ABSTRACT

Title of dissertation: CONNECTIVITY ANALYSIS OF WIRELESS AD-HOC NETWORKS

Guang Han, Doctor of Philosophy, 2007

Dissertation directed by: Professor Armand Makowski
Department of Electrical and Computer Engineering

Connectivity is one of the most fundamental properties of wireless ad-hoc networks as most network functions are predicated upon the network being connected. Although increasing node transmission power will improve network connectivity, too large a power level is not feasible as energy is a scarce resource in wireless ad-hoc networks. Thus, it is crucial to identify the minimum node transmission power that will ensure network connectivity with high probability.

It is known that there exists a critical level transmission power such that a suitably larger power will ensure network connectivity with high probability. A small variation across this threshold level will lead to a sharp transition of the probability that the network is connected. Thus, in order to precisely estimate the minimum node transmission power, not only do we need to identify this critical threshold, but also how fast this transition takes place. To characterize the sharpness of transition, we define weak, strong and very strong critical thresholds associated with increasing transition speeds.

In this dissertation, we seek to estimate the minimum node transmission power

for large scale one-dimensional wireless ad-hoc networks under the Geometric Random Graph (GRG) models. Unlike in previous works where nodes are taken to be uniformly distributed, we assume a more general node distribution. Using the methods of first and second moments, we theoretically prove the existence of a very strong critical threshold when the density function is everywhere positive. On the other hand, only weak thresholds are shown to exist when the density function contains vanishing densities.

We also study the connectivity of two-dimensional wireless ad-hoc networks under the random connection model, which accounts for statistical channel variations. With the help of the Stein-Chen method, we derive a closed form formula for the limiting probability that there are no isolated nodes under a very general assumption of channel variations. The node transmission power to ensure the absence of isolated nodes provides a tight lower bound on the transmission power needed to ensure network connectivity.

Connectivity Analysis of Wireless Ad-hoc Networks

by

Guang Han

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Commmittee:

Professor Armand M. Makowski, Chair/Advisor
Professor Alexander Barg
Professor Anthony Ephremides
Professor Richard J. La
Professor Aravind Srinivasan

DEDICATION

To my parents and Xiong

ACKNOWLEDGMENTS

First and foremost, I must thank Professor Armand M. Makowski, for being a great advisor, for his tremendous guidance and support on my research, and for painstakingly teaching me how to write scientific papers and give well-organized presentations. I always feel fortunate that I have been his student and have taken advantage of his wisdom during my graduate studies. His enthusiasm for research and teaching is admirable. He is not only an advisor, but also a role model to me.

I would also like to thank Professor Richard J. La for his kind help and enlightening suggestions. I also wish to thank the other committee members, Professor Alexander Barg, Professor Anthony Ephremides, and Professor Aravind Srinivasan for their constructive feedbacks in this work.

I am grateful for the funding support from the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011.

Last but not least, I would like to thank my parents and my wife. Without their constant love and support, it is impossible for me to accomplish this enormous task.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Motivation and objectives

It is envisioned that we shall soon inhabit a world where myriads of wireless devices tightly interact with the physical world as well as with human beings [18, 55]. However, current cellular wireless network architectures are not suitable to support such an exceedingly large number of users because the number of base stations can not scale accordingly for environmental and regulatory reasons. In addition, infrastructure deployment is barely feasible in some inhospitable regions such as battlefields, deserts and disaster areas.

Such constraints on infrastructure deployment make wireless ad-hoc networks a promising technology. Wireless ad-hoc networks are large-scale infrastructureless networks. Network nodes are usually randomly deployed and their locations may evolve in time. Without the help of base stations, two far away nodes communicate only if some intermediate nodes relay their packets. The basic question for such networks is whether intermediate relays exist for any pair of nodes, a fact which indicates network connectivity.

Connectivity is one of the most fundamental properties of wireless ad-hoc networks. As a matter of fact, most network functions are predicated upon the network being connected. Although increasing node transmission power will improve network connectivity, large power levels are not desirable as energy is a scarce resource in wireless ad-hoc networks [16]. Thus, it is needed to answer the following question: What is the minimum transmission power to ensure network connectivity? [1]

Because nodes are randomly distributed, this critical transmission power is a random variable, whose distribution is usually hard to obtain. Fortunately, we are mainly interested in networks with a large number of nodes (e.g. wireless sensor networks). In this case, we can instead try to identify a critical scaling (with respect to the number of nodes), or threshold, for the transmission power. Such a critical threshold indicates a boundary in the space of scalings with respect to network connectivity: A suitably larger (resp. smaller) power will ensure network connectivity with probability close to one (resp. zero). For a large-scale network, a small variation across this critical threshold will lead to a sharp transition in the probability of network connectivity. In addition, in order to precisely estimate the minimum node transmission power, not only do we need to identify the critical threshold, but also how fast the transition takes place. To characterize the sharpness of this transition, we define weak, strong and very strong critical thresholds associated with increasing transition speed.

[1] In this dissertation, we assume homogeneous nodes with equal transmission power.

Whether a solution could be useful in practice depends on how well the adopted network model realistically captures the basic characteristics of wireless ad-hoc networks. In most existing studies [14, 20, 23, 28, 38, 45], two assumptions are made which are too simplistic to model wireless ad-hoc networks. The first one assumes that network nodes are uniformly distributed over a certain geographical area. The second assumption is the disk connection model [24], where two nodes are assumed to communicate if and only if their distance is below an artificial transmission range. The first assumption is not valid if network nodes are mobile, e.g., the stationary node distribution under the random waypoint mobility model without pause [51]. Moreover, if nodes (e.g., wireless sensors) are scattered from an airplane, their distribution can not be uniform. The second assumption is questionable because of shadowing and fading, typical phenomena in wireless communications. Since shadowing and fading lead to random channel variations, the existence of a link between two nodes is a random event. Even if they are very close to each other, the availability of this link is not guaranteed.

In this dissertation, we focus mostly on identifying the critical threshold of node transmission power, and on the sharpness of transition under general node placement distributions under disk connection model. The connectivity of wireless ad-hoc networks under the random connection model is discussed in Chapter 8.

1.2 Related work

Many of the papers [14, 20, 23, 28, 38, 45] exploring the connectivity of wireless ad-hoc networks are based on two simplistic assumptions, namely uniform node distribution and disk connection model. The exact probability of connectivity of one-dimensional finite networks is computed in [14, 20, 23]; critical thresholds for transmission range when the number of nodes in the network tends to infinity are identified in [28, 38, 45].

Prior works on network connectivity under the assumptions of general node distribution and disk connection model include: Foh et al. [21] study the connectivity of one-dimensional MANETs where users follow the random waypoint mobility model. However, their theoretical analysis (section II) involves some unverified approximations which need further investigation. Santi [51] investigates the critical transmission range (CTR) for connectivity of mobile ad-hoc networks, with nonuniform stationary user distribution. When the density of user distribution is strictly positive, their results indicate that the critical transmission range is a constant factor larger than that obtained in the uniform distribution case. However, when the user density vanishes, they only obtain the loose result that the CTR is much larger than that of the uniform case without showing how large it should be. Deheuvels derives upper and lower bounds on the critical range for the case of one-dimensional graphs under non-uniform node distribution [13]. However, he does not identify the strong critical threshold for graph connectivity.

As the deterministic channel assumption is not realistic, many researchers have started to investigate the connectivity of ad-hoc networks under a random channel model. Hekmat and Mieghem analyze the connectivity problem under the lognormal shadowing channel model through extensive simulations [36]. Their results show that larger channel variations will improve network connectivity. In [8], Bettstetter and Hartmann study the network connectivity under the same channel model. They derive a closed form formula of the minimum node density to ensure that no node is isolated; this density provides a tight lower bound on the node density required to ensure network connectivity. In [44], Miorandi and Altman compute the node isolation probability under a more generic random channel. However, the analysis in both [8] and [44] are based on an unnecessary approximation that the isolation of nodes are independent events. Avin studies a model of distance graphs [4], where nodes are uniformly distributed in a unit disk, and the probability of edge presence between any pair of nodes is a function of their distance. However, he only considers a very special type of connection function.

In summary, the existing literature on the connectivity of wireless ad-hoc networks is mostly based on the two simplistic assumptions. Although some attempts have been made to explore generalized models, results to date are either based on simulation or derived under some arbitrary approximations.

1.3 Main contributions

One of the main contributions of this dissertation is to introduce systematic approaches to study connectivity of wireless ad-hoc networks. More specifically, all our results are rigorously established with the help of three efficient tools: Spacings, the method of the first and second moments, and the Stein-Chen method.

Assuming the disk connection model, we study the connectivity of one-dimensional networks under three types of density functions $f : [0,1] \to \mathbb{R}_+$: Firstly, the minimum density $f_\star = 1$, which corresponds to the case of uniform node distribution. Secondly, $0 < f_\star < 1$, which corresponds to the case of nonuniform node distribution with non-vanishing density. Lastly, $f_\star = 0$, which corresponds to the case of nonuniform node distribution with vanishing density.

For the first case where nodes are uniformly distributed, we use the method of the first and second moments to identify $\tau_n^\star = \frac{\log n}{n}$ as a very strong threshold.

When nodes are placed according to the continuous density function f:

$$f(x) = c + a|x - x_\star|^r + h(x), \quad 0 \le x \le 1 \tag{1.1}$$

for some parameters $r > 0$, $a > 0$ and $c > 0$, and for some function $h : [0,1] \to \mathbb{R}$ such that

$$\lim_{x \to x_\star} \frac{h(x)}{|x - x_\star|^r} = 0, \tag{1.2}$$

it can be shown that $c = f_\star$ and we identify

$$\tau_n^\star = \frac{\log n - \frac{1}{r}\log\log n}{n f_\star} \tag{1.3}$$

as a very strong threshold via a variation of the method of the first and second moments. It should be noted that when node distribution is nonuniform, the very strong threshold does not only depend on the minimum density $f_\star$, but also the smoothness of the density function around the minimum density, a feature captured by r.

If the density function f is given by

$$f(x) = (p+1)x^p, \quad x \in [0,1], \tag{1.4}$$

we prove the existence of a weak threshold function

$$\tau^{\star}_{p,n} = n^{-\frac{1}{p+1}} \quad n = 1, 2, \ldots \tag{1.5}$$

by the method of spacings. We also prove that a strong threshold does not exist for this case. This implies a much slower transition from network disconnectivity to network connectivity as the transmission range varies across this weak threshold.

Finally, we user the Stein-Chen method to compute the limiting probability that there are no isolated nodes for two-dimensional networks under some random connection models with bounded support. More importantly, we show that this limiting probability only depends on the expected node degree, and the details of channel variations do not affect the result.

1.4 Organization

The structure of the dissertation is as follows: Chapter 2 describes models of wireless ad-hoc network connectivity. In Chapter 3, we introduce some mathematical tools that will be used in this dissertation. From Chapter 4 to Chapter 7, we study the connectivity of one-dimensional networks under the disk connection model. The case of uniform node distribution is investigated in Chapter 4. We prove the existence of strong and very strong thresholds in Chapter 5 and Chapter 6, respectively, for general node distributions with non-vanishing densities. In

Chapter 7, we discuss the case of node distributions with vanishing densities. Finally in Chapter 8, we compute the limiting probability that there are no isolated nodes in two-dimensional ad-hoc networks under some random connection models with bounded support.

Chapter 2

Models of network connectivity

2.1 Model description

First a word on the notation and conventions used throughout this dissertation: We assume that all the rvs under consideration are defined on the *same* probability triple $(\Omega, \mathcal{F}, \mathbb{P})$, possibly by enlarging it to accommodate these rvs. All probabilistic statements are made with respect to this probability measure $\mathbb{P}$. The notation $\xrightarrow{P}_n$ (resp. $\Longrightarrow_n$) is used to signify convergence in probability (resp. convergence in distribution) with n going to infinity.

In our model, we consider wireless ad-hoc networks under the following assumptions: Firstly, the thermal noise power remains the same throughout the network. Secondly, we assume the existence of a highly efficient MAC layer protocol, which eliminates most packet collisions. Thus, interferences from other users can be ignored. Thirdly, all nodes have equal transmission power. Finally, for any transmitter-receiver pair, there exists a mapping that maps the transmission power and transmitter-receiver distance onto the received power. This mapping is assumed to be fixed for a given network.

Before introducing the network model, we first explain our link model. A wireless link exists between two users i and j if and only if

$$\frac{P_r^i}{N_0^i + I_i} \geq \beta \quad \text{and} \quad \frac{P_r^j}{N_0^j + I_j} \geq \beta, \tag{2.1}$$

where β is a certain threshold, N_0^i (resp. N_0^j) is the thermal noise power at user i (resp. j), I_i (resp. I_j) is the interference power at user i (resp. j) and P_r^i (resp. P_r^j) is the received power at user i (resp. j) assuming that user j (resp. i) transmits with power P_t^j (resp. P_t^i). According to the aforementioned assumptions, we have

$$N_0^i = N_0^j = N_0, \tag{2.2}$$

$$I_i = I_j = 0, \tag{2.3}$$

$$P_t^i = P_t^j = P_t, \tag{2.4}$$

and

$$P_r^i = P_r^j = \Phi(P_t, d_{ij}), \tag{2.5}$$

where Φ is the aforementioned mapping and d_{ij} is the distance between user i and user j. Therefore, it is clear that (2.1) holds if and only if

$$\Phi(P_t, d_{ij}) \geq N_0\beta. \tag{2.6}$$

In (2.6), $N_0\beta$ is fixed, d_{ij} is random since users i and j are randomly placed, and Φ is predetermined for a given network, so that the node transmission power P_t is the key parameter to be computed.

We thus define the probability that a link exists between users i and j by

$$g_{P_t}(d_{ij}) := \mathbb{P}[\Psi(P_t, d_{ij}) \geq N_0\beta] \tag{2.7}$$

where $\{g_{P_t}, P_t > 0\}$ is a family of link probability functions with $g_{P_t} : \mathbb{R}_+ \rightarrow [0,1] : r \rightarrow g_{P_t}(r)$ [1] being a non-increasing function which characterizes the random channel variations when nodes transmit with power P_t.

The network model is built upon the link model. We model the network as a geometric graph $G(V, E)$. The vertex set $V = \{1, \ldots, n\}$ consists of n vertices (users) randomly placed on $[0,1]^d$, where $d = 1$ (resp. $d = 2$) indicates one-dimensional (resp. two-dimensional) networks. For $n = 2, 3, \ldots$, vertex locations $X_1, \ldots, X_n$ are i.i.d. rvs which are distributed in $[0,1]^d$ according to some common distribution $F : [0,1]^d \rightarrow [0,1]$. Given $X_1, \ldots, X_n$, the edge set E is constructed through a link probability function g as follows. We introduce i.i.d. $\{0,1\}$-valued rvs $\{L_{i,j} : i, j = 1, 2, \ldots, n, i < j\}$ to indicate the existence of edges with nodes i and j linked by an edge if and only if $L_{i,j} = 1$. Because interferences are assumed to be zero, the indicator rvs $\{L_{i,j} : i, j = 1, 2, \ldots, n, i < j\}$ are independent given $X_1, \ldots, X_n$. We postulate that

$$\mathbb{P}[L_{i,j} = 1 | X_1, \ldots, X_n] = g\big(||X_i - X_j||\big),$$

where $||X_i - X_j||$ is the Euclidean distance between X_i and X_j. Finally, network connectivity is taken to be equivalent to the connectivity of the graph $G(V, E)$, i.e., any two vertices in V are linked by a path over edges in E.

[1] We use P_t as a subscript to emphasize that the link functions depend on the node transmission power P_t. This notation will be ignored thereafter for simplicity.

Let $\mathcal{G}$ denote the collection of graphs on the vertex set V, i.e., an element of $\mathcal{G}$ is denoted as G with edge set E. Let $\mathbb{G}$ be the graph-valued rv $\mathbb{G} : \Omega \to \mathcal{G}$. For any $G = (V, E)$ in $\mathcal{G}$, we have

$$\begin{aligned} \mathbb{P}[\mathbb{G} = G] &= \mathbb{E}\Big[\mathbb{E}\Big[\mathbf{1}[\mathbb{G} = G]\Big| X_1, \ldots, X_n\Big]\Big] \\ &= \mathbb{E}\Bigg[\prod_{\substack{(i,j)\in E \\ i<j \\ i,j=1,\ldots,n}} g\big(||X_i - X_j||\big) \prod_{\substack{(i,j)\notin E \\ i<j \\ i,j=1,\ldots,n}} \Big(1 - g\big(||X_i - X_j||\big)\Big)\Bigg]. \end{aligned} \tag{2.8}$$

If $\mathcal{G}_{con}$ denotes the collection of connected graphs, then the probability that $\mathbb{G}$ is connected is given by:

$$\mathbb{P}[\mathbb{G} \in \mathcal{G}_{con}] = \sum_{G\in\mathcal{G}_{con}} \mathbb{P}[\mathbb{G} = G]. \tag{2.9}$$

From (2.8) and (2.9), we observe that the probability of graph connectivity is determined solely by the user distribution function F and by the link probability function g. Past work on network connectivity mainly employs the uniform user distribution model and the disk connection model (see Section 2.2.2), which can not capture all wireless network scenarios. Thus the network connectivity problem needs to be investigated under more generalized models.

We assume that F admits a density function $f : [0, 1]^d \to \mathbb{R}_+$ which is *continuous* on $[0, 1]^d$, and we write

$$f_\star = \inf\big(f(x),\ x \in [0, 1]^d\big) \tag{2.10}$$

and

$$f^\star = \sup\left(f(x),\ x \in [0, 1]\right). \tag{2.11}$$

The continuity of f on the compact $[0, 1]^d$ guarantees that this infimum is achieved by at least one element $x_\star$ in $[0, 1]^d$.

According to the value of $f_\star$, F falls in one of the following three classes: $f_\star = 1$, $0 < f_\star < 1$ and $f_\star = 0$. Clearly, $f_\star = 1$ corresponds to the uniform node distribution.

Since channel variations are very complicated, no single link model fits all scenarios. In the next section, we introduce four specific graph models based on different link probability functions.

2.2 Specific random graph models

2.2.1 The Erdös Rényi graph (ERG) model

The Erdös Rényi graph model was introduced by Erdös and Rényi in their groundbreaking paper [17]. With $0 < p < 1$, its link function is

$$g_{ER}(r) = p \quad r \geq 0.$$

In the ERG model, the probability of link presence between two users is independent of their distance. This assumption ignores the pathloss phenomenon of the wireless communications.

2.2.2 The geometric random graph (GRG) model

The geometric random graph (GRG) model [46] is a basic model for wireless ad-hoc networks. Its link function has the form

$$g_G(r) = \mathbf{1}[r \leq \tau], \quad r \geq 0 \tag{2.12}$$

where τ denotes the transmission range which is determined by the transmission power P_t. Thus, two vertices are linked by an edge if and only if their distance is less than τ. Such a link function is also called the disk connection model [28], based on the fact that a user can only communicate with the neighbors within a disk of radius τ. This notion of connectivity gives rise to the undirected geometric random graph $\mathbb{G}_d(n;\tau)$. We use $P_{con,d}(n;\tau)$ to denote the probability that $\mathbb{G}_d(n;\tau)$ is connected. For notational convenience, the one-dimensional GRG is simply denoted by $\mathbb{G}(n;\tau)$, whose probability of connectivity is given by $P(n;\tau)$.

The basic assumption of the GRG model is that the received power P_r is deterministic and decreasing with transmitter-receiver distance r. Thus, there exists a boundary distance τ at which the received power $P_r = N_0\beta$, so that $P_r \geq N_0\beta$ if and only if $r \leq \tau$. In the GRG model, the received power is often given by a power law function [49, p. 107] of the form

$$P_r = Ar^{-\eta}, \quad r > 0 \tag{2.13}$$

where η is the pathloss exponent and A is a constant that is mainly determined by the transmission power, antenna gains and signal wavelength. According to (2.13) and (2.6), a link exists between a transmitter-receiver pair with distance d if and only if

$$r \leq \left(\frac{A}{N\beta}\right)^{\frac{1}{\eta}}. \tag{2.14}$$

By comparing (2.12) and (2.14), the transmission range τ is equal to $\left(\frac{A}{N\beta}\right)^{\frac{1}{\eta}}$. Any deterministic channel can be captured by the GRG model; the power law function (2.13) only provides an example of such a deterministic channel model.

Examples of connected and disconnected GRGs are shown in Fig. 2.1 and Fig. 2.2, respectively. In both graphs, there are $n = 1000$ users with communication range $\tau = 0.04924$ for the connected graph and $\tau = 0.04455$ for the disconnected graph.

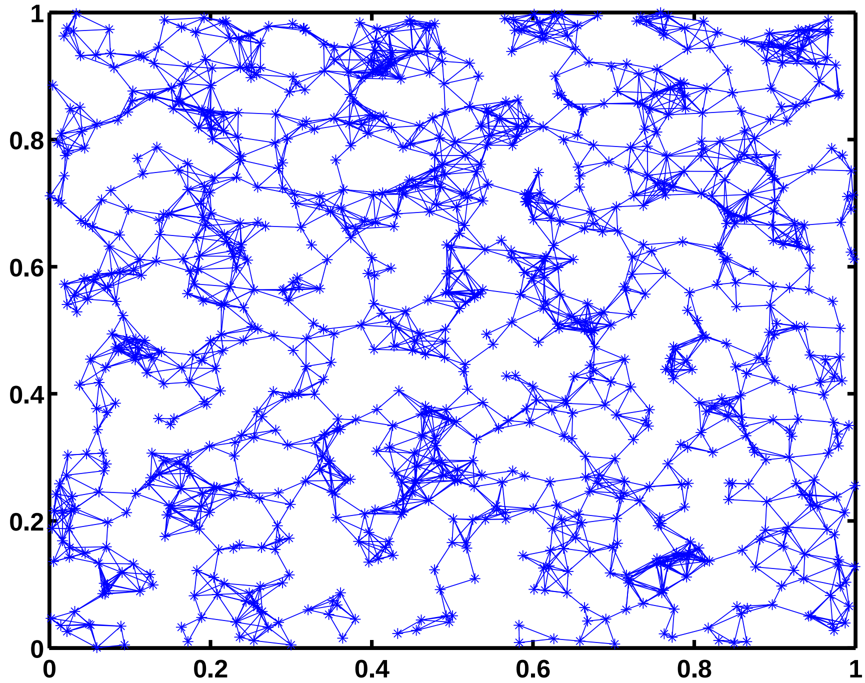


Figure 2.1: Example of a connected geometric random graph.

2.2.3 The lognormal connection graph (LCG) model

Hekmat and Mieghem introduce a radio model [36], where the logarithm of the received signal power follows the normal distribution. The mean value of this normal distribution is determined by the power law function (2.13). Based on this lognormal distribution assumption, the edge function of the LCG model is defined

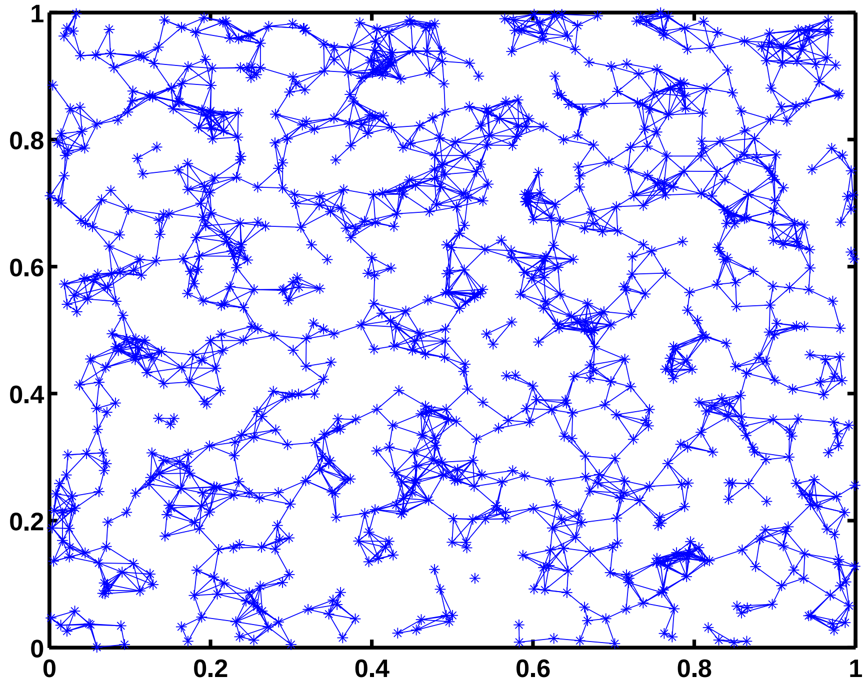


Figure 2.2: Example of a disconnected geometric random graph.

by

$$g_L(r) = \frac{1}{2}\left[1 - \mathrm{erf}\left(\alpha \frac{\log(r/\tau)}{\xi}\right)\right], \quad r \geq 0 \tag{2.15}$$

where $\alpha = \frac{10}{\sqrt{2}\log 10}$, τ is the transmission range defined in the GRG model and the parameter ξ captures channel variations. In practice, ξ is usually between 0 and 6. As usual, the error function (erf) used in (2.15) is given by

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt. \quad x \geq 0$$

2.2.4 The bounded connection graph (BCG) model

In the bounded connection graph (BCG) model, the probability that two users are pairwise connected is positive only if their distance is less than the boundary range ρ which is determined by the transmission power P_t. The edge function of the

BCG model is given by

$$g_B(r) = \mathbf{1}[r < \rho]h_\rho(r), \quad r \geq 0 \tag{2.16}$$

where $h_\rho : [0,1] \to [0,1]$ is a right continuous non-increasing function where

$$h_\rho(r) \begin{cases} = 1 & r = 0 \\ \in (0,1) & 0 < r < \rho \\ = 0 & \rho \leq r \leq 1 \end{cases} \tag{2.17}$$

The BCG model gives rise to the undirected graph $\mathbb{G}_d(n;\rho)$. We use $P_{con,d}(n;\rho)$ to denote the probability that $\mathbb{G}_d(n;\rho)$ is connected.

The form of the function h_ρ is almost arbitrary with few constraints. In contrast to specific link functions (such as (2.15)), this generalized function can be used to model almost all kinds of wireless channels. One may question the validity of the link function being bounded, which is the main constraint imposed on this model. However, due to the fast signal attenuation with distance, two far-away users can hardly communicate effectively. Even if a link exists, it can hardly be utilized, and it is therefore reasonable to simply ignore the very long links.

2.3 Discussion

The key parameters for the GRG and BCG models are τ and ρ, respectively. Each parameter is assumed to be an increasing function of the user transmission power P_t. Thus in order to estimated the minimum transmission power to ensure network connectivity, it is equivalent to estimate the minimum values of these parameters to ensure graph connectivity. In other words, both parameters act as proxies for the transmission power.

