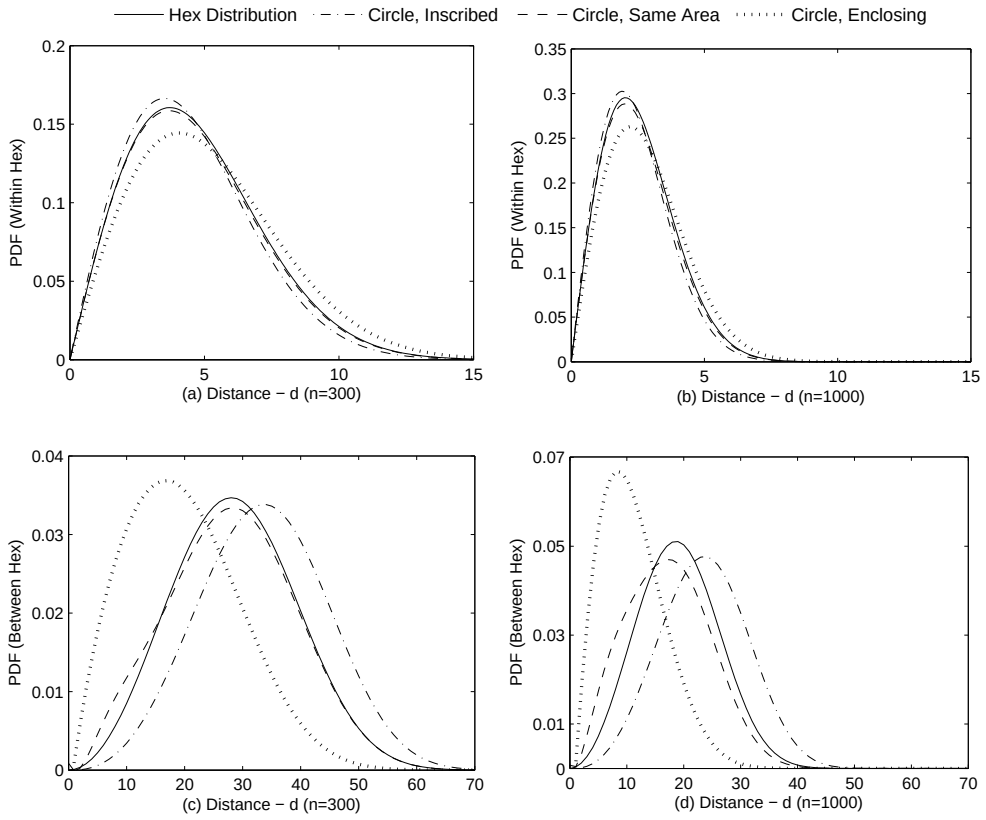


Figure 6.13: Nearest Neighbor Distribution.

PDF of δ is thus

$$f_{\Delta}(\delta) = (n-1) f_D(\delta) \left[1 - F_D(\delta)\right]^{n-2}. \tag{6.14}$$

Similarly, the k-th nearest neighbor to node i can be derived by order statistics.

(1) Distribution of the Distance to the Nearest Neighbor

In the following we use the hexagonal topology since it is commonly used in communication networks. Figure 6.13(a)–(d) show the corresponding PDFs in (6.14) using the hexagon distribution functions we derived. It also shows those three circular approximations, of which the distance distribution functions are given in [69] for non-overlapping circles. When circles overlap, results from Monte Carlo simulations are used. No closed-form PDFs for overlapping circles are available in the literature, except for the empirical approximations on CDFs [16], which cannot provide any result for (6.14).

Figure 6.13(a) and (b) plot the PDFs in (6.14) using the distance distribution *within* a hexagon, when the side length of the hexagon is 100 m, and the total number of nodes is $n = 300$ and $1,000$, respectively. When the network size is fixed, the increased number of nodes makes the nearest neighbor closer to node i. Also note that, the tails of the PDFs are truncated at 15 m in these two figures. Although the longest distance between nodes can be as far as 200 m, it is very likely that the nearest neighbor is only a few meters away. Figure 6.13(c) and (d) use the distribution *between* two adjacent hexagons, with the same parameters as Figure 6.13(a) and (b), and are truncated at 70 m. In these two figures, none of the circular approximations give a good fit to the hexagon distributions. Intuitively, inscribed circles skew nodes apart from the hexagon boundary, whereas same-area circles and enclosing circles overlap and cause nodes to be physically closer.

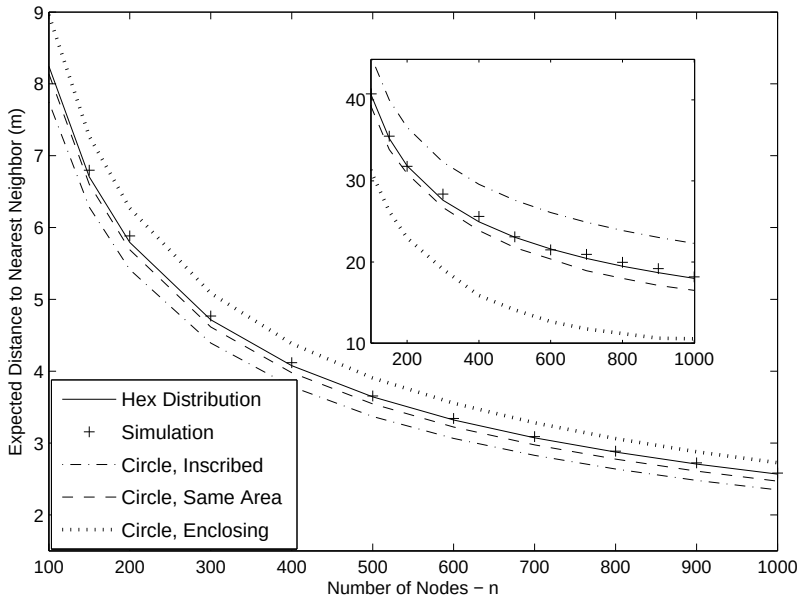
(2) Expectation of the Distance to the Nearest Neighbor

The expectation of the distance to the nearest neighbor can be obtained by integrating (6.14) over the hexagonal area. The results are given in Figure 6.14(a) and (b). In both figures, the zoom-in subfigures are between adjacent hexagons. In Figure 6.14(a), the hexagon size is fixed at 100 m, and Figure 6.14(b) fixes the number of nodes at $1,000$. From Figure 6.14(a), if the underlying routing protocol chooses the nearest neighbor within a hexagon of size 100 m, the expected distance is less than 8 m with more than 100 nodes, whereas an average distance of 82.6 m is needed to reach an arbitrary node according to Table 6.1 in Section 6.3.1.

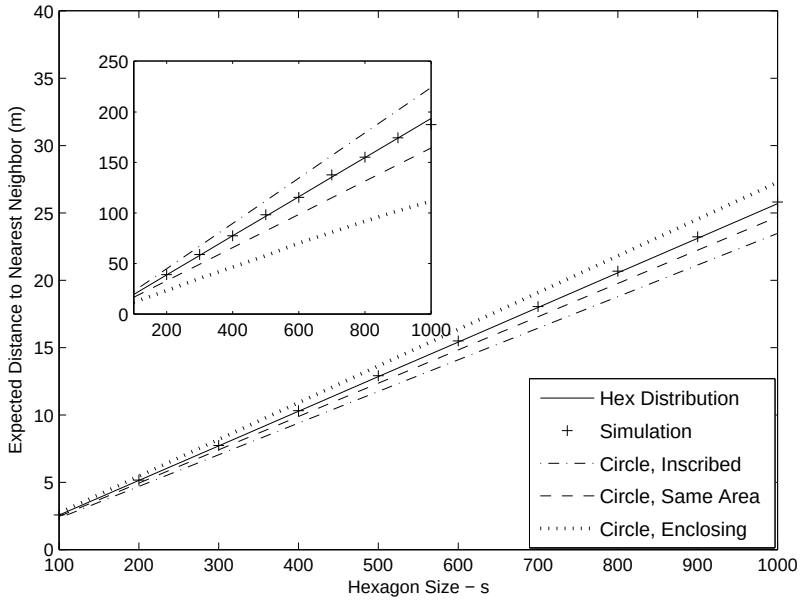
We may also observe that in a single hexagon, circles can approximate relatively well; but when nodes communicating between hexagons, the error of circular approximations grows nonlinearly. It is also interesting to see in Figure 6.14(b) that the distance to the nearest neighbor increases linearly with the network size because of the uniform distribution.

6.4.2 Dense Network Scenario: the Farthest Neighbor

In a sparse network, it is highly desirable to choose the nearest neighbor for improving energy efficiency. In a small, densely deployed network, on the other hand, it is common for a node to have several neighbors simultaneously. As a result, the number of transmissions and routing overhead can be minimized by choosing the farthest node



(a) Distance vs. No. of Nodes.



(b) Distance vs. Network Size.

Figure 6.14: Expected Distance to Nearest Neighbor.

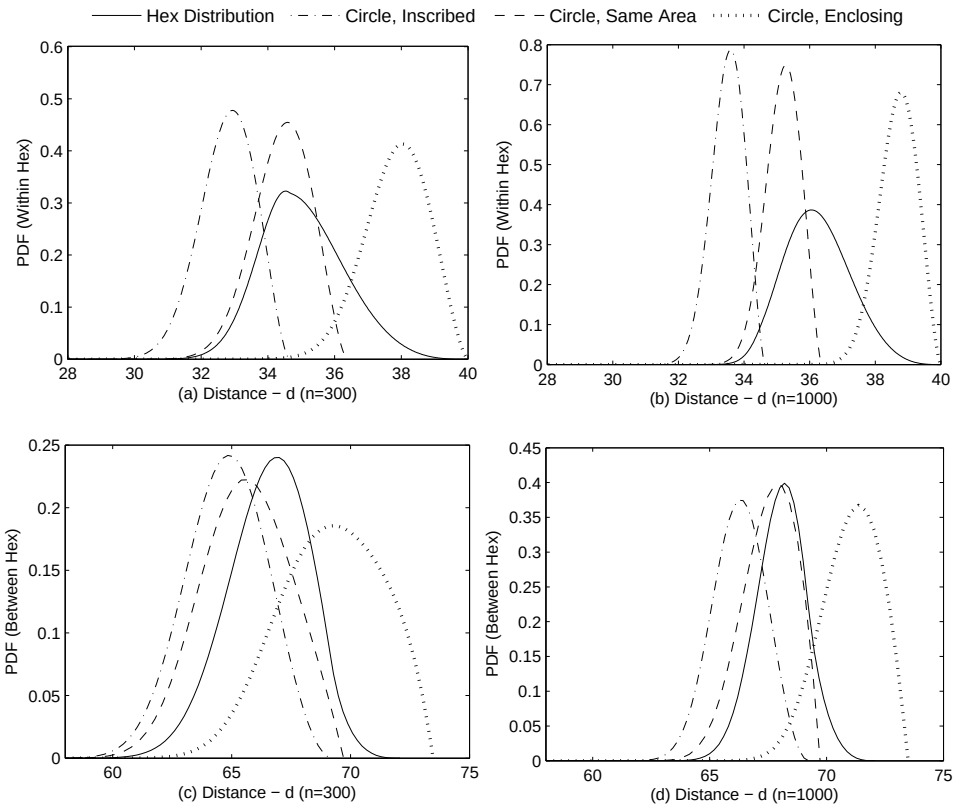


Figure 6.15: Farthest Neighbor Distribution.

as relay. Suppose the total number of nodes is n, for node i define

$$\gamma = \max\{d_1, d_2, \cdots, d_{n-1}\}. \tag{6.15}$$

The distribution of γ has CDF

$$F_\Gamma(\gamma) = P(d_i \leq \gamma)^{n-1} = F_D^{n-1}(\gamma),$$

and the corresponding PDF

$$f_\Gamma(\gamma) = (n-1) f_D(\gamma) F_D^{n-2}(\gamma). \tag{6.16}$$

(1) Distribution of the Distance to the Farthest Neighbor

Figure 6.15(a) and (b) plot the PDFs in (6.16) using the distance distribution within a hexagon, of which the side length is 20 m, and the total number of nodes is $n = 300$

and $1,000$, respectively. The increased number of nodes brings the farthest neighbor closer to the network boundary at 40 m. Thus the left tails of the PDFs are truncated from 28 m. Figure 6.15(c) and (d) show the distribution between two adjacent hexagons with the same parameters, and are truncated from 58 m. In Figure 6.15(a)–(d), it is obvious that none of the circular topologies give a good approximation for hexagons. Therefore, when the communication happens between two adjacent hexagons, the circular approximations become even less accurate. Thus, it is important to capture the location-critical performance metrics with a geometrical probability approach for hexagons.

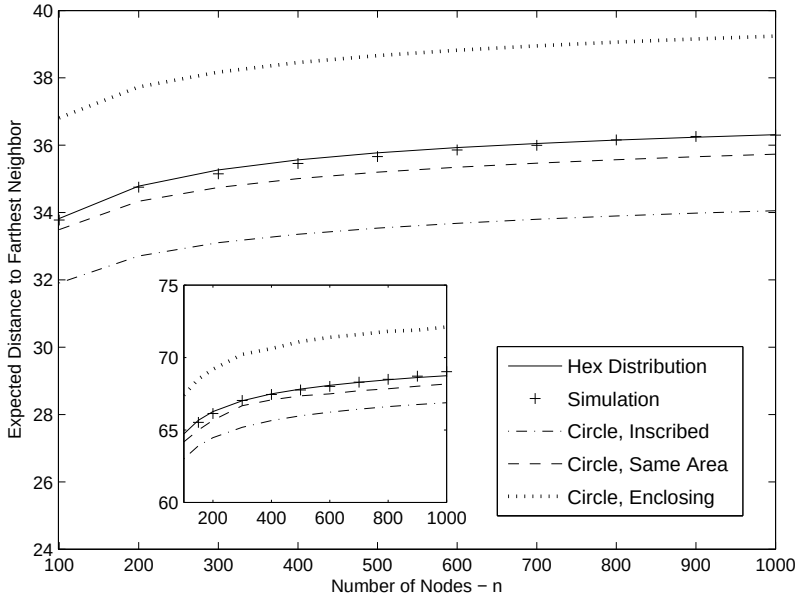
(2) Expectation of the Distance to the Farthest Neighbor

Figure 6.16(a) shows the expected distance to the farthest neighbor, with an increasing number of nodes, and the size of the hexagon is fixed at 20 m. The zoom-in subfigure shows the results between two adjacent hexagons. Compared with the nearest neighbor, the increase in the number of nodes does not increase the expected distance much, since the farthest node is already around the network boundary. We can anticipate that in a dense network, single hop communication is very likely to cover the neighbors that are far-away. The underlying routing protocol thus can achieve the minimal routing overhead. Figure 6.16(b) shows the expected distance when the number of nodes is fixed at $1,000$ and the cell size changes from 20 m to 50 m. Similar to Figure 6.14(b), the distance to the farthest neighbor also increases linearly with the network size, because of the uniform node distribution.