In this dissertation, we will mainly focus on the GRG and the BCG models, which are representatives of the deterministic and random channel models, respectively.

Chapter 3

Approaches

3.1 The basic idea

Graph connectivity is a global property that can not be determined by studying local properties defined on subsets of nodes. For example, the triangle containment is a local property. Whether a graph has such a property can be found by checking all the 3-node subsets. Similarly, the absence of node isolation is also a local property. A local property is usually much easier to study than a global property.

Fortunately, for the one-dimensional GRG model, the connectivity of $\mathbb{G}(n;\tau)$ is actually a local property. To clarify this point, we introduce the notion of breakpoint nodes.

Definition 3.1 *Fix $n = 2, 3, \dots$ and τ in $(0, 1)$. For each $i = 1, \dots, n$, node i is said to be a breakpoint node in $\mathbb{G}(n;\tau)$ whenever (i) it is* not *the leftmost node in $[0, 1]$ and (ii) there is no node in the random interval $[X_i - \tau, X_i]$.*

It is clear that $\mathbb{G}(n;\tau)$ is connected if and only if no node is a breakpoint node, and the connectivity of $\mathbb{G}(n;\tau)$ is therefore a local property. If we denote the number of breakpoint nodes in $\mathbb{G}(n;\tau)$ by $C_n(\tau)$, then we have the representation

$$P(n;\tau) = \mathbb{P}[C_n(\tau) = 0]. \tag{3.1}$$

For ease of analysis, we often represent $C_n(\tau)$ as a sum of indicator rvs. We first define spacings as follows:

Definition 3.2 *The spacings between consecutive order statistics are given by*

$$L_{n,k} := X_{n,k} - X_{n,k-1}, \quad k = 1, \ldots, n+1 \tag{3.2}$$

where $X_{n,1}, \ldots, X_{n,n}$ is the ordered sequence of $X_1, \ldots, X_n$ such that

$$X_{n,1} \leq X_{n,2} \leq \ldots \leq X_{n,n}.$$

By convention, $X_{n,0} = 0$ and $X_{n,n+1} = 1$.

It is now clear that

$$C_n(\tau) = \sum_{k=2}^{n} \chi_{n,k}(\tau),$$

where the $\{0,1\}$-valued rvs $\chi_{n,1}(\tau), \ldots, \chi_{n,n+1}(\tau)$ are given by

$$\chi_{n,k}(\tau) := \mathbf{1}\left[L_{n,k} > \tau\right], \quad k = 1, \ldots, n+1. \tag{3.3}$$

Since $C_n(\tau)$ can be represented as a sum of indicator rvs, the probability of $C_n(\tau)$ being zero can be estimated using the methods of the first and second moments and the Stein-Chen method. The details are given in Section 3.2.2 and Section 3.2.3.

Another way to study the connectivity of $\mathbb{G}(n;\tau)$ is through the maximal spacing M_n, which is defined by

$$M_n := \max(L_{n,k}, k = 2, \ldots, n). \tag{3.4}$$

Clearly we have

$$P(n;\tau) = \mathbb{P}[M_n \leq \tau]. \tag{3.5}$$

Although the maximal spacing has been extensively studied in the past decades [13, 39], it has never gained attention in the network community. We will discuss some existing facts of spacings for $\mathbb{G}(n;\tau)$ under uniform node distribution in Section 3.2.1. Based on these facts, we obtain some new results for the case of nonuniform node distribution in Chapters 5 and 7.

Unfortunately, the notions of breakpoint nodes and maximal spacing do not exist for $\mathbb{G}_2(n;\tau)$, $\mathbb{G}_1(n;\rho)$ and $\mathbb{G}_2(n;\rho)$, and new methods are therefore required to study their connectivity.

Penrose [47] solved the connectivity problem for $\mathbb{G}_2(n;\tau)$ by proving the following asymptotic equality

$$\lim_{n\to\infty} P_{con,2}(n;\tau_n) = \lim_{n\to\infty} P_{iso,2}(n;\tau_n) = e^{-e^{-\alpha}} \tag{3.6}$$

for a range function $\tau : \mathbb{N} \to \mathbb{R}_+$ in the form

$$\tau_n = \sqrt{\frac{\log n + \alpha}{\pi n}}$$

for some α in $\mathbb{R}$, and $P_{iso,2}(n;\tau_n)$ is the probability of the absence of isolated nodes in $\mathbb{G}_2(n;\tau_n)$. According to (3.6), for a two-dimensional GRG $\mathbb{G}_2(n;\tau_n)$, the absence of isolated nodes can be viewed as equivalent to connectivity when n is large enough. Thus, we can instead study the absence of isolated nodes in $\mathbb{G}_2(n;\tau)$. The detailed analysis can be found in [47].

Similarly, we will study the connectivity of $\mathbb{G}_2(n;\rho)$ under the conjecture that the graph connectivity and the absence of isolated nodes are asymptotically equivalent events. For $\mathbb{G}_1(n;\rho)$, however, connectivity and the absence of isolated nodes can not be regarded as equivalent events even if n is very large. Up to now, we have not found a satisfactory approach to study the connectivity of $\mathbb{G}_1(n;\rho)$.

In this dissertation, we will mainly investigate the connectivity of $\mathbb{G}(n;\tau)$ and $\mathbb{G}_2(n;\rho)$, using approaches introduced in the rest of this Chapter.

3.2 Mathematical tools

3.2.1 Spacings

In this Section, we consider the one-dimensional graph $\mathbb{G}(n;\tau)$. We begin with a useful fact concerning the distributional properties of spacings [12, Eq. (6.4.3), p. 135].

Lemma 3.1 *For any fixed subset $I \subseteq \{1,\ldots,n\}$, we have*

$$\mathbb{P}\left[L_{n,k} > t_k,\ k \in I\right] = \left(1 - \sum_{k\in I} t_k\right)_+^n, \quad t_k \in [0,1]$$

with the notation $x_+^n = x^n$ if $x \geq 0$ and $x_+^n = 0$ if $x \leq 0$.

From Lemma 3.1, the distribution of the maximal spacing can be derived by the inclusion-exclusion principle.

Lemma 3.2 *It holds*

$$\mathbb{P}[M_n \le x] = \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} (1-kx)_+^n. \quad x \in [0,1] \tag{3.7}$$

This result has been rediscovered by several authors, e.g., Godehardt and Jaworski [25, Cor. 1, p. 146], and Desai and Manjunath [14] (as Eqn (8) with $z = 1$ and $r = \tau$). According to (3.5), we can see that (3.7) actually provides an explicit expression for $P(n;\tau)$. Its usefulness and constraints will be discussed in Chapter 4.

An elegant relationship exists between the spacings and exponential distributions [48, p. 404]:

Lemma 3.3 *Let $\xi_1, \ldots, \xi_{n+1}$ be i.i.d exponential rvs with unit parameter. We have the stochastic equivalence*

$$(L_{n,1}, \ldots, L_{n,n+1}) =_{st} \left(\frac{\xi_1}{T_{n+1}}, \ldots, \frac{\xi_{n+1}}{T_{n+1}} \right)$$

where $T_{n+1} = \xi_1 + \ldots + \xi_{n+1}$.

With the help of Lemma 3.3, we are now ready to present some useful properties of maximal spacings. The following is a variation of a result given by Lévy [39].

Theorem 3.1 *It holds that*

$$\frac{nM_n}{\log n} \xrightarrow{P}_n 1 \tag{3.8}$$

and

$$nM_n - \log n \Longrightarrow_n \Lambda \tag{3.9}$$

where Λ denotes a Gumbel rv with distribution

$$P(\Lambda \le x) = e^{-e^{-x}}, \quad x \in \mathbb{R}.$$

Proof. First note that (3.9) implies (3.8) since

$$\frac{nM_n - \log n}{\log n} = \left(\frac{nM_n}{\log n} - 1 \right), \quad n = 2, 3, \ldots$$

and we need only establish (3.9).

Fix $n = 2, 3, \ldots$. By Lemma 3.3, we have

$$\begin{aligned} M_n &= \max_{k=2,\ldots,n} L_{n,k} \\ &=_{st} \max_{k=2,\ldots,n} \left(\frac{\xi_k}{T_{n+1}} \right) \\ &= \frac{1}{T_{n+1}} \left(\max_{k=2,\ldots,n} \xi_k \right). \end{aligned}$$

Therefore,

$$\begin{aligned} nM_n - \log n &=_{st} \frac{n}{T_{n+1}} \left(\max_{k=2,\ldots,n} \xi_k \right) - \log n \\ &= \frac{n}{T_{n+1}} \left(\max_{k=2,\ldots,n} \xi_k - \log n \right) + \delta_n \end{aligned} \tag{3.10}$$

with

$$\delta_n := \frac{n}{T_{n+1}} \sqrt{n} \left(1 - \frac{T_{n+1}}{n} \right) \frac{\log n}{\sqrt{n}}.$$

The Strong Law of Large Numbers [19] gives

$$\lim_{n \to \infty} \frac{T_{n+1}}{n} = 1 \quad a.s.,$$

while the Central Limit Theorem [19] yields

$$\sqrt{n}\left(\frac{T_{n+1}}{n} - 1\right) \Longrightarrow_n U$$

with $U =_{st} N(0,1)$. It is now easy to check that $\delta_n \Longrightarrow_n 0$, whence $\delta_n \xrightarrow{P}_n 0$.

Therefore it follows from (3.10) that in order to prove (3.9), we only need to prove

$$\max_{k=2,\ldots,n} \xi_k - \log n \Longrightarrow_n \Lambda.$$

For each x in $\mathbb{R}$ and $n = 2, 3, \ldots$, we have

$$\begin{aligned}
\mathbb{P}\left[\max_{k=2,\ldots,n} \xi_k - \log n \le x\right] &= \mathbb{P}[\xi_k \le x + \log n, k = 2, \ldots, n] \\
&= \prod_{k=2}^{n} \mathbf{P}[\xi_k \le x + \log n] \\
&= \left(1 - e^{-(x+\log n)}\right)^{n-1} \\
&= \left(1 - \frac{1}{n}e^{-x}\right)^{n-1},
\end{aligned}$$

so that

$$\lim_{n\to\infty} \mathbf{P}\left[\max_{k=2,\ldots,n} \xi_k - \log n \le x\right] = e^{-e^{-x}}$$

as needed.

■

For the case of a general distribution F [13, Theorem 4, p. 1183], we have the following result on the maximal spacing.

Theorem 3.2 *Assume that F admits a continuous probability density function $f : [0,1] \to \mathbb{R}_+$ which satisfies the following conditions:*

(i) There exists an isolated minimum $x_\star$ of f on $[0,1]$.

(ii) There exist positive constants r, d_r and D_r such that

$$d_r := \liminf_{h\to 0} \frac{f(x_\star + h) - f(x_\star)}{|h|^r}$$

and

$$D_r := \limsup_{h\to 0} \frac{f(x_\star + h) - f(x_\star)}{|h|^r}.$$

Then, we have the bounds

$$-\frac{1}{r} = \liminf_{n\to\infty} \frac{nM_n f_\star - \log n}{\log\log n} < \limsup_{n\to\infty} \frac{nM_n f_\star - \log n}{\log\log n} = 2 - \frac{1}{r} \quad a.s.$$

While Theorem 3.2 implies $nM_nf_\star - \log n = \Theta(\log\log n)$ a.s., this result does not establish an exact asymptotic relationship between $nM_nf_\star - \log n$ and $\log\log n$. We will return to this issue in Chapter 6.

3.2.2 Methods of the first and second moments

The method of first moment is a special case of Markov's inequality.

Lemma 3.4 *For any $\mathbb{N}$-valued rv Z with $E[Z] < \infty$, we have*

$$P[Z = 0] \geq 1 - E[Z].$$

The method of second moment is a simple corollary of the Cauchy-Schwartz inequality.

Lemma 3.5 *For any* $\mathbb{N}$*-valued rv* Z *with* $0 < \mathbb{E}[Z^2] < \infty$*, we have*

$$\mathbb{P}[Z = 0] \leq 1 - \frac{\mathbb{E}[Z]^2}{\mathbb{E}[Z^2]}.$$

Proof. By the Cauchy-Schwartz inequality,

$$\mathbb{E}[Z]^2 = \mathbb{E}[\mathbf{1}[Z \neq 0]Z]^2 \leq \mathbb{E}[\mathbf{1}[Z \neq 0]^2]\mathbb{E}[Z^2],$$

so that

$$\frac{\mathbb{E}[Z]^2}{\mathbb{E}[Z^2]} \leq \mathbb{P}[Z \neq 0].$$

■

The following result can be deduced from Lemma 3.4 and Lemma 3.5, and holds the key to many of the proofs given in the thesis.

Theorem 3.3 *Let* $\{Z_n, n = 1, 2, \ldots\}$ *be a sequence of* $\mathbb{N}$*-valued rvs with* $\mathbb{E}[Z_n^2] < \infty$ *for each* $n = 1, 2, \ldots$*. Then, the convergence statements*

$$\lim_{n\to\infty} \mathbb{P}[Z_n = 0] = 1 \quad \textit{if} \quad \lim_{n\to\infty} \mathbb{E}[Z_n] = 0,$$

and

$$\lim_{n\to\infty} \mathbb{P}[Z_n = 0] = 0 \quad \textit{if} \quad \lim_{n\to\infty} \frac{\mathbb{E}[Z_n]^2}{\mathbb{E}[Z_n^2]} = 1$$

hold.

In the forthcoming chapters, we will have the need to evaluate the distribution of the rv Z_n, where Z_n is the sum of indicator rvs $\{I_{n,\alpha}, \alpha \in \Gamma_n\}$ with $\Gamma_n \subseteq \{1, 2, \ldots, n\}$ and $\lim_{n\to\infty} |\Gamma_n| = \infty$. In particular, we will want to know the limiting probability of the event $Z_n = 0$. Lemma 3.4 and Lemma 3.5 only provide lower and upper bounds to $\mathbb{P}[Z_n = 0]$, respectively, and we can not determine $\lim_{n\to\infty} \mathbb{P}[Z_n = 0]$ unless the upper (resp. lower) bound tends to zero (resp. one), cases which are handled by Theorem 3.3. If neither bound approaches its extreme value, the method of first and second moments can not be used.

If the indicator rvs $\{I_{n,\alpha}, \alpha \in \Gamma_n\}$ are mutually independent, and

$$\lim_{n\to\infty} \sup_{\alpha\in\Gamma_n} \mathbb{E}[I_{n,\alpha}] = 0 \quad \text{with} \quad \lim_{n\to\infty} \mathbb{E}[Z_n] = \lambda < +\infty,$$

we can approximate the distribution of Z_n by a Poisson distribution with mean λ, and conclude $\lim_{n\to\infty} \mathbb{P}[Z_n = 0] = e^{-\lambda}$. If, however, the indicator rvs are not independent, we have to resort to the Stein-Chen method [5] described in the next section.

3.2.3 The Stein-Chen method

Throughout this section, we define $Z := \sum_{\alpha\in\Gamma} I_\alpha$ for some finite index set Γ, where $\{I_\alpha, \alpha \in \Gamma\}$ are $\{0, 1\}$-valued rvs, and set $\lambda := \mathbb{E}[Z]$. We denote by $\Pi(\lambda)$ a Poisson rv with parameter λ.

Essentially, the Stein-Chen method computes upper bounds on the *total variation distance* between the rv Z and the Poisson rv $\Pi(\lambda)$. The definition of total variation distance is stated as follows:

Definition 3.3 *Let X and Y be $\mathbb{N}$-valued rvs. The total variation distance between X and Y is defined by*

$$d_{TV}(X,Y) := \frac{1}{2}\sum_{k=0}^{\infty}\Big|P[X=k]-P[Y=k]\Big|.$$

The Stein-Chen method is advantageous over other Poisson convergence methods (e.g., Brun's sieve [1, p. 119]) for the following reasons:

1. It does not only establish Poisson convergence, but also often leads to a rate of convergence; and

2. We only need to compute the first two moments of Z.

The following property of the total variation distance is immediate from its definition.

Lemma 3.6 *For $\mathbb{N}$-valued rvs X and Y, we have*

$$\Big|\mathbb{P}[X=t]-\mathbb{P}[Y=t]\Big| \leq d_{TV}(X,Y), \quad t=0,1,\ldots.$$

Proof. Indeed, for each $t=0,1,\ldots$, we have

$$\begin{aligned} d_{TV}(X,Y) &= \frac{1}{2}\sum_{k=0}^{\infty}\Big|P[X=k]-P[Y=k]\Big| \\ &= \frac{1}{2}\left(\Big|\mathbb{P}[X=t]-\mathbb{P}[Y=t]\Big| + \sum_{k\geq 0, k\neq t}\Big|P[X=k]-P[Y=k]\Big|\right) \\ &\geq \frac{1}{2}\left(\Big|\mathbb{P}[X=t]-\mathbb{P}[Y=t]\Big| + \Big|\mathbb{P}[X\neq t]-\mathbb{P}[Y\neq t]\Big|\right) \\ &= \Big|\mathbb{P}[X=t]-\mathbb{P}[Y=t]\Big|. \end{aligned}$$

$\blacksquare$

A general result on the Stein-Chen method is given in Theorem 3.4 [5, Theorem 2. A, p. 23].

Theorem 3.4 *Assume that for each α in Γ, the rvs U_α and V_α can be constructed on a common probability space such that*

$$U_\alpha =_{st} Z \quad \text{and} \quad (1+V_\alpha) =_{st} (Z|I_\alpha = 1).$$

Then, the bound

$$d_{TV}(Z, \Pi(\lambda)) \leq \left(\frac{1-e^{\lambda}}{\lambda}\right) \sum_{\alpha\in\Gamma} \mathbb{E}[I_\alpha]\mathbb{E}[|U_\alpha - V_\alpha|]$$

holds.

The upper bound of Theorem 3.4 can be greatly simplified if some form of negative or positive dependence holds amongst the indicator rvs $\{I_\alpha, \alpha \in \Gamma\}$. The definition of negatively (positively) related indicators [5, Definition 2. 1. 1, p. 24] is given below.

Definition 3.4 *The indicator rvs $\{I_\alpha,\ \alpha \in \Gamma\}$ are said to be negatively related if for each α in Γ, there exists rvs $\{J_{\beta\alpha},\ \beta \in \Gamma_n\}$ defined on the same probability space as $\{I_\beta,\ \beta \in \Gamma\}$ such that*

$$(J_{\beta\alpha}, \beta \in \Gamma\backslash\{\alpha\}) =_{st} (I_\beta, \beta \in \Gamma\backslash\{\alpha\}|I_\alpha = 1) \tag{3.11}$$

and

$$J_{\beta\alpha} \leq I_\beta, \quad \beta \in \Gamma\backslash\{\alpha\}. \tag{3.12}$$

The rvs $\{I_\alpha,\, \alpha \in \Gamma\}$ are positively related when (3.12) is replaced by

$$J_{\beta\alpha} \geq I_\beta, \quad \beta \in \Gamma \backslash \{\alpha\}. \tag{3.13}$$

If we set $U_\alpha = Z$ and $V_\alpha = \sum_{\beta \in \Gamma \backslash \{\alpha\}} J_{\beta\alpha}$ in Theorem 3.4, the following corollaries [5, Cor 2. C. 2, Cor 2. C. 4, p. 26] can be derived after some simple algebraic manipulations.

Corollary 3.1 *If the rvs $\{I_\alpha,\, \alpha \in \Gamma\}$ are negatively related, then we have*

$$d_{TV}(Z, \Pi(\lambda)) \leq \frac{1 - e^{-\lambda}}{\lambda} \left(\lambda - Var[Z]\right).$$

Corollary 3.2 *If rvs $\{I_\alpha,\, \alpha \in \Gamma\}$ are positively related, then we have*

$$d_{TV}(Z, \Pi(\lambda)) \leq \frac{1 - e^{-\lambda}}{\lambda} \left(Var[Z] - \lambda + 2 \sum_{\alpha \in \Gamma} \mathbb{E}[I_\alpha]^2 \right).$$

It is often difficult to show that the indicator rvs $\{I_\alpha,\, \alpha \in \Gamma\}$ are negatively (resp., positively) related. However, the upper bound of Theorem 3.4 can also be simplified if there exists *neighborhoods of dependence* in the following sense.

Definition 3.5 *For α in Γ, consider a subset $B_\alpha \subset \Gamma$ such that α is an element of B_α. We say that B_α is a neighborhood of dependence for α if I_α is independent of the rvs $\{I_\beta, \beta \in \Gamma \backslash B_\alpha\}$.*

With this definition, we have the following result [5, Cor 2. C. 5, p. 26]:

Corollary 3.3 *For each α in Γ, let B_α be a neighborhood of dependence for α. Then, it holds*

$$d_{TV}(Z, \Pi(\lambda)) \leq \frac{1-e^{-\lambda}}{\lambda}(A_1 + A_2) \tag{3.14}$$

where

$$A_1 = \sum_{\alpha\in\Gamma} \sum_{\beta\in B_\alpha} \mathbb{E}[I_\alpha]\mathbb{E}[I_\beta]$$

and

$$A_2 = \sum_{\alpha\in\Gamma} \sum_{\beta\in B_\alpha\setminus\{\alpha\}} E[I_\alpha I_\beta].$$

In many applications of Corollary 3.3, the size of the neighborhoods $\{B_\alpha, \alpha \in \Gamma\}$ is much smaller than the size n of the index set Γ, and the upper bound in (3.14) will go to zero as n tends to infinity. This is not the case, however, if the size of the neighborhoods is of the same order of n.

Chapter 4

Network connectivity under the GRG model I : Uniform user distribution

4.1 Introduction

The GRG model with uniform node placement is the simplest scenario to be considered in this dissertation. This model, despite its simplicity, deserves attention for the following reasons: Firstly, it captures the essential feature that the existence of a link between two users is a *random* event whose success probability is mainly determined by their *distance*. Secondly, the approaches used in this model can be easily adapted to analyze more complicated models.

In order to find the minimum transmission range to ensure graph connectivity with high probability, it is desirable to derive a closed form expression of $P(n;\tau)$. Under uniform node placement, such an expression is given by

$$P(n;\tau) = \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} (1-k\tau)_+^n, \tag{4.1}$$

which is immediate from Lemma 3.2 by replacing x with τ in (3.7).

For any given n and target probability of graph connectivity p_{con} (e.g. 0.99), we can numerically calculate the minimum transmission range using (4.1). Such a numerical result, albeit useful, can not be obtained for the general scenario. When the graph setting is slightly changed (e.g. from uniform distribution to non-uniform distribution), it becomes extremely hard to find out a similar closed form expression.

On the other hand, as we will see in the rest of this dissertation, approaches introduced in Chapter 3 can be more widely used. Rather than simply obtaining numerical results, we will establish an explicit expression for the transmission range as a function of n and p_{con}. Moreover, we identify a critical threshold of the transmission range such that a slightly larger (resp. smaller) value will lead to a connected (resp. disconnected) graph.

To get some intuitive impression of the relationship between the transmission range and graph connectivity, we first look at some simulation results in Fig. 4.1. Fig. 4.1 displays the probability of graph connectivity $P(n;\tau)$ as a function of τ (in base 10 logarithm) for $n = 10$, $1,000$ and $100,000$. For each n, we have generated $10,000$ independent topologies of n nodes uniformly and independently distributed on the interval $[0,1]$. The value of $P(n;\tau)$ is estimated as the percentage of topologies that result in a connected graph when the transmission range is τ.

We observe from Fig. 4.1 that there exist *sharp transitions* from $P(n;\tau) = 0$ to $P(n;\tau) = 1$ as τ varies across some critical threshold. The larger n, the sharper the transition. If we can identify this critical threshold, then a suitably larger (resp. smaller) range will lead to graph connectivity (resp. disconnectivity) with high probability. Thus in order to find the minimum transmission range ensuring

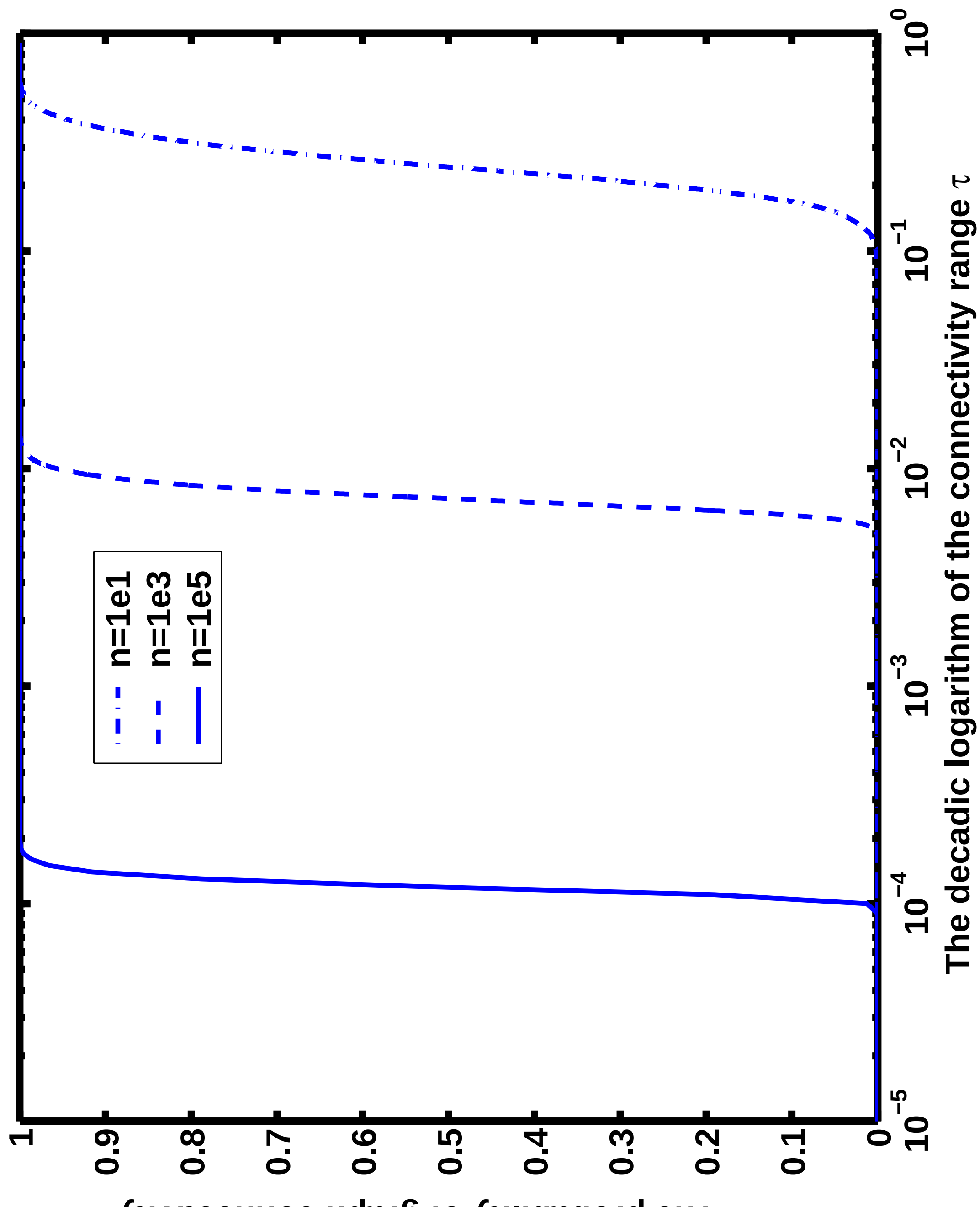


Figure 4.1: Existence of sharp phase transitions.

graph connectivity, we need to know both the critical threshold and the sharpness of the transition. According to the sharpness of the transition around this critical threshold, we introduce the notion of threshold functions in the next Section.

4.2 Threshold functions for graph connectivity

4.2.1 Definitions

A range function τ is defined as any mapping $\tau : \mathbb{N}_0 \to \mathbb{R}_+$. A range function τ^* is said to be a *weak threshold* [42, p. 376] if

$$\lim_{n\to\infty} P(n;\tau_n) = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \frac{\tau_n}{\tau_n^*} = 0 \\ \\ 1 & \text{if } \lim_{n\to\infty} \frac{\tau_n}{\tau_n^*} = +\infty. \end{cases} \tag{4.2}$$

A range function τ^* is said to be a *strong threshold* [42, p. 376] if

$$\lim_{n\to\infty} P(n;c\tau_n^*) = \begin{cases} 0 & \text{if } 0 < c < 1 \\ \\ 1 & \text{if } 1 < c. \end{cases} \tag{4.3}$$

Finally, we say that τ^* is a *very strong threshold* if

$$\lim_{n\to\infty} P(n;\tau_n) = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \alpha_n = -\infty \\ \\ 1 & \text{if } \lim_{n\to\infty} \alpha_n = +\infty \end{cases} \tag{4.4}$$

for any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form

$$\tau_n = \tau_n^* + \frac{\alpha_n}{n}, \quad n = 1, 2, \ldots. \tag{4.5}$$

for some deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}_+$. Please note that there is no loss of generality in writing a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ in the form of (4.5).

4.2.2 Threshold functions and maximal spacings

Let the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ be considered as a candidate threshold function. We will show in the following two Lemmas that maximal spacings can be used to check whether the range function $\tau^\star$ is indeed a weak or a strong threshold.

Lemma 4.1 *Under the enforced assumptions, if there exists an $\mathbb{R}_+$-valued rv L such that*

$$\frac{M_n}{\tau_n^\star} \Longrightarrow_n L \tag{4.6}$$

holds with $\mathbb{P}[L = 0] = 0$, then the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ is a weak threshold.

Proof. Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ which satisfies

$$\lim_{n\to\infty} \frac{\tau_n}{\tau_n^\star} = \infty,$$

so that for each $B > 0$, there exists an integer $n^\star(B)$ such that $\tau_n > B\tau_n^\star$ whenever $n \geq n^\star(B)$. Thus we find

$$P(n; \tau_n) = \mathbb{P}(M_n \leq \tau_n) \geq \mathbb{P}\left[\frac{M_n}{\tau_n^\star} \leq B\right], \quad n \geq n^\star(B). \tag{4.7}$$

Letting n go to infinity in this last inequality yields

$$\liminf_{n\to\infty} P(n; \tau_n) \geq \liminf_{n\to\infty} \mathbb{P}\left[\frac{M_n}{\tau_n^\star} \leq B\right]. \tag{4.8}$$

With B a point of continuity for the distribution of L, we can invoke (4.6) in order to strengthen (4.8) as

$$\liminf_{n\to\infty} P(n;\tau_n) \geq \mathbb{P}\left[L \leq B\right].$$

The points of continuity of the distribution of L form a dense set in $\mathbb{R}_+$. Therefore, letting B go to infinity along such points of continuity, we get

$$\liminf_{n\to\infty} P(n;\tau_n) \geq \lim_{B\to\infty} \mathbb{P}\left[L \leq B\right] = 1$$

since L is $\mathbb{R}_+$-valued, whence $\lim_{n\to\infty} P(n;\tau_n) = 1$.

Next, consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ such that

$$\lim_{n\to\infty} \frac{\tau_n}{\tau_n^\star} = 0.$$

This time, for each $\varepsilon > 0$, there exists an integer $n^\star(\varepsilon)$ such that $\tau_n < \varepsilon\tau_n^\star$ whenever $n \geq n^\star(\varepsilon)$, and we conclude to

$$P(n;\tau_n) = \mathbb{P}(M_n \leq \tau_n) \leq \mathbb{P}\left[\frac{M_n}{\tau_n^\star} \leq \varepsilon\right], \quad n \geq n^\star(\varepsilon). \tag{4.9}$$

Letting n go to infinity in this last inequality yields

$$\limsup_{n\to\infty} P(n;\tau_n) \leq \limsup_{n\to\infty} \mathbb{P}\left[\frac{M_n}{\tau_n^\star} \leq \varepsilon\right]. \tag{4.10}$$

If we pick ε to be a point of continuity for the distribution of L, we can invoke (4.6) in order to strengthen (4.10) as

$$\limsup_{n\to\infty} P(n;\tau_n) \leq \mathbb{P}\left[L \leq \varepsilon\right].$$

Letting ε go to zero along the points of continuity of the distribution of L, we get

$$\limsup_{n\to\infty} P(n;\tau_n) \leq \lim_{\varepsilon\to 0} \mathbb{P}\left[L \leq \varepsilon\right] = \mathbb{P}[L = 0],$$

and we conclude to $\lim_{n\to\infty} P(n;\tau_n) = 0$ as desired since $L > 0$ a.s.. This completes the proof that the range function $\tau^\star$ is indeed a weak threshold for $\mathbb{G}(n;\tau)$. ■

Lemma 4.2 *Under the enforced assumptions, the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ is a strong threshold if and only if*

$$\frac{M_n}{\tau^\star_n} \xrightarrow{P}_n 1. \tag{4.11}$$

Proof. First, we note that (5.1) is equivalent to

$$\frac{M_n}{\tau^\star_n} \Longrightarrow_n 1 \tag{4.12}$$

since the modes of convergence in distribution and in probability are equivalent when the limit is a constant. However, the convergence (4.12) amounts to

$$\lim_{n\to\infty} P(n; c\tau^\star_n) = \begin{cases} 0 & \text{if } 0 < c < 1 \\ \\ 1 & \text{if } 1 < c. \end{cases} \tag{4.13}$$

■

4.2.3 Operational interpretation of threshold functions

For each $n = 2, 3, \ldots$, the *critical transmission range* for the n node network is defined as the rv R_n given by

$$R_n := \min \left(\tau > 0 : \mathbb{G}(n;\tau) \text{ is connected} \right).$$

In short, R_n is the smallest transmission range that ensures that the node set $X_1, \ldots, X_n$ forms a connected network. The obvious identity

$$R_n = M_n$$

leads to the following operational interpretation of threshold functions: By Lemma 4.2, the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ is a strong threshold function if and only if $R_n \sim \tau_n^\star$ for n large in some appropriate distributional sense (formalized at (4.11)). On the other hand, if $\tau^\star$ is a weak threshold function, then Lemma 4.1 only states that $R_n \sim \tau_n^\star L$ for n large with a non-zero (possibly non-degenerate) rv L. In either case, but with different degrees of accuracy, the threshold function serves as a proxy or estimate of the critical transmission range for the many node networks.

4.3 A very strong threshold function for graph connectivity

4.3.1 The main result and preliminary analysis

Theorem 4.1 *Under the assumption that $X_1, \ldots, X_n$ are i.i.d rvs distributed according to $U_{[0,1]}$, the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ given by $\tau_n^* = \frac{\log n}{n}, n = 1, 2, \ldots$ is* *a very strong threshold for graph connectivity.*

The fact that $\frac{\log n}{n}$ is a strong threshold has been discovered by several authors, e.g., see [2, p. 352, Theorem 1] and [45, Thm. 2.2]. From (4.3), it follows that a perturbation of $\frac{(c-1)\log n}{n}$ from τ_n^* will yield the one law (resp. zero law) of $P(n;\tau_n)$ when $c > 1$ (resp. $0 < c < 1$). However, the very strong threshold established by Theorem 4.1 indicates that "smaller" deviation $\frac{\alpha_n}{n}$ from τ_n^* can have the same effect on $P(n;\tau_n)$, with the only constraint $\lim_{n\to\infty}\alpha_n = \infty$ (resp. $\lim_{n\to\infty}\alpha_n = -\infty$). For example, Theorem 4.1 implies that the deviation $\frac{\log\log n}{n}$ from τ_n^* will lead to zero-one laws of $P(n;\tau_n)$, a fact which can not be established from (4.3).

We introduce some easy convergence facts to be used in the proof of Theorem 4.1: With $0 \le x < 1$, it is a simple matter to check that

$$\log(1-x) = -\int_0^x \frac{1}{1-t}dt = -x - \Psi(x) \tag{4.14}$$

where we have set

$$\Psi(x) := \int_0^x \frac{t}{1-t}dt, \quad 0 \le x < 1. \tag{4.15}$$

The mapping $x \to \Psi(x)$ is increasing and convex on the interval $[0,1)$ with

$$0 < \Psi(x) \le \frac{x^2}{2(1-x)}, \quad 0 \le x < 1. \tag{4.16}$$

Now consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ in the form

$$\tau_n = \frac{1}{n}\left(\log n + \alpha_n\right), \quad n = 1, 2, \dots \tag{4.17}$$

for some deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$. Note that the range function in (4.17) has the same form as that in (4.5) by replacing τ_n^* with $\frac{\log n}{n}$.

For each $p > 0$, provided $p\tau_n < 1$, the decomposition (A.1) yields

$$\begin{aligned} (1 - p\tau_n)_+^n &= e^{-n(p\tau_n + \Psi(p\tau_n))} \\ &= e^{-p(\log n + \alpha_n)} e^{-n\Psi(p\tau_n)} \\ &= n^{-p} e^{-p\alpha_n} e^{-n\Psi(p\tau_n)}. \end{aligned} \tag{4.18}$$

The next two technical lemmas rely on this observation; they will be useful in a number of places.

Lemma 4.3 *For any range function* $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ *in the form (4.17) with* $\lim_{n\to\infty} \alpha_n = -\infty$*, we have*

$$\lim_{n\to\infty} \frac{(1 - p\tau_n)_+^n}{n^{-p} e^{-p\alpha_n}} = 1, \quad p > 0. \tag{4.19}$$

Proof. Fix $p > 0$. From the assumption $\lim_{n\to\infty} \alpha_n = -\infty$, we note that $\alpha_n < 0$ for large enough n and the form (4.17) therefore implies both $\tau_n \leq \frac{\log n}{n}$ and $\frac{|\alpha_n|}{n} \leq \frac{\log n}{n}$ on that range, whence

$$\lim_{n\to\infty} \tau_n = \lim_{n\to\infty} \frac{\alpha_n}{n} = 0$$

since $\lim_{n\to\infty} \frac{\log n}{n} = 0$. This already establishes that

$$p\tau_n < 1 \quad \text{for all sufficiently large } n \tag{4.20}$$

Still on that range, the monotonicity of Ψ yields

$$n\Psi(p\tau_n) \leq n\Psi\left(p\frac{\log n}{n}\right)$$

so that

$$n\Psi(p\tau_n) \leq \frac{p^2}{2} \cdot \left(1 - p\frac{\log n}{n}\right)^{-1} \cdot \frac{(\log n)^2}{n}$$

by invoking the bound (A.2). It is now plain that

$$\lim_{n\to\infty} n\Psi(p\tau_n) = 0. \tag{4.21}$$

To conclude, condition (4.20) ensures the validity of (4.18) for large enough n, and (4.21) readily implies (4.19) via (4.18).

■

Lemma 4.4 *Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ in the form (4.17). It holds that*

$$\lim_{n\to\infty} n(1-\tau_n)_+^n = \begin{cases} \infty & \textit{if } \lim_{n\to\infty} \alpha_n = -\infty \\ \\ 0 & \textit{if } \lim_{n\to\infty} \alpha_n = +\infty. \end{cases} \tag{4.22}$$

Proof. First, we note that

$$n\,(1-\tau_n)_+^n = e^{-\alpha_n} \cdot \frac{(1-\tau_n)_+^n}{n^{-1}e^{-\alpha_n}}, \quad n = 1, 2, \ldots \tag{4.23}$$

and Lemma 4.3 (with $p = 1$) readily yields the conclusion $\lim_{n\to\infty} n(1-\tau_n)_+^n = \infty$ when $\lim_{n\to\infty} \alpha_n = -\infty$.

We also have $n(1-\tau_n)_+^n = 0$ if $1 \leq \tau_n$, while when $\tau_n \leq 1$, the relation (4.18) yields $n(1-\tau_n)_+^n \leq e^{-\alpha_n}$ by the non-negativity of Ψ. It is now immediate that

$\lim_{n\to\infty} n(1-\tau_n)_+^n = 0$ when $\lim_{n\to\infty} \alpha_n = +\infty$. ■

4.3.2 A proof of Theorem 4.1

The basic idea of the proof is to leverage the representation (3.1) in order to provide lower and upper bounds on the probability of graph connectivity through moments of the counting variable $C_n(\tau)$. This is achieved through the method of first and second moments: According to Theorem 3.3, the one-law and zero-law follow if we show that

$$\lim_{n\to\infty} \mathbb{E}[C_n(\tau_n)] = 0 \quad \text{if} \quad \lim_{n\to\infty} \alpha_n = \infty \tag{4.24}$$

and

$$\lim_{n\to\infty} \frac{\mathbb{E}[C_n(\tau_n)^2]}{\mathbb{E}[C_n(\tau_n)]^2} = 1 \quad \text{if} \quad \lim_{n\to\infty} \alpha_n = -\infty \tag{4.25}$$

where τ is any range function in the form (4.17).

From Lemma 4.4, we readily see that

$$\lim_{n\to\infty} \mathbb{E}[C_n(\tau_n)] = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \alpha_n = +\infty \\ \\ \infty & \text{if } \lim_{n\to\infty} \alpha_n = -\infty, \end{cases} \tag{4.26}$$

and (4.24) is established.

Next, from Lemma 3.1, for each $n = 2, 3, \ldots$, it is a simple matter to derive the closed-form expressions

$$\mathbb{E}[C_n(\tau)] = (n-1)(1-\tau)_+^n \tag{4.27}$$

and

$$\mathbb{E}[C_n(\tau)^2] = \mathbb{E}[C_n(\tau)] + (n-1)(n-2)(1-2\tau)_+^n. \tag{4.28}$$

Thus as τ ranges over $(0,1)$, we conclude that

$$\frac{\mathbb{E}[C_n(\tau_n)^2]}{\mathbb{E}[C_n(\tau_n)]^2} = \mathbb{E}[C_n(\tau_n)]^{-1} + \frac{(n-2)}{(n-1)}\frac{(1-2\tau_n)_+^n}{(1-\tau_n)_+^{2n}}. \tag{4.29}$$

We have already shown that $\lim_{n\to\infty}\mathbb{E}[C_n(\tau_n)] = \infty$ whenever $\lim_{n\to\infty}\alpha_n = -\infty$. From Lemma 4.3 (first with $p=2$ and then $p=1$) under this last condition, we also get

$$\lim_{n\to\infty}\frac{(1-2\tau_n)_+^n}{n^{-2}e^{-2\alpha_n}} = 1 = \lim_{n\to\infty}\frac{(1-\tau_n)_+^n}{n^{-1}e^{-\alpha_n}}.$$

It is now a simple matter to check from these facts that

$$\lim_{n\to\infty}\frac{(1-2\tau_n)_+^n}{(1-\tau_n)_+^{2n}} = \lim_{n\to\infty}\frac{(1-2\tau_n)_+^n}{n^{-2}e^{-2\alpha_n}}\left[\frac{n^{-1}e^{-\alpha_n}}{(1-\tau_n)_+^n}\right]^2 = 1$$

and (4.25) follows upon letting n go to infinity in (4.29).

4.4 How fast does the transition take place?

4.4.1 Definitions

The threshold function only identifies when the phase transition takes place. A natural question consists in estimating how quickly this transition takes place. To address this issue, we introduce the following definitions: For each $n = 2, 3, \ldots$, the mapping $\tau \to P(n;\tau)$ can be shown to be continuous and strictly monotone increasing. Given fixed a in $(0,1)$, this property guarantees the existence and uniqueness of solutions to the equation

$$P(n;\tau)=a,\quad \tau\in(0,1). \tag{4.30}$$

Let $\tau_n(a)$ denote this unique solution, and whenever a lies in the interval $(0,\frac{1}{2})$, we set

$$\delta_n(a):=\tau_n(1-a)-\tau_n(a).$$

The transition width $\delta_n(a)$ measures how quickly $P(n;\tau)$ climbs from level a to level $1-a$, thereby giving an indication of the sharpness of the phase transition. Given the rather complex dependence of $\delta_n(a)$ on n and a, it is desirable to find asymptotic bounds (if nothing else) for large n.

4.4.2 The main result

The main result concerning the behavior of $\tau_n(a)$ for large n is given first.

Theorem 4.2 *For every a in the interval $(0,1)$, it holds that*

$$\tau_n(a)=\frac{\log n}{n}-\frac{1}{n}\log\left(\log\left(\frac{1}{a}\right)\right)+o\left(n^{-1}\right). \tag{4.31}$$

Theorem 4.2 is established in Section 4.4.5. The desired result on the width of the transition interval flows as an easy corollary.

Corollary 4.1 *For every a in the interval $(0,\frac{1}{2})$, we have*

$$\delta_n(a)=\frac{C(a)}{n}+o\left(n^{-1}\right) \tag{4.32}$$

with constant $C(a)$ given by

$$C(a) = \log\left(\frac{\log a}{\log(1-a)}\right). \tag{4.33}$$

It is a simple matter to check that $a \to C(a)$ is decreasing on the interval $(0, \frac{1}{2})$ with $\lim_{a\downarrow 0} C(a) = \infty$ and $\lim_{a\uparrow \frac{1}{2}} C(a) = 0$. These qualitative features are in line with one's intuition.

4.4.3 How to guess the result

We now present a plausibility argument which allows us to guess the validity of Theorem 4.2, and which eventually paves the way to its proof.

For each x in $\mathbb{R}$, define the $[0,1]$-valued sequence $\{\sigma_n(x),\ n = 1,2,\ldots\}$ by

$$\sigma_n(x) = \min\left(1, \left(\frac{\log n + x}{n}\right)_+\right), \quad n = 1,2,\ldots \tag{4.34}$$

so that

$$\sigma_n(x) = \frac{\log n + x}{n} \tag{4.35}$$

for n large enough. The next result can be easily extracted from (3.5) and Theorem 3.1. A similar result has also been obtained by Godehardt and Jaworski [25, Theorem. 2, p. 157]

Corollary 4.2 *For each x in $\mathbb{R}$, it holds that*

$$\lim_{n\to\infty} P(n;\sigma_n(x)) = p(x) \tag{4.36}$$

with

$$p(x) = e^{-e^{-x}}. \tag{4.37}$$

To see how the convergence (4.36) underpins Theorem 4.2, consider the following heuristic arguments: For each x in $\mathbb{R}$, the convergence (4.36) yields the approximation

$$P(n; \sigma_n(x)) \simeq p(x)$$

for large enough n. The mapping $p : \mathbb{R} \to \mathbb{R}_+ : x \to p(x)$ is strictly monotone and continuous with $\lim_{x\to-\infty} p(x) = 0$ and $\lim_{x\to\infty} p(x) = 1$. Therefore, for each a in the interval $(0, 1)$, there exists a unique scalar, denoted x_a, such that $p(x_a) = a$. In fact,

$$x_a = -\log\left(-\log a\right). \tag{4.38}$$

Given a in the interval $(0, 1)$, we find that

$$P(n; \sigma_n(x_a)) \simeq a$$

for large n, whence $P(n; \sigma_n(x_a)) \simeq P(n; \tau_n(a))$. This suggests (but not quite yet proves) that $\sigma_n(x_a)$ and $\tau_n(a)$ behave in tandem asymptotically, thereby laying the grounds for the validity of (4.31) – Just insert (4.38) into (4.35) and (4.37). These ideas form the basis for the proof of Theorem 4.2 found in Section 4.4.5.

4.4.4 Numerical validation

Below we present some limited numerical results validating the asymptotic results obtained here. We consider n users which are uniformly and independently distributed in the interval $[0,1]$, with n ranging from $n = 1,000$ to $n = 9,000$ in increments of $1,000$.

According to (3.7), for fixed a in $(0,1)$, the threshold $\tau_n(a)$ is calculated by solving the following equation

$$\sum_{k=0}^{k(\tau)} (-1)^k \binom{n-1}{k} (1-k\tau)^n = a \tag{4.39}$$

with $k(\tau) = \min(n-1, \lfloor \frac{1}{\tau} \rfloor)$. In these calculations, some care needs to be exercised owing to possible buffer overflows associated with the evaluation of combinatorial coefficients. To avoid computing directly the coefficients $\binom{n-1}{k}$, $k = 0, 1, \ldots, k(\tau)$, we focus instead on evaluating the quantities $b_k = \binom{n-1}{k}(1-k\tau)^n$, $k = 0, 1, \ldots, k(\tau)$ through the simple recursion

$$b_0 = 1; \quad b_{k+1} = \frac{n-k-1}{k+1} \left(1 - \frac{\tau}{1-k\tau}\right)^n \cdot b_k$$

for $k = 0, 1, \ldots, k(\tau) - 1$.

The asymptotics (4.31) and (4.32) suggest that we approximate $\tau_n(a)$ and $\delta_n(a)$ by the quantities

$$\tau_n^*(a) := \frac{\log n}{n} - \frac{1}{n} \log\left(\log\left(\frac{1}{a}\right)\right) \tag{4.40}$$

and

$$\delta_n^*(a) := \frac{C(a)}{n}. \tag{4.41}$$

The accuracy of these approximations is measured by the error variables

$$\xi_n(a) := |\tau_n(a) - \tau_n^*(a)| \text{ and } \varepsilon_n(a) := |\delta_n(a) - \delta_n^*(a)|.$$

The numerical results are computed for $a = 0.1$. The quantities $\tau_n(a)$ and $\tau_n^*(a)$ are plotted in Fig. 4.2(a). The results for $\delta_n(a)$ and $\delta_n^*(a)$ are displayed in Fig.

4.2(b). The symbols represent the numerical results (as per computations explained above) and the lines represent the approximations calculated by (4.40) and (4.41). It is plain that the approximations are highly accurate.

By virtue of Theorem 4.2 and Corollary 4.1, the approximation errors, namely $\xi_n(a)$ and $\varepsilon_n(a)$ should be of order $o(n^{-1})$. This is indeed reflected by Table 4.1 upon noting that $n\xi_n(0.1)$ and $n\varepsilon_n(0.1)$ all go to zero as n grows large.

Table 4.1: The asymptotic behavior of error variables

n	100	1000	10000	100000
$n\xi_n(0.1)$	0.1022	0.0245	0.0046	0.0007
$n\varepsilon_n(0.1)$	0.3213	0.0630	0.0106	0.0016

4.4.5 A proof of Theorem 4.2

Fix x in $\mathbb{R}$. We restate (4.36) by noting that for each $\varepsilon > 0$, there exists a finite integer $n^\star(\varepsilon, x)$ such that

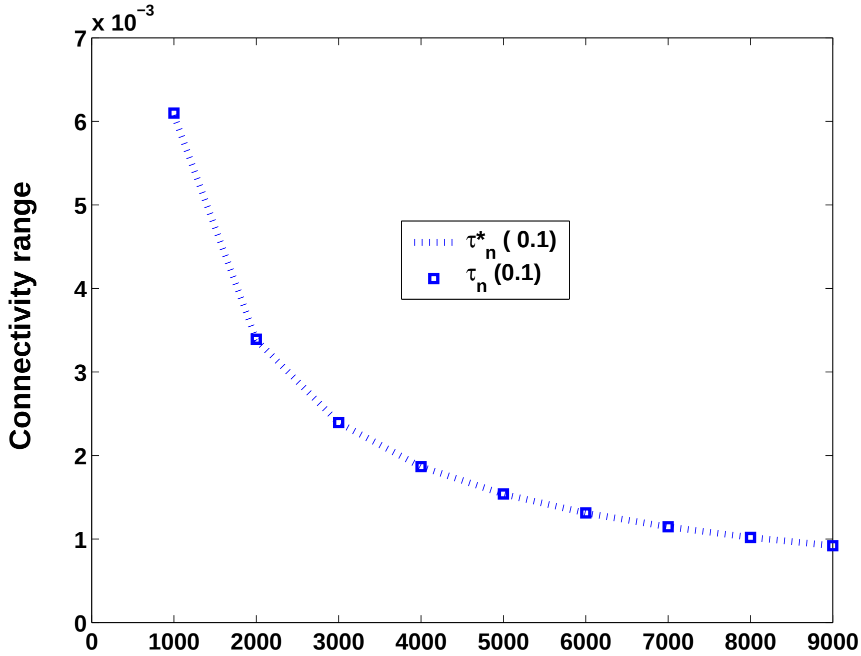
$$p(x) - \varepsilon < P(n; \sigma_n(x)) < p(x) + \varepsilon, \quad n \geq n^\star(\varepsilon, x). \tag{4.42}$$

Now fix a in the interval $(0, 1)$, and pick ε sufficiently small such that $0 < 2\varepsilon < a$ and $a + 2\varepsilon < 1$. Repeatedly applying (7.9) with $x = x_{a+\varepsilon}$ and $x = x_{a-\varepsilon}$, we get

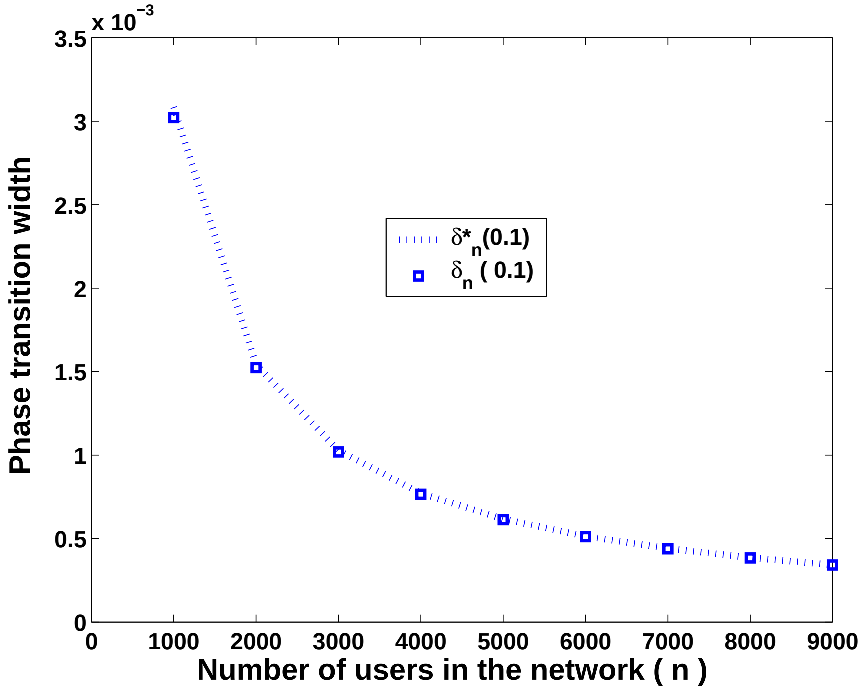
$$p(x_{a+\varepsilon}) - \varepsilon < P(n; \sigma_n(x_{a+\varepsilon})) < p(x_{a+\varepsilon}) + \varepsilon \tag{4.43}$$

whenever $n \geq n^\star(\varepsilon, x_{a+\varepsilon})$, and

$$p(x_{a-\varepsilon}) - \varepsilon < P(n; \sigma_n(x_{a-\varepsilon})) < p(x_{a-\varepsilon}) + \varepsilon \tag{4.44}$$



(a)



(b)

Figure 4.2: Connectivity range and phase transition width when $a = 0.1$

whenever $n \geq n^\star(\varepsilon, x_{a-\varepsilon})$. In the remainder of this proof, all inequalities are now understood to hold for $n \geq n^\star(a;\varepsilon)$ where we have set

$$n^\star(a;\varepsilon) = \max\left(n^\star(x_a), n^\star(\varepsilon, x_{a+\varepsilon}), n^\star(\varepsilon, x_{a-\varepsilon})\right)$$

with $n^\star(x)$ denoting the finite integer beyond which the representation (4.35) holds.