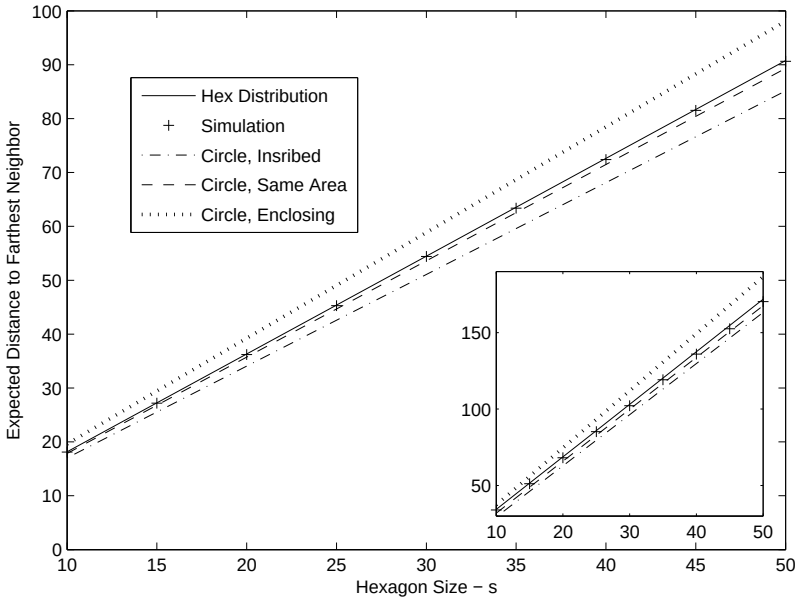
6.4.3 Transmission Power Control

Knowing the neighbor statistics, it is then possible to use transmission power control to maintain the network connectivity while conserving the energy of each device. This section uses the nearest neighbor as a case study. Given a receiver power threshold P_r, below which the received signal cannot be successfully decoded, the minimum transmission power to overcome the path loss between nearest neighbors is proportional to $P_r\delta^\alpha$ [48], where δ is defined in (6.13). Without loss of generality, we assume P_r is 1 power unit, and let $H = \delta^\alpha$. With the distribution in (6.14) we have

$$P(H \leq h) = P(\delta \leq h^{\frac{1}{\alpha}}) = F_\Delta(h^{\frac{1}{\alpha}}). \tag{6.17}$$

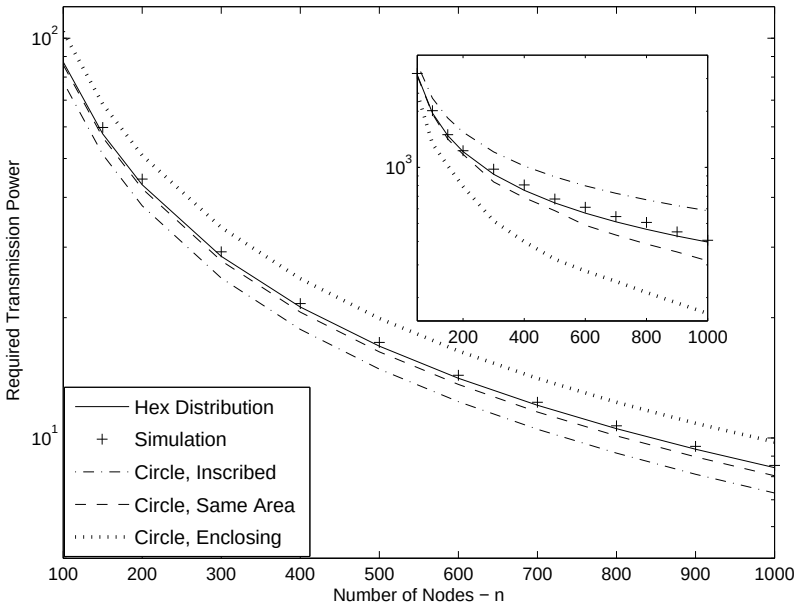


(a) Distance vs. No. of Nodes.

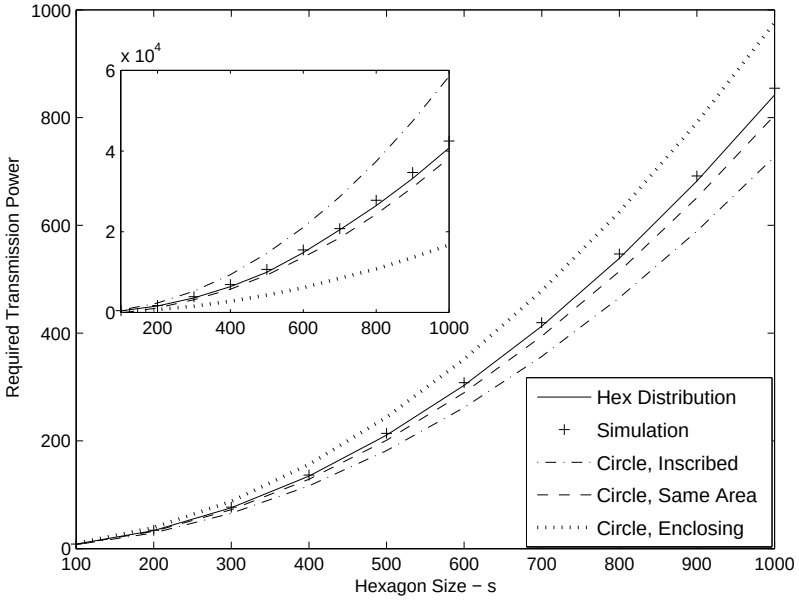


(b) Distance vs. Network Size.

Figure 6.16: Expected Distance to Farthest Neighbor.



(a) Transmission Power vs. No. of Nodes.



(b) Transmission Power vs. Network Size.

Figure 6.17: Expected Transmission Power to the Nearest Neighbor.

Therefore,

$$f_H(h) = F'_\Delta(h^{\frac{1}{\alpha}}) = \frac{1}{\alpha} h^{\frac{1}{\alpha}-1} f_\Delta(h^{\frac{1}{\alpha}}). \tag{6.18}$$

Figure 6.17 shows the expected transmission power to reach the nearest neighbor, by integrating (6.18). As the number of nodes and network size increase, the error of circular approximations also increases. If one needs to analyze the power control schemes which adjust the transmission power to reach an arbitrary neighbor, it is sufficient to replace $f_\Delta(\cdot)$ with $f_D(\cdot)$ in (6.18). Furthermore, the distribution of path loss can be obtained in a very similar way by replacing H with $\delta^{-\alpha}$ in (6.17) and (6.18). The analysis on path loss, interference, etc, with a fixed reference point in hexagonal networks is given in our previous work [103].

6.5 Summary

In Chapters 5 and 6, a unified approach is developed to deriving the distribution of random distances, by a quadratic product of the probability function and the probability density. The closed-form probability density functions of the random distances associated with rhombuses and hexagons are derived for the first time in the literature. These results are further utilized to investigate the nearest and farthest neighbor statistics in both sparse and dense network scenarios, with transmission power control as a case study. The analytical and simulation results show the high accuracy and promising potentials of this approach.

6.6 Discussions and Future Work

The future work includes deriving the distance distributions under general point distributions, such as bivariate Gaussian distribution and Beta distribution, as discussed in Section 5.6.2.

The *conditional* probability of the distance distributions associated with rhombuses and hexagons is useful, as it can be utilized to model the location-dependent interference, wireless propagation channel, hidden terminals and cooperative communications. The detailed discussions will be given in Section 7.2. We believe the probabilistic models presented in this chapter and their future extensions will provide important guidelines for a more accurate network dimensioning and a better protocol design process.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

The field of wireless communication has been experiencing tremendous growth. However, how to model and analyze these communication networks by combining the emerging wireless communication and networking technologies with a realistic mathematical model, is still an open issue and a demanding task. Throughout this dissertation, the location-critical performance metrics in various wireless communication networks are studied from a geometrical point of view, where the network coverage and node locations are of profound importance for the operation of these networks. Each of the location-critical performance metrics, such as transmission power and path loss, the k-th nearest neighbor and traveling distances, etc., can be modeled as a nonlinear function of random nodal distances in the relevant network topology. Such random distances between network devices, characterized by geometrical probability, play a fundamental role in determining the crucial performance metrics in communication networks.

7.1.1 Application of One-Dimensional Random Distances

In Chapter 3, the Poisson point process is applied to vehicular ad-hoc networks, where the network geometry is constrained by highway and road structures. A time and location-critical framework is proposed for a one-dimensional highway scenario, where the message propagation is always limited to a certain distance from the message source. Utilizing physical layer scalable modulation and coding scheme, vehicles at different distances to an accident site can receive information with different levels of

details. Based on the memory-less property of a Poisson process, the fundamental limits of message propagation in a highway scenario is investigated, by deriving the distribution of the vehicle cluster size.

The extension of this model is applied to a two-dimensional Manhattan-like city in Chapter 4. Different from one dimension, there exists a critical threshold above which the entire network is connected with a high probability, which is similar to the percolation phenomenon. The network connectivity is further analyzed by two different message forwarding schemes by showing the performance tradeoff.

Important location-critical performance metrics, such as the "carry-and-forward" delay in one dimension, and connectivity probability in two dimensions, are obtained through simple approaches than the current literature [75, 100], without using mathematical approximation and average analysis. Extensive simulations also show that the derived results are highly accurate. This is the contribution of this dissertation from the application aspect, where the distance distribution along a line is applied to a vehicular ad-hoc network scenario.

7.1.2 Random Distances Associated with Complex Geometries

In Chapter 5 and 6, a unified approach to deriving the distribution of random distances by a quadratic product is proposed, which is applicable to both elementary and complex geometries. This is the contribution of this dissertation from the theory aspect, where the closed-form distributions of random distance are derived for for hexagons and rhombuses, the topological shapes that are the most suitable for cellular systems, and the sectorized cells with directional antennas. Previously, conducting the distribution analysis on network performance metrics has been intractable for complex network geometries.

The analytical models based on hexagon distributions are applied to the analysis of the nearest neighbor distribution in a sparse network for improving energy efficiency, and the farthest neighbor distribution in a dense network for minimizing routing overhead. The corresponding analysis on transmission power control also show that using the explicit distance distributions, the analytical models are highly accurate when compared with the latest approximation methods. It is therefore believed that the theory of geometrical probability is essential to the successful network protocol and system design.

7.2 Future Work

Extensions of the current work, including a study on hybrid vehicle-to-vehicle and vehicle-to-infrastructure communication structure (Section 3.4), collisions, throughput and vehicle mobility in vehicular ad-hoc networks (Section 4.4), nonuniform point distribution (Section 5.6), and the conditional distribution of the distance distributions (Section 6.6) are our ongoing work. Furthermore, the future research plans beyond this dissertation, revolve around the probabilistic distance distributions under the impact of human and vehicle mobility, and the stochastic models of network performance. The conditional distance distributions also have profound impact on the wireless channel models and cooperative communications. Both measurement studies and theoretical analysis are essential to the successful network protocol and system design. We anticipate that the applications of this research will be numerous and diverse.

7.2.1 Vehicular Ad Hoc Networks with Real-World Traces

(1) Inter-Vehicle Distance Models

The challenges in VANET are mainly due to the dynamic mobility and network topology: high vehicle speed, short sessions of connectivity with frequent disruptions, etc. The exponential inter-vehicle distances model in [13] etc., showed that such distributions are different with respect to geographical location, direction, lane number/type, and time of the day. Implicit parameters in those findings which were not characterized are vehicle speeds in lanes, driving habits and vehicle types. Given real-world vehicle traces provided by both academic and industry supporters, future extension of this dissertation is to characterize the complicated relationship by the measured and predicted vehicle trajectory, and develop distribution models of inter-vehicle distances which capture the variable speed and vehicle behaviors on the road.

(2) Stochastic Models of Network Performance Metrics

Although vehicle mobility is highly dynamic, in many cases, it is also predictable because of the road constraints and traffic conditions. Given the time/location-criticality of safety applications and the popularity of infotainment applications, future target of this dissertation also includes a probabilistic and stochastic understanding of network topology and mobility, and the implication on upper-layer network

performance. With the inter-vehicle distance in such a complicated environment, the capacity of a drive-thru Internet can be analyzed, which depends on the stochastic vehicle arrival and departure processes [111, 112].

7.2.2 Wireless Channel Models

In a typical wireless radio propagation environment, the radio signals emitted from the transmitters can be reflected and diffracted by scatterers in the surrounding environment before arriving at the receivers. Any object and obstacle between the transceivers can cause such reflection, diffraction or scattering of signals, each of which leading to a certain loss of power over the distance between the scatterers and transceivers. Moreover, the Doppler shift takes effect if any transceivers are moving, such as in a vehicular network, resulting in an even more complicated communication channel.

Signal can go through multiple paths from the transmitter to the receiver, either by a direct line-of-sight (LOS) path, or after being scattered through non-line-of-sight (NLOS) paths. The signal observed by a receiver or a scatterer is the sum of the signals arriving via all incoming paths. The scatterers sum up the incoming signals and re-emit the sum-signal via the outgoing paths. The receivers are sinks with only incoming paths. Whenever a signal propagates along a path, it experiences path loss, propagation delay, and the resultant phase shift, all of which depend on the path length, or the distance between different objects. In a LOS path, the distance can be obtained given the distribution of objects and the geometry of the network. When the path is NLOS, however, the length of incoming and outgoing paths through the same scatterer is non-independent. This poses a mathematical challenge of deriving the joint distance distribution given the common point of two links, in both elementary and complex geometries. Note that this is also applicable to the relay systems for cooperative communications, where the scatterer described above becomes a relay station.

7.2.3 Inter-Disciplinary Research

Besides the application in the networking area, the geometrical probability research can also be applied to a wide variety of science and engineering fields. Below are some of the examples.

(1) Vehicle Dispatching and Transportation

The travel time and cost in an urban area are substantially affected by the traveled distance, which can be modeled by a continuous statistical distribution. For instance, [28, 95] considered the distance between two cities or zones. [49] in particular, considered different types of journeys, such as internal journeys and cross-cordon journeys, etc.