Since $p(x_{a\pm\varepsilon}) = a \pm \varepsilon$, the two chains of inequalities (4.43) and (4.44) can be rewritten as

$$a < P(n; \sigma_n(x_{a+\varepsilon})) < a + 2\varepsilon$$

and

$$a - 2\varepsilon < P(n; \sigma_n(x_{a-\varepsilon})) < a.$$

Thus,

$$P(n; \tau(n; a)) < P(n; \sigma_n(x_{a+\varepsilon})) < P(n; \tau_n(a + 2\varepsilon))$$

and

$$P(n; \tau_n(a - 2\varepsilon)) < P(n; \sigma_n(x_{a-\varepsilon})) < P(n; \tau(n; a)),$$

and the strict monotonicity of $\tau \to P(n;\tau)$ yields

$$\tau_n(a) < \sigma_n(x_{a+\varepsilon}) < \tau_n(a + 2\varepsilon)$$

and

$$\tau_n(a - 2\varepsilon) < \sigma_n(x_{a-\varepsilon}) < \tau_n(a).$$

Combining these last two inequalities, we conclude that

$$\sigma_n(x_{a-\varepsilon}) < \tau_n(a) < \sigma_n(x_{a+\varepsilon}). \tag{4.45}$$

Upon writing

$$\zeta_n(a) = \tau_n(a) - \sigma_n(x_a), \quad n = 2, 3, \ldots \tag{4.46}$$

we obtain from (4.45) that

$$\sigma_n(x_{a-\varepsilon}) - \sigma_n(x_a) < \zeta_n(a) < \sigma_n(x_{a+\varepsilon}) - \sigma_n(x_a)$$

with

$$\sigma_n(x_{a\pm\varepsilon}) - \sigma_n(x_a) = n^{-1}(x_{a\pm\varepsilon} - x_a)$$

As a result, $x_{a-\varepsilon} - x_a \leq \liminf_{n\to\infty}(n\zeta_n(a))$ and $\limsup_{n\to\infty}(n\zeta_n(a)) \leq x_{a+\varepsilon} - x_a$. Given that ε can be taken to be arbitrary small, it follows that

$$\liminf_{n\to\infty}(n\zeta_n(a)) = \limsup_{n\to\infty}(n\zeta_n(a)) = 0$$

since $\lim_{\varepsilon\downarrow 0}(x_{a\pm\varepsilon} - x_a) = 0$.

Thus, $\lim_{n\to\infty}(n\xi_n(a)) = 0$, whence $\zeta_n(a) = o(n^{-1})$. Reporting into (4.46) leads to

$$\tau_n(a) = \sigma_n(x_a) + o(n^{-1}), \quad n = 2, 3, \ldots$$

and the desired result readily follows from (4.34) and (4.38).

4.5 Finite node analysis

Up to now, we have only derived asymptotic results as the number of nodes tends to infinity. However, since the number of nodes in a network is finite, it is desirable to estimate $P(n;\tau)$ for given n and τ. Although we can compute $P(n;\tau)$ through (4.1), we can not find a similar expression for more general scenarios. Thus we resort to the Stein-Chen method introduced in Section 3.2.3. Our results are presented in Theorem 4.3 and Corollary 4.3. Similar issues have been discussed by Barbour and Holst [6, Theorem 6.1, p. 83].

We begin with a simple technical fact concerning binary valued rvs. For some $n = 2, 3, \ldots$, consider a collection of $\{0,1\}$-valued rvs $\xi_1, \ldots, \xi_n$ defined on the same probability space. Next, with $\mathcal{P}_n^\star$ denoting the collection of all non-empty subsets of $\{1, \ldots, n\}$, we define

$$P(K) := \mathbb{P}\left[\xi_k = 1,\ k \in K\right], \quad K \in \mathcal{P}_n^\star.$$

Lemma 4.5 *The probabilities $\{P(K),\ K \in \mathcal{P}_n^\star\}$ collectively determine the joint pmf of the $\{0,1\}^n$-valued rv $(\xi_1, \ldots, \xi_n)$.*

Proof. Pick an arbitrary non-zero element $\boldsymbol{a} = (a_1, \ldots, a_n)$ in $\{0,1\}^n$, and write $K(\boldsymbol{a}) = \{k = 1, \ldots, n : \ a_k = 1\}$ and $K^c(\boldsymbol{a}) = \{k = 1, \ldots, n : \ a_k = 0\}$. Direct inspection yields

$$\begin{aligned} \mathbb{P}\left[\xi_k = a_k,\ k = 1, \ldots, n\right] &= \mathbb{E}\left[\prod_{k \in K(\boldsymbol{a})} \xi_k \cdot \prod_{k \in K^c(\boldsymbol{a})} (1 - \xi_k)\right] \\ &= \sum_{K \in \mathcal{P}_n(\boldsymbol{a})} c(K) P(K) \end{aligned} \tag{4.47}$$

for some appropriate collection $\mathcal{P}_n(\boldsymbol{a})$ of subsets of $\{1, \ldots, n\}$, and coefficients $\{c(K),\ K \in \mathcal{P}_n(\boldsymbol{a})\}$ taking values ± 1. We have used the convention $\prod_{k \in K^c(\boldsymbol{a})}(1 - \xi_k) = 1$ when $K(\boldsymbol{a})$ is empty. The proof is completed upon noting that

$$\mathbb{P}\left[\xi_1 = \ldots = \xi_n = 0\right] = 1 - \sum_{K \in \mathcal{P}_n^\star} P(K).$$

■

Theorem 4.3 *For each* $n = 2, 3, \ldots$ *and* τ *in the interval* $(0, 1)$, *it holds that*

$$d_{TV}\Big(C_n(\tau); \Pi(\lambda_n(\tau))\Big) \leq \frac{1 - e^{-\lambda_n(\tau)}}{\lambda_n(\tau)}\Big(\lambda_n(\tau) - \eta_n(\tau)\Big) \tag{4.48}$$

with

$$\lambda_n(\tau) = \mathbb{E}[C_n(\tau)] \quad \textit{and} \quad \eta_n(\tau) = Var[C_n(\tau)]$$

Proof. Fix $n = 2, 3, \ldots$, let Γ be the index set $\{2, 3, \ldots, n\}$. We construct the indicator rvs

$$I_\alpha := \chi_{n,\alpha}(\tau) \quad \text{and} \quad J_{\beta\alpha} := \chi_{n,\beta}\left(\frac{\tau}{1-\tau}\right), \quad \alpha, \beta \in \Gamma.$$

It can be shown that for each α in Γ,

$$J_{\beta\alpha} \leq I_\beta, \quad \beta \in \Gamma \backslash \{\alpha\}. \tag{4.49}$$

Pick α in Γ and let K be a nonempty subset of $\Gamma\backslash\{\alpha\}$. From Lemma 3.1, we have

$$\mathbb{P}\Big[J_{\beta\alpha} = 1, \beta \in K\Big] = \left(1 - \frac{|K|\tau}{1-\tau}\right)_+^n$$

where $|K|$ refers to the cardinality of set K. On the other hand, we also have

$$\begin{aligned}
\mathbb{P}\Big[I_\beta = 1, \beta \in K \Big| I_\alpha = 1\Big] &= \frac{\mathbb{P}\Big[I_\beta = 1, \beta \in K, I_\alpha = 1\Big]}{\mathbb{P}[I_\alpha = 1]} \\
&= \frac{\Big(1 - (|K|+1)\tau\Big)_+^n}{(1-\tau)_+^n} \\
&= \left(1 - \frac{|K|\tau}{1-\tau}\right)_+^n .
\end{aligned}$$

Thus

$$\mathbb{P}\Big[I_\beta = 1, \beta \in K \Big| I_\alpha = 1\Big] = \mathbb{P}\Big[J_{\beta\alpha} = 1, \beta \in K\Big] \tag{4.50}$$

and it follows from (4.50) and Lemma 4.5 that

$$\Big(I_\beta, \beta \in \Gamma\backslash\{\alpha\} \Big| I_\alpha = 1\Big) =_{st} \Big(J_{\beta\alpha}, \beta \in \Gamma\backslash\{\alpha\}\Big). \tag{4.51}$$

According to (4.49), (4.51) and Definition 3.4, the rvs $\{I_\alpha,\ \alpha \in \Gamma\}$ are negatively related, therefore the rvs $\{\chi_{n,\alpha}(\tau),\ \alpha \in \Gamma\}$ are negatively related. Since $C_n(\tau) = \sum_{k=2}^n \chi_{n,k}(\tau)$, the total variation distance between $C_n(\tau)$ and $\Pi(\lambda_n(\tau))$ can be bounded by Corollary 3.1, and (4.48) follows. ■

Corollary 4.3 *For each $n = 2, 3, \ldots$ and x in $\mathbb{R}$, let $\tau = \frac{\log n + x}{n}$. It holds that*

$$\Big|\mathbb{P}[C_n(\tau) = 0] - e^{-e^{-x}}\Big| \le \frac{1 - e^{-\lambda_n(\tau)}}{\lambda_n(\tau)}\Big(\lambda_n(\tau) - \eta_n(\tau)\Big) + \Big|e^{-\lambda_n(\tau)} - e^{-e^{-x}}\Big|.$$

provided τ is in the interval $(0, 1)$.

Proof. By the triangle inequality, we have

$$\begin{aligned}
\Big|\mathbb{P}[C_n(\tau) = 0] - e^{-e^{-x}}\Big| &\le \Big|\mathbb{P}[C_n(\tau) = 0] - \mathbb{P}[\Pi(\lambda_n(\tau)) = 0]\Big| \\
&+ \Big|\mathbb{P}[\Pi(\lambda_n(\tau)) = 0] - e^{-e^{-x}}\Big|.
\end{aligned}$$

The second component on the right hand side of the expression above is simply equal to $|e^{-\lambda_n(\tau)} - e^{-e^{-x}}|$, while the first component satisfies the bounds:

$$\begin{aligned}
\Big|\mathbb{P}[C_n(\tau) = 0] - \mathbb{P}[\Pi(\lambda_n(\tau)) = 0]\Big| &\le d_{TV}(C_n(\tau); \Pi(\lambda_n(\tau))) \\
&\le \frac{1 - e^{-\lambda_n(\tau)}}{\lambda_n(\tau)}\Big(\lambda_n(\tau) - \eta_n(\tau)\Big)
\end{aligned}$$

as can be seen from Theorem 4.3 and Lemma 3.6, respectively. ■

From (4.27) and (4.28), we immediately find the expression

$$\lambda_n(\tau) = (n-1)(1-\tau)^n$$

and

$$\eta_n(\tau) = (n-1)(1-\tau)^n + (n-1)(n-2)(1-2\tau)_+^n - (n-1)^2(1-\tau)^{2n}.$$

4.6 Summary

In this Chapter, we have estimated the probability of the event $C_n(\tau) = 0$ with the help of the tools introduced in Chapter 3. By computing the first and second moments of $C_n(\tau)$, we show that $\lim_{n\to\infty} \mathbb{P}[C_n(\tau) = 0] = 1$ (resp. $= 0$) if the connectivity range τ deviates from the critical range $\tau^* = \frac{\log n}{n}$ by $\frac{\alpha_n}{n}$ where $\lim_{n\to\infty} \alpha_n = \infty$ (resp. $\lim_{n\to\infty} \alpha_n = -\infty$). Moreover, $C_n(\tau)$ is approximately Poisson distributed if $\lim_{n\to\infty} \alpha_n$ is finite. Such Poisson approximations explain the sharp phase transition of the probability of graph connectivity when τ deviates from the critical range τ^*. Actually,

$$P(n;\tau) = \mathbb{P}[C_n(\tau) = 0] \simeq e^{-E[C_n(\tau)]} = e^{-(n-1)\gamma_n}$$

with $\gamma_n = (1-\tau)^n$. The last equality is due to (4.27). Roughly speaking, when τ is around $\tau_n^* = \frac{\log n}{n}$, a small change of τ yields a moderate change in γ_n, the change in γ_n is then magnified by $n-1$, and as a result, the probability of graph connectivity varies significantly.

Chapter 5

Network connectivity under the GRG model II : A strong threshold for general user distribution with non-vanishing densities

5.1 The main result

In this Chapter, we prove that the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ given by

$$\tau_n^\star = \frac{\log n}{f_\star n}, \quad n = 1, 2, \ldots$$

is a strong threshold for graph connectivity.

Theorem 5.1 *Under the enforced assumptions, the range function $\tau_n^\star$ is a strong threshold for graph connectivity.*

According to Lemma 4.2, we readily have

Lemma 5.1 *Under the enforced assumptions, the range function $\tau_n^\star$ is a strong threshold for graph connectivity if and only if*

$$\frac{M_n}{\tau_n^\star} \xrightarrow{P}_n 1. \tag{5.1}$$

5.2 Preliminaries

We introduce some easy facts concerning F and f: Since $f_\star > 0$, the mapping $F : [0,1] \to [0,1]$ is strictly increasing, hence invertible. Let $F^{-1} : [0,1] \to [0,1]$ denote the inverse mapping of F. This inverse mapping is strictly increasing and continuous since F is itself strictly increasing and continuous. Also the differentiability of F implies that of F^{-1}. Therefore, differentiating both sides of the identity $F^{-1}(F(t)) = t$ on $[0,1]$ and making use of the chain rule, we get

$$\begin{aligned} \frac{d}{dt}F^{-1}(t) &= \frac{1}{f(F^{-1}(t))} \\ &= \frac{1}{g(t)}, \quad 0 \le t \le 1 \end{aligned} \tag{5.2}$$

where the mapping $g : [0,1] \to \mathbb{R}_+$ is defined by

$$g(t) = f(F^{-1}(t)), \quad 0 \le t \le 1.$$

As a result, we can write

$$F^{-1}(x) = \int_0^x \frac{1}{g(t)} dt, \quad 0 \le x \le 1$$

since $F(0) = 0$.

Consider any $x_\star$ in $[0,1]$ which achieves the minimum of f. By the strict monotonicity of F, there exists a unique $t_\star$ in $[0,1]$ such that $F^{-1}(t_\star) = x_\star$, namely $F(x_\star) = t^\star$. Note that $x_\star = 0$ (resp. $0 < x_\star < 1$, $x_\star = 1$) if and only if $t_\star = 0$ (resp. $0 < t_\star < 1$, $t_\star = 1$). Moreover, as the composition of two continuous mappings, the mapping g is also continuous and we have the bound

$$g(t) \geq g(t_\star) = f(x_\star) = f_\star, \quad 0 \leq t \leq 1. \tag{5.3}$$

5.3 An outline of the proof of Theorem 5.1

In addition to the i.i.d. $[0,1]$-valued rvs $\{X_i,\ i = 1, 2, \ldots\}$, consider a second collection of i.i.d., rvs $\{U_i,\ i = 1, 2, \ldots\}$ which are all uniformly distributed on $[0,1]$. In analogy with the notation introduced above, for each $n = 2, 3, \ldots$, we introduce the order statistics $U_{n,1}, \ldots, U_{n,n}$ associated with the n i.i.d. rvs $U_1, \ldots, U_n$ and we again use the convention $U_{n,0} = 0$ and $U_{n,n+1} = 1$.

Key to our approach is the well-known stochastic equivalence

$$(X_1, \ldots, X_n) =_{st} (F^{-1}(U_1), \ldots, F^{-1}(U_n)) \tag{5.4}$$

which leads to the representation

$$(X_{n,1}, \ldots, X_{n,n}) =_{st} (F^{-1}(U_{n,1}), \ldots, F^{-1}(U_{n,n})). \tag{5.5}$$

It is now plain that

$$\begin{aligned} M_n &= \max\left(L_{n,k},\ k = 2, \ldots, n\right) \\ &=_{st} \max\left(\int_{U_{n,k-1}}^{U_{n,k}} \frac{1}{g(t)} dt,\ k = 2, \ldots, n \right) \end{aligned}$$

as we note that

$$F^{-1}(U_{n,k}) - F^{-1}(U_{n,k-1}) = \int_{U_{n,k-1}}^{U_{n,k}} \frac{1}{g(t)} dt, \quad k = 1, \ldots, n+1.$$

These observations suggest that the convergence (5.1) is likely to emerge as a consequence of limiting properties of the rvs $\{U_{n,k},\ k = 0, \ldots, n+1\}$ and of properties of the function f (via g). As we shall see shortly, this is indeed the case. To help us along down this road, we shall find it convenient to write

$$M_n^u := \max\left(L_{n,k}^u,\ k = 2, \ldots, n\right) \tag{5.6}$$

with

$$L_{n,k}^u := U_{n,k} - U_{n,k-1}, \quad k = 1, \ldots, n+1. \tag{5.7}$$

The quantities defined at (5.7) and (5.6) coincide with the quantities defined at (3.2) and (3.7), respectively, when F is the uniform distribution on $[0,1]$.

For each $n = 1, 2, \ldots$, define the rv $\widetilde{M}_n$ by

$$\widetilde{M}_n := \max\left(\frac{L_{n,k}^u}{g(U_{n,k-1})},\ k = 2, \ldots, n\right).$$

The next result shows that when establishing (5.1) we can replace M_n by the simpler quantity $\widetilde{M}_n$.

Proposition 5.1 *Under the enforced assumptions, it holds that*

$$\frac{M_n - \widetilde{M}_n}{\tau_n^\star} \xrightarrow{P}_n 0. \tag{5.8}$$

Proposition 5.1 is established in Section 5.5. We next show that the convergence (5.1) indeed holds when M_n is replaced by $\widetilde{M}_n$.

Proposition 5.2 *Under the enforced assumptions, it holds that*

$$\frac{\widetilde{M}_n}{\tau_n^\star} \xrightarrow{P}_n 1. \tag{5.9}$$

We give a proof of Proposition 5.2 in Section 5.6. Combining Proposition 5.1 and Proposition 5.2 we immediately conclude to the following desired generalization of Lévy's result.

Proposition 5.3 *Under the enforced assumptions, the convergence statement (5.1) holds.*

Theorem 5.1 is now within easy reach: Just combine Lemma 5.1, and Proposition 5.3.

5.4 A useful representation

The starting point in proving Propositions 5.1 and 5.2 resides in the representation (5.5). We shall leverage it by relying on a useful representation of the order statistics $\{U_{n,k},\ k = 0, 1, \ldots, n+1\}$ via i.i.d. exponential rvs $\{\xi_j,\ j = 1, 2, \ldots\}$ with unit parameter: Set

$$T_0 = 0,\ T_k = \xi_1 + \ldots + \xi_k, \quad k = 1, 2, \ldots.$$

For all $n = 1, 2, \ldots$, the stochastic equivalence

$$(U_{n,1}, \ldots, U_{n,n}) =_{st} \left(\frac{T_1}{T_{n+1}}, \ldots, \frac{T_n}{T_{n+1}} \right) \tag{5.10}$$

is immediate from Lemma 3.3.

This representation makes it possible to provide an elementary proof for a technical fact to be used repeatedly in what follows. For each $n = 1, \ldots$, let K_n denote a non-empty subset of $\{1, \ldots, n+1\}$, and let $|K_n|$ denote its cardinality.

Also set

$$M(K_n) := \max\left(\xi_k,\ k \in K_n\right).$$

Lemma 5.2 *The convergence*

$$\frac{M(K_n)}{\log n} \xrightarrow{P}_n 1 \tag{5.11}$$

takes place whenever there exists some θ *in* $(0,1]$ *such that*

$$\lim_{n\to\infty} \frac{|K_n|}{n} = \theta. \tag{5.12}$$

Proof. Fix $n = 1, 2, \ldots$ and $t \geq 0$. By independence, we get

$$\begin{aligned}
\mathbb{P}\left[M(K_n) \leq t\right] &= \mathbb{P}\left[\xi_k \leq t,\ k \in K_n\right] \\
&= \left(1 - e^{-t}\right)^{|K_n|}
\end{aligned}$$

so that

$$\begin{aligned}
\mathbb{P}\left[\frac{M(K_n)}{\log n} \leq t\right] &= \left(1 - e^{-t\log n}\right)^{|K_n|} \\
&= \left(1 - \frac{n^{1-t}}{n}\right)^{|K_n|}.
\end{aligned}$$

With the help of (5.12) it is straightforward to check that

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{M(K_n)}{\log n} \leq t\right] = \begin{cases} 0 & \text{if } 0 \leq t < 1 \\ & \\ 1 & \text{if } 1 < t. \end{cases}$$

As this last convergence implies

$$\frac{M(K_n)}{\log n} \Longrightarrow_n \ 1,$$

the convergence (5.11) follows from the equivalence of convergence in distribution and in probability when the limit is a constant. ∎

Lemma 5.3 *Under the assumptions of Lemma 5.2 we also have*

$$\frac{1}{\tau_n^\star}\left(\max\left(L_{n,k}^u,\ k \in K_n\right)\right) \xrightarrow{P}_n 1. \tag{5.13}$$

Proof. By virtue of (5.7) and the stochastic identity (5.10), we need only show that

$$\frac{1}{\tau_n^\star}\left(\max_{k\in K_n}\left(\frac{\xi_k}{T_{n+1}}\right)\right) \xrightarrow{P}_n 1, \tag{5.14}$$

a convergence statement which is equivalent to

$$\frac{n}{T_{n+1}}\frac{M(K_n)}{\log n} \xrightarrow{P}_n 1. \tag{5.15}$$

The validity of this convergence statement follows from Lemma 5.2 and from the fact that

$$\lim_{n\to\infty}\frac{T_{n+1}}{n} = 1 \quad a.s. \tag{5.16}$$

by the Strong Law of Large Numbers. ∎

Specializing this last result to $K_n = \{2, \ldots, n\}$, we get

$$\frac{M_n^u}{\tau_n^\star} \xrightarrow{P}_n 1. \tag{5.17}$$

This result was already obtained in Theorem 3.1, and establishes Theorem 5.1 when F is the uniform distribution (since then $f_\star = 1$).

5.5 A proof of Proposition 5.1

Fix $n = 2, 3, \ldots$ and pick $k = 2, \ldots, n$. Upon writing

$$\begin{aligned}\Delta_{n,k} &:= \int_{U_{n,k-1}}^{U_{n,k}} \frac{1}{g(t)} dt - \frac{U_{n,k} - U_{n,k-1}}{g(U_{n,k-1})} \\ &= \int_{U_{n,k-1}}^{U_{n,k}} \left(\frac{1}{g(t)} - \frac{1}{g(U_{n,k-1})} \right) dt,\end{aligned}$$

we find

$$|\Delta_{n,k}| \leq f_\star^{-2} \int_{U_{n,k-1}}^{U_{n,k}} |g(t) - g(U_{n,k-1})| \, dt.$$

Recalling the definition (5.7), we then get

$$|\Delta_{n,k}| \leq f_\star^{-2} G_{n,k} \cdot L^u_{n,k}$$

where we have set

$$G_{n,k} := \max \left(|g(t) - g(U_{n,k-1})| \, , \; U_{n,k-1} \leq t \leq U_{n,k} \right).$$

These facts lead to

$$\begin{aligned}|\widetilde{M}_n - M_n| &\leq \max \left(|\Delta_{n,k}|, \; k = 2, \ldots, n \right) \\ &\leq f_\star^{-2} \max \big(G_{n,k} \cdot L^u_{n,k}, \; k = 2, \ldots, n \big) \\ &\leq f_\star^{-2} G_n \cdot M^u_n\end{aligned}$$

where M^u_n is defined at (5.6) and

$$G_n := \max \left(G_{n,k}, \; k = 2, \ldots, n \right).$$

The bound

$$\frac{|\widetilde{M}_n - M_n|}{\tau_n^\star} \le f_\star^{-2} G_n \cdot \frac{M_n^u}{\tau_n^\star}$$

is now immediate. Thus, from (5.17) we see that (5.8) holds if we show that $G_n \xrightarrow{P}_n 0$. In other words, for arbitrary $\varepsilon > 0$, we need to show that

$$\lim_{n\to\infty} \mathbb{P}\left[G_n > \varepsilon\right] = 0. \tag{5.18}$$

To do so, we recall that the mapping g is continuous on the compact $[0,1]$, hence uniformly continuous on $[0,1]$. Thus, for every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that with x and y in $[0,1]$,

$$|g(x) - g(y)| \le \varepsilon \tag{5.19}$$

whenever $|x - y| \le \delta$.

Fix $\varepsilon > 0$ and consider an arbitrary integer $n = 2, 3, \ldots$. Obviously, $G_n \le \varepsilon$ if and only if $G_{n,k} \le \varepsilon$ for all $k = 2, \ldots, n$. In view of the comments at (5.19), this will occur if $L_{n,k}^u \le \delta$ for all $k = 2, \ldots, n$, a condition equivalent to $M_n^u \le \delta$. Consequently,

$$\mathbb{P}\left[G_n \le \varepsilon\right] \ge \mathbb{P}\left[M_n^u \le \delta\right].$$

In other words,

$$\mathbb{P}\left[G_n > \varepsilon\right] \le \mathbb{P}\left[M_n^u > \delta\right], \quad n = 1, 2, \ldots \tag{5.20}$$

Now observe that $M_n^u \xrightarrow{P}_n 0$ by virtue of (5.17) since $\lim_{n\to\infty} \tau_n^\star = 0$, whence $\lim_{n\to\infty} \mathbb{P}\left[M_n^u > \delta\right] = 0$. We readily get (5.18) upon letting n go to infinity in the inequality (5.20). This completes the proof of Proposition 5.1. ∎

5.6 A proof of Proposition 5.2

Fix $n = 2, 3, \ldots$. By virtue of (5.10), we have the representation

$$\widetilde{M}_n =_{st} \max\left(\frac{\xi_k}{T_{n+1} \cdot g(\frac{T_{k-1}}{T_{n+1}})},\ k = 2, \ldots, n\right)$$

so that

$$\frac{\widetilde{M}_n}{\tau_n^\star} =_{st} \frac{n}{T_{n+1}} \cdot f_\star \frac{\widehat{M}_n}{\log n}$$

where we have used the notation

$$\widehat{M}_n := \max\left(\frac{\xi_k}{g(\frac{T_{k-1}}{T_{n+1}})},\ k = 2, \ldots, n\right).$$

By the Strong Law of Large Numbers (5.16), the convergence (5.9) will be established if we show that

$$f_\star \frac{\widehat{M}_n}{\log n} \xrightarrow{P}_n 1. \tag{5.21}$$

Thus, we need to show that for every $\varepsilon > 0$, we have

$$\lim_{n\to\infty} \mathbb{P}\left[\left| f_\star \frac{\widehat{M}_n}{\log n} - 1 \right| \geq \varepsilon\right] = 0 \tag{5.22}$$

and this is equivalent to establishing the simultaneous validity of the two convergence statements

$$\lim_{n\to\infty} \mathbb{P}\left[1 + \varepsilon \leq f_\star \frac{\widehat{M}_n}{\log n}\right] = 0 \tag{5.23}$$

and

$$\lim_{n\to\infty} \mathbb{P}\left[f_\star \frac{\widehat{M}_n}{\log n} \leq 1 - \varepsilon\right] = 0. \tag{5.24}$$

To do so, we start with the easy upper bound

$$f_\star \frac{\widehat{M}_n}{\log n} \leq \frac{M(\{2, \ldots, n\})}{\log n}, \quad n = 2, 3, \ldots \tag{5.25}$$

so that the convergence (5.23) now follows readily from (5.11) (specialized to $K_n = \{2, \ldots, n\}$).

The proof of (5.24), given next, is somewhat more involved. It will require the introduction of a family of lower bounds (in contrast with the proof of (5.23) which relied on the single upper bound (5.25)): Pick any element $x_\star$ in $[0, 1]$ which achieves the minimum of g. It will be easier to structure the forthcoming discussion according to whether $x_\star = 0$, $0 < x_\star < 1$ and $x_\star = 1$. Here, we give a complete discussion for the case $0 < x_\star < 1$, as the two other cases can be handled *mutatis mutandi.*

Thus, with $0 < x_\star < 1$, let $t_\star = F(x_\star)$ and note that $0 < t_\star < 1$. Now pick θ such that

$$0 < \theta < \min(t_\star, 1 - t_\star). \tag{5.26}$$

For each $n = 2, 3, \ldots$, we introduce $K_n(\theta)$ as the subset of $\{1, \ldots, n+1\}$ defined by

$$K_n(\theta) := \{\lceil n(t_\star - \theta)\rceil, \ldots, \lceil n(t_\star + \theta)\rceil\}.$$

Since we are interested in limiting results, we need only consider $n \geq n^\star(\theta)$ with $n^\star(\theta) = 2(t_\star - \theta)^{-1}$ (as we do from now on), in which case $\lceil n(t_\star - \theta)\rceil \geq 2$ and $K_n(\theta) \subseteq \{2, \ldots, n\}$. The lower bound

$$\widehat{M_n}(\theta) \leq \widehat{M_n} \tag{5.27}$$

is then immediate where we have set

$$\widehat{M_n}(\theta) := \max\left(\frac{\xi_k}{g(\frac{T_{k-1}}{T_{n+1}})},\ k \in K_n(\theta)\right).$$

To proceed, we observe the following elementary facts: For each $a = 0, \pm 1$, it is plain that

$$\lim_{n\to\infty} \frac{\lceil n(t_\star + a\theta)\rceil}{n} = t_\star + a\theta,$$

so that

$$\lim_{n\to\infty} \frac{T_{\lceil n(t_\star+a\theta)\rceil-1}}{T_{n+1}} = t_\star + a\theta \quad a.s. \tag{5.28}$$

by the Strong Law of Large Numbers. Building on this observation, with $\eta > 0$, we introduce for each $n \geq n^\star(\theta)$, the events

$$\Omega_n^a(\theta;\eta) := \left[\left|\frac{T_{\lceil n(t_\star+a\theta)\rceil-1}}{T_{n+1}} - (t_\star + a\theta)\right| \leq \eta\right], \quad a = 0, \pm 1.$$

If we set

$$\Omega_n(\theta;\eta) := \cap_{a=0,\pm 1}\ \Omega_n^a(\theta;\eta),$$

then the convergence (5.28) implies

$$\lim_{n\to\infty} \mathbb{P}\left[\Omega_n(\theta;\eta)\right] = 1, \quad \eta > 0. \tag{5.29}$$

Fix $n \geq n^\star(\theta)$ and pick $\eta > 0$ such that $\theta + \eta < t_\star < 1 - (\theta + \eta)$. Such a choice of η is possible under (5.26), in which case on the event $\Omega_n(\theta;\eta)$, the inequalities

$$\left|\frac{T_{k-1}}{T_{n+1}} - t_\star\right| \leq (\theta + \eta), \quad k \in K_n(\theta) \tag{5.30}$$

all hold.

We are now in position to complete the proof: Fix $\zeta > 0$ and set $\delta = \delta(\zeta)$ where $\delta(\zeta)$ insures (5.19) (with ε replaced by ζ) as a result of the uniform continuity of g. Pick θ in $(0, 1)$ and $\eta > 0$ such that $\theta + \eta \leq \delta$, By selecting θ and η sufficiently small, the constraints (5.26) and $\theta + \eta < t_\star < 1 - (\theta + \eta)$ can also be satisfied simultaneously. With this choice, it follows from (5.30) that the inequalities

$$\left| g\left(\frac{T_{k-1}}{T_{n+1}}\right) - g(t_\star) \right| \le \zeta, \quad k \in K_n(\theta)$$

all hold on the event $\Omega_n(\theta;\eta)$. Therefore,

$$f_\star \le g\left(\frac{T_{k-1}}{T_{n+1}}\right) \le f_\star + \zeta, \quad k \in K_n(\theta)$$

since $g(t_\star) = f(x_\star) = f_\star$, and we obtain the inequality

$$(f_\star + \zeta)^{-1} \cdot M(K_n(\theta)) \le \widehat{M}_n(\theta). \tag{5.31}$$

We now return to the lower bound (5.27). On the event $\Omega_n(\theta;\eta)$, for a given $\varepsilon > 0$, the inequality $f_\star \frac{\widehat{M}_n}{\log n} \le 1 - \varepsilon$, when coupled with (5.31), readily implies

$$\frac{M(K_n(\theta))}{\log n} \le a(\varepsilon;\zeta) \tag{5.32}$$

with

$$a(\varepsilon;\zeta) := (1-\varepsilon) \cdot \frac{f_\star + \zeta}{f_\star}.$$

As a result, by standard bounding and decomposition arguments, we get

$$\begin{aligned} \mathbb{P}\left[f_\star \frac{\widehat{M}_n}{\log n} \le 1 - \varepsilon\right] &\le \mathbb{P}\left[\left[\frac{M(K_n(\theta))}{\log n} \le a(\varepsilon;\zeta)\right] \cap \Omega_n(\theta;\eta)\right] + \mathbb{P}\left[\Omega_n(\theta;\eta)^c\right] \\ &\le \mathbb{P}\left[\frac{M(K_n(\theta))}{\log n} \le a(\varepsilon;\zeta)\right] + 1 - \mathbb{P}\left[\Omega_n(\theta;\eta)\right]. \end{aligned} \tag{5.33}$$

Note that (5.24) needs to be established only for $0 < \varepsilon < 1$ for otherwise the convergence is trivially true. Thus, pick $0 < \varepsilon < 1$ and note that $\zeta > 0$ can be selected sufficiently small such that $a(\varepsilon;\zeta) < 1$. Indeed this last condition is equivalent to

$$\zeta < \frac{\varepsilon}{1-\varepsilon} \cdot f_\star.$$

With such a selection of ζ, Lemma 5.2 (with $K_n = K_n(\theta)$) implies

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{M(K_n(\theta))}{\log n} \le a(\varepsilon;\zeta)\right] = 0. \tag{5.34}$$

Let n go to infinity in (5.33). The desired result (5.24) follows from (5.29) and (5.34).

The cases $x_\star = 0$ and $x_\star = 1$ can be analyzed in a similar way: Now, still with $t_\star = F(x_\star)$, we have $t_\star = 0$ and $t_\star = 1$, respectively. As a result we need only change the definition of $K_n(\theta)$ to read $\{2, \ldots, \lceil n(t_\star + \theta)\rceil\}$ and $\{\lceil n(t_\star - \theta)\rceil, \ldots, n\}$, respectively, for n large enough in order to ensure $K_n(\theta) \subset \{2, \ldots, n\}$. This completes the proof of Proposition 5.2. ■

5.7 Discussion

5.7.1 Strong versions of Lévy's result

The convergence (5.1) is compatible with a multi-dimensional result obtained by Penrose [47]: Formally setting $d = 1$ in Theorem 1.1 [47, p. 247] (discussed under the dimensional assumption $d \ge 2$), we obtain (5.1) in its a.s. form.

Slud [52, Thm. 2.1, p. 343] has shown that

$$nM_n^u - \log n = O(\log\log n) \quad a.s. \tag{5.35}$$

so that the convergence (5.17) also holds in the a.s. sense.

5.7.2 Connections with earlier results

In principle, Proposition 5.3 would follow from Theorem 3.2. However, such a sharper result is given under additional stronger conditions than the one used here. As a result of this trade-off, we are able to give a simple and direct proof of the convergence (5.1).

Chapter 6

Network connectivity under the GRG model III : A very strong threshold for general user distribution with non-vanishing densities

6.1 Model and assumptions

In this Chapter, we consider a fairly general representation for the density f:

$$f(x) = c + a|x - x_\star|^r + h(x), \quad 0 \le x \le 1 \tag{6.1}$$

for some parameters $r > 0$, $a > 0$ and $c > 0$, and for some function $h : [0,1] \to \mathbb{R}$ such that

$$\lim_{x \to x_\star} \frac{h(x)}{|x - x_\star|^r} = 0. \tag{6.2}$$

We have $\lim_{x \to x_\star} h(x) = 0$ by virtue of (6.2), whence $h(x_\star) = 0$ since the continuity of f implies that of h. As a result, we necessarily have $c = f_\star$.

A particularly useful special case occurs when the density f has the form

$$f(x) = c + a|x - x_\star|^r, \quad 0 \le x \le 1 \tag{6.3}$$

for some parameters $r > 0$, $a > 0$ and $c > 0$ with $x_\star$ in $[0,1]$. There obviously exists a relationship between the four parameters $x_\star$, r, a and c, namely

$$c + \frac{a}{r+1}\left[x_\star^{r+1} + (1-x_\star)^{r+1}\right] = 1$$

by virtue of the requirement $\int_0^1 f(t)dt = 1$.

The conditions (6.1) and (6.2) are not overly restrictive, For instance, they do hold when the density function f admits $2\ell+1$ *bounded* derivatives $f^{(1)}, \ldots f^{(2\ell+1)} : [0,1] \to \mathbb{R}$ such that

$$f^{(1)}(x_\star) = \ldots = f^{(2\ell-1)}(x_\star) = 0, \;\; f^{(2\ell)}(x_\star) > 0$$

for some positive integer ℓ when $x_\star$ a unique global minimum for f in the open interval $(0,1)$. In that case, the existence of a Taylor series expansion at $x = x_\star$ leads to taking $r = 2\ell$ and

$$a = \frac{1}{(2\ell)!} f^{(2\ell)}(x_\star)$$

so that

$$h(x) = f(x) - f(x_\star) - a(x-x_\star)^{2\ell}, \quad 0 \le x \le 1,$$

and (6.2) holds.

6.2 The main results

The results take a more symmetric form if we write any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ in the form

$$\tau_n = \frac{1}{f_\star} \cdot \frac{1}{n}\left(\log n - \frac{1}{r}\log\log n + \alpha_n\right) \tag{6.4}$$

for all $n = 1, 2, \ldots$, for some function $\alpha : \mathbb{N}_0 \to \mathbb{R}$. We refer to such a function α as a deviation function.[1] There is no loss of generality in using the representation (6.4).

Theorem 6.1 *Under the enforced assumptions, for any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$, it holds that*

$$\lim_{n\to\infty} P(n; \tau_n) = \begin{cases} 0 & \textit{if} \ \lim_{n\to\infty} \alpha_n = -\infty \\ \\ 1 & \textit{if} \ \lim_{n\to\infty} \alpha_n = +\infty. \end{cases}$$

Assumptions weaker than (6.1) and (6.2) can be handled at the cost of some technicalities. The reader is referred to Section 6.9 for comments and pointers on some of the possibilities.

The proof of Theorem 6.1 is divided in two parts with the one law and the zero law being given in Sections 6.6 and 6.7, respectively. To simplify the exposition we shall present the arguments only when $x_\star$ is an interior point, i.e., $x_\star$ belongs to the open interval $(0, 1)$. The boundary cases $x_\star = 0, 1$ are outlined in Section 6.9.

At this point the reader may wonder as to the appropriate version of Theorem 6.1 when the density f achieves its minimum value $f_\star$ at *non*-isolated points. This situation is formalized next.

Theorem 6.2 *Under the enforced assumptions, assume also that there exists a non-empty open interval $I \subseteq (0, 1)$ such that $f(x) = f_\star$ for all x in I. Then, (6.9.5)*

[1] By convention we use $\log \log n = 0$ for $n = 1, 2$.

still holds for any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form

$$\tau_n = \frac{1}{f_\star} \cdot \frac{1}{n} \left(\log n + \alpha_n \right), \quad n = 1, 2, \ldots . \tag{6.5}$$

with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$.

In other words, we need only set $r = \infty$ in Theorem 6.1, as expected. The proof of Theorem 6.2 follows the same pattern as the one used in the proof of Theorem 6.1; details are available in Section 6.9.3.

6.3 Breakpoint users, connectivity and zero-one laws

Fix $n = 2, 3, \ldots$ and $\tau > 0$. For each $k = 1, \ldots, n$, we say that node k is a *breakpoint* node [2] in the random graph $\mathbb{G}(n; \tau)$ if the interval $[X_k, X_k + \tau]$ does *not* contain any other node of the graph. The event $E_{n,k}(\tau)$ that node k is a breakpoint node in $\mathbb{G}(n; \tau)$ can be expressed as

$$E_{n,k}(\tau) = \cap_{j=1,\ell \neq k}^{n} [X_j \notin [X_k, X_k + \tau]]$$

and its indicator rv $\chi_{n,k}(\tau)$ is the $\{0, 1\}$-valued rv given by

$$\begin{aligned} \chi_{n,k}(\tau) &= \mathbf{1}\left[E_{n,k}(\tau)\right] \\ &= \prod_{j=1, j\neq k}^{n} \mathbf{1}\left[X_j \notin [X_k, X_k + \tau]\right]. \end{aligned} \tag{6.6}$$

[2]This definition of breakpoint node is different from the definition in Chapter 3.

The number of breakpoint nodes in $\mathbb{G}(n;\tau)$ is given by

$$C_n(\tau) = \sum_{k=1}^{n} \chi_{n,k}(\tau). \tag{6.7}$$

Interest in this quantity arises from the following observations: First we note that $C_n(\tau) \geq 1$ since the right-most node is always a breakpoint node. Moreover, the graph $\mathbb{G}(n;\tau)$ is connected if and only if that right-most node is the only breakpoint node. Therefore, $\mathbb{G}(n;\tau)$ is connected if and only if $C_n(\tau) = 1$, so that

$$P(n;\tau) = \mathbb{P}\left[C_n(\tau) = 1\right]. \tag{6.8}$$

For a given range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$, we now give conditions for the validity of either $\lim_{n\to\infty} P(n;\tau_n) = 1$ or $\lim_{n\to\infty} P(n;\tau_n) = 0$ in terms of the limiting behavior of the first moment of the sequence $\{C_n(\tau_n),\ n = 2,3,\ldots\}$. We begin with conditions for the one law.

Lemma 6.1 *For any range function* $\tau : \mathbb{N}_0 \to \mathbb{R}_+$, *we have* $\lim_{n\to\infty} P(n;\tau_n) = 1$ *whenever* $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = 1$.

Proof. Fix $n = 2,3,\ldots$ and $\tau > 0$. Since $C_n(\tau) \geq 1$, it is plain that

$$\begin{aligned}
\mathbb{E}\left[C_n(\tau)\right] &= \sum_{k=0}^{\infty} \mathbb{P}\left[C_n(\tau) > k\right] \\
&= 1 + \sum_{k=2}^{\infty} \mathbb{P}\left[C_n(\tau) \geq k\right] \\
&\geq 1 + \mathbb{P}\left[C_n(\tau) \geq 2\right].
\end{aligned} \tag{6.9}$$

Thus, for any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$, we find

$$\mathbb{P}\left[C_n(\tau_n) \geq 2\right] \leq \mathbb{E}\left[C_n(\tau_n)\right] - 1, \quad n = 2,3,\ldots$$

We now let n go to infinity in this last inequality. Because we assume $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = 1$, we get $\lim_{n\to\infty} \mathbb{P}\left[C_n(\tau_n) \geq 2\right] = 0$, hence the desired conclusion $\lim_{n\to\infty} \mathbb{P}\left[C_n(\tau_n) = 1\right] = 1$. ■

Conditions for the zero law are given next.

Lemma 6.2 *For any range function* $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ *such that* $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = \infty$, *we have* $\lim_{n\to\infty} P(n; \tau_n) = 0$ *whenever*

$$\frac{C_n(\tau_n)}{\mathbb{E}\left[C_n(\tau_n)\right]} \xrightarrow{P}_n 1. \tag{6.10}$$

Proof. Pick ε in the interval $(0, 1)$. Under the condition $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = \infty$, there exists a positive integer $n^\star(\varepsilon)$ such that

$$2(1-\varepsilon)^{-1} \leq \mathbb{E}\left[C_n(\tau_n)\right], \quad n \geq n^\star(\varepsilon).$$

On that range, we then find

$$\begin{aligned} \mathbb{P}\left[\left|\frac{C_n(\tau)}{\mathbb{E}\left[C_n(\tau)\right]} - 1\right| \leq \varepsilon\right] &\leq \mathbb{P}\left[2 \leq C_n(\tau_n) \leq (1+\varepsilon)\mathbb{E}\left[C_n(\tau_n)\right]\right] \\ &\leq \mathbb{P}\left[2 \leq C_n(\tau_n)\right]. \end{aligned} \tag{6.11}$$

We now let n go to infinity in this last inequality. The convergence (6.10) yields $\lim_{n\to\infty} \mathbb{P}\left[C_n(\tau_n) \geq 2\right] = 1$, and the desired conclusion $\lim_{n\to\infty} \mathbb{P}\left[C_n(\tau_n) = 1\right] = 0$ follows. ■

We complement Lemma 6.2 with a sufficient condition for (6.10) to hold.

Lemma 6.3 *For any range function* $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ *such that* $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = \infty$, *we have (6.10) whenever*

$$\limsup_{n\to\infty} \frac{\mathrm{Cov}[\chi_{n,1}(\tau_n), \chi_{n,2}(\tau_n)]}{\mathbb{E}\left[\chi_{n,1}(\tau_n)\right]\mathbb{E}\left[\chi_{n,2}(\tau_n)\right]} \leq 0. \tag{6.12}$$

Proof. Fix $n = 2, 3, \ldots$ and τ in the interval $(0,1)$. For arbitrary $\varepsilon > 0$, Chebyshev's inequality yields

$$\mathbb{P}\left[\left|\frac{C_n(\tau)}{\mathbb{E}\left[C_n(\tau)\right]} - 1\right| > \varepsilon\right] \leq \varepsilon^{-2}\frac{\mathrm{Var}[C_n(\tau)]}{\mathbb{E}\left[C_n(\tau)\right]^2} \tag{6.13}$$

where

$$\begin{aligned}\mathrm{Var}[C_n(\tau)] &= \sum_{k,\ell=1,k\neq\ell}^{n} \mathrm{Cov}[\chi_{n,k}(\tau), \chi_{n,\ell}(\tau)] \\ &\quad + \sum_{k=1}^{n} \mathrm{Var}[\chi_{n,k}(\tau)].\end{aligned}$$

Upon exploiting the binary nature of the rvs $\chi_{n,1}(\tau), \ldots, \chi_{n,n}(\tau)$, we obtain

$$\begin{aligned}\mathrm{Var}[\chi_{n,k}(\tau)] &= \mathbb{E}\left[\chi_{n,k}(\tau)\right] - \mathbb{E}\left[\chi_{n,k}(\tau)\right]^2 \\ &\leq \mathbb{E}\left[\chi_{n,k}(\tau)\right] \end{aligned} \tag{6.14}$$

for each $k = 1, \ldots, n$. Therefore,

$$\begin{aligned}\sum_{k=1}^{n} \mathrm{Var}[\chi_{n,k}(\tau)] &\leq \sum_{k=1}^{n} \mathbb{E}\left[\chi_{n,k}(\tau)\right] \\ &= \mathbb{E}\left[C_n(\tau)\right],\end{aligned}$$

so that

$$\mathrm{Var}[C_n(\tau)] \;\leq\; \sum_{k,\ell=1,k\neq\ell}^{n} \mathrm{Cov}[\chi_{n,k}(\tau),\chi_{n,\ell}(\tau)] + \;\mathbb{E}\left[C_n(\tau)\right]. \tag{6.15}$$

The exchangeability of the rvs $\chi_{n,1}(\tau),\ldots,\chi_{n,n}(\tau)$ implies the relations

$$\mathbb{E}\left[C_n(\tau)\right] = n\mathbb{E}\left[\chi_{n,k}(\tau)\right], \quad k = 1,\ldots,n$$

and

$$\sum_{k,\ell=1,k\neq\ell}^{n} \mathrm{Cov}[\chi_{n,k}(\tau),\chi_{n,\ell}(\tau)] \;=\; n(n-1)\cdot \mathrm{Cov}[\chi_{n,1}(\tau),\chi_{n,2}(\tau)].$$

Reporting this information into the bound (6.15) yields

$$\frac{\mathrm{Var}[C_n(\tau)]}{\mathbb{E}\left[C_n(\tau)\right]^2} \;\leq\; \frac{n-1}{n}\cdot\frac{\mathrm{Cov}[\chi_{n,1}(\tau),\chi_{n,2}(\tau)]}{\mathbb{E}\left[\chi_{n,1}(\tau)\right]\mathbb{E}\left[\chi_{n,2}(\tau)\right]} + \;\mathbb{E}\left[C_n(\tau)\right]^{-1}. \tag{6.16}$$

Now consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ such that $\lim_{n\to\infty}\mathbb{E}\left[C_n(\tau_n)\right] = \infty$. Substitute in (6.16) the free variable τ by τ_n, and let n go to infinity in the resulting inequality. The condition (6.12) and the fact $\lim_{n\to\infty}\mathbb{E}\left[C_n(\tau_n)\right] = \infty$ together imply

$$\lim_{n\to\infty}\frac{\mathrm{Var}[C_n(\tau_n)]}{\mathbb{E}\left[C_n(\tau_n)\right]^2} = 0$$

via a standard limsup argument coupled with the non-negativity of the variance. From (6.13), the conclusion

$$\lim_{n\to\infty}\mathbb{P}\left[\left|\frac{C_n(\tau_n)}{\mathbb{E}\left[C_n(\tau_n)\right]} - 1\right| > \varepsilon\right] = 0$$

is now immediate and this establishes (6.10). ■

6.4 An outline of the proof of Theorem 6.1

Together Lemmas 6.1, 6.2 and 6.3 provide the basic ingredients for manufacturing a proof of Theorem 6.1. These results naturally point to the importance of determining conditions that ensure either $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = 1$ or $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = \infty$. This issue is taken on in the next two results below, namely Propositions 6.1 and 6.2, respectively.

Proposition 6.1, which paves the way for the one law, is proved in Section 6.6.

Proposition 6.1 *For any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$, we have*

$$\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = 1 \tag{6.17}$$

whenever

$$\lim_{n\to\infty} \alpha_n = \infty. \tag{6.18}$$

Establishing the zero law will make use of Proposition 6.2, a proof of which can be found in Section 6.7.

Proposition 6.2 *For any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$, we have*

$$\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = \infty \tag{6.19}$$

whenever

$$\lim_{n\to\infty} \alpha_n = -\infty. \tag{6.20}$$

The next technical result shows that (6.20) already implies the condition (6.12). A proof is available in Section 6.8.

Proposition 6.3 *Consider any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$ satisfying (6.20). The condition (6.12) holds under the additional condition*

$$\lim_{n\to\infty} \frac{|\alpha_n|}{\log n} = 0. \tag{6.21}$$

A proof of Theorem 6.1 is now within easy reach. Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$:

If (6.18) holds, then $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = 1$ by Proposition 6.1, and the one law in Theorem 6.1 follows from Lemma 6.1.

If (6.20) holds, then $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = \infty$ by Proposition 6.2. Two cases are possible: If the condition (6.21) is in place, then (6.12) is seen to hold by Proposition 6.3 and the validity of (6.10) follows from Lemma 6.3. We conclude to the zero law by making use of Lemma 6.2.

Consider now the case when (6.21) fails to hold. With the given range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$, we associate another range function $\tau' : \mathbb{N}_0 \to \mathbb{R}_+$, also written in the form (6.4) but with deviation function $\alpha' : \mathbb{N}_0 \to \mathbb{R}$ given by

$$\alpha'_n = \max\left(\alpha_n, -\sqrt{\log n}\right), \quad n = 3, 4, \ldots$$

Obviously, $\lim_{n\to\infty} \alpha'_n = -\infty$ under (6.20) and

$$\lim_{n\to\infty} \frac{|\alpha'_n|}{\log n} = 0.$$

The first part of the proof shows that $\lim_{n\to\infty} P(n;\tau'_n) = 0$, and the desired conclusion $\lim_{n\to\infty} P(n;\tau_n) = 0$ is now immediate upon noting that

$$P(n;\tau_n) \le P(n;\tau'_n), \quad n = 2, 3, \ldots$$

by monotonicity since $\tau_n \le \tau'_n$ by construction. ■

The remainder of the paper is devoted to establishing Propositions 6.1, 6.2 and 6.3. in Sections 6.6, 6.7 and 6.8, respectively.