(2) VLSI Design

In [59], the geometrical distance models characterize the wire length needed to connect random terminals on a chip. This problem arises in VLSI layout, computing the expected cost of minimum-cost Euclidean spanning trees, and the analysis of rectangle heuristics for minimum weighted Euclidean matching [82].

(3) City Planning

The practical problems such as the optimal placement of depots in city planning [27], the control of environmental hazards [65], such as vehicle emissions and noise pollution, and the road surface wear and tear, are all essentially reduced to the problem of the distances between random points. Consequently, the calculation of pollution intensities and optimal road patterns can be examined.

(4) Natural Sciences

Similar research challenges appear in natural sciences, such as the study of the density of plants in forestry, the crystal growth in chemistry [66], and the investigation on gamma-rays in a nuclear reactor [50].

The variety of results and approaches taken in this field illustrate the way in which distance distributions and their moments can be applied to practical problems. We believe that as the classic theories being extended to new scenarios, academic research will eventually lead to successful applications in real life.

Bibliography

[1] Affine transformation [online]. http://en.wikipedia.org/wiki/Affine_transformation.

[2] Almost surely [online]. http://en.wikipedia.org/wiki/Almost_surely.

[3] Dedicated short range communications (DSRC) home [online]. http://www.leearmstrong.com/DSRC/DSRCHomeset.htm.

[4] An introduction to hexagonal geometry [online]. http://hexnet.org/library/hexnet/hexagonal-geometry.

[5] Least squares [online]. http://en.wikipedia.org/wiki/Least_squares.

[6] List of exits on highway 401 (Ontario) [online]. http://en.wikipedia.org/wiki/List_of_exits_on_Highway_401_(Ontario).

[7] Poisson distribution [online]. http://en.wikipedia.org/wiki/Poisson_distribution.

[8] M. Abuelela and S. Olariu. Soda: A smart opportunistic data dissemination approach for vanet. In *Proceeding of the 6th International Workshop on Intelligent Transportation*, pages 211–216, March 2009.

[9] V.S. Alagar. The distribution of the distance between random points. *Journal of Applied Probability*, pages 558–566, 1976.

[10] J. Andrews, F. Baccelli, and R. Ganti. A tractable approach to coverage and rate in cellular networks. *IEEE Transactions on Communications*, (99):1–13, November 2010.

[11] F. Baccelli and B. Blaszczyszyn. *Stochastic Geometry and Wireless Networks, Part I: Theory*. Now Publishers Inc, 2009.

[12] F. Baccelli and B. Blaszczyszyn. *Stochastic Geometry and Wireless Networks, Part II: Applications*. Now Publishers Inc, 2009.

[13] F. Bai and B. Krishnamachari. Spatio-temporal variations of vehicle traffic in vanets: Facts and implications. In *Proceedings of the Sixth ACM International Workshop on VehiculAr InterNETworking*, (VANET 2009), pages 43–52, New York, NY, USA, 2009.

[14] K.B. Baltzis. Empirical description of node-to-node distance density in non-overlapping wireless networks. *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, 9(1):57–68, 2010.

[15] K.B. Baltzis. Analytical and closed-form expressions for the distribution of path loss in hexagonal cellular networks. *Wireless Personal Communications*, pages 1–12, October 2011.

[16] K.B. Baltzis. A geometric method for computing the nodal distance distribution in mobile networks. *Progress In Electromagnetics Research*, 114:159–175, 2011.

[17] K.B. Baltzis and J.N. Sahalos. On the statistical description of the aoa of the uplink interfering signals in a cellular communication system. *European Transactions on Telecommunications*, 21(2):187–194, March 2010.

[18] C. Bettstetter and J. Eberspacher. Hop distances in homogeneous ad hoc networks. In *Proceeding of IEEE Vehicular Technology Conference, (VTC 2003-Spring)*, volume 4, pages 2286–2290, 2003.

[19] C. Bettstetter, H. Hartenstein, and X. Pérez-Costa. Stochastic properties of the random waypoint mobility model. *Wireless Networks*, 10(5):555–567, September 2004.

[20] Y. Bi, H. Zhao, and X. Shen. A directional broadcast protocol for emergency message exchange in inter-vehicle communications. In *Proceeding of IEEE International Conference on Communications (ICC 2009)*, pages 1–5, June 2009.

[21] B. Blaszczyszyn, M.K. Karray, and F.X. Klepper. Impact of the geometry, path-loss exponent and random shadowing on the mean interference factor in wireless cellular networks. In *Proceeding of IFIP Wireless and Mobile Networking Conference (WMNC 2010)*, pages 1–6, October 2010.

[22] J.J. Blum, A. Eskandarian, and L.J. Hoffman. Challenges of intervehicle ad hoc networks. *IEEE Transactions on Intelligent Transportation Systems*, 5(4):347–351, December 2004.

[23] H. Cai, X. Jia, and M. Sha. Critical sensor density for partial connectivity in large area wireless sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 7(4):35, February 2011.

[24] L. Cai, Y. Luo, S. Xiang, and J. Pan. Scalable modulation for scalable wireless videocast. In *Proceedings of IEEE INFOCOM*, pages 1–5, San Diego, CA, USA, March 2010.

[25] L.C. Chen and F.Y. Wu. Directed percolation in two dimensions: An exact solution. In *Differential geometry and physics: Proceedings of the 23rd International Conference of Differential Geometric Methods in Theoretical Physics, Tianjin, China, 20-26 August 2005*, volume 10, page 160. World Scientific Pub Co Inc, 2006.

[26] Y.C. Cheng and T.G. Robertazzi. Critical connectivity phenomena in multihop radio models. *IEEE Transactions on Communications*, 37(7):770–777, 1989.

[27] N. Christofides and S. Eilon. Expected distances in distribution problems. *Operational Research*, pages 437–443, 1969.

[28] D.P. Chu. Distance between random points in two rectangular cities. *Communications in Statistics—Simulation and Computation®*, 35(2):257–276, 2006.

[29] Y. Debessu, H. We, S.Y. Chang, and S. Huang. Lifetime analysis for wireless sensor network with hexagonal clustering. In *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Houston, TX, USA, December 2011.

[30] M. Desai and D. Manjunath. On the connectivity in finite ad hoc networks. *IEEE Communications Letters*, 6(10):437–439, 2002.

[31] H.S. Dhillon, R.K. Ganti, F. Baccelli, and J.G. Andrews. Modeling and analysis of K-tier downlink heterogeneous cellular networks. *Arxiv preprint arXiv:1103.2177*, 2011.

[32] O. Dousse. *Asymptotic Properties of Wireless Multi-hop Networks*. EPFL Ph. D. Thesis, 2005.

[33] J.S. Evans and D. Everitt. On the teletraffic capacity of CDMA cellular networks. *IEEE Transactions on Vehicular Technology*, 48(1):153–165, January 1999.

[34] M. Evans, N. Hastings, and B. Peacock. Triangular distribution. *Statistical distributions*, 3:187–188, 2000.

[35] P.A. Ferrari and N.L. Garcia. One-dimensional loss networks and conditioned M/G/ queues. *Journal of Applied Probability*, pages 963–975, 1998.

[36] D.J. Gates. Asymptotics of two integrals from optimization theory and geometric probability. *Advances in Applied Probability*, pages 908–910, 1985.

[37] B. Ghosh. On the distribution of random distances in a rectangle. *Science and Culture*, 8(9):388, 1943.

[38] B. Ghosh. Random distances within a rectangle and between two rectangles. *Bulletin of the Calcutta Mathematical Society*, 43(1):17–24, 1951.

[39] E.N. Gilbert. Random plane networks. *Journal of the Society for Industrial and Applied Mathematics*, 9(4):533–543, 1961.

[40] A. Goldsmith. *Wireless Communications*. Cambridge Univ Press, 2005.

[41] M.J. Grace and R.B. Potts. A theory of the diffusion of traffic platoons. *Operations Research*, pages 255–275, 1964.

[42] G. Grimmett. *Percolation*, volume 321. Springer Verlag, 1999.

[43] T.S. Hale. *A Taxonomy of Expected Distance Functions with Applications to Facility Location Problems*. PhD thesis, Texas A & M University, 1997.

[44] J. Hansen and M. Reitzner. Efficient indoor radio channel modeling based on integral geometry. *IEEE Transactions on Antennas and Propagation*, 52(9):2456–2463, 2004.

[45] J. Hansen and M. Reitzner. Electromagnetic wave propagation and inequalities for moments of chord lengths. *Advances in Applied Probability*, pages 987–995, 2004.

[46] M. Hassan-Ali and K. Pahlavan. A new statistical model for site-specific indoor radio propagation prediction based on geometric optics and geometric probability. *IEEE Transactions on Wireless Communications*, 1(1):112–124, 2002.

[47] L. He, Y. Zhuang, J. Pan, and J. Xu. Evaluating on-demand data collection with mobile elements in wireless sensor networks. In *Proceedings of IEEE Vehicular Technology Conference Fall (VTC 2010-Fall)*, pages 1–5, Ottawa, ON, Canada, September 2010.

[48] W.B. Heinzelman, A.P. Chandrakasan, and H. Balakrishnan. An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1(4):660–670, 2002.

[49] E. M. Holroyd. Theoretical average journey lengths in circular towns with various routing systems. *Road Research Laboratory Report*, (43), 1966.

[50] M. Horowitz. Probability of random paths across elementary geometrical shapes. *Journal of Applied Probability*, 2(1):169–177, 1965.

[51] K. Huang and J. Andrews. A stochastic-geometry approach to coverage in cellular networks with multi-cell cooperation. In *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Houston, TX, USA, December 2011.

[52] D. Jiang and L. Delgrossi. IEEE 802.11 p: Towards an international standard for wireless access in vehicular environments. In *Proceedings of IEEE Vehicular Technology Conference, 2008 (VTC Spring 2008)*, pages 2036–2040, 2008.

[53] M. Kafsi, P. Papadimitratos, O. Dousse, T. Alpcan, and J.P. Hubaux. Vanet connectivity analysis. *Arxiv preprint arXiv:0912.5527*, 2009.

[54] F.P. Kelly. Stochastic models of computer communication systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 379–395, 1985.

[55] F.P. Kelly. One-dimensional circuit-switched networks. *The Annals of Probability*, pages 1166–1179, 1987.

[56] G. Korkmaz, E. Ekici, F. Özgüner, and Ü. Özgüner. Urban multi-hop broadcast protocol for inter-vehicle communication systems. In *Proceedings of the 1st*

ACM international workshop on Vehicular ad hoc networks, (VANET 2004), pages 76–85, New York, NY, USA, 2004. ACM.

[57] C. Kurrer and K. Schulten. Dependence of percolation thresholds on lattice connectivity. *Physical Review E*, 48(1):614, 1993.

[58] J.D. Lawrence. *A Catalog of Special Plane Curves*. Dover Publications, 1972.

[59] D.M. Lazo and A.T. Sherman. An exact formula for the expected wire length between two randomly chosen terminals. 1994.

[60] Z. Lei, D.J. Goodman, and N.B. Mandayam. Location-dependent other-cell interference and its effect on the uplink capacity of a cellular CDMA system. In *Vehicular Technology Conference, (VTC 1999)*, volume 3, pages 2164–2168. IEEE, 1999.

[61] L. Liu and D.H. Shi. Busy period in $\mathrm{GI}^X/\mathrm{G}/\infty$. *Journal of Applied Probability*, pages 815–829, 1996.

[62] V. Wong M. Cheung, F. Hou and J. Huang. Dora: Dynamic optimal random access for vehicle-to-roadside communications. *IEEE Journal on Selected Areas in Communications*, 2012.

[63] W.S. Wong M. Cheung, F. Hou and J. Huang. Dynamic optimal random access for vehicle-to-roadside communications. In *Proceeding of IEEE International Conference on Communications (ICC 2011)*, pages 1–6. IEEE, 2011.

[64] F.B. Martin. 349. note: Beehive designs for observing variety competition. *Biometrics*, pages 397–402, 1973.

[65] A.M. Mathai. Pollution by vehicular travels from the suburbs to the city core. *Environmetrics*, 9(6):617–628, 1998.

[66] A.M. Mathai. Random distances associated with triangles. *International Journal of Mathematical and Statistical Sciences*, 7(1):77–96, 1998.

[67] A.M. Mathai. *An Introduction to Geometrical Probability: Distributional Aspects with Applications*, volume 1. CRC, 1999.

[68] A.M. Mathai, P. Moschopoulos, and G. Pederzoli. Random points associated with rectangles. *Rendiconti del Circolo Matematico di Palermo*, 48(1):163–190, 1999.

[69] A.M. Mathai and G. Pederzoli. Random points with reference to a circle revisited. *Rendiconti del circolo matematico di Palermo, Serie II, Suppl*, 50:235–258, 1997.

[70] A.M. Mathai and S.B. Provost. *Quadratic Forms in Random Variables: Theory and Applications*. M. Dekker, 1992.

[71] M. Mauve, A. Widmer, and H. Hartenstein. A survey on position-based routing in mobile ad hoc networks. *Network, IEEE*, 15(6):30–39, 2001.

[72] R. Mazumdar. *Performance modeling, loss networks, and statistical multiplexing*. Morgan and Claypool Publishers, 2010.

[73] L.E. Miller. Distribution of link distances in a wireless network. *Journal of Research-National Institute of Standards and Technology*, 106(2):401–412, 2001.

[74] L.E. Miller. Joint distribution of link distances. In *Proc. Conference on Information Science and Systems, The John Hopkins University*. Citeseer, 2003.

[75] D. Miorandi and E. Altman. Connectivity in one-dimensional ad hoc networks: A queueing theoretical approach. *Wireless Networks*, 12(5):573–587, 2006.

[76] J.D. Murchland and D.J. Daley. Problem 75-12. *SIAM Review*, pages 497–501, 1976.

[77] S.C. Ng, W. Zhang, Y. Zhang, Y. Yang, and G. Mao. Analysis of access and connectivity probabilities in vehicular relay networks. *IEEE Journal on Selected Areas in Communications*, 29(1):140, 2011.

[78] T. Novlan, R.K. Ganti, S.Y. Chang, and J. Andrews. Coverage in two-tier cellular networks with fractional frequency reuse. In *Global Telecommunications Conference, 2011. (GLOBECOM 2011). IEEE*, 2011.

[79] P. Pirinen. Cellular topology and outage evaluation for ds-uwb system with correlated lognormal multipath fading. In *IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2006)*, pages 1–5, 2006.

[80] P. Pirinen. Outage analysis of ultra-wideband system in lognormal multipath fading and square-shaped cellular configurations. *EURASIP Journal on Wireless Communications and Networking*, 2006(2):63–63, 2006.