6.5 Some basic bounds

The basic idea behind the proof of Propositions 6.1 and 6.2 is quite simple: Lower and upper bounds on $\mathbb{E}\left[C_n(\tau)\right]$ are introduced in terms of an auxiliary quantity (defined at (6.28) below). The form of this auxiliary quantity allows for an easier analysis of its asymptotic behavior, in the process leading to (6.17) and (6.19), respectively, under the appropriate conditions.

The first step consists in obtaining expressions for the first moments involved: Fix $\tau > 0$ and set

$$b(x;\tau) := \int_x^{x+\tau} f(t)dt, \quad x \ge 0. \tag{6.22}$$

In this last expression we have conveniently extended the definition of the density function f to the entire positive line, namely $f : \mathbb{R}_+ \to \mathbb{R}_+$ with $f(x) = 0$ whenever $x > 1$. The boundary effects are automatically taken into account with this convention on f.

Fix $n = 2, 3, \ldots$. Under the enforced assumptions, the rvs $\chi_{n,1}(\tau), \ldots, \chi_{n,n}(\tau)$ are exchangeable. Moreover, for each $k = 1, \ldots, n$, the definition (6.6) readily yields

$$\begin{aligned} \mathbb{E}\left[\chi_{n,k}(\tau)\right] &= \mathbb{E}\left[\prod_{j=1, j\neq k}^{n} \mathbf{1}\left[X_j \notin [X_k, X_k + \tau]\right]\right] \\ &= \mathbb{E}\left[(1 - b(X_1; \tau))^{n-1}\right] \\ &= \int_0^1 (1 - b(x; \tau))^{n-1} f(x)\, dx. \end{aligned} \tag{6.23}$$

This expression follows by first preconditioning with respect to X_1, and then using the fact that the rvs $X_1, \ldots, X_n$ are i.i.d. rvs. The expression

$$\mathbb{E}\left[C_n(\tau)\right] := n \int_0^1 (1 - b(x; \tau))^{n-1} f(x)\, dx \tag{6.24}$$

is now immediate.

The easy calculations given next will help us identify the appropriate bounds: Fix $n = 1, 2, \ldots$ and $\tau > 0$. If $1 \leq \tau$, then

$$b(x; \tau) = \int_x^1 f(t) dt = 1 - F(x), \quad x \in [0, 1]$$

since $1 \leq x + \tau$ for all x in $[0, 1]$, and a straightforward integration yields

$$\mathbb{E}\left[C_n(\tau)\right] = \int_0^1 n F(x)^{n-1} f(x) dx = 1 \tag{6.25}$$

as we make use of the conditions $F(1) = 1$ and $F(0) = 0$. On the other hand, if $0 < \tau < 1$, then we still have

$$b(x; \tau) = \int_x^1 f(t) dt = 1 - F(x), \quad 1 - \tau \leq x \leq 1$$

but now we find

$$\begin{aligned}\mathbb{E}\left[C_n(\tau)\right] &= \int_0^{1-\tau} n(1-b(x;\tau))^{n-1} f(x)dx + \int_{1-\tau}^1 nF(x)^{n-1}f(x)dx \\ &= \tilde{K}(n;\tau) + 1 - F(1-\tau)^n \qquad (6.26)\end{aligned}$$

where we have set

$$\tilde{K}(n;\tau) := \int_0^{1-\tau} n(1-b(x;\tau))^{n-1} f(x)dx \qquad (6.27)$$

for all τ in the interval $[0,1]$ and all $n = 1, 2, \ldots$.

For any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ such that $\tau_n < 1$ for large n, the asymptotic behavior of $\mathbb{E}\left[C_n(\tau_n)\right]$ is determined by that of the terms $\tilde{K}(n;\tau_n)$ and $1 - F(1-\tau_n)^n$. As will become apparent shortly, the asymptotics for the first term are best studied through those of the integral expression

$$K(n;\tau) := \int_0^{1-\tau} ne^{-nb(x;\tau)}dx \qquad (6.28)$$

defined for all τ in the interval $[0,1]$ and all $n = 1, 2, \ldots$. This is clarified by the following basic bounds, and the discussion that follows.

Proposition 6.4 *Consider any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ which satisfies the condition*

$$\lim_{n\to\infty} \tau_n = 0. \qquad (6.29)$$

(i) There exists a positive integer $n^\star$ such that for all $n \geq n^\star$, we have $\tau_n < 1$ and

$$\tilde{K}(n;\tau_n) \leq f^\star e^{f^\star} K(n;\tau_n); \qquad (6.30)$$

(ii) Assume the additional condition

$$\lim_{n\to\infty} n(\tau_n)^2 = 0. \tag{6.31}$$

For every ε *in the unit interval* $(0,1)$, *there exists a finite integer* $n^\star(\varepsilon)$ *such that*

$$f_\star(1-\varepsilon)K(n;\tau_n) \leq \tilde{K}(n;\tau_n), \quad n \geq n^\star(\varepsilon). \tag{6.32}$$

Proposition 6.4 is proved in Appendix A. Its usefulness lies in pointing out that under appropriate conditions on the range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$, we will have $\lim_{n\to\infty} \tilde{K}(n;\tau_n) = 0$ (resp. $\lim_{n\to\infty} \tilde{K}(n;\tau_n) = \infty$) if $\lim_{n\to\infty} K(n;\tau_n) = 0$ (resp. $\lim_{n\to\infty} K(n;\tau_n) = \infty$). To pursue this idea further, we shall make use of the following observation which shows that the special case (6.3) is in some sense generic.

Lemma 6.4 *For any continuous density function* f *which satisfies Assumptions* **??**-**??**, *there always exist positive constants* a_- *and* a_+ *such that*

$$f_-(x) \leq f(x) \leq f_+(x), \quad 0 \leq x \leq 1 \tag{6.33}$$

where

$$f_\pm(x) = c + a_\pm |x - x_\star|^r, \quad 0 \leq x \leq 1. \tag{6.34}$$

A proof of Lemma 6.4 is available in Appendix B. The specific values of $a_\pm$ are not important. While f is a probability density function, there is no guarantee that f_- and f_+ are themseleves probability density functions.

We close this section with an easy byproduct of Lemma 6.4: Fix $\tau > 0$ and with the function $f_\pm$ appearing at (6.34), write

$$b_\pm(x;\tau) := \int_x^{x+\tau} f_\pm(t)dt, \quad x \geq 0. \tag{6.35}$$

As in the definition (6.22), we have conveniently extended the definition of the functions $f_\pm$ to the entire positive line, namely $f_\pm : \mathbb{R}_+ \to \mathbb{R}_+$ with $f_\pm(x) = 0$ whenever $x > 1$. Next, in analogy with (6.28) we introduce the quantities

$$K_\pm(n;\tau) := \int_0^{1-\tau} ne^{-nb_\pm(x;\tau)}dx \tag{6.36}$$

defined for all τ in the interval $[0,1]$ and all $n = 1,2,\ldots$. In view of (6.33) it is plain that the bounds

$$K_+(n;\tau) \leq K(n;\tau) \leq K_-(n;\tau) \tag{6.37}$$

hold for all $n = 1,2,\ldots$.

The advantage of working with (6.36), instead of (6.28), is purely analytical as should be apparent from the calculations below: Indeed, for all τ in $[0,1]$, we get

$$\begin{aligned} b_\pm(x;\tau) &= c\tau + a_\pm \int_x^{x+\tau} |t - x_\star|^r dt \\ &= c\tau + a_\pm \int_0^{\tau} |x + t - x_\star|^r dt \end{aligned} \tag{6.38}$$

on the range $0 \leq x \leq 1-\tau$, whence

$$K_\pm(n;\tau) = ne^{-nc\tau} \int_0^{1-\tau} e^{-na_\pm \int_0^\tau |x+t-x_\star|^r dt} dx \tag{6.39}$$

for all $n = 1,2,\ldots$.

6.6 A proof of Proposition 6.1

In the following proof of Proposition 6.1, we make an additional assumption on the parameter r: $r \geq 1$. In Section 6.9.2, we will show that Proposition 6.1 can be similarly proved when $0 < r < 1$.

Pick a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4). We need to show that

$$\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = 1 \tag{6.40}$$

whenever its deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$ satisfies the condition (6.18). To that end, we note from the upper bound in Proposition 6.4, with the help of (6.26), that (6.40) will hold provided *both* convergence statements

$$\lim_{n\to\infty} K(n;\tau_n) = 0 \tag{6.41}$$

and

$$\lim_{n\to\infty} F(1-\tau_n)^n = 0 \tag{6.42}$$

hold.

We address these issues in turn, sometimes with additional assumptions on the range function τ, notably (6.29) and (6.51). This is done mostly for technical reasons in that it leads to simpler proofs. In due course these additional conditions will be removed to ensure the desired final result. We begin by discussing (6.42).

Lemma 6.5 *For any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ which satisfies the condition (6.29), we always have (6.42).*

Proof. Fix $n = 2, 3, \ldots$ and τ in the interval $(0, 1)$. The bound (A.3) gives

$$F(1-\tau)^n \leq e^{-n(1-F(1-\tau))} \tag{6.43}$$

where the exponent can be written as

$$n(1 - F(1-\tau)) = n\tau \cdot \frac{1}{\tau} \int_{1-\tau}^{1} f(x)dx. \tag{6.44}$$

Condition (6.18) ensures $\lim_{n\to\infty} n\tau_n = \infty$ while (6.29) leads to

$$\lim_{n\to\infty} \frac{1}{\tau_n} \int_{1-\tau_n}^{1} f(x)dx = f(1) > 0.$$

Therefore, we get $\lim_{n\to\infty} n(1 - F(1-\tau_n)) = \infty$ via (6.44), and (6.42) holds by virtue of (6.43). ■

As we return to establishing (6.41) under (6.18) (with the additional condition (6.29)), we observe from the inequalities (6.37) that it suffices to establish

$$\lim_{n\to\infty} K_-(n; \tau_n) = 0. \tag{6.45}$$

Here, and in other places later in the paper, we find it convenient to write the range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ in the more compact form

$$\tau_n = \frac{1}{f_\star} \cdot \frac{1}{n} \left(\log n + \beta_n\right), \quad n = 1, 2, \ldots \tag{6.46}$$

for some deviation function $\beta : \mathbb{N}_0 \to \mathbb{R}$. The two representations (6.4) and (6.46) are related by

$$\beta_n = -\frac{1}{r} \log\log n + \alpha_n, \quad n = 2, 3, \ldots. \tag{6.47}$$

In establishing (6.45) we shall rely on the following lemma.

Lemma 6.6 *Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.46) with deviation function $\beta : \mathbb{N}_0 \to \mathbb{R}$. Under the additional condition (6.29), we have*

$$K_-(n;\tau_n) \le e^{-\Delta_n} \cdot \int_{-\infty}^{\infty} e^{-a_-|z|^r} dz \tag{6.48}$$

for all n sufficiently large where we have set

$$\Delta_n = \beta_n + \frac{1}{r}\log(n\tau_n) \tag{6.49}$$

for each $n = 1, 2, \ldots$.

Proof. Fix τ in $(0,1)$. For each $x \ge 0$, Jensen's inequality yields

$$\begin{aligned} \frac{1}{\tau}\int_0^{\tau} |x - x_\star + t|^r dt &\ge \left| \frac{1}{\tau}\int_0^{\tau} (x - x_\star + t) dt \right|^r \\ &= \left| x - x_\star + \frac{\tau}{2} \right|^r . \end{aligned} \tag{6.50}$$

Now pick a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.46) with deviation function $\beta : \mathbb{N}_0 \to \mathbb{R}$. Under (6.29), we have $\tau_n < 1$ for all n large enough and on that range the expression for $K_-(n;\tau_n)$ at (6.39) becomes

$$\begin{aligned} K_-(n;\tau_n) &= e^{-\beta_n} \int_0^{1-\tau_n} e^{-na_- \int_0^{\tau_n} |x - x_\star + t|^r dt} dx \\ &\le e^{-\beta_n} \cdot \int_0^{1-\tau_n} e^{-na_- \tau_n \left| x - x_\star + \frac{\tau_n}{2} \right|^r} dx \\ &= \frac{e^{-\beta_n}}{(n\tau_n)^{\frac{1}{r}}} \cdot \int_{(n\tau_n)^{\frac{1}{r}}(-x_\star + \frac{\tau_n}{2})}^{(n\tau_n)^{\frac{1}{r}}(1 - x_\star - \frac{\tau_n}{2})} e^{-a_-|z|^r} dz \end{aligned}$$

where the last step follows from the change of variable

$$z = (n\tau_n)^{\frac{1}{r}} \left(x - x_\star + \frac{\tau_n}{2} \right).$$

The desired conclusion (6.48)–(6.49) follows by direct inspection. ■

The first factor in the lower bound at (6.48) is discussed next.

Lemma 6.7 *Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$ satisfying (6.18). Under the additional condition*

$$\lim_{n\to\infty} \frac{\alpha_n}{\log n} = 0, \tag{6.51}$$

we have

$$\lim_{n\to\infty} \Delta_n = \infty. \tag{6.52}$$

Proof. Pick a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$. We write the range function in the more compact form (6.46) with deviation function $\beta : \mathbb{N}_0 \to \mathbb{R}$ given at (6.47). With this change of notation, set $\gamma^\star = -\frac{1}{r}\log f_\star$, and note by direct inspection that

$$\begin{aligned}
\Delta_n &= \beta_n + \frac{1}{r}\log(f_\star n\tau_n) + \gamma^\star \\
&= \beta_n + \frac{1}{r}\log(\log n + \beta_n) + \gamma^\star \\
&= \beta_n + \frac{1}{r}\log\log n + \frac{1}{r}\log\left(1 + \frac{\beta_n}{\log n}\right) + \gamma^\star \\
&= \alpha_n + \frac{1}{r}\log\left(1 + \frac{\beta_n}{\log n}\right) + \gamma^\star
\end{aligned}$$

for all $n = 2, 3, \ldots$. While $\beta_n < 0$ possibly for some $n = 2, 3, \ldots$, it is still the case that $\log n + \beta_n \geq 0$ for all $n = 2, 3, \ldots$.

Under (6.51), we obviously have

$$\lim_{n\to\infty} \frac{\beta_n}{\log n} = \lim_{n\to\infty} \frac{\alpha_n - \frac{1}{r}\log\log n}{\log n} = 0,$$

and the conclusion (6.52) follows under the condition (6.18). ■

We now complete the proof of Proposition 6.1: Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) where the deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$ satisfies the condition (6.18).

If (6.51) holds, then so does (6.29) automatically. We then get (6.42) by Lemma 6.5, while Lemmas 6.6 and 6.7 readily lead to (6.45), hence (6.41), by the finiteness of the integral

$$\int_{-\infty}^{\infty} e^{-a-|z|^r} dz.$$

As argued earlier this suffices to establish (6.40).

Consider now the situation where (6.51) fails to hold. In that case, with the given range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$, we associate an auxiliary range function $\tau' : \mathbb{N}_0 \to \mathbb{R}_+$, also written in the form (6.4) but with deviation function $\alpha' : \mathbb{N}_0 \to \mathbb{R}$ given by

$$\alpha'_n = \min\left(\alpha_n, \sqrt{\log n}\right), \quad n = 3, 4, \ldots$$

Obviously, $\lim_{n\to\infty} \alpha'_n = \infty$ under (6.18) and

$$\lim_{n\to\infty} \frac{\alpha'_n}{\log n} = 0.$$

Therefore, by the first part of the proof we already have $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau'_n)\right] = 1$. The desired conclusion $\lim_{n\to\infty} \mathbb{E}[C_n(\tau_n)] = 1$ is now immediate once we note that

$$1 \le \mathbb{E}[C_n(\tau_n)] \le \mathbb{E}[C_n(\tau_n')], \quad n = 2, 3, \ldots$$

by monotonicity since $\tau_n' \le \tau_n$ by construction.

6.7 A proof of Proposition 6.2

Again, we assume $r \ge 1$ in this Section. In Section 6.9.2, we will prove Proposition 6.2 in a similar manner when $0 < r < 1$.

Pick a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4). We need to show that

$$\lim_{n\to\infty} \mathbb{E}[C_n(\tau_n)] = \infty \tag{6.53}$$

whenever its deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$ satisfies the condition (6.20). The point of departure is the observation, derived from the lower bound in Proposition 6.4 with the help of (6.26), that (6.53) will hold provided we can show that

$$\lim_{n\to\infty} K(n; \tau_n) = \infty.$$

In view of the inequalities (6.37), this will be achieved if we show that

$$\lim_{n\to\infty} K_+(n; \tau_n) = \infty. \tag{6.54}$$

As in Section 6.6, we find it convenient to view the range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ in the more compact form (6.46) for some appropriate deviation function $\beta : \mathbb{N}_0 \to \mathbb{R}$ given at (6.47). Note that (6.20) automatically implies

$$\lim_{n\to\infty} \beta_n = -\infty. \tag{6.55}$$

Under (6.55), we see that $\beta_n < 0$ for large enough n (whence $\beta_n = -|\beta_n|$) and (6.46) therefore implies $0 \le \tau_n \le \frac{1}{f_\star} \frac{\log n}{n}$ and $|\beta_n| \le \log n$ on that range, whence

$$\lim_{n\to\infty} \tau_n = \lim_{n\to\infty} \frac{\beta_n}{n} = 0. \tag{6.56}$$

Furthermore, for each $p > 0$ it holds that

$$\lim_{n\to\infty} n(\tau_n)^{1+p} = 0 \tag{6.57}$$

since

$$n(\tau_n)^{1+p} \le f_\star^{-(p+1)} \cdot \frac{(\log n)^{1+p}}{n^p}$$

for n sufficiently large. In short, both conditions (6.29) and (6.31) are satisfied under (6.55).

Our first step towards establishing (6.54) is contained in the next technical lemma. For each $\tau > 0$ we introduce the quantities

$$B(\tau) = 2^{r-1}\tau \quad \text{and} \quad C(\tau) = \frac{2^{r-1}}{r+1}\tau^{r+1}$$

and for each $\lambda > 0$, we set

$$z_n(\lambda) = \lambda \left(nB(\tau_n)\right)^{\frac{1}{r}}, \quad n = 1, 2, \ldots.$$

Lemma 6.8 *Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.46) with deviation function $\beta : \mathbb{N}_0 \to \mathbb{R}$ satisfying the condition (6.55). For any $\lambda > 0$ such that $(x_\star - \lambda, x_\star + \lambda) \subseteq (0,1)$, we have*

$$K_+(n;\tau_n) \ge e^{\Gamma_n} \cdot \int_{-z_n(\lambda)}^{z_n(\lambda)} e^{-a_+|z|^r} dz \tag{6.58}$$

for all n sufficiently large where we have set

$$\Gamma_n = -\beta_n - na_+ C(\tau_n) - \frac{1}{r}\log\left(nB(\tau_n)\right)$$

for each $n = 1, 2, \ldots$.

Lemma 6.8 is predicated on $x_\star$ being an element of the open interval $(0, 1)$. The appropriate versions for the boundary cases $x_\star = 0, 1$ are given in Section 6.9.

Proof. Pick a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.46) with deviation function $\beta : \mathbb{N}_0 \to \mathbb{R}$. The expression at (6.39) specializes to

$$K_+(n;\tau_n) = e^{-\beta_n} \int_0^{1-\tau_n} e^{-na_+ \int_0^{\tau_n} |x - x_\star + t|^r dt} dx$$

whenever $\tau_n < 1$. For any $\lambda > 0$ as in the statement of Lemma 6.8, there exists a finite integer $n^\star(\lambda)$ such that $x_\star + \lambda < 1 - \tau_n$ for all $n \geq n^\star(\lambda)$ (since here $\lim_{n\to\infty} \tau_n = 0$ as pointed out at (6.56)). Consequently, on that range we find

$$\begin{aligned} K_+(n;\tau_n) &\geq e^{-\beta_n} \int_{x_\star-\lambda}^{x_\star+\lambda} e^{-na_+ \int_0^{\tau_n} |x+t-x_\star|^r dt} dx \\ &= e^{-\beta_n} \int_{-\lambda}^{\lambda} e^{-na_+ \int_0^{\tau_n} |y+t|^r dt} dy \end{aligned} \qquad (6.59)$$

upon making the change of variable $y = x - x_\star$.

Next, a standard convexity argument gives

$$|y + t|^r \leq 2^{r-1}\left(|y|^r + |t|^r\right), \quad y, t \in \mathbb{R} \qquad (6.60)$$

and for each y in $\mathbb{R}$, we get the upper bound

$$\begin{aligned} \int_0^{\tau_n} |y + t|^r dt &\leq 2^{r-1}\left(\tau_n |y|^r + \int_0^{\tau_n} t^r dt\right) \\ &= B(\tau_n)|y|^r + C(\tau_n). \end{aligned}$$

Reporting this last fact into (6.59) yields

$$\begin{aligned}\int_{-\lambda}^{\lambda} e^{-na_+ \int_0^{\tau_n} |y+t|^r dt} dy &\geq e^{-na_+ C(\tau_n)} \int_{-\lambda}^{\lambda} e^{-na_+ B(\tau_n)|y|^r} dy \\ &= \frac{e^{-na_+ C(\tau_n)}}{(nB(\tau_n))^{\frac{1}{r}}} \cdot \int_{-z_n(\lambda)}^{z_n(\lambda)} e^{-a_+|z|^r} dz \qquad (6.61)\end{aligned}$$

where the last equality follows from the change of variable

$$z = (nB(\tau_n))^{\frac{1}{r}} y.$$

The proof is completed by noting that the first factor in (6.61) can be written as e^{Γ_n}. ■

Next, we focus on the first factor in the lower bound at (6.58).

Lemma 6.9 *Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$ satisfying (6.20). Then, we have*

$$\lim_{n\to\infty} \Gamma_n = \infty. \qquad (6.62)$$

Proof. Pick a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$. We write the range function in the more compact form (6.46) for some appropriate deviation function $\beta : \mathbb{N}_0 \to \mathbb{R}$ given at (6.47). With this change of notation, we check by direct substitution that

$$\begin{aligned}\Gamma_n &= -\beta_n - \frac{a_+ 2^{r-1}}{r+1} \cdot n\tau_n^{r+1} - \frac{1}{r}\log(2^{r-1} n\tau_n) \\ &= -\beta_n - \frac{1}{r}\log(f_\star n\tau_n) - \gamma_n \qquad (6.63)\end{aligned}$$

for each $n = 1, 2, \ldots$ where

$$\gamma_n = \frac{a_+ 2^{r-1}}{r+1} \cdot n\tau_n^{r+1} + \frac{r-1}{r} \log 2 - \frac{1}{r} \log f_\star.$$

As pointed out earlier, (6.20) automatically implies $\lim_{n\to\infty} \beta_n = -\infty$. Thus, from (6.57) (with $p = r$) we get

$$\lim_{n\to\infty} \gamma_n = \frac{r-1}{r} \log 2 - \frac{1}{r} \log f_\star =: \gamma_\star$$

The discussion leading to (6.57) also shows that $\beta_n = -|\beta_n|$ and $|\beta_n| \leq \log n$ for n sufficiently large. On that range, we then have

$$\begin{aligned}
\Gamma_n &= -\beta_n - \frac{1}{r} \log(\log n + \beta_n) - \gamma_\star + o(1) \\
&= |\beta_n| - \frac{1}{r} \log(\log n - |\beta_n|) - \gamma_\star + o(1)
\end{aligned}$$

with

$$\begin{aligned}
\log(\log n - |\beta_n|) &= \log\log n + \log\left(1 - \frac{|\beta_n|}{\log n}\right) \\
&\leq \log\log n.
\end{aligned}$$

Therefore,

$$\Gamma_n \geq -\beta_n - \frac{1}{r} \log\log n - \gamma_\star + o(1),$$

or equivalently, $\Gamma_n \geq -\alpha_n - \gamma_\star + o(1)$, and the condition $\lim_{n\to\infty} \alpha_n = -\infty$ readily yields (6.62). ■

We are now poised to complete the proof of Proposition 6.2: Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$

satisfying (6.20). Two cases naturally emerge depending on the value of

$$T := \liminf_{n\to\infty} (n\tau_n),$$

namely $T = 0$ or $T > 0$.

If $T > 0$ (with T possibly infinite), then

$$z(\lambda) := \liminf_{n\to\infty} z_n(\lambda) = 2^{\frac{r-1}{r}} \lambda T^{\frac{1}{r}} > 0$$

with $\lambda > 0$ as specified in Lemma 6.8. By Fatou's Lemma it is plain that

$$\liminf_{n\to\infty} \int_{-z_n(\lambda)}^{z_n(\lambda)} e^{-a_+|z|^r} dz \geq \int_{-z(\lambda)}^{z(\lambda)} e^{-a_+|z|^r} dz > 0.$$

Let n go to infinity in (6.58). Combining this last observation with Lemma 6.9 implies $\liminf_{n\to\infty} K_+(n;\tau_n) = \infty$ and the desired conclusion (6.54) is obtained. This establishes (6.53).

Next, consider the case $T = 0$. With $b > 0$, define the auxiliary range function $\tau_b : \mathbb{N}_0 \to \mathbb{R}_+$ given by

$$\tau_{b,n} = \frac{1}{f_\star} \cdot \frac{1}{n} \left(\log n - \frac{1}{r} \log\log n + \alpha_n + b \right)$$

for all $n = 2, 3, \ldots$. By monotonicity, it is plain that

$$\mathbb{E}\left[C_n(\tau_{b,n})\right] \leq \mathbb{E}\left[C_n(\tau_n)\right], \quad n = 2, 3, \ldots$$

since $\tau_n \leq \tau_{b,n}$ by construction. Now, observe that the deviation function of τ_b satisfies (6.20) since that of τ does. Moreover,

$$\liminf_{n\to\infty} (n\tau_{b,n}) = T + \frac{b}{f_\star} = \frac{b}{f_\star} > 0,$$

so that $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_{b,n})\right] = \infty$ by the arguments given earlier, whence $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = \infty$ as well.

6.8 A proof of Proposition 6.3

Pick a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$. We need to show under the appropriate conditions that (6.12) holds, or equivalently,

$$\limsup_{n\to\infty} \frac{\mathbb{E}\left[\chi_{n,1}(\tau_n)\chi_{n,2}(\tau_n)\right]}{\mathbb{E}\left[\chi_{n,1}(\tau_n)\right]\mathbb{E}\left[\chi_{n,2}(\tau_n)\right]} \leq 1. \tag{6.64}$$

Our first step is to provide expressions for the quantities involved. Thus, fix τ in $(0,1)$ and $n = 3, 4, \ldots$, and write

$$g_n(\boldsymbol{x};\tau) := (1 - b(x_1;\tau))^{n-1}(1 - b(x_2;\tau))^{n-1}$$

and

$$h_n(\boldsymbol{x};\tau) := (1 - b(x_1;\tau) - b(x_2;\tau))^{n-2}$$

with $\boldsymbol{x} = (x_1, x_2)$ ranging over the unit square $[0,1]^2$.