[81] L. Quadrifoglio, R.W. Hall, and M.M. Dessouky. Performance and design of mobility allowance shuttle transit services: bounds on the maximum longitudinal velocity. *Transportation science*, 40(3):351–363, 2006.

[82] E.M. Reingold and K.J. Supowit. Probabilistic analysis of divide-and-conquer heuristics for minimum weighted euclidean matching. *Networks*, 13(1):49–66, 1983.

[83] G. Resta, P. Santi, and J. Simon. Analysis of multi-hop emergency message propagation in vehicular ad hoc networks. In *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 2007)*, pages 140–149, 2007.

[84] H. Saito. Performance analysis of combined vehicular communication. *IEICE Transactions on Communications*, 89(5):1486–1494, 2006.

[85] P. Santi and D.M. Blough. The critical transmitting range for connectivity in sparse wireless ad hoc networks. *IEEE Transactions on Mobile Computing*, pages 25–39, 2003.

[86] M. Schönhof, A. Kesting, M. Treiber, and D. Helbing. Coupled vehicle and information flows: Message transport on a dynamic vehicle network. *Physica A: Statistical Mechanics and its Applications*, 363(1):73–81, 2006.

[87] T.K. Sheng. The distance between two random points in plane regions. *Advances in Applied Probability*, pages 748–773, 1985.

[88] S. Shioda, J. Harada, Y. Watanabe, T. Goi, H. Okada, and K. Mase. Fundamental characteristics of connectivity in vehicular ad hoc networks. In *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2008)*, pages 1–6. IEEE, 2008.

[89] S. Srinivasa and M. Haenggi. Modeling interference in finite uniformly random networks. In *International Workshop on Information Theory for Sensor Networks (WITS 2007)*.

[90] S. Srinivasa and M. Haenggi. Distance distributions in finite uniformly random networks: Theory and applications. *IEEE Transactions on Vehicular Technology*, 59(2):940–949, 2010.

[91] I. Stojmenovic. Honeycomb networks: Topological properties and communication algorithms. *IEEE Transactions on Parallel and Distributed Systems*, 8(10):1036–1042, 1997.

[92] R.E. Stone. Technical note—some average distance results. *Transportation Science*, 25(1):83–90, 1991.

[93] D. Stoyan, W.S. Kendall, J. Mecke, and L. Ruschendorf. *Stochastic Geometry and Its Applications*, volume 313. Wiley New York, 1995.

[94] S. Ukkusuri and L. Du. Geometric connectivity of vehicular ad hoc networks: Analytical characterization. *Transportation Research Part C: Emerging Technologies*, 16(5):615–634, 2008.

[95] R. Vaughan. Approximate formulas for average distances associated with zones. *Transportation science*, 18(3):231, 1984.

[96] P. Vellore, P. Gillard, and R. Venkatesan. Probability distribution of multihop multipath connection in a random network. In *Global Telecommunications Conference, 2009. (GLOBECOM 2009). IEEE*, pages 1–5, 2009.

[97] W. Viriyasitavat, O.K. Tonguz, and F. Bai. Network connectivity of VANETs in urban areas. In *6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON 2009)*, pages 1–9. IEEE, 2009.

[98] X. Wang and L. Cai. Interference analysis of co-existing wireless body area networks. In *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Houston, TX, USA, December 2011.

[99] Eric W. Weisstein. Quadratic surface. from mathworld–a wolfram web resource [online]. `http://mathworld.wolfram.com/QuadraticSurface.html`.

[100] N. Wisitpongphan, F. Bai, P. Mudalige, V. Sadekar, and O. Tonguz. Routing in sparse vehicular ad hoc wireless networks. *IEEE Journal on Selected Areas in Communications*, 25(8):1538–1556, 2007.

[101] W. Zhang, Y. Chen, Y. Yang, X. Wang, Y. Zhang, X. Hong, and G. Mao. Multi-hop connectivity probability in infrastructure-based vehicular networks. *IEEE Journal on Selected Areas in Communications*, 2012.

[102] H. Zhu, M. Li, L. Fu, G. Xue, Y. Zhu, and L.M. Ni. Impact of traffic influxes: Revealing exponential intercontact time in urban vanets. *IEEE Transactions on Parallel Distributed Systems*, 22:1258–1266, August 2011.

[103] Y. Zhuang, Y. Luo, L. Cai, and J. Pan. A geometric probability model for capacity analysis and interference estimation in wireless mobile cellular systems. In *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Houston, TX, USA, December 2011.

[104] Y. Zhuang and J. Pan. Probabilistic energy optimization in wireless sensor networks with variable size griding. In *Proceeding of IEEE International Conference on Communications (ICC 2010)*, pages 1–5, Cape Town, South Africa, May 2010. IEEE.

[105] Y. Zhuang and J. Pan. Random distances associated with hexagons. *Arxiv preprint arXiv:1106.2200*, 2011.

[106] Y. Zhuang and J. Pan. Random distances associated with rhombuses. *Arxiv preprint arXiv:1106.1257*, 2011.

[107] Y. Zhuang and J Pan. A geometrical probability approach to location-critical network performance metrics. In *Proceedings of IEEE INFOCOM 2012*, pages 1–9, Orlando, FL, USA, March 2012.

[108] Y. Zhuang, J. Pan, and L. Cai. Minimizing energy consumption with probabilistic distance models in wireless sensor networks. In *Proceedings of IEEE INFOCOM 2010*, pages 1–9, San Diego, CA, USA, March 2010. IEEE.

[109] Y. Zhuang, J. Pan, and L. Cai. A probabilistic model for message propagation in two-dimensional vehicular ad-hoc networks. In *Proceedings of the seventh ACM international workshop on VehiculAr InterNETworking (VANET 2010)*, pages 31–40, Chicago, IL, USA, 2010.

[110] Y. Zhuang, J. Pan, Y. Luo, and L. Cai. Time and location-critical emergency message dissemination for vehicular ad-hoc networks. *IEEE Journal on Selected Areas in Communications*, 29(1):187–196, 2011.

[111] Y. Zhuang, V. Viswanathan, J. Pan, and L. Cai. Upload capacity analysis for drive-thru internet. *Engine*, 2010.

[112] Y. Zhuang, V. Viswanathan, J. Pan, and L. Cai. On the uplink mac performance of a drive-thru internet. *IEEE Transactions on Vehicular Technology*, 2012.

[113] I. Ziedins. Quasi-stationary distributions and one-dimensional circuit-switched networks. *Journal of Applied Probability*, pages 965–977, 1987.

Appendix A

Derivation of the Distance Distribution between Two Adjacent Rhombuses

A.1 Long-Diagonal Case

Section 5.3.3 gives the result for the distance distribution of two adjacent rhombuses when they have a common long diagonal. The derivation of (5.12) is given as follows.

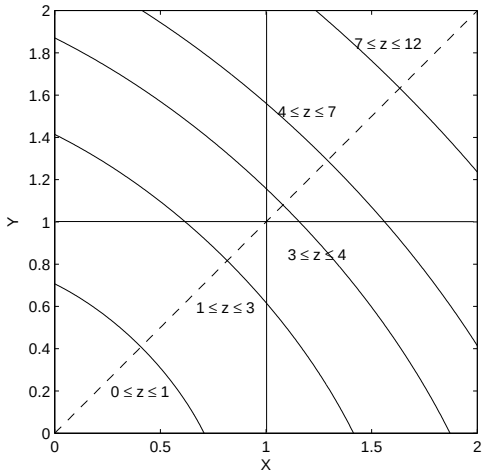


Figure A.1: Five Sub-cases for $Z = |P'Q'|^2$.

When two adjacent rhombuses are sharing a common diagonal, we first consider points P′ and Q′ as shown in Figure 5.3, i.e., the long-diag case. Thus $X = X_1 + X_2$

and $Y = Y_1 + Y_2$ are to be used in (5.6), and their distributions are

$$f_X(x) = \begin{cases} x & 0 \le x \le 1 \\ 2 - x & 1 \le x \le 2 \end{cases}, \quad \text{and} \quad f_Y(y) = \begin{cases} y & 0 \le y \le 1 \\ 2 - y & 1 \le y \le 2 \end{cases}. \tag{A.1}$$

The effective ellipse $Z = X^2 + XY + Y^2$ on the $\mathcal{X}$-$\mathcal{Y}$ plane is now within the boundaries $\mathcal{X} \in [0, 2]$ and $\mathcal{Y} \in [0, 2]$, symmetric with respect to $\mathcal{Y} = \mathcal{X}$, as shown in Figure A.1. Different from the effective ellipses shown in Sections 5.3.1 and 5.3.2, the transitional points when $\mathcal{X} = 1$ and $\mathcal{Y} = 1$ become the tangent lines of the ellipse, are out of the boundary $\mathcal{X} \in [0, 2]$ and $\mathcal{Y} \in [0, 2]$. Thus we have these transitional values of Z: $z_1 = 0$ where the ellipse is a point at the origin, $z_2 = 1$ where the semimajor axis of the ellipse is equal to $\sqrt{2}$, $z_3 = 3$ where the semiminor axis of the ellipse is equal to $\sqrt{2}$, $z_4 = 4$ where the semimajor axis of the ellipse is equal to $\sqrt{8}$, $z_5 = 7$ where the transitional point on the boundary of $f(\omega)$, either $(2, 1)$ or $(1, 2)$, becomes one point on the ellipse, and $z_6 = 12$ where the semiminor axis of the ellipse is equal to $\sqrt{8}$. This leads to the five effective ranges of $[0, 1]$, $[1, 3]$, $[3, 4]$, $[4, 7]$ and $[7, 12]$ as shown in Figure A.1.

$0 \le z \le 1$:

$$F_Z(z) = \int_0^{\sqrt{z}} \int_0^{-\frac{y}{2} + \sqrt{z - \frac{3}{4}y^2}} x dx y dy = \left(\frac{1}{6} - \frac{\pi}{18\sqrt{3}} \right) z^2,$$

and the PDF is

$$f_Z(z) = \left(\frac{1}{3} - \frac{\pi}{9\sqrt{3}} \right) z.$$

$1 \le z \le 3$:

$$\begin{aligned} F_Z(z) &= 2 \left[\int_0^{-\frac{1}{2} + \sqrt{z - \frac{3}{4}}} \int_y^1 x dx y dy + \int_{-\frac{1}{2} + \sqrt{z - \frac{3}{4}}}^{\sqrt{\frac{z}{3}}} \int_y^{-\frac{y}{2} + \sqrt{z - \frac{3}{4}y^2}} x dx y dy \right] \\ &\quad + 2 \int_1^{\sqrt{z}} \int_0^{-\frac{x}{2} + \sqrt{z - \frac{3}{4}x^2}} y dy (2 - x) dx \\ &= \frac{z^2}{3\sqrt{3}} \left(\sin^{-1} \frac{\sqrt{3}}{4} \frac{\sqrt{4z - 3} - 1}{\sqrt{z}} - \sin^{-1} \frac{\sqrt{3}}{2\sqrt{z}} \right) + \left(\frac{\pi}{18\sqrt{3}} - \frac{1}{2} \right) z^2 \\ &\quad + \frac{16}{9} z^{3/2} - z - \frac{10z - 3}{36} \sqrt{4z - 3} + \frac{1}{12}, \end{aligned}$$

and

$$f_Z(z) = \frac{2z}{3\sqrt{3}}\left(\sin^{-1}\frac{\sqrt{3}}{4}\frac{\sqrt{4z-3}-1}{\sqrt{z}} - \sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}}\right) + \left(\frac{\pi}{9\sqrt{3}} - 1\right)z + \frac{8}{3}\sqrt{z} - \frac{\sqrt{4z-3}}{3} - 1.$$

$3 \le z \le 4$:

$$\begin{aligned} F_Z(z) &= \frac{1}{4} + 2\int_0^1 \int_1^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (2-x)dxydy + 2\int_1^{\sqrt{\frac{z}{3}}} \int_y^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (2-x)dx(2-y)dy \\ &= \left(\frac{2z^2}{3\sqrt{3}} - \frac{8z}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{1}{6} - \frac{\pi}{18\sqrt{3}}\right)z^2 + \frac{16}{9}z^{3/2} + \left(1 + \frac{4\pi}{3\sqrt{3}}\right)z \\ &\quad -\frac{6z+3}{4}\sqrt{4z-3} + \frac{19}{12}, \end{aligned}$$

and

$$f_Z(z) = \frac{4}{\sqrt{3}}\left(\frac{z}{3} - 2\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{1}{3} - \frac{\pi}{9\sqrt{3}}\right)z + \frac{8}{3}\sqrt{z} - \frac{7}{3}\sqrt{4z-3} + \frac{4\pi}{3\sqrt{3}} + 1.$$

$4 \le z \le 7$:

$$\begin{aligned} F_Z(z) &= \frac{1}{4} + 2\left[\int_0^{-1+\sqrt{z-3}} \int_1^2 (2-x)dxydy + \int_{-1+\sqrt{z-3}}^1 \int_1^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (2-x)dxydy\right] \\ &\quad +2\int_1^{\sqrt{\frac{z}{3}}} \int_y^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (2-x)dx(2-y)dy \\ &= \left(\frac{2z^2}{3\sqrt{3}} - \frac{8z}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} - \frac{z^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2}\frac{\sqrt{z-3}-1}{\sqrt{z}} + \left(\frac{1}{2} - \frac{\pi}{18\sqrt{3}}\right)z^2 \\ &\quad + \left(3 + \frac{4\pi}{3\sqrt{3}}\right)z - \frac{6z+3}{4}\sqrt{4z-3} + \frac{5z-6}{9}\sqrt{z-3} + \frac{11}{12}, \end{aligned}$$

and

$$\begin{aligned} f_Z(z) &= \frac{4}{\sqrt{3}}\left(\frac{z}{3} - 2\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} - \frac{2z}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2}\frac{\sqrt{z-3}-1}{\sqrt{z}} + \left(1 - \frac{\pi}{9\sqrt{3}}\right)z \\ &\quad -\frac{7}{3}\sqrt{4z-3} + \frac{2}{3}\sqrt{z-3} + \frac{4\pi}{3\sqrt{3}} + 3. \end{aligned}$$

$7 \le z \le 12$:

$$\begin{aligned} F_Z(z) &= \frac{3}{4} + 2\int_1^{-1+\sqrt{z-3}} \int_y^2 (2-x)dx(2-y)dy + 2\int_{-1+\sqrt{z-3}}^{\sqrt{\frac{z}{3}}} \int_y^{-\frac{y}{2}+\sqrt{z-\frac{3}{4}y^2}} (2-x)dx(2-y)dy \\ &= \left(\frac{z^2}{3\sqrt{3}} - \frac{8z}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2}\frac{\sqrt{z-3}-1}{\sqrt{z}} - \left(\frac{1}{6} + \frac{\pi}{18\sqrt{3}}\right)z^2 + \left(\frac{4\pi}{3\sqrt{3}} - 2\right)z \\ &\quad + \frac{11z+30}{9}\sqrt{z-3} - 5, \end{aligned}$$

and

$$f_Z(z) = \frac{2}{\sqrt{3}}\left(\frac{z}{3} - 4\right)\sin^{-1}\frac{\sqrt{3}}{2}\frac{\sqrt{z-3}-1}{\sqrt{z}} - \left(\frac{\pi}{9\sqrt{3}} + \frac{1}{3}\right)z + 2\sqrt{z-3} + \frac{4\pi}{3\sqrt{3}} - 2.$$

The final result in (5.12) is then derived by (5.8).

A.2 Short-Diagonal Case

Section 5.3.3 also gives the result for the distance distribution of two adjacent rhombuses when they have a common short diagonal. The derivation of (5.13) is given as follows.

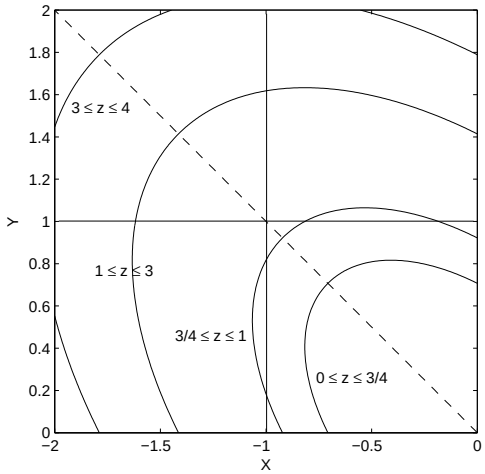


Figure A.2: Four Sub-cases for $Z = |M'N'|^2$.

When two adjacent rhombuses are sharing a common diagonal, the second case is points M′ and N′ as shown in Figure 5.3, i.e., the short-diag case. Thus $X = -X_1 - X_2$

and $Y = Y_1 + Y_2$ are to be used in (5.6), and

$$f_X(x) = \begin{cases} 2+x & -2 \le x \le -1 \\ -x & -1 \le x \le 0 \end{cases}, \quad \text{and} \quad f_Y(y) = \begin{cases} y & 0 \le y \le 1 \\ 2-y & 1 \le y \le 2 \end{cases}. \quad \text{(A.2)}$$

The effective ellipse on the $\mathcal{X}$-$\mathcal{Y}$ plane, in this case, is in $\mathcal{X} \in [-2, 0]$ and $\mathcal{Y} \in [0, 2]$, symmetric with respect to $\mathcal{Y} = -\mathcal{X}$. Similar to Section 5.3.2 but with one less sub-case, there are $z_1 = 0$ where the ellipse is a point at the origin, $z_2 = \frac{3}{4}$ where $\mathcal{X} = -1$ and $\mathcal{Y} = 1$ are the tangent lines of the ellipse, $z_3 = 1$ where the semimajor axis of the ellipse is equal to $\sqrt{2}$, $z_4 = 3$ where the semiminor axis of the ellipse is equal to $\sqrt{2}$, and $z_5 = 4$ where the semimajor axis of the ellipse is equal to $\sqrt{8}$. Therefore depending on the value of Z, we have four sub-cases as shown in Figure A.2.

$0 \le z \le \frac{3}{4}$:

$$F_Z(z) = 2\int_{-\sqrt{z}}^{0}\int_{-x}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} y dy(-x)dx = \left(\frac{1}{6} + \frac{\pi}{9\sqrt{3}}\right) z^2,$$

and the PDF is

$$f_Z(z) = \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right) z.$$

$\frac{3}{4} \le z \le 1$:

$$\begin{aligned} F_Z(z) &= 2\left[\int_{-\sqrt{z}}^{0}\int_{-x}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} y dy(-x)dx - \int_{-\frac{1}{2}-\sqrt{z-\frac{3}{4}}}^{-\frac{1}{2}+\sqrt{z-\frac{3}{4}}}\int_{1}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} y dy(-x)dx\right] \\ &\quad +2\int_{-\frac{1}{2}-\sqrt{z-\frac{3}{4}}}^{-\frac{1}{2}+\sqrt{z-\frac{3}{4}}}\int_{1}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} (2-y)dy(-x)dx \\ &= -\frac{2z^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2}\frac{\sqrt{4z-3}}{z} + \left(\frac{1}{6} + \frac{\pi}{9\sqrt{3}}\right) z^2 + \frac{10z-3}{18}\sqrt{4z-3}, \end{aligned}$$

and

$$f_Z(z) = -\frac{4z}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2}\frac{\sqrt{4z-3}}{z} + \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right) z + \frac{2}{3}\sqrt{4z-3}.$$

$1 \le z \le 3$:

$$\begin{aligned} F_Z(z) &= \frac{1}{4} + 2\int_{-1}^{0}\int_{1}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} (2-y)dy(-x)dx + 2\int_{-\sqrt{z}}^{-1}\int_{-x}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}} (2-y)dy(2+x)dx \\ &= -\left(\frac{2z^2}{3\sqrt{3}} + \frac{8z}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{1}{6} + \frac{\pi}{9\sqrt{3}}\right)z^2 + \frac{16}{9}z^{3/2} + \left(\frac{8\pi}{3\sqrt{3}} + 1\right)z \\ &\quad -\frac{74z+21}{36}\sqrt{4z-3} + \frac{1}{4}, \end{aligned}$$

and

$$f_Z(z) = -\left(\frac{4z}{3\sqrt{3}} + \frac{8}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2\sqrt{z}} + \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right)z + \frac{8}{3}\sqrt{z} - 3\sqrt{4z-3} + \frac{8\pi}{3\sqrt{3}} + 1.$$

$3 \le z \le 4$:

$$\begin{aligned} F_Z(z) &= \frac{1}{4} + 2\left[\int_{-1}^{-1+\sqrt{z-3}}\int_{1}^{2}(2-y)dy(-x)dx + \int_{-1+\sqrt{z-3}}^{0}\int_{1}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}}(2-y)dy(-x)dx\right] \\ &\quad +2\left[\int_{-\sqrt{z}}^{-1-\sqrt{z-3}}\int_{-x}^{-\frac{x}{2}+\sqrt{z-\frac{3}{4}x^2}}(2-y)dy(2+x)dx + \int_{-1-\sqrt{z-3}}^{-1}\int_{-x}^{2}(2-y)dy(2+x)dx\right] \\ &= -\frac{8z}{\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2}\frac{\sqrt{z-3}+1}{\sqrt{z}} - \frac{z^2}{2} + \frac{16}{9}z^{3/2} + \left(\frac{8\pi}{3\sqrt{3}} - 4\right)z + \frac{16z+24}{9}\sqrt{z-3} + 1, \end{aligned}$$

and

$$f_Z(z) = -\frac{8}{\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2}\frac{\sqrt{z-3}+1}{\sqrt{z}} - z + \frac{8}{3}\sqrt{z} + \frac{8}{3}\sqrt{z-3} + \frac{8\pi}{3\sqrt{3}} - 4.$$

The final result in (5.13) is then derived by (5.8).

Appendix B

Validation of Rhombus-Related Distributions by Recursion

Here the purpose is to verify that $\frac{1}{4}f_{D_{\rm I}}(d)+\frac{1}{2}f_{D_{\rm P}}(d)+\frac{1}{8}f_{D_{\rm LD}}(d)+\frac{1}{8}f_{D_{\rm SD}}(d) = \frac{1}{2}f_{D_{\rm I}}(\frac{d}{2})$, where $f_{D_{\rm I}}(d)$, $f_{D_{\rm P}}(d)$, $f_{D_{\rm LD}}(d)$ and $f_{D_{\rm SD}}(d)$ are given in (5.9), (5.11), (5.12) and (5.13) for distances associated with rhombuses, respectively. To confirm that the above two equalities of $f_{2D_{\rm I}}(d)$ are equivalent, we have verified them mathematically as follows.

$0 \le d \le \frac{\sqrt{3}}{2}$:

$$\frac{1}{4}f_{D_{\rm I}}(d) = \frac{d}{2}\left[\left(\frac{4}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{16}{3}d + \frac{2\pi}{\sqrt{3}}\right], \quad \frac{1}{2}f_{D_{\rm P}}(d) = d\left[\frac{4}{3}d - \left(\frac{2}{3}+\frac{\pi}{9\sqrt{3}}\right)d^2\right],$$

$$\frac{1}{8}f_{D_{\rm LD}}(d) = \frac{d}{4}\left(\frac{1}{3}-\frac{\pi}{9\sqrt{3}}\right)d^2, \quad \frac{1}{8}f_{D_{\rm SD}}(d) = \frac{d}{4}\left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2.$$

Thus,

$$\begin{aligned} f_{2D_{\rm I}}(d) &= \frac{1}{4}f_{D_{\rm I}}(d) + \frac{1}{2}f_{D_{\rm P}}(d) + \frac{1}{8}f_{D_{\rm LD}}(d) + \frac{1}{8}f_{D_{\rm SD}}(d) \\ &= \frac{d}{2}\left[\left(\frac{4}{3}+\frac{2\pi}{9\sqrt{3}}\right)\left(\frac{d}{2}\right)^2 - \frac{16}{3}\left(\frac{d}{2}\right) + \frac{2\pi}{\sqrt{3}}\right] = \frac{1}{2}f_{D_{\rm I}}(\frac{d}{2}). \end{aligned}$$

$\frac{\sqrt{3}}{2} \leq d \leq 1$:

$$\frac{1}{4} f_{D_{\mathrm{I}}}(d) = \frac{d}{2}\left[\frac{8}{\sqrt{3}}\left(1+\frac{d^2}{3}\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{4}{3} - \frac{10\pi}{9\sqrt{3}}\right)d^2 - \frac{16}{3}d + \frac{10}{3}\sqrt{4d^2-3} - \frac{2\pi}{\sqrt{3}}\right],$$

$$\frac{1}{2} f_{D_{\mathrm{P}}}(d) = d\left[-\frac{2}{\sqrt{3}}(d^2+2)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{8\pi}{9\sqrt{3}} - \frac{2}{3}\right)d^2 + \frac{4}{3}d - \frac{11}{6}\sqrt{4d^2-3} + \frac{2\pi}{\sqrt{3}}\right],$$

$$\frac{1}{8} f_{D_{\mathrm{LD}}}(d) = \frac{d}{4}\left(\frac{1}{3} - \frac{\pi}{9\sqrt{3}}\right)d^2, \quad \frac{1}{8} f_{D_{\mathrm{SD}}}(d) = \frac{d}{4}\left[\frac{8d^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{1}{3} - \frac{10\pi}{9\sqrt{3}}\right)d^2 + \frac{2}{3}\sqrt{4d^2-3}\right].$$

Thus,

$$\begin{aligned} f_{2D_{\mathrm{I}}}(d) &= \frac{1}{4} f_{D_{\mathrm{I}}}(d) + \frac{1}{2} f_{D_{\mathrm{P}}}(d) + \frac{1}{8} f_{D_{\mathrm{LD}}}(d) + \frac{1}{8} f_{D_{\mathrm{SD}}}(d) \\ &= \frac{d}{2}\left[\left(\frac{4}{3} + \frac{2\pi}{9\sqrt{3}}\right)\left(\frac{d}{2}\right)^2 - \frac{16}{3}\left(\frac{d}{2}\right) + \frac{2\pi}{\sqrt{3}}\right] = \frac{1}{2} f_{D_{\mathrm{I}}}(\frac{d}{2}). \end{aligned}$$

$1 \leq d \leq \sqrt{3}$:

$$\frac{1}{4} f_{D_{\mathrm{I}}}(d) = \frac{d}{2}\left[\frac{4}{\sqrt{3}}\left(1-\frac{d^2}{3}\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{2}{3} - \frac{2\pi}{9\sqrt{3}}\right)d^2 + \sqrt{4d^2-3} - \frac{2\pi}{3\sqrt{3}} - 1\right],$$

$$\frac{1}{2} f_{D_{\mathrm{P}}}(d) = d\left[\frac{4d^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{2}{3} - \frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{8}{3}d + \frac{\sqrt{4d^2-3}}{3} + \frac{2\pi}{3\sqrt{3}} + \frac{1}{2}\right],$$

$$\frac{1}{8} f_{D_{\mathrm{LD}}}(d) = \frac{d}{4}\left[-\frac{4d^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{\pi}{3\sqrt{3}} - 1\right)d^2 + \frac{8}{3}d - \frac{\sqrt{4d^2-3}}{3} - 1\right],$$

$$\frac{1}{8} f_{D_{\mathrm{SD}}}(d) = \frac{d}{4}\left[-\left(\frac{4d^2}{3\sqrt{3}} + \frac{8}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{1}{3} + \frac{2\pi}{9\sqrt{3}}\right)d^2 + \frac{8}{3}d - 3\sqrt{4d^2-3} + \frac{8\pi}{3\sqrt{3}} + 1\right].$$

Thus,

$$\begin{aligned} f_{2D_{\rm I}}(d) &= \frac{1}{4}f_{D_{\rm I}}(d)+\frac{1}{2}f_{D_{\rm P}}(d)+\frac{1}{8}f_{D_{\rm LD}}(d)+\frac{1}{8}f_{D_{\rm SD}}(d) \\ &= \frac{d}{2}\left[\left(\frac{4}{3}+\frac{2\pi}{9\sqrt{3}}\right)\left(\frac{d}{2}\right)^2-\frac{16}{3}\left(\frac{d}{2}\right)+\frac{2\pi}{\sqrt{3}}\right]=\frac{1}{2}f_{D_{\rm I}}(\frac{d}{2}). \end{aligned}$$