As discussed in Section 6.5, we already have

$$\begin{aligned}
\mathbb{E}\left[\chi_{n,1}(\tau)\right]\mathbb{E}\left[\chi_{n,2}(\tau)\right] &= \left(\int_0^1 (1 - b(x_1,\tau))^{n-1} f(x_1)dx_1\right)^2 \\
&= \int_0^1\int_0^1 g_n(\boldsymbol{x};\tau)f(x_1)f(x_2)dx_1dx_2. \qquad (6.65)
\end{aligned}$$

Next, with (6.6) in mind, we note that the simultaneous validity of the conditions $X_2 \notin [X_1, X_1+\tau]$ and $X_1 \notin [X_2, X_2+\tau]$ is equivalent to either $X_1 < X_2 + \tau$ or $X_2 < X_1 + \tau$, in which case for each $j = 3, \ldots, n$, we get

$$\begin{aligned}
&\mathbf{1}\left[X_j \notin [X_1, X_1+\tau]\right]\mathbf{1}\left[X_j \notin [X_2, X_2+\tau]\right] \\
=\ & 1 - \mathbf{1}\left[X_j \in [X_1, X_1+\tau]\right] - \mathbf{1}\left[X_j \in [X_2, X_2+\tau]\right].
\end{aligned}$$

These observations and simple conditioning arguments yield

$$\mathbb{E}\left[\chi_{n,1}(\tau)\chi_{n,2}(\tau)\right] = \mathbb{E}\left[\mathbf{1}\left[|X_1 - X_2| > \tau\right] h_n(X_1, X_2;\tau)\right]$$

under the enforced independence assumptions. Therefore, with the triangles $R_+(\tau)$ and $R_-(\tau)$ defined by

$$R_+(\tau) = \left\{ \boldsymbol{x} \in [0,1]^2 : \begin{array}{c} 0 \leq x_1 \leq 1-\tau \\ \\ x_1 + \tau \leq x_2 \leq 1 \end{array} \right\}$$

and

$$R_-(\tau) = \left\{ \boldsymbol{x} \in [0,1]^2 : \begin{array}{c} \tau \leq x_1 \leq 1 \\ \\ 0 \leq x_2 \leq x_1 - \tau \end{array} \right\},$$

we conclude that

$$\mathbb{E}\left[\chi_{n,k}(\tau)\chi_{n,2}(\tau)\right] = J_n^-(\tau) + J_n^+(\tau) \tag{6.66}$$

where

$$J_n^{\pm}(\tau) = \int_{R_{\pm}(\tau)} h_n(\boldsymbol{x};\tau) f(x_1) f(x_2) dx_1 dx_2. \tag{6.67}$$

A straightforward probabilistic interpretation yields the bounds

$$0 \leq b(x_1;\tau) + b(x_2;\tau) \leq 1, \quad \boldsymbol{x} \in R_{\pm}(\tau). \tag{6.68}$$

The desired result (6.64) would be easily established if the inequalities

$$h_n(\boldsymbol{x};\tau) \leq g_n(\boldsymbol{x};\tau), \quad \boldsymbol{x} \in R_{\pm}(\tau) \tag{6.69}$$

were to hold since then a pointwise comparison of the integrands at (6.65) and (6.67) would lead to

$$\mathbb{E}\left[\chi_{n,1}(\tau)\chi_{n,2}(\tau)\right] \leq \mathbb{E}\left[\chi_{n,1}(\tau)\right]\mathbb{E}\left[\chi_{n,2}(\tau)\right]$$

in a straightforward manner. Unfortunately, the pointwise inequalities (6.69) do *not* hold on the entire range $R_{\pm}(\tau)$. However, not all is lost due to the fact that the inequalities in question can be made to hold on increasingly larger portions of

$R_{\pm}(\tau_n)$ as n increases when τ_n is scaled according to (6.4) under (6.20). This is a consequence of Lemma C.1 to be found in Appendix C.

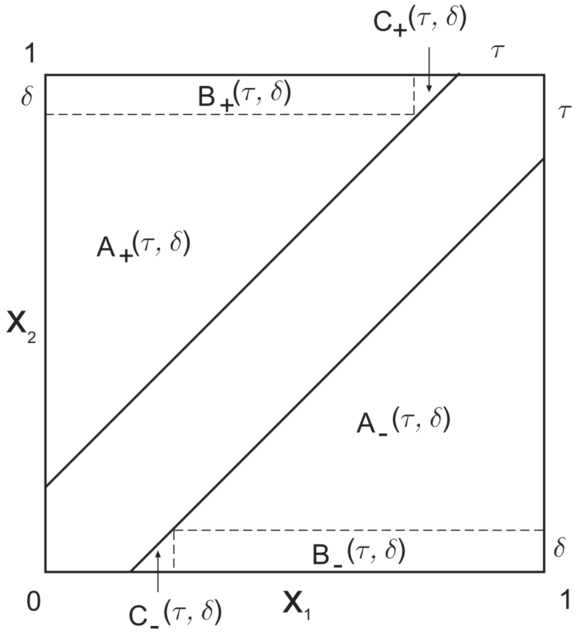


Figure 6.1: Partition of triangle $R_{\pm}(\tau)$ into three regions.

To exploit this observation, we partition the triangle $R_{\pm}(\tau)$ into three regions as described in Figure 6.1: An auxiliary parameter δ is first selected in the interval $(0, 1)$ under the constraints

$$0 < \delta < \min(\tau, 1-\tau) \quad \text{with} \quad 0 < \tau < \frac{1}{2}. \tag{6.70}$$

Later on both parameters τ and δ will be scaled with n.

The three regions associated with $R_+(\tau)$ are defined by

$$A_+(\tau,\delta) = \left\{ \boldsymbol{x} \in R_+(\tau) : \begin{array}{c} 0 \le x_1 \le 1-\tau-\delta \\ \\ x_1+\tau \le x_2 \le 1-\delta \end{array} \right\},$$

$$B_+(\tau,\delta) = \left\{ \boldsymbol{x} \in R_+(\tau) : \begin{array}{c} 0 \le x_1 \le 1-\tau-\delta \\ \\ 1-\delta \le x_2 \le 1 \end{array} \right\}$$

and

$$C_+(\tau,\delta) = \left\{ \boldsymbol{x} \in R_+(\tau) : \begin{array}{c} 1-\tau-\delta \le x_1 \le 1-\tau \\ \\ x_1+\tau \le x_2 \le 1 \end{array} \right\}.$$

In a symmetric manner, the three regions associated with $R_-(\tau)$ are defined by

$$A_-(\tau,\delta) = \left\{ \boldsymbol{x} \in R_-(\tau) : \begin{array}{c} \tau+\delta \le x_1 \le 1 \\ \\ 0 \le x_2 \le x_1-\tau \end{array} \right\},$$

$$B_-(\tau,\delta) = \left\{ \boldsymbol{x} \in R_-(\tau) : \begin{array}{c} \tau+\delta \le x_1 \le 1 \\ \\ 0 \le x_2 \le \delta \end{array} \right\}$$

and

$$C_-(\tau,\delta) = \left\{ \boldsymbol{x} \in R_-(\tau) : \begin{array}{c} \tau \le x_1 \le \tau+\delta \\ \\ x_1-\delta \le x_2 \le 1 \end{array} \right\}.$$

The basic idea for establishing (6.64) consists in showing that δ can be suitably scaled with n, say δ_n, so that (6.69) holds on $A_\pm(\tau_n,\delta_n)$ while the contributions from $B_\pm(\tau_n,\delta_n)$ and $C_\pm(\tau_n,\delta_n)$ become negligible as n becomes large. The next two lemmas will help address this technical point.

Lemma 6.10 *Assume the constraints (6.70) on τ and δ to hold. Whenever $f^\star\tau < 1$, the bounds*

$$\left|B_n^\pm(\tau,\delta)\right| \le \frac{\delta}{1-f^\star\tau} \cdot \mathbb{E}\left[\chi_{n,1}(\tau)\right] \tag{6.71}$$

hold where we have set

$$B_n^\pm(\tau,\delta) := \int_{B_\pm(\tau,\delta)} h_n(\boldsymbol{x};\tau) f(x_1) f(x_2) dx_1 dx_2.$$

Proof. We establish (6.71) only for the region $B_+(\tau, \delta)$ as the arguments for the region $B_-(\tau, \delta)$ are identical, and therefore omitted.

As we recall (6.68), we see that the easy bound

$$0 \leq h_n(\boldsymbol{x}; \tau) \leq (1 - b(x_1; \tau))^{n-2}, \boldsymbol{x} \in R_\pm(\tau)$$

implies the inequality

$$\left| \int_{B_+(\tau,\delta)} h_n(\boldsymbol{x}; \tau) f(x_1) f(x_2) dx_1 dx_2 \right| \leq \int_{B_+(\tau,\delta)} (1 - b(x_1; \tau))^{n-2} f(x_1) f(x_2) dx_1 dx_2.$$

Since the region $B_+(\tau, \delta)$ coincides with the rectangle $[0, 1 - (\tau + \delta)] \times [1 - \delta, 1]$, this last integral can be rewritten as the product of two one-dimensional integrals, namely

$$\int_0^{1-(\tau+\delta)} (1 - b(x_1; \tau))^{n-2} f(x_1) dx_1 \tag{6.72}$$

and

$$\int_{1-\delta}^1 f(x_2) dx_2. \tag{6.73}$$

Obviously,

$$\int_{1-\delta}^1 f(x_2) dx_2 \leq f^\star \delta. \tag{6.74}$$

Next, upon making use of (6.23), we find

$$\begin{aligned} \int_0^{1-(\tau+\delta)} (1 - b(x_1; \tau))^{n-2} f(x_1) dx_1 &\leq \int_0^1 \frac{(1 - b(x_1; \tau))^{n-1}}{1 - f^\star \tau} f(x_1) dx_1 \\ &= \frac{\mathbb{E}\left[\chi_{n,1}(\tau)\right]}{1 - f^\star \tau}. \end{aligned} \tag{6.75}$$

We readily get (6.71) if we apply the bounds (6.74) and (6.75) on the appropriate factors (6.72) and (6.73) in the aforementionned product form bound. ■

Lemma 6.11 *Under the constraints (6.70) on τ and δ, the bounds*

$$\left|C_n^{\pm}(\tau,\delta)\right| \le \frac{(f^\star\delta)^2}{2}$$

hold where we have set

$$C_n^{\pm}(\tau,\delta) := \int_{C_\pm(\tau,\delta)} h_n(\boldsymbol{x};\tau) f(x_1) f(x_2) dx_1 dx_2.$$

Proof. The bounds (6.68) yield

$$\begin{aligned}\left|\int_{C_\pm(\tau,\delta)} h_n(\boldsymbol{x};\tau) f(x_1) f(x_2) dx_1 dx_2\right| &\le \int_{C_\pm(\tau,\delta)} f(x_1) f(x_2) dx_1 dx_2 \\ &\le (f^\star)^2 \cdot |C_\pm(\tau,\delta)| \qquad (6.76)\end{aligned}$$

with $|C_\pm(\tau,\delta)|$ denoting the area of $C_\pm(\tau,\delta)$. The region $C_\pm(\tau,\delta)$ being a right isoceles triangle with identical sides of length δ, we have $|C_\pm(\tau,\delta)| = \frac{1}{2}\delta^2$ and the desired inequality is now immediate. ■

Next, we give conditions on δ and τ that ensure the pointwise comparison (6.69) on the triangles $A_\pm(\tau,\delta)$.

Lemma 6.12 *Assume that τ and δ satisfy the constraints (6.70). The comparison*

$$h_n(\boldsymbol{x};\tau) \le g_n(\boldsymbol{x};\tau), \quad \boldsymbol{x} \in A_\pm(\tau,\delta) \qquad (6.77)$$

holds under the additional condition

$$\frac{1}{f_\star} \cdot \frac{2}{n-2} \leq \delta. \tag{6.78}$$

Proof. From Lemma C.1 in Appendix C we see that (6.77) holds if the condition (6.78) garantees that

$$\frac{2}{n-2} \leq \min\left(b(x_1;\tau), b(x_2;\tau)\right) \tag{6.79}$$

for *all* $\boldsymbol{x}$ in $A_\pm(\tau,\delta)$.

It is plain that

$$f_\star \min\left(\tau, 1-x\right) \leq b(x;\tau), \quad x \in [0,1].$$

Therefore, for any $\boldsymbol{x}$ in $A_+(\tau,\delta)$, (6.79) holds provided

$$\frac{1}{f_\star} \cdot \frac{2}{n-2} \leq \min\left(\tau, 1-x_1, 1-x_2\right). \tag{6.80}$$

Membership of $\boldsymbol{x}$ in $A_+(\tau,\delta)$ amounts to $0 \leq x_1 \leq 1-\tau-\delta$ and $x_1+\tau \leq x_2 \leq 1-\delta$. This in turn implies $\tau+\delta \leq 1-x_1$ and $\delta \leq 1-x_2$, so that (6.80) will hold if

$$\frac{1}{f_\star} \cdot \frac{2}{n-2} \leq \min\left(\tau, \tau+\delta, \delta\right). \tag{6.81}$$

That this is the case is a simple consequence of the conditions (6.70) and (6.77).

This establishes (6.77) for the region $A_+(\tau,\delta)$. The arguments for the region $A_-(\tau,\delta)$ are identical, and are therefore omitted. ■

By now the cumulative effect of Lemmas 6.10, 6.11 and 6.12 should have become clear: Assume that n, τ and δ satisfy the conditions of *all* three lemmas simultaneously. Then, upon writing

$$A_n^{\pm}(\tau,\delta) := \int_{A_{\pm}(\tau,\delta)} h_n(\boldsymbol{x};\tau) f(x_1) f(x_2) dx_1 dx_2,$$

we first get from Lemma 6.12 that

$$\begin{aligned} A_n^{\pm}(\tau,\delta) &\leq \int_{A_{\pm}(\tau,\delta)} g_n(\boldsymbol{x};\tau) f(x_1) f(x_2) dx_1 dx_2 \\ &\leq \mathbb{E}\left[\chi_{n,1}(\tau)\right] \mathbb{E}\left[\chi_{n,2}(\tau)\right]. \end{aligned} \tag{6.82}$$

Next, Lemma 6.10 yields

$$\begin{aligned} \frac{|B_n^{\pm}(\tau,\delta)|}{\mathbb{E}\left[\chi_{n,1}(\tau)\right]\mathbb{E}\left[\chi_{n,2}(\tau)\right]} &\leq \frac{\delta}{1-f^\star\tau} \cdot \mathbb{E}\left[\chi_{n,1}(\tau)\right]^{-1} \\ &= \frac{1}{1-f^\star\tau} \cdot \frac{n\delta}{\mathbb{E}\left[C_n(\tau)\right]}, \end{aligned} \tag{6.83}$$

while Lemma 6.11 leads to

$$\begin{aligned} \frac{|C_n^{\pm}(\tau,\delta)|}{\mathbb{E}\left[\chi_{n,1}(\tau)\right]\mathbb{E}\left[\chi_{n,2}(\tau)\right]} &\leq \frac{(f^\star\delta)^2}{2\mathbb{E}\left[\chi_{n,1}(\tau)\right]\mathbb{E}\left[\chi_{n,2}(\tau)\right]} \\ &= \frac{(f^\star)^2}{2} \cdot \left(\frac{n\delta}{\mathbb{E}\left[C_n(\tau)\right]}\right)^2. \end{aligned} \tag{6.84}$$

In both cases, we used the fact that $\mathbb{E}\left[C_n(\tau)\right] = n\mathbb{E}\left[\chi_{n,k}(\tau)\right]$ for all $k = 1,\ldots,n$. Finally, a straightforward decomposition argument gives

$$\begin{aligned} \frac{\mathbb{E}\left[\chi_{n,1}(\tau)\chi_{n,2}(\tau)\right]}{\mathbb{E}\left[\chi_{n,1}(\tau)\right]\mathbb{E}\left[\chi_{n,2}(\tau)\right]} &\leq 1 + \frac{2}{1-f^\star\tau} \cdot \frac{n\delta}{\mathbb{E}\left[C_n(\tau)\right]} \\ &\quad + (f^\star)^2 \cdot \left(\frac{n\delta}{\mathbb{E}\left[C_n(\tau)\right]}\right)^2 \end{aligned} \tag{6.85}$$

upon combining the inequalities (6.82), (6.83) and (6.84).

We are now ready to complete the proof of Proposition 6.3: Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) whose deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$ satisfies (6.20). Then, $\lim_{n\to\infty} \tau_n = 0$ as pointed out already at (6.56), whence $\tau_n < \frac{1}{2}$ and $\tau_n f^\star < 1$ for all n sufficiently large. Next, we scale δ with n according to

$$\delta_n = \frac{1}{f_\star} \cdot \frac{2}{n-2}, \quad n = 3, 4, \ldots$$

With this choice, the constraint $\delta_n < \tau_n$ is equivalent to

$$\frac{2n}{n-2} < \log n - \frac{1}{r} \log\log n - |\alpha_n|,$$

a condition which is clearly satisfied under (6.21) for all n sufficiently large. In short, for all n sufficiently large, it is appropriate in the inequality (6.85) to replace the parameters τ and δ by τ_n and δ_n as specified earlier. Finally, let n go to infinity in the resulting inequality: We readily get $\lim_{n\to\infty} \mathbb{E}\left[C_n(\tau_n)\right] = \infty$ by Proposition 6.2, while $\lim_{n\to\infty} n\delta_n = \frac{2}{f_\star}$.. Combining these facts yields

$$\lim_{n\to\infty} \frac{n\delta_n}{\mathbb{E}\left[\chi_{n,k}(\tau_n)\right]} = 0, \quad k = 1, \ldots, n$$

and the desired convergence (6.64) is established.

6.9 Discussion

6.9.1 The boundary cases $x_\star = 0, 1$

Some extra needs to be exercised when dealing with the boundary cases $x_\star = 0, 1$. The discussion of Sections 6.6 and 6.7 indicates that only Lemma 6.8 needs to be modified. We review the needed changes for a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.46) with deviation function $\beta : \mathbb{N}_0 \to \mathbb{R}$ satisfying the condition (6.55):

For the case $x_\star = 0$, the bound (6.58) cannot be given anymore in a symmetric form. However, the arguments leading to (6.58) can be easily modified to show that for any λ in $(0,1)$, we now have

$$K_+(n;\tau_n) \geq e^{\Gamma_n} \cdot \int_0^{z_n(\lambda)} e^{-a_+|z|^r} dz \tag{6.86}$$

for all n sufficiently large.

The case $x_\star = 1$ is slightly more involved: For any λ in $(0,1)$, we have now

$$K_+(n;\tau_n) \geq e^{\Gamma_n} \cdot \int_{z_n(\lambda)}^{z_n(1)} e^{-a_+|z|^r} dz \tag{6.87}$$

for all n sufficiently large. Indeed, for such λ there exists a finite integer $n^\star(\lambda)$ such that $\tau_n < \lambda$ for all $n \geq n^\star(\lambda)$ (since here $\lim_{n\to\infty} \tau_n = 0$). On that range, the expression at (6.39) yields

$$\begin{aligned} K_+(n;\tau_n) &\geq e^{-\beta_n} \int_{1-\lambda}^{0} e^{-na_+ \int_0^{\tau_n} |x+t-1|^r dt} dx \\ &= e^{-\beta_n} \int_{\lambda}^{1} e^{-na_+ \int_0^{\tau_n} |t-y|^r dt} dy \end{aligned}$$

upon making the change of variable $y = 1 - x$. The subsequent convexity argument and change of variable remain unchanged as in the derivation of the bound (6.58), ultimately yielding (6.87).

We conclude by observing that in their respective cases, the bounds (6.86) and (6.87) are sufficient to allow the concluding arguments of Section 6.6 to proceed.

6.9.2 From $r \geq 1$ to $0 < r < 1$

We first prove Proposition 6.1 under the assumption that $0 < r < 1$. The key idea is to replace (6.50) by the following inequality

$$\frac{1}{\tau}\int_0^\tau |x - x_\star + t|^r dt > \frac{1}{2^r(r+1)}\left|x - x_\star + \frac{\tau}{2}\right|^r. \tag{6.88}$$

This inequality holds due to the following facts:

When $x - x_\star > \frac{\tau}{2}$ or $x - x_\star < -\frac{3\tau}{2}$, we have

$$\frac{1}{\tau}\int_0^\tau |x - x_\star + t|^r dt > |x - x_\star|^r > \frac{1}{2^r(r+1)}\left|x - x_\star + \frac{\tau}{2}\right|^r;$$

while when $-\frac{3\tau}{2} \leq x - x_\star \leq \frac{\tau}{2}$, we see that

$$\frac{1}{\tau}\int_0^\tau |x - x_\star + t|^r dt \geq \frac{\tau^r}{2^r(r+1)} \geq \frac{1}{2^r(r+1)}\left|x - x_\star + \frac{\tau}{2}\right|^r$$

where the two equalities hold when $x - x_\star = -\frac{\tau}{2}$ and when $x - x_\star = \frac{\tau}{2}, -\frac{3\tau}{2}$, respectively.

The only difference between (6.50) and (6.88) is the constant factor $\frac{1}{2^r(r+1)}$. Furthermore, (6.50) is the only expression in Section 6.6 whose establishment requires $r \geq 1$. As a result, Proposition 6.1 can be shown for the case $0 < r < 1$ in a very similar way as the case $r \geq 1$.

The basic idea to prove Proposition 6.2 under the assumption $0 < r < 1$ is similar as we can replace (6.60) by the following inequality

$$|y + t|^r \leq |y|^r + |t|^r, \quad y, t \in \mathbb{R}. \tag{6.89}$$

Note that (6.89) is not valid for $r > 1$.

6.9.3 Non-isolated minima

A proof of Theorem 6.2 can easily be cobbled from the discussion of Theorem 6.1. Going back to to the relation (6.26) we need only show that $\lim_{n\to\infty} \tilde{K}(n;\tau_n) = 0$ (resp. $\lim_{n\to\infty} \tilde{K}(n;\tau_n) = \infty$) under appropriate conditions on the range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$. This is easily done by noting the following bounds: First, it is always the case that

$$\tilde{K}(n;\tau) \leq nf^\star(1-\tau f_\star)^{n-1}, \quad n = 1, 2, \ldots$$

for all τ in $(0,1)$ (as we observe that $f_\star\tau < 1$ since $f_\star < 1$). Next, since I is non-empty, there always exists a non-empty interval $J \subseteq I$ such that $J + \tau \subseteq I$ for all τ in $(0,1)$ small enough. For such values, we find that

$$b(x;\tau) = f_\star\tau, \quad x \in J$$

and the lower bound

$$\tilde{K}(n;\tau) \geq n|J|(1-\tau f_\star)^{n-1} f_\star, \quad n = 1, 2, \ldots$$

follows where $|J|$ denotes the length of the interval J.

It is now a simple matter to check that $\lim_{n\to\infty} n(1-\tau_n f_\star)^{n-1} = 0$ (resp. $\lim_{n\to\infty} n(1-\tau_n f_\star)^{n-1} = \infty$) for any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.5) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$ satisfying $\lim_{n\to\infty} \alpha_n = \infty$ (resp. $\lim_{n\to\infty} \alpha_n = -\infty$).

6.9.4 Extensions

It can be shown that the results obtained in this Chapter still hold if the density function f represented in the form

$$f(x) = \begin{cases} c + b_-|x - x_\star|^r + h_-(x) & \text{if } 0 \le x \le x_\star \\ \\ c + b_+|x - x_\star|^r + h_+(x) & \text{if } x_\star \le x \le 1 \end{cases}$$

for some parameters $r > 0$, $b_- > 0$, $b_+ > 0$ and $c > 0$, and for some functions $h_\pm : [0,1] \to \mathbb{R}$ such that

$$\lim_{x \uparrow x_\star} \frac{h_-(x)}{|x - x_\star|^r} = 0 \tag{6.90}$$

and

$$\lim_{x \downarrow x_\star} \frac{h_+(x)}{|x - x_\star|^r} = 0. \tag{6.91}$$

6.9.5 Earlier results of Deheuvels

According to Theorem 3.2, Deheuvels identifies the following asymptotic bounds:

$$-\frac{1}{r} = \lim\inf_{n\to\infty} \frac{nM_n f_\star - \log n}{\log\log n} < \lim\sup_{n\to\infty} \frac{nM_n f_\star - \log n}{\log\log n} = 2 - \frac{1}{r} \quad a.s.. \tag{6.92}$$

On the other hand, our results indicate that for any range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ written in the form (6.4) with deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$, it holds that

$$\lim_{n\to\infty} P(M_n \le \tau_n) = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \alpha_n = -\infty \\ \\ 1 & \text{if } \lim_{n\to\infty} \alpha_n = +\infty. \end{cases}$$

After simple algebraic manipulations, we get

$$\lim_{n\to\infty} P\left(\frac{nM_n f(x_\star) - \log n}{\log\log n} + \frac{1}{r} \leq \frac{\alpha_n}{\log\log n}\right) = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \alpha_n = -\infty \\ \\ 1 & \text{if } \lim_{n\to\infty} \alpha_n = +\infty. \end{cases}$$

Choosing $\alpha_n = o(\log\log n)$, the one-law indicates that

$$\frac{nM_n f_\star - \log n}{\log\log n} \xrightarrow{P}{}_n \; -\frac{1}{r}. \tag{6.93}$$

Although the convergence in (6.93) is not almost sure convergence, it suggests that the lower bound in (6.92) is a tighter bound than the upper bound. We expect

$$\lim_{n\to\infty} \frac{nM_n f(x_\star) - \log n}{\log\log n} = -\frac{1}{r} \quad a.s.,$$

but it is not clear how to establish this result.

Chapter 7

Network connectivity under the GRG model IV : A weak threshold for general user distribution with vanishing densities

7.1 The main result

When $f_\star = 0$, a blind application of Theorem 5.1 yields $\tau_n^\star = \infty$ for all $n = 1, 2, \ldots$. This begs the question as to what is the appropriate analog of Theorem 5.1 when the density f vanishes. We explore this issue through the following simple example: With $p > 0$, consider the probability distribution F given by

$$F(x) = x^{p+1}, \quad x \in [0, 1] \tag{7.1}$$

so that

$$f(x) = (p+1)x^p, \quad x \in [0, 1]. \tag{7.2}$$

Theorem 5.1 needs to be replaced by the following result.

Theorem 7.1 *Under (7.1), the property of graph connectivity admits only weak* *critical threshold functions, and the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ given by*

$$\tau^\star_n = n^{-\frac{1}{p+1}}, \quad n = 1, 2, \ldots \tag{7.3}$$

is such a weak threshold function.