$\sqrt{3}\le d\le 2$:

$$f_{D_{\rm I}}(d)=0,$$

$$\begin{aligned} \frac{1}{2}f_{D_{\rm P}}(d) &= d\left[\left(\frac{2}{\sqrt{3}}-\frac{d^2}{3\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d}+\left(\frac{2}{\sqrt{3}}+\frac{d^2}{3\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{d}+\left(\frac{1}{3}-\frac{\pi}{9\sqrt{3}}\right)d^2-\frac{8}{3}d\right. \\ &\quad \left.+\sqrt{d^2-3}+\frac{7}{12}\sqrt{4d^2-3}+\frac{3}{4}-\frac{2\pi}{3\sqrt{3}}\right], \end{aligned}$$

$$\frac{1}{8}f_{D_{\rm LD}}(d)=\frac{d}{4}\left[\frac{4}{\sqrt{3}}\left(\frac{d^2}{3}-2\right)\sin^{-1}\frac{\sqrt{3}}{2d}+\left(\frac{1}{3}-\frac{\pi}{9\sqrt{3}}\right)d^2-\frac{7}{3}\sqrt{4d^2-3}+\frac{8}{3}d+\frac{4\pi}{3\sqrt{3}}+1\right],$$

$$\frac{1}{8}f_{D_{\rm SD}}(d)=\frac{d}{4}\left[\frac{8}{\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{d}-d^2+\frac{8}{3}d+\frac{8}{3}\sqrt{d^2-3}-\frac{8\pi}{3\sqrt{3}}-4\right].$$

Therefore,

$$\begin{aligned} f_{2D_{\rm I}}(d) &= \frac{1}{2}f_{D_{\rm P}}(d)+\frac{1}{8}f_{D_{\rm LD}}(d)+\frac{1}{8}f_{D_{\rm SD}}(d) \\ &= \frac{d}{2}\left[\frac{8}{\sqrt{3}}\left(1+\frac{(d/2)^2}{3}\right)\sin^{-1}\frac{\sqrt{3}}{2(d/2)}+\left(\frac{4}{3}-\frac{10\pi}{9\sqrt{3}}\right)\left(\frac{d}{2}\right)^2-\frac{16}{3}\left(\frac{d}{2}\right)\right. \\ &\quad \left.+\frac{10}{3}\sqrt{4\left(\frac{d}{2}\right)^2-3}-\frac{2\pi}{\sqrt{3}}\right]=\frac{1}{2}f_{D_{\rm I}}(\frac{d}{2}). \end{aligned}$$

$2\le d\le\sqrt{7}$:

$$f_{D_{\rm I}}(d)=f_{D_{\rm SD}}(d)=0,$$

$$\frac{1}{2}f_{D_{\rm P}}(d) = d\left[\left(\frac{2}{\sqrt{3}}-\frac{d^2}{3\sqrt{3}}\right)\left(\sin^{-1}\frac{\sqrt{3}}{2d}+\sin^{-1}\frac{\sqrt{3}}{d}\right)+\left(\frac{\pi}{9\sqrt{3}}-\frac{1}{3}\right)d^2+\frac{7}{12}\sqrt{4d^2-3}\right.$$
$$\left.+\frac{\sqrt{d^2-3}}{3}-\frac{2\pi}{3\sqrt{3}}-\frac{5}{4}\right],$$

$$\frac{1}{8}f_{D_{\rm LD}}(d) = \frac{d}{4}\left[\frac{4}{\sqrt{3}}\left(\frac{d^2}{3}-2\right)\sin^{-1}\frac{\sqrt{3}}{2d}+\frac{2d^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{d}+\left(1-\frac{\pi}{3\sqrt{3}}\right)d^2-\frac{7}{3}\sqrt{4d^2-3}\right.$$
$$\left.+\frac{2}{3}\sqrt{d^2-3}+\frac{4\pi}{3\sqrt{3}}+3\right].$$

Therefore,

$$\begin{aligned} f_{2D_{\rm I}}(d) &= \frac{1}{2}f_{D_{\rm P}}(d)+\frac{1}{8}f_{D_{\rm LD}}(d) \\ &= \frac{d}{2}\left[\frac{4}{\sqrt{3}}\left(1-\frac{(d/2)^2}{3}\right)\sin^{-1}\frac{\sqrt{3}}{d}-\left(\frac{2}{3}-\frac{2\pi}{9\sqrt{3}}\right)\left(\frac{d}{2}\right)^2+\sqrt{4\left(\frac{d}{2}\right)^2-3}-\frac{2\pi}{3\sqrt{3}}-1\right] \\ &= \frac{1}{2}f_{D_{\rm I}}(\frac{d}{2}). \end{aligned}$$

$\sqrt{7} \le d \le 2\sqrt{3}$:

$$f_{D_{\rm I}}(d) = f_{D_{\rm P}}(d) = f_{D_{\rm SD}}(d) = 0,$$

$$\begin{aligned} f_{2D_{\rm I}}(d) &= \frac{1}{8}f_{D_{\rm LD}}(d) = \frac{d}{4}\left[\frac{2}{\sqrt{3}}\left(4-\frac{d^2}{3}\right)\sin^{-1}\frac{\sqrt{3}}{d}+\left(\frac{\pi}{9\sqrt{3}}-\frac{1}{3}\right)d^2+2\sqrt{d^2-3}-\frac{4\pi}{3\sqrt{3}}-2\right] \\ &= \frac{d}{2}\left[\frac{4}{\sqrt{3}}\left(1-\frac{(d/2)^2}{3}\right)\sin^{-1}\frac{\sqrt{3}}{d}-\left(\frac{2}{3}-\frac{2\pi}{9\sqrt{3}}\right)\left(\frac{d}{2}\right)^2+\sqrt{4\left(\frac{d}{2}\right)^2-3}-\frac{2\pi}{3\sqrt{3}}-1\right] \\ &= \frac{1}{2}f_{D_{\rm I}}(\frac{d}{2}). \end{aligned}$$

In summary, we have $f_{2D_{\rm I}}(d) = \frac{1}{2}f_{D_{\rm I}}(\frac{d}{2})$ by scaling, and the probabilistic sum $\frac{1}{4}f_{D_{\rm I}}(d)+\frac{1}{2}f_{D_{\rm P}}(d)+\frac{1}{8}f_{D_{\rm LD}}(d)+\frac{1}{8}f_{D_{\rm SD}}(d)$ is equal to $\frac{1}{2}f_{D_{\rm I}}(\frac{d}{2})$ in all the cases discussed above. The above results are a strong validation of the approach that we proposed, and the correctness of the distance distributions that we derived for distances associ-

ated with rhombuses.

Appendix C

Validation of Hexagon-Related Distributions by Recursion

In this appendix, the mathematical verification of all the distribution functions derived in Chapter 6 is given, through the scaling $f_{2D}(d) = \frac{1}{2} f_D(\frac{d}{2})$, and the recursive probabilistic sum in (6.12):

$$\begin{aligned} f_{2D_{\mathrm{H_I}}}(d) &= \frac{3}{16} f_{D_{\mathrm{H_I}}}(d) + \frac{3}{8} f_{D_{\mathrm{H_A}}}(d) + \frac{1}{12} \left[f_{D_{\mathrm{PX}}}(d) + f_{D_2}(d) + f_{D_1}(d) + f_{D_3}(d) \right] \\ &\quad + \frac{1}{24} \left[f_{D_4}(d) + f_{D_5}(d) \right] + \frac{1}{48} f_{D_{\mathrm{I}}}(d), \end{aligned}$$

$f_{D_5}(d)$ is the distance density function of $|AL|$ in Figure 6.11, which is the only distance distribution function that has not been given yet. We will do so in the immediate following and the result in give in (C.1). The rest of the above notations for distance distributions, and the corresponding equation number in this dissertation, are listed in Table C.1.

Table C.1: Distance Distributions and the Corresponding Equation Number

Distance Distribution	$f_{D_{\mathrm{H_I}}}(d)$	$f_{D_{\mathrm{H_A}}}(d)$	$f_{D_{\mathrm{PX}}}(d)$	$f_{D_{\mathrm{I}}}(d)$
Equation Number	(6.5)	(6.11)	(6.4)	(5.9)
Distance Distribution	$f_{D_1}(d)$	$f_{D_2}(d)$	$f_{D_3}(d)$	$f_{D_4}(d)$
Equation Number	(6.7)	(6.8)	(6.9)	(6.10)

$|AL|$

$$f_{D_5}(d) = 2d \begin{cases} -\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+4\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{10}{3}\sqrt{d^2-3} & \\ +\frac{8\pi}{3\sqrt{3}}+2 & 2 \le d \le \sqrt{7} \\ \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+8\right)\sin^{-1}\frac{\sqrt{3}}{d} + \frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+6\right)\sin^{-1}\frac{3\sqrt{3}}{2d} & \\ \quad -\left(1+\frac{2\pi}{3\sqrt{3}}\right)d^2 + 6\sqrt{d^2-3} + \frac{11}{3}\sqrt{4d^2-27} - \frac{40\pi}{3\sqrt{3}} - 11 & \sqrt{7} \le d \le 3 \\ \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+8\right)\sin^{-1}\frac{\sqrt{3}}{d} - \frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+12\right)\sin^{-1}\frac{3\sqrt{3}}{2d} & \\ \quad +\left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 + 6\sqrt{d^2-3} - \frac{19}{3}\sqrt{4d^2-27} + \frac{32\pi}{3\sqrt{3}} + 7 & 3 \le d \le 2\sqrt{3} \\ -\frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+12\right)\sin^{-1}\frac{3\sqrt{3}}{2d} - \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+8\right)\sin^{-1}\frac{2\sqrt{3}}{d} & \\ \quad +\left(1+\frac{2\pi}{3\sqrt{3}}\right)d^2 - \frac{19}{3}\sqrt{4d^2-27} - 4\sqrt{d^2-12} + \frac{64\pi}{3\sqrt{3}} + 17 & 2\sqrt{3} \le d \le \sqrt{13} \\ \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+16\right)\sin^{-1}\frac{2\sqrt{3}}{d} - \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 + \frac{20}{3}\sqrt{d^2-12} & \\ \quad -\frac{32\pi}{3\sqrt{3}} - 8 & \sqrt{13} \le d \le 4 \\ 0 & \text{otherwise} \end{cases} \quad \text{(C.1)}$$

Validation

In order to confirm that the two definitions of $f_{2D_{\mathrm{H_I}}}(d)$ at the beginning of this section are equivalent, i.e., $\frac{3}{16}f_{D_{\mathrm{H_I}}}(d) + \frac{3}{8}f_{D_{\mathrm{H_A}}}(d) + \frac{1}{12}\left[f_{D_{\mathrm{PX}}}(d) + f_{D_2}(d) + f_{D_1}(d) + f_{D_3}(d)\right] + \frac{1}{24}\left[f_{D_4}(d) + f_{D_5}(d)\right] + \frac{1}{48}f_{D_{\mathrm{I}}}(d)$ is equal to $\frac{1}{2}f_{D_{\mathrm{H_I}}}(\frac{d}{2})$, we verify them mathematically as follows.

i) $0 \le d \le \frac{\sqrt{3}}{2}$:

$$\frac{3}{16}f_{D_{\mathrm{H_I}}}(d) = \frac{d}{8}\left[\left(\frac{2}{3}-\frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{8}{3}d + \frac{2\pi}{\sqrt{3}}\right], \quad \frac{3}{8}f_{D_{\mathrm{H_A}}}(d) = \frac{d}{12}\left[\left(\frac{\pi}{9\sqrt{3}}-\frac{1}{3}\right)d^2 + \frac{4}{3}d\right],$$

$$\frac{1}{12}f_{D_{\mathrm{PX}}}(d) = \frac{d}{6}\left[\frac{4}{3}d - \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2\right], \quad \frac{1}{12}f_{D_2}(d) = \frac{d}{6}\left[\left(\frac{2\pi}{9\sqrt{3}}-\frac{1}{6}\right)d^2\right],$$

$$\frac{1}{48}f_{D_{\mathrm{I}}}(d) = \frac{d}{24}\left[\left(\frac{4}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{16}{3}d + \frac{2\pi}{\sqrt{3}}\right],$$

while the probability density functions are 0 for all other cases. Thus,

$$\begin{aligned} f_{2D_{\mathrm{H_I}}}(d) &= \frac{3}{16}f_{D_{\mathrm{H_I}}}(d) + \frac{3}{8}f_{D_{\mathrm{H_A}}}(d) + \frac{1}{12}\left[f_{D_{\mathrm{PX}}}(d) + f_{D_2}(d)\right] + \frac{1}{48}f_{D_{\mathrm{I}}}(d) \\ &= \frac{d}{6}\left[\left(\frac{1}{6} - \frac{\pi}{18\sqrt{3}}\right)d^2 - \frac{4}{3}d + \frac{2\pi}{\sqrt{3}}\right] \\ &= \frac{d}{6}\left[\left(\frac{2}{3} - \frac{2\pi}{9\sqrt{3}}\right)\left(\frac{d}{2}\right)^2 - \frac{8}{3}\left(\frac{d}{2}\right) + \frac{2\pi}{\sqrt{3}}\right] = \frac{1}{2}f_{D_{\mathrm{H_I}}}(\frac{d}{2}). \end{aligned}$$

ii) $\frac{\sqrt{3}}{2} \le d \le 1$:

$$\frac{3}{16}f_{D_{\mathrm{H_I}}}(d) = \frac{d}{8}\left[\left(\frac{2}{3} - \frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{8}{3}d + \frac{2\pi}{\sqrt{3}}\right], \quad \frac{3}{8}f_{D_{\mathrm{H_A}}}(d) = \frac{d}{12}\left[\left(\frac{\pi}{9\sqrt{3}} - \frac{1}{3}\right)d^2 + \frac{4}{3}d\right],$$

$$\frac{1}{12}f_{D_{\mathrm{PX}}}(d) = \frac{d}{6}\left[-\frac{4}{\sqrt{3}}\left(\frac{d^2}{3} + 1\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{4\pi}{9\sqrt{3}} - \frac{1}{3}\right)d^2 + \frac{4}{3}d - \frac{5}{3}\sqrt{4d^2 - 3} + \frac{2\pi}{\sqrt{3}}\right],$$

$$\frac{1}{12}f_{D_2}(d) = \frac{d}{6}\left[\frac{2}{\sqrt{3}}\left(\frac{2d^2}{3} + 1\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{4\pi}{9\sqrt{3}} + \frac{1}{6}\right)d^2 + \sqrt{4d^2 - 3} - \frac{\pi}{\sqrt{3}}\right],$$

$$\frac{1}{12}f_{D_1}(d) = \frac{d}{6}\left[-\frac{2d^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2d} + \frac{\pi}{3\sqrt{3}}d^2 - \frac{\sqrt{4d^2 - 3}}{6}\right],$$

$$\frac{1}{48}f_{D_{\mathrm{I}}}(d) = \frac{d}{24}\left[\frac{8}{\sqrt{3}}\left(1 + \frac{d^2}{3}\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{4}{3} - \frac{10\pi}{9\sqrt{3}}\right)d^2 - \frac{16}{3}d + \frac{10}{3}\sqrt{4d^2 - 3} - \frac{2\pi}{\sqrt{3}}\right].$$

Thus,

$$\begin{aligned} f_{2D_{\mathrm{H_I}}}(d) &= \frac{3}{16}f_{D_{\mathrm{H_I}}}(d) + \frac{3}{8}f_{D_{\mathrm{H_A}}}(d) + \frac{1}{12}\left[f_{D_{\mathrm{PX}}}(d) + f_{D_2}(d) + f_{D_1}(d)\right] + \frac{1}{48}f_{D_{\mathrm{I}}}(d) \\ &= \frac{d}{6}\left[\left(\frac{2}{3} - \frac{2\pi}{9\sqrt{3}}\right)\left(\frac{d}{2}\right)^2 - \frac{8}{3}\left(\frac{d}{2}\right) + \frac{2\pi}{\sqrt{3}}\right] = \frac{1}{2}f_{D_{\mathrm{H_I}}}(\frac{d}{2}). \end{aligned}$$

iii) $1 \le d \le \sqrt{3}$:

$$\frac{3}{16} f_{D_{\mathrm{H_I}}}(d) = \frac{d}{8}\left[-\frac{4}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \frac{2\pi}{3\sqrt{3}}d^2 - 2\sqrt{4d^2-3} + \frac{10\pi}{3\sqrt{3}}\right],$$

$$\frac{3}{8} f_{D_{\mathrm{H_A}}}(d) = \frac{d}{12}\left[\frac{2}{\sqrt{3}}(d^2+2)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{1}{3}+\frac{5\pi}{9\sqrt{3}}\right)d^2 + \frac{11}{6}\sqrt{4d^2-3} - \frac{4\pi}{3\sqrt{3}} - \frac{1}{2}\right],$$

$$\frac{1}{12} f_{D_{\mathrm{PX}}}(d) = \frac{d}{6}\left[-\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{2\pi}{9\sqrt{3}}+\frac{1}{3}\right)d^2 - \frac{3}{2}\sqrt{4d^2-3} + \frac{2\pi}{\sqrt{3}} + \frac{1}{2}\right],$$

$$\frac{1}{12} f_{D_2}(d) = \frac{d}{6}\left[\frac{1}{\sqrt{3}}\left(\frac{4d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{1}{3}-\frac{4\pi}{9\sqrt{3}}\right)d^2 + \frac{2}{3}\sqrt{4d^2-3} - \frac{2}{3}d - \frac{2\pi}{3\sqrt{3}} + \frac{1}{2}\right],$$

$$\frac{1}{12} f_{D_1}(d) = \frac{d}{6}\left[\frac{d^2+4}{\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{2}{3}d + \frac{19}{12}\sqrt{4d^2-3} - \frac{4\pi}{3\sqrt{3}} - \frac{3}{4}\right],$$

$$\frac{1}{12} f_{D_3}(d) = \frac{d}{6}\left[-\left(\frac{d^2}{3\sqrt{3}}+\frac{1}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{1}{6}+\frac{\pi}{9\sqrt{3}}\right)d^2 - \frac{5}{12}\sqrt{4d^2-3} + \frac{\pi}{3\sqrt{3}} + \frac{1}{4}\right],$$

$$\frac{1}{48} f_{D_{\mathrm{I}}}(d) = \frac{d}{24}\left[\frac{4}{\sqrt{3}}\left(1-\frac{d^2}{3}\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{2}{3}-\frac{2\pi}{9\sqrt{3}}\right)d^2 + \sqrt{4d^2-3} - \frac{2\pi}{3\sqrt{3}} - 1\right].$$

Thus,

$$\begin{aligned} f_{2D_{\mathrm{H_I}}}(d) &= \frac{3}{16} f_{D_{\mathrm{H_I}}}(d) + \frac{3}{8} f_{D_{\mathrm{H_A}}}(d) + \frac{1}{12}\left[f_{D_{\mathrm{PX}}}(d) + f_{D_2}(d) + f_{D_1}(d) + f_{D_3}(d)\right] \\ &\quad + \frac{1}{48} f_{D_{\mathrm{I}}}(d) = \frac{d}{6}\left[\left(\frac{2}{3}-\frac{2\pi}{9\sqrt{3}}\right)\left(\frac{d}{2}\right)^2 - \frac{8}{3}\left(\frac{d}{2}\right) + \frac{2\pi}{\sqrt{3}}\right] = \frac{1}{2} f_{D_{\mathrm{H_I}}}(\frac{d}{2}). \end{aligned}$$

iv) $\sqrt{3} \le d \le 2$:

$$\frac{3}{16} f_{D_{\mathrm{H_I}}}(d) = \frac{d}{8}\left[\frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+4\right)\sin^{-1}\frac{\sqrt{3}}{d} - \left(\frac{4\pi}{9\sqrt{3}}+\frac{2}{3}\right)d^2 + \frac{20}{3}\sqrt{d^2-3} - \frac{16\pi}{3\sqrt{3}} - 4\right],$$

$$\frac{3}{8}f_{D_{\mathrm{H_A}}}(d) = \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}-2\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+4\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(1+\frac{\pi}{3\sqrt{3}}\right)d^2 - \frac{7}{6}\sqrt{4d^2-3} - \frac{20}{3}\sqrt{d^2-3} + \frac{8\pi}{\sqrt{3}} + \frac{9}{2}\right],$$

$$\frac{1}{12}f_{D_{\mathrm{PX}}}(d) = \frac{d}{6}\left[\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+4\right)\sin^{-1}\frac{\sqrt{3}}{d} - \left(\frac{2\pi}{9\sqrt{3}}+\frac{1}{3}\right)d^2 + \frac{10}{3}\sqrt{d^2-3} - \frac{8\pi}{3\sqrt{3}} - 2\right],$$

$$\frac{1}{12}f_{D_2}(d) = \frac{d}{6}\left[\frac{5}{\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{4\pi}{9\sqrt{3}}-\frac{1}{3}\right)d^2 + \frac{5}{3}\sqrt{4d^2-3} - 4\sqrt{d^2-3} - \frac{2}{3}d + \frac{8\pi}{3\sqrt{3}} - \frac{1}{2}\right],$$

$$\frac{1}{12}f_{D_1}(d) = \frac{d}{6}\left[-\left(\frac{d^2}{3\sqrt{3}}+\frac{10}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{d} - \left(\frac{2d^2}{3\sqrt{3}}+2\sqrt{3}\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{4}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{13}{6}\sqrt{4d^2-3} - \frac{11}{3}\sqrt{d^2-3} - \frac{2}{3}d + \frac{16\pi}{3\sqrt{3}} + \frac{11}{2}\right],$$

$$\frac{1}{12}f_{D_3}(d) = \frac{d}{6}\left[\left(\frac{2d^2}{3\sqrt{3}}+\frac{4}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{2}{\sqrt{3}}-\frac{d^2}{3\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 + \frac{7}{12}\sqrt{4d^2-3} + 2\sqrt{d^2-3} - \frac{13\pi}{6\sqrt{3}} - \frac{5}{4}\right],$$

$$\frac{1}{24}f_{D_4}(d) = \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \frac{4}{\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{d} - \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 + \sqrt{4d^2-3} + \frac{4}{3}\sqrt{d^2-3} - \frac{7\pi}{3\sqrt{3}} - 2\right].$$

Thus,

$$f_{2D_{\mathrm{H_I}}}(d) = \frac{3}{16}f_{D_{\mathrm{H_I}}}(d) + \frac{3}{8}f_{D_{\mathrm{H_A}}}(d) + \frac{1}{12}\left[f_{D_{\mathrm{PX}}}(d) + f_{D_2}(d) + f_{D_1}(d) + f_{D_3}(d)\right] + \frac{1}{24}f_{D_4}(d) = \frac{d}{6}\left[\left(\frac{2}{3}-\frac{2\pi}{9\sqrt{3}}\right)\left(\frac{d}{2}\right)^2 - \frac{8}{3}\left(\frac{d}{2}\right) + \frac{2\pi}{\sqrt{3}}\right] = \frac{1}{2}f_{D_{\mathrm{H_I}}}(\frac{d}{2}).$$

v) $2 \leq d \leq \frac{3\sqrt{3}}{2}$:

$$\frac{3}{8} f_{D_{\mathrm{H_A}}}(d) = \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}-2\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{1}{3}-\frac{\pi}{9\sqrt{3}}\right)d^2 - \frac{7}{6}\sqrt{4d^2-3} + \frac{8\pi}{3\sqrt{3}} + \frac{1}{2}\right],$$

$$\begin{aligned}\frac{1}{12} f_{D_2}(d) &= \frac{d}{6}\left[\frac{5}{\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{2\pi}{9\sqrt{3}}-\frac{1}{6}\right)d^2\right. \\ &\quad \left. + \frac{5}{3}\sqrt{4d^2-3} - 2\sqrt{d^2-3} - 2d + \frac{4\pi}{3\sqrt{3}} - \frac{1}{2}\right],\end{aligned}$$

$$\begin{aligned}\frac{1}{12} f_{D_1}(d) &= \frac{d}{6}\left[\left(\frac{d^2}{3\sqrt{3}}+2\sqrt{3}\right)\sin^{-1}\frac{\sqrt{3}}{d} - \left(\frac{2d^2}{3\sqrt{3}}+2\sqrt{3}\right)\sin^{-1}\frac{\sqrt{3}}{2d}\right. \\ &\quad \left. - \frac{13}{6}\sqrt{4d^2-3} + \frac{7}{3}\sqrt{d^2-3} + 2d - \frac{1}{2}\right],\end{aligned}$$

$$\begin{aligned}\frac{1}{12} f_{D_3}(d) &= \frac{d}{6}\left[\frac{d^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{2}{\sqrt{3}}-\frac{d^2}{3\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{1}{6}+\frac{\pi}{9\sqrt{3}}\right)d^2\right. \\ &\quad \left. + \frac{7}{12}\sqrt{4d^2-3} + \frac{\sqrt{d^2-3}}{3} - \frac{5\pi}{6\sqrt{3}} - \frac{1}{4}\right],\end{aligned}$$

$$\begin{aligned}\frac{1}{24} f_{D_4}(d) &= \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d}\right. \\ &\quad \left. + \sqrt{4d^2-3} - 2\sqrt{d^2-3} + \frac{\pi}{3\sqrt{3}}\right],\end{aligned}$$

$$\frac{1}{24} f_{D_5}(d) = \frac{d}{12}\left[-\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+4\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{10}{3}\sqrt{d^2-3} + \frac{8\pi}{3\sqrt{3}} + 2\right].$$

Thus,

$$\begin{aligned}
f_{2D_{\mathrm{H_I}}}(d) &= \frac{3}{8}f_{D_{\mathrm{H_A}}}(d) + \frac{1}{12}\left[f_{D_2}(d) + f_{D_1}(d) + f_{D_3}(d)\right] + \frac{1}{24}\left[f_{D_4}(d) + f_{D_5}(d)\right] \\
&= \frac{d}{6}\left[-\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} + \frac{\pi}{6\sqrt{3}}d^2 - 2\sqrt{d^2-3} + \frac{10\pi}{3\sqrt{3}}\right] \\
&= \frac{d}{6}\left[-\frac{4}{\sqrt{3}}\left(\frac{2(d/2)^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2(d/2)} + \frac{2\pi}{3\sqrt{3}}\left(\frac{d}{2}\right)^2 - 2\sqrt{4\left(\frac{d}{2}\right)^2-3} + \frac{10\pi}{3\sqrt{3}}\right] \\
&= \frac{1}{2}f_{D_{\mathrm{H_I}}}(\frac{d}{2}).
\end{aligned}$$

vi) $\frac{3\sqrt{3}}{2} \le d \le \sqrt{7}$:

$$\frac{3}{8}f_{D_{\mathrm{H_A}}}(d) = \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}-2\right)\sin^{-1}\frac{\sqrt{3}}{2d} + \left(\frac{1}{3}-\frac{\pi}{9\sqrt{3}}\right)d^2 - \frac{7}{6}\sqrt{4d^2-3} + \frac{8\pi}{3\sqrt{3}} + \frac{1}{2}\right],$$

$$\begin{aligned}
\frac{1}{12}f_{D_2}(d) &= \frac{d}{6}\left[\frac{5}{\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{2\pi}{9\sqrt{3}}-\frac{1}{6}\right)d^2\right. \\
&\quad \left. + \frac{5}{3}\sqrt{4d^2-3} - 2\sqrt{d^2-3} - 2d + \frac{4\pi}{3\sqrt{3}} - \frac{1}{2}\right],
\end{aligned}$$

$$\begin{aligned}
\frac{1}{12}f_{D_1}(d) &= \frac{d}{6}\left[\left(\frac{2d^2}{3\sqrt{3}}+4\sqrt{3}\right)\sin^{-1}\frac{3\sqrt{3}}{2d} + \left(\frac{d^2}{3\sqrt{3}}+2\sqrt{3}\right)\sin^{-1}\frac{\sqrt{3}}{d}\right. \\
&\quad - \left(\frac{2d^2}{3\sqrt{3}}+2\sqrt{3}\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{\pi}{3\sqrt{3}}d^2 - \frac{13}{6}\sqrt{4d^2-3} + \frac{7}{3}\sqrt{d^2-3} \\
&\quad \left. + \frac{11}{6}\sqrt{4d^2-27} + 2d - 2\sqrt{3}\pi - \frac{1}{2}\right],
\end{aligned}$$

$$\begin{aligned}
\frac{1}{12}f_{D_3}(d) &= \frac{d}{6}\left[\frac{d^2}{3\sqrt{3}}\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{2}{\sqrt{3}}-\frac{d^2}{3\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \left(\frac{2d^2}{3\sqrt{3}}+3\sqrt{3}\right)\sin^{-1}\frac{3\sqrt{3}}{2d}\right. \\
&\quad \left. + \left(\frac{2\pi}{9\sqrt{3}}-\frac{1}{6}\right)d^2 + \frac{7}{12}\sqrt{4d^2-3} + \frac{\sqrt{d^2-3}}{3} - \frac{3}{2}\sqrt{4d^2-27} + \frac{11\pi}{3\sqrt{3}} - \frac{1}{4}\right],
\end{aligned}$$

$$\frac{1}{24}f_{D_4}(d) = \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d} - \frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} - 2\sqrt{3}\sin^{-1}\frac{3\sqrt{3}}{2d} + \sqrt{4d^2-3} - 2\sqrt{d^2-3} - \frac{2}{3}\sqrt{4d^2-27} + \frac{10\pi}{3\sqrt{3}}\right],$$

$$\frac{1}{24}f_{D_5}(d) = \frac{d}{12}\left[-\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+4\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2 - \frac{10}{3}\sqrt{d^2-3} + \frac{8\pi}{3\sqrt{3}} + 2\right].$$

Thus,

$$\begin{aligned} f_{2D_{\mathrm{H_I}}}(d) &= \frac{3}{8}f_{D_{\mathrm{H_A}}}(d) + \frac{1}{12}\left[f_{D_2}(d)+f_{D_1}(d)+f_{D_3}(d)\right] + \frac{1}{24}\left[f_{D_4}(d)+f_{D_5}(d)\right] \\ &= \frac{d}{6}\left[-\frac{4}{\sqrt{3}}\left(\frac{2(d/2)^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2(d/2)} + \frac{2\pi}{3\sqrt{3}}\left(\frac{d}{2}\right)^2 - 2\sqrt{4\left(\frac{d}{2}\right)^2-3} + \frac{10\pi}{3\sqrt{3}}\right] \\ &= \frac{1}{2}f_{D_{\mathrm{H_I}}}(\frac{d}{2}). \end{aligned}$$

vii) $\sqrt{7} \le d \le 3$:

$$\frac{3}{8}f_{D_{\mathrm{H_A}}}(d) = \frac{d}{12}\left[-\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+6\right)\sin^{-1}\frac{3\sqrt{3}}{2d} - \frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d} + \left(\frac{1}{3}+\frac{5\pi}{9\sqrt{3}}\right)d^2 - 4\sqrt{d^2-3} - \frac{11}{6}\sqrt{4d^2-27} + \frac{28\pi}{3\sqrt{3}} + \frac{9}{2}\right],$$

$$\frac{1}{12}f_{D_2}(d) = \frac{d}{6}\left[\left(\frac{2d^2}{3\sqrt{3}}+3\sqrt{3}\right)\sin^{-1}\frac{3\sqrt{3}}{2d} + \left(\frac{1}{6}-\frac{2\pi}{9\sqrt{3}}\right)d^2 + \frac{3}{2}\sqrt{4d^2-27} - 2d - \sqrt{3}\pi\right],$$

$$\frac{1}{12}f_{D_1}(d) = \frac{d}{6}\left[-\frac{d^2}{3\sqrt{3}}\sin^{-1}\frac{3\sqrt{3}}{2d} + \left(\frac{\pi}{9\sqrt{3}}-\frac{1}{3}\right)d^2 + 2d - \frac{\sqrt{4d^2-27}}{4} - \frac{9}{4}\right],$$

$$\begin{aligned}\frac{1}{12}f_{D_3}(d) &= \frac{d}{6}\left[\left(\frac{d^2}{3\sqrt{3}}-\frac{4}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{d}-\left(\frac{d^2}{3\sqrt{3}}+\frac{1}{2\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d}\right.\\ &\quad -\left(\frac{2d^2}{3\sqrt{3}}+\frac{7\sqrt{3}}{2}\right)\sin^{-1}\frac{3\sqrt{3}}{2d}+\left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2-\frac{\sqrt{4d^2-3}}{4}-\sqrt{d^2-3}\\ &\quad \left.-\frac{5}{3}\sqrt{4d^2-27}+\frac{11\pi}{2\sqrt{3}}+\frac{13}{4}\right],\end{aligned}$$

$$\begin{aligned}\frac{1}{24}f_{D_4}(d) &= \frac{d}{12}\left[\frac{1}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d}-\frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d}-3\sqrt{3}\sin^{-1}\frac{3\sqrt{3}}{2d}\right.\\ &\quad \left.+\left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2+\frac{\sqrt{4d^2-3}}{2}-4\sqrt{d^2-3}-\sqrt{4d^2-27}+\frac{17\pi}{3\sqrt{3}}+\frac{9}{2}\right],\end{aligned}$$

$$\begin{aligned}\frac{1}{24}f_{D_5}(d) &= \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+8\right)\sin^{-1}\frac{\sqrt{3}}{d}+\frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+6\right)\sin^{-1}\frac{3\sqrt{3}}{2d}-\left(1+\frac{2\pi}{3\sqrt{3}}\right)d^2\right.\\ &\quad \left.+6\sqrt{d^2-3}+\frac{11}{3}\sqrt{4d^2-27}-\frac{40\pi}{3\sqrt{3}}-11\right].\end{aligned}$$

Thus,

$$\begin{aligned}f_{2D_{\mathrm{H_I}}}(d) &= \frac{3}{8}f_{D_{\mathrm{H_A}}}(d)+\frac{1}{12}\left[f_{D_2}(d)+f_{D_1}(d)+f_{D_3}(d)\right]+\frac{1}{24}\left[f_{D_4}(d)+f_{D_5}(d)\right]\\ &= \frac{d}{6}\left[-\frac{4}{\sqrt{3}}\left(\frac{2(d/2)^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2(d/2)}+\frac{2\pi}{3\sqrt{3}}\left(\frac{d}{2}\right)^2-2\sqrt{4\left(\frac{d}{2}\right)^2-3}+\frac{10\pi}{3\sqrt{3}}\right]\\ &= \frac{1}{2}f_{D_{\mathrm{H_I}}}(\frac{d}{2}).\end{aligned}$$

viii) $3 \le d \le 2\sqrt{3}$:

$$\begin{aligned}\frac{3}{8}f_{D_{\mathrm{H_A}}}(d) &= \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+12\right)\sin^{-1}\frac{3\sqrt{3}}{2d}-\frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d}\right.\\ &\quad \left.+\left(\frac{\pi}{9\sqrt{3}}-\frac{1}{3}\right)d^2-4\sqrt{d^2-3}+\frac{19}{6}\sqrt{4d^2-27}-\frac{8\pi}{3\sqrt{3}}-\frac{9}{2}\right],\end{aligned}$$

$$\frac{1}{12}f_{D_3}(d) = \frac{d}{6}\left[\left(\frac{d^2}{3\sqrt{3}}-\frac{4}{\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{d}-\left(\frac{d^2}{3\sqrt{3}}+\frac{1}{2\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d}+\frac{5\sqrt{3}}{2}\sin^{-1}\frac{3\sqrt{3}}{2d}\right.$$
$$\left.-\frac{\sqrt{4d^2-3}}{4}-\sqrt{d^2-3}+\frac{5}{6}\sqrt{4d^2-27}-\frac{\pi}{2\sqrt{3}}-\frac{5}{4}\right],$$

$$\frac{1}{24}f_{D_4}(d) = \frac{d}{12}\left[\frac{1}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d}-\frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+2\right)\sin^{-1}\frac{\sqrt{3}}{d}\right.$$
$$\left.+\left(\frac{2d^2}{3\sqrt{3}}+3\sqrt{3}\right)\sin^{-1}\frac{3\sqrt{3}}{2d}+\frac{\sqrt{4d^2-3}}{2}-4\sqrt{d^2-3}+\frac{3}{2}\sqrt{4d^2-27}-\frac{\pi}{3\sqrt{3}}\right],$$

$$\frac{1}{24}f_{D_5}(d) = \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+8\right)\sin^{-1}\frac{\sqrt{3}}{d}-\frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+12\right)\sin^{-1}\frac{3\sqrt{3}}{2d}+\left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2\right.$$
$$\left.+6\sqrt{d^2-3}-\frac{19}{3}\sqrt{4d^2-27}+\frac{32\pi}{3\sqrt{3}}+7\right].$$

Thus,

$$f_{2D_{\mathrm{H_I}}}(d) = \frac{3}{8}f_{D_{\mathrm{H_A}}}(d)+\frac{1}{12}f_{D_3}(d)+\frac{1}{24}\left[f_{D_4}(d)+f_{D_5}(d)\right]$$
$$= \frac{d}{6}\left[-\frac{4}{\sqrt{3}}\left(\frac{2(d/2)^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2(d/2)}+\frac{2\pi}{3\sqrt{3}}\left(\frac{d}{2}\right)^2-2\sqrt{4\left(\frac{d}{2}\right)^2-3}+\frac{10\pi}{3\sqrt{3}}\right]$$
$$= \frac{1}{2}f_{D_{\mathrm{H_I}}}(\frac{d}{2}).$$

ix) $2\sqrt{3}\le d\le\sqrt{13}$:

$$\frac{3}{8}f_{D_{\mathrm{H_A}}}(d) = \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+12\right)\left(\sin^{-1}\frac{3\sqrt{3}}{2d}+\sin^{-1}\frac{2\sqrt{3}}{d}\right)-\left(\frac{2}{3}+\frac{4\pi}{9\sqrt{3}}\right)d^2\right.$$
$$\left.+\frac{19}{6}\sqrt{4d^2-27}+\frac{16}{3}\sqrt{d^2-12}-\frac{16\pi}{\sqrt{3}}-\frac{25}{2}\right],$$

$$\frac{1}{12}f_{D_3}(d) = \frac{d}{6}\left[\left(\frac{d^2}{3\sqrt{3}}+\frac{8}{\sqrt{3}}\right)\sin^{-1}\frac{2\sqrt{3}}{d}-\left(\frac{d^2}{3\sqrt{3}}+\frac{1}{2\sqrt{3}}\right)\sin^{-1}\frac{\sqrt{3}}{2d}+\frac{5\sqrt{3}}{2}\sin^{-1}\frac{3\sqrt{3}}{2d}\right.$$
$$\left.-\left(\frac{1}{6}+\frac{\pi}{9\sqrt{3}}\right)d^2-\frac{\sqrt{4d^2-3}}{4}+\frac{5}{6}\sqrt{4d^2-27}+2\sqrt{d^2-12}-\frac{31\pi}{6\sqrt{3}}-\frac{9}{4}\right],$$

$$\frac{1}{24}f_{D_4}(d) = \frac{d}{12}\left[\frac{1}{\sqrt{3}}\left(\frac{2d^2}{3}+1\right)\sin^{-1}\frac{\sqrt{3}}{2d}+\left(\frac{2d^2}{3\sqrt{3}}+3\sqrt{3}\right)\sin^{-1}\frac{3\sqrt{3}}{2d}+\frac{8}{\sqrt{3}}\sin^{-1}\frac{2\sqrt{3}}{d}\right.$$
$$\left.-\left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2+\frac{\sqrt{4d^2-3}}{2}+\frac{3}{2}\sqrt{4d^2-27}+\frac{4}{3}\sqrt{d^2-12}-\frac{17\pi}{3\sqrt{3}}-8\right],$$

$$\frac{1}{24}f_{D_5}(d) = \frac{d}{12}\left[-\frac{4}{\sqrt{3}}\left(\frac{d^2}{3}+12\right)\sin^{-1}\frac{3\sqrt{3}}{2d}-\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+8\right)\sin^{-1}\frac{2\sqrt{3}}{d}\right.$$
$$\left.+\left(1+\frac{2\pi}{3\sqrt{3}}\right)d^2-\frac{19}{3}\sqrt{4d^2-27}-4\sqrt{d^2-12}+\frac{64\pi}{3\sqrt{3}}+17\right].$$

Thus,

$$f_{2D_{\mathrm{H_I}}}(d) = \frac{3}{8}f_{D_{\mathrm{H_A}}}(d)+\frac{1}{12}f_{D_3}(d)+\frac{1}{24}\left[f_{D_4}(d)+f_{D_5}(d)\right]$$
$$= \frac{d}{6}\left[\frac{1}{\sqrt{3}}\left(\frac{d^2}{3}+16\right)\sin^{-1}\frac{2\sqrt{3}}{d}-\left(\frac{\pi}{9\sqrt{3}}+\frac{1}{6}\right)d^2+\frac{10}{3}\sqrt{d^2-12}-\frac{16\pi}{3\sqrt{3}}-4\right]$$
$$= \frac{d}{6}\left[\frac{4}{\sqrt{3}}\left(\frac{(d/2)^2}{3}+4\right)\sin^{-1}\frac{\sqrt{3}}{d/2}-\left(\frac{4\pi}{9\sqrt{3}}+\frac{2}{3}\right)\left(\frac{d}{2}\right)^2+\frac{20}{3}\sqrt{\left(\frac{d}{2}\right)^2-3}\right.$$
$$\left.-\frac{16\pi}{3\sqrt{3}}-4\right] = \frac{1}{2}f_{D_{\mathrm{H_I}}}(\frac{d}{2}).$$

x) $\sqrt{13} \le d \le 4$:

$$\frac{1}{24}f_{D_5}(d) = \frac{d}{12}\left[\frac{2}{\sqrt{3}}\left(\frac{d^2}{3}+16\right)\sin^{-1}\frac{2\sqrt{3}}{d}-\left(\frac{1}{3}+\frac{2\pi}{9\sqrt{3}}\right)d^2+\frac{20}{3}\sqrt{d^2-12}-\frac{32\pi}{3\sqrt{3}}-8\right].$$

Thus,

$$f_{2D_{\mathrm{H_I}}}(d) = \frac{1}{24}f_{D_5}(d)$$
$$= \frac{d}{6}\left[\frac{1}{\sqrt{3}}\left(\frac{d^2}{3}+16\right)\sin^{-1}\frac{2\sqrt{3}}{d}-\left(\frac{\pi}{9\sqrt{3}}+\frac{1}{6}\right)d^2+\frac{10}{3}\sqrt{d^2-12}-\frac{16\pi}{3\sqrt{3}}-4\right]$$
$$= \frac{d}{6}\left[\frac{4}{\sqrt{3}}\left(\frac{(d/2)^2}{3}+4\right)\sin^{-1}\frac{\sqrt{3}}{d/2}-\left(\frac{4\pi}{9\sqrt{3}}+\frac{2}{3}\right)\left(\frac{d}{2}\right)^2+\frac{20}{3}\sqrt{\left(\frac{d}{2}\right)^2-3}\right.$$
$$\left.-\frac{16\pi}{3\sqrt{3}}-4\right] = \frac{1}{2}f_{D_{\mathrm{H_I}}}(\frac{d}{2}).$$

In summary, we have $f_{2D_{\mathrm{H_I}}}(d) = \frac{1}{2} f_{D_{\mathrm{H_I}}}(\frac{d}{2})$ by scaling, and the probabilistic sum $\frac{3}{16} f_{D_{\mathrm{H_I}}}(d) + \frac{3}{8} f_{D_{\mathrm{H_A}}}(d) + \frac{1}{12} [f_{D_{\mathrm{PX}}}(d) + f_{D_2}(d) + f_{D_1}(d) + f_{D_3}(d)] + \frac{1}{24} [f_{D_4}(d) + f_{D_5}(d)] + \frac{1}{48} f_{D_{\mathrm{I}}}(d)$ is equal to $\frac{1}{2} f_{D_{\mathrm{H_I}}}(\frac{d}{2})$ in all the cases discussed above. The results are a strong validation of the correctness of the distance distribution functions that we have derived for distances associated with hexagons.