The random graph $\mathbb{G}(n;\tau)$ under (7.1) provides yet another situation where a strong critical threshold does not exist for a monotone graph property [42, Thm. 5.1, p. 382]. The remainder of this Chapter is devoted to establishing Theorem 7.1.

7.2 A representation of the maximal spacing

Consider the order statistics $U_{n,1}, \ldots, U_{n,n}$ associated with the n i.i.d. rvs $U_1, \ldots, U_n$ which are all uniformly distributed on $[0,1]$. Since $F^{-1}(t) = t^{\frac{1}{p+1}}$, $0 \leq t \leq 1$, it is plain from (5.5) that

$$\begin{aligned}(L_{n,k},\ k = 2, \ldots, n) &=_{st} \left(F^{-1}(U_{n,k}) - F^{-1}(U_{n,k-1}),\ k = 2, \ldots, n\right) \\ &= \left((U_{n,k})^{\frac{1}{p+1}} - (U_{n,k-1})^{\frac{1}{p+1}},\ k = 2, \ldots, n\right).\end{aligned}$$

In order to take advantage of this last equivalence, we introduce a collection of $\{\xi_j,\ j = 1, 2, \ldots\}$ of i.i.d. rvs which are exponentially distributed with unit parameter, and set

$$T_0 = 0,\ T_k = \xi_1 + \ldots + \xi_k, \quad k = 1, 2, \ldots.$$

Upon defining

$$V_k := (T_k)^{\frac{1}{p+1}} - (T_{k-1})^{\frac{1}{p+1}}, \quad k = 1, 2, \ldots,$$

we get

$$\left(F^{-1}(U_{n,k}) - F^{-1}(U_{n,k-1}),\ k = 2,\dots,n\right) \quad =_{st} \quad \left(\frac{V_k}{(T_{n+1})^{\frac{1}{p+1}}},\ k = 2,\dots,n\right)$$

according to the stochastic equivalence established in (5.10).

Consequently, the distributional equivalence

$$M_n =_{st} \frac{M_n^\star}{(T_{n+1})^{\frac{1}{p+1}}} \tag{7.4}$$

holds where we have defined

$$M_n^\star := \max\left(V_k,\ k = 2,\dots,n\right). \tag{7.5}$$

7.3 A proof of Theorem 7.1

Throughout this section the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ is the one given by (7.3). We start with the following key representation that flows from (7.4)–(7.5), namely

$$\frac{M_n}{\tau_n^\star} =_{st} \left(\frac{n}{T_{n+1}}\right)^{\frac{1}{p+1}} \cdot M_n^\star \tag{7.6}$$

for all $n = 1, 2 \dots$. The proof proceeds according to three distinct steps.

7.3.1 The range function $\tau^\star$ is a weak threshold

In view of Lemma 4.1, the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ is a weak threshold if we show that (4.6) holds for some $\mathbb{R}_+$-valued rv L with $L > 0$ a.s. By the Strong Law of Large Numbers, we already have

$$\lim_{n\to\infty} \frac{T_{n+1}}{n} = 1 \quad a.s. \tag{7.7}$$

Moreover, the sequence $\{M^\star_n,\ n = 2, 3, \ldots\}$ being monotone, we have the a.s. convergence

$$\lim_{n\to\infty} M^\star_n = \sup\left(V_k,\ k = 2, \ldots\right) =: M^\star. \tag{7.8}$$

We shall show that $M^\star$ is a.s. finite with $M^\star > 0$ a.s.

First, we note that $M^\star \geq V_2$. But $V_2 = 0$ if and only if $T_2 = T_1$, which occurs if and only if $\xi_2 = 0$, this last event occuring with zero probability. Consequently $V_2 > 0$ a.s. and $M^\star > 0$ a.s., as needed.

Next, fix $k = 2, 3, \ldots$ and for notational convenience, set $q = \frac{p}{p+1}$ and $r = \frac{p+1}{p} = q^{-1}$. It is plain that

$$\begin{aligned}
V_k &= (T_k)^{\frac{1}{p+1}} - (T_{k-1})^{\frac{1}{p+1}} \\
&= \frac{1}{p+1} \int_{T_{k-1}}^{T_k} t^{-q} dt \\
&\leq \frac{1}{p+1} \int_{T_{k-1}}^{T_k} (T_{k-1})^{-q}\, dt \\
&= \frac{1}{p+1} \cdot (T_{k-1})^{-q} \cdot \xi_k
\end{aligned} \tag{7.9}$$

with

$$(T_{k-1})^{-q} \cdot \xi_k = \left(\frac{k}{T_{k-1}} \cdot \frac{\xi_k^r}{k} \right)^q.$$

The Strong Law of Large Numbers immediately implies

$$\lim_{k\to\infty} \frac{k}{T_{k-1}} = 1 \quad a.s.$$

as pointed out earlier. Applying again the Strong Law of Large Numbers, this time to the sequence of i.i.d. rvs $\{\xi_k^r,\ k = 1, 2, \ldots\}$, we find

$$\lim_{k\to\infty} \frac{1}{k} \sum_{\ell=1}^{k} \xi_\ell^r = \mathbb{E}\left[\xi_1^r\right] \quad a.s.$$

The exponential distribution having finite moments of all orders, we obviously have $\mathbb{E}\left[\xi_1^r\right]$ finite, whence

$$\lim_{k\to\infty} \frac{\xi_k^r}{k} = 0 \quad a.s.$$

according to a standard argument.

With the help of these observations, we conclude that

$$\lim_{k\to\infty} \left(T_{k-1}\right)^{-q} \cdot \xi_k = 0 \quad a.s.$$

whence $\lim_{k\to\infty} V_k = 0$ a.s. Therefore, there exists a positive integer (sample dependent) ν which is a.s. *finite* such that $M^\star = V_\nu$ and $M^\star$ is a.s. finite.

Making use of the convergence statements (7.7) and (7.8), we readily see from (7.6) that

$$\frac{M_n}{\tau_n^\star} \Longrightarrow_n M^\star \tag{7.10}$$

and (4.6) therefore holds with $L =_{st} M^\star$ as desired.

7.3.2 The range function $\tau^\star$ is not a strong threshold

Pick ε in $(0,1)$ and $n = 2, 3, \ldots$. Obviously, $M_n^\star \geq V_2$, so that

$$\mathbb{P}\left[M_n^\star > 1 + \varepsilon\right] \geq \mathbb{P}\left[V_2 > 1 + \varepsilon\right] > 0$$

and $M^\star > 1$ with positive probability! Thus, (4.11) fails and by Lemma 4.2 the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ is *not* a strong threshold for the property of graph connectivity in $\mathbb{G}(n;\tau)$.

7.3.3 There exists no strong threshold

The argument proceeds by contradiction: Assume that a strong threshold function does exist, say $\sigma : \mathbb{N}_0 \to \mathbb{R}_+$, in which case we have $\frac{M_n}{\sigma_n} \xrightarrow{P}{}_n 1$ by Lemma 4.2. Using (7.10), we readily conclude

$$\frac{\sigma_n}{\tau_n^\star} \Longrightarrow_n M^\star \tag{7.11}$$

as we note

$$\frac{\sigma_n}{\tau_n^\star} = \frac{\sigma_n}{M_n} \cdot \frac{M_n}{\tau_n^\star}, \quad n = 2, 3, \ldots$$

The limit $\lim_{n\to\infty} \frac{\sigma_n}{\tau_n^\star}$ being *deterministic* we have a contradiction since $M^\star$ is not a degenerate rv. Consequently, there cannot be any strong threshold function for the property of graph connectivity.

7.4 Discussion

It is easy to check from Theorem 5.1 that the threshold function $n \to \frac{\log n}{n}$ is a weak threshold function, a *robust*, albeit weak, conclusion which holds across *all* distributions F with non-vanishing density. However, with F given by (7.1), the critical threshold given by (7.3) is now of a much larger order since

$$\frac{\log n}{n} = o\left(n^{-\frac{1}{p+1}}\right).$$

Implications for resource dimensioning in two-dimensional ad-hoc networks were already discussed in the references [50, 51], and take here the following form: As will become apparent from the comments following Lemma 4.2, critical thresholds serve as proxy for the critical transmission range when n is large. Thus, under a node placement with a vanishing density such as (7.1), we see that the critical transmission range is orders of magnitude *larger* than would otherwise have been the case when the density function does not vanish, resulting in *higher* minimum power levels to ensure connectivity. Similar *qualitative* conclusions were already pointed out by Santi [51, Thm. 4] for two-dimensional networks under the random waypoint mobility model without pause. In one dimension, the corresponding stationary spatial node density is given by

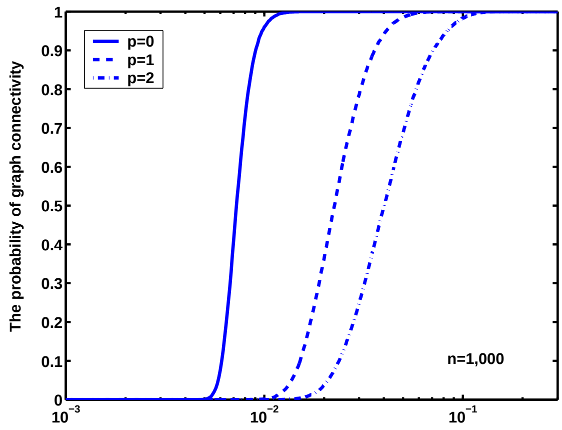
$$f_{\mathrm{RWP}}(x) = 6\ x(1-x), \quad 0 \le x \le 1. \tag{7.12}$$

Here, under (7.1) we can go beyond qualitative statements and give *precise* information on the *order* of the asymptotics for the critical transmission range.

Although the distribution (7.1) was selected because its simpler form facilitated the analysis, it is nevertheless representative of vanishing densities such as (7.12). Indeed, both Theorems 5.1 and 7.1 derive from limiting properties of the maximal spacing under F. Such properties are influenced by the behavior of the density in the vicinity of its minimum point [37, p. 519.]: The densities (7.2) (with $p = 1$) and (7.12) have similar behavior near $x = 0$ since $f_{\mathrm{RWP}}(x) \sim 6x$ as $x \simeq 0$. Thus, the results obtained here suggest that this model requires a much larger critical transmission range function given by

$$\tau^{\star}_{\mathrm{RWP},n} = \frac{1}{\sqrt{n}}, \quad n = 1, 2, \ldots.$$

According to Theorem 4.3, the number of breakpoint users under uniform node placement will converge to a Poisson rv under the appropriate critical scaling. This



property crisply captures the fact that the phase transition usually associated with strong zero-one laws is a very sharp one indeed. However, the absence of strong critical thresholds under (7.1) precludes such Poisson convergence, and essentially rules out the possibility that the corresponding phase transition will be sharp in this case.

These conclusions are already apparent from the limited simulation results presented above where nodes are placed according to F_p with $p = 0, 1, 2$; the case $p = 0$ corresponds to the uniform distribution. For each $p = 0, 1, 2$, the figure displays the corresponding plot of $P(n, \tau)$ as a function of τ (in base 10 log-scale) for $n = 1,000$. In each case we generated $K = 10,000$ mutually independent configurations of n points on the interval $[0, 1]$ drawn independently according to F_p. We compute the value $P(n, \tau)$ as the ratio $X_K(n, \tau)/K$ where $X_K(n, \tau)$ records the number of configurations among these K configurations which result in a connected graph when the transmission range is τ. As expected, the phase transition is much sharper for $p = 0$ than for positive p. These displays also suggest that the sharpness of the phase transition decreases with increasing p. However, at the time of this writing, we are not in a position to offer precise quantitative results along these lines.

Chapter 8

Network connectivity under the BCG model

In this final Chapter, we study network connectivity in the context of the bounded connection graph (BCG) model. In contrast with the GRG model, the BCG model takes into account random radio signal variations, which are unavoidable in wireless communication networks. Moreover, as discussed in Chapter 2, the BCG model could capture a broad range of radio propagation models. Although the BCG model covered here does not include the lognormal case, such a generalization is a big advantage over the lognormal connection graph (LCG) model.

Our main contribution in this Chapter is to identify the critical scaling (with respect to the number of nodes) of the boundary communication range for the absence of isolated nodes. We prove that if the boundary communication range is around the critical scaling, the distribution of the number of isolated nodes converges to a Poisson distribution as the number of nodes tends to infinity.

Our proof is composed of two steps: Our main efforts are devoted to prove Proposition 8.1 under the assumption that nodes are placed according to a homogeneous Poisson point process on $[0,1]^2$. In Theorem 8.1, we extend this result to the case that nodes are uniformly and independently distributed on $[0,1]^2$.

8.1 Notation and definitions

We will introduce some of the notation and definitions to be used throughout Chapter 8. We want to emphasize that all these notation and definitions are based on two assumptions, namely that nodes follow a homogeneous Poisson point process with density n and that the probability that a link exists between a pair of nodes (i.e. the two nodes are connected) is computed by the BCG model with parameter ρ. We use $\mathbb{G}_{2,P}(n;\rho)$ to denote the two-dimensional bounded connection graph formed under these two assumptions.

Fix $\rho > 0$. With the definition of h_ρ in (2.17), we define

$$s_\rho := 2\pi \int_0^\rho h_\rho(r) r \, dr \tag{8.1}$$

and

$$\kappa_\rho := \frac{2\pi \int_0^\rho h_\rho^2(r) r \, dr}{s_\rho}. \tag{8.2}$$

Assuming that the boundary effects are ignored, it can be shown that s_ρ is the probability that a node with given coordinates is connected with another node that is uniformly distributed on $[0,1]^2$. Moreover, according to (2.17), $h_\rho(r) < 1$ when $r > 0$, whence $\kappa_\rho < 1$ as can be easily derived.

Fix $n = 2, 3, \ldots$. With $m_n = n^2$, we define the squarelets

$$\Sigma_{n,ij} = \left[\frac{i-1}{m_n}, \frac{i}{m_n}\right] \times \left[\frac{j-1}{m_n}, \frac{j}{m_n}\right], \quad i, j = 1, \ldots, m_n.$$

The length of the squarelet's side is denoted by $\ell_n = \frac{1}{m_n}$. Set $t_n = \lceil \frac{\rho}{\ell_n} \rceil$.

In order to facilitate the forthcoming analysis, we divide these m_n^2 squares into 9 groups according to their locations. Specifically, we have 4 groups at the corner, 4 groups on the edge, and 1 group in the center. Two division patterns are demonstrated in Fig. 8.1 and Fig. 8.2. In Fig. 8.1, the $m_n^2(n^4)$ squarelets are divided into 9 groups. The group R_1 in the center contains $(m_n - 2t_n)^2$ squarelets. Each of the groups $\{R_{2i}, \quad i = 1, \ldots, 4\}$ located in the corners contains t_n^2 squarelets. Each of the groups $\{R_{3i}, \quad i = 1, \ldots, 4\}$ located in the edges contains $t_n(m_n - 2t_n)$ squarelets. We define $R_2 = \cup_{i=1}^4 R_{2i}$ and $R_3 = \cup_{i=1}^4 R_{3i}$. In Fig. 8.2, the $m_n^2(n^4)$ squarelets are divided into 9 groups. The group R_4 in the center contains $(m_n - 4t_n)^2$ squarelets. Each of the groups $\{R_{5i}, \quad i = 1, \ldots, 4\}$ located in the corners contains $4t_n^2$ squarelets. Each of the groups $\{R_{6i}, \quad i = 1, \ldots, 4\}$ located in the edges contains $2t_n(m_n - 4t_n)$ squarelets. We define $R_5 = \cup_{i=1}^4 R_{5i}$ and $R_6 = \cup_{i=1}^4 R_{6i}$.

Fix $\rho > 0$. With $i, j = 1, \ldots, m_n$, we use the symbols $N_{ij}(n)$ and $J_{ij}(n;\rho)$ to denote the number of users, and the number of isolated users in the squarelet $\Sigma_{n,ij}$, respectively. We also write

$$J'_{ij}(n;\rho) := \mathbf{1}[N_{ij}(n) = 1] J_{ij}(n;\rho), \quad i, j = 1, \ldots, m_n$$

and

$$J''_{ij}(n;\rho) := \mathbf{1}[N_{ij}(n) > 1] J_{ij}(n;\rho), \quad i, j = 1, \ldots, m_n.$$

Their corresponding sums are given by

$$C'(n;\rho) := \sum_{(i,j)\in\Gamma_n} J'_{ij}(n;\rho) \quad \text{and} \quad C''(n;\rho) := \sum_{(i,j)\in\Gamma_n} J''_{ij}(n;\rho)$$

where $\Gamma_n = \{1, \ldots, n\} \times \{1, \ldots, n\}$. We denote the total number of isolated nodes by $C(n;\rho)$. We clearly have

$$J_{ij}(n;\rho) = J'_{ij}(n;\rho) + J''_{ij}(n;\rho), \quad i,j = 1,\ldots,m_n$$

and

$$C(n;\rho) = C'(n;\rho) + C''(n;\rho).$$

The probability $P_{iso,2,P}(n;\rho)$ that $\mathbb{G}_{2,P}(n;\rho)$ contains no isolated users is given by

$$P_{iso,2,P}(n;\rho) = \mathbb{P}[C(n;\rho) = 0]. \tag{8.3}$$

Another frequently used notation is $D(A,B)$, a rv that represents the Euclidean distance between users A and B that are randomly placed. We also use $d(A,B)$ to denote the Euclidean distance between users A and B with given coordinates.

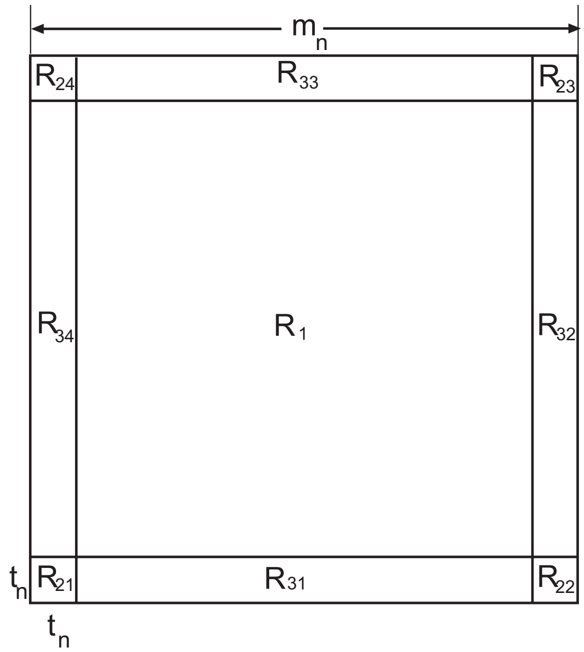


Figure 8.1: Squarelet division pattern one.

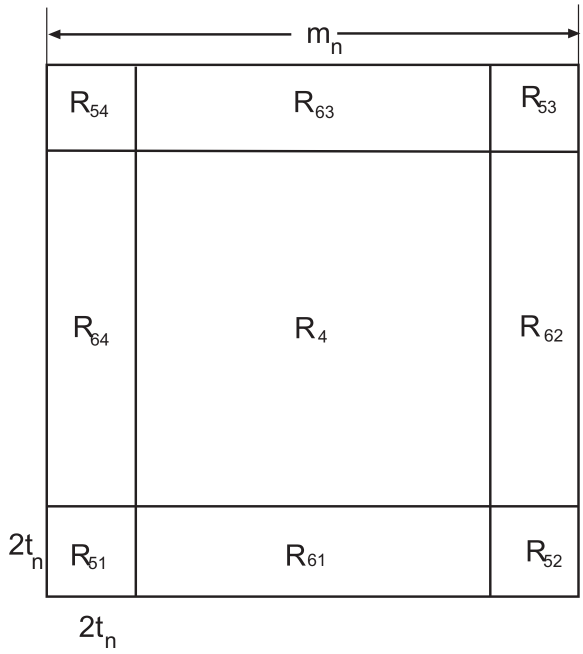


Figure 8.2: Squarelet division pattern two.

8.2 Preliminary results

Lemma 8.1 *For any boundary range function $\rho : \mathbb{N}_0 \to \mathbb{R}_+$, we have*

$$\lim_{n\to\infty} \mathbb{P}[C''(n;\rho_n) > 0] = 0. \tag{8.4}$$

Proof. Using the union bound, it is plain that

$$\begin{aligned} \mathbb{P}[C''(n;\rho_n) > 0] &\leq \sum_{(i,j)\in\Gamma_n} \mathbb{P}[J''_{ij}(n;\rho_n) > 0] \\ &< \sum_{(i,j)\in\Gamma_n} \mathbb{P}[N_{ij}(n) \geq 2]. \end{aligned}$$

Recall that

$$\lim_{\lambda\downarrow 0} \lambda^{-2}\left(\sum_{l=2}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda}\right) = 2,$$

thus we have

$$\sum_{(i,j)\in\Gamma_n} \mathbb{P}[N_{ij}(n) \geq 2] = n^4\Theta((n\ell_n^2)^2)$$
$$= \Theta(n^{-2}).$$

Finally

$$\lim_{n\to\infty} \mathbb{P}[C''(n;\rho_n) > 0] = 0.$$

■

Lemma 8.2 *For any (i,j) in Γ_n, $B_{n,ij} = \{(k,l) : |k-i| \leq 2t_n$ and $|l-j| \leq 2t_n, k,l = 1,\ldots,m_n\}$ is a neighborhood of dependence for (i,j) with respect to the indicator rvs $\{J'_{ij}(n;\rho),(i,j)\in\Gamma_n\}$.*

Proof. Based on its definition, the indicator rv $J'_{ij}(n;\rho)$ is determined by the Poisson point process in

$$L_{n,ij} = \cup_{\substack{|k-i|\leq t_n \\ |l-j|\leq t_n \\ (k,l)\in\Gamma_n}} \Sigma_{n,kl}.$$

On the other hand, for any (i',j') outside $B_{n,ij}$, the indicator rv $J'_{i'j'}(n;\rho)$ is determined by the Poisson point process in

$$L_{n,i'j'} = \cup_{\substack{|k-i'|\leq t_n \\ |l-j'|\leq t_n \\ (k,l)\in\Gamma_n}} \Sigma_{n,kl}.$$

Based on the definition of $B_{n,ij}$, it is clear that $L_{n,ij}$ and $L_{n,i'j'}$ are non-overlapping. Since nodes are placed according to a homogeneous Poisson point process, the point processes in $L_{n,ij}$ and $L_{n,i'j'}$ are mutually independent, whence the rvs $J'_{ij}(n;\rho)$ and $J'_{i'j'}(n;\rho)$ are mutually independent, and this establishes the desired result.

$\blacksquare$

Lemma 8.3 *Assume user A with coordinates (x_A, y_A) is the only user in $\Sigma_{n,ij}$, and user B is uniformly distributed in $[0,1]^2 - \Sigma_{n,ij}$. Under these assumptions, if $T_{n,ij}$ be the indicator rv that users A and B are not connected, then*

$$\mathbb{E}[T_{n,ij}] \begin{cases} \in (1 - s_\rho, 1 - s_\rho + \ell_n^2] & \text{if} \quad (i,j) \in R_1 \\ \\ \le 1 - 0.25 s_\rho + \ell_n^2 & \text{if} \quad (i,j) \in R_2 \\ \\ \le 1 - 0.5 s_\rho - \frac{(w_{n,ij}-1)\ell_n}{\rho\pi} s_\rho + \ell_n^2 & \text{if} \quad (i,j) \in R_3 \end{cases} \tag{8.5}$$

where $w_{n,ij} = \min(i, j, m_n - i, m_n - j)$.

Lemma 8.3 is proved in Appendix D. Note that (8.5) holds irrespective of the specific coordinates of user A.

Lemma 8.4 *Fix $k = 1, 2, \ldots$. Assume user A and users $B_1, \ldots, B_k$ are independently and uniformly distributed on $[0,1]^2$. We construct a graph on these $k+1$ users based on the BCG model with parameter ρ. If $p_{iso,k}(\rho)$ denotes the probability that user A is isolated, then the bounds*

$$\begin{aligned} \left(1 - s_\rho\right)^k \;\le\; p_{iso,k}(\rho) &\le \frac{(m_n - 2t_n)^2}{m_n^2}\left(1 - s_\rho\right)^k + \frac{4t_n^2}{m_n^2}\left(1 - 0.25 s_\rho\right)^k \\ &+ \frac{4(m_n - 2t_n)}{m_n^2} \sum_{w_{n,ij}=1}^{t_n} \left(1 - 0.5 s_\rho - \frac{(w_{n,ij}-1)\ell_n}{\rho\pi} s_\rho\right)^k \end{aligned} \tag{8.6}$$

hold, and the expected degree $D_k(\rho)$ of user A satisfies

$$k(1-2\rho)^2 s_\rho < D_k(\rho) < k s_\rho. \tag{8.7}$$

Proof. Conditioning on the event $E_{n,ij}$ that user A is located in the squarelet $\Sigma_{n,ij}$, we have

$$\begin{aligned}
p_{iso,k}(\rho) &= \mathbb{E}\Big[\prod_{i=1}^{k}\Big(1-h_\rho\big(D(A,B_i)\big)\Big)\Big] \\
&= \sum_{(i,j)\in\Gamma_n}\mathbb{E}\Big[\prod_{i=1}^{k}\Big(1-h_\rho\big(D(A,B_i)\big)\Big)\Big|E_{n,ij}\Big]\mathbb{P}[E_{n,ij}] \\
&= \sum_{(i,j)\in\mathbf{R}_1}\mathbb{E}\Big[\prod_{i=1}^{k}\Big(1-h_\rho\big(D(A,B_i)\big)\Big)\Big|E_{n,ij}\Big]\mathbb{P}[E_{n,ij}] \\
&\quad+\sum_{(i,j)\in\mathbf{R}_2}\mathbb{E}\Big[\prod_{i=1}^{k}\Big(1-h_\rho\big(D(A,B_i)\big)\Big)\Big|E_{n,ij}\Big]\mathbb{P}[E_{n,ij}] \\
&\quad+\sum_{(i,j)\in\mathbf{R}_3}\mathbb{E}\Big[\prod_{i=1}^{k}\Big(1-h_\rho\big(D(A,B_i)\big)\Big)\Big|E_{n,ij}\Big]\mathbb{P}[E_{n,ij}].
\end{aligned}$$

Then it can be shown that (8.6) is a simple corollary from Lemma 8.3.

Similarly, (8.7) follows because Conditioning on the event $C(\rho)$ that user A is located in $[\rho,1-\rho]\times[\rho,1-\rho]$, we have

$$\begin{aligned}
D_k &= k\mathbb{E}\Big[h_\rho\big(D(A,B_1)\big)\Big] \\
&= k\mathbb{E}\Big[h_\rho\big(D(A,B_1)\big)\Big|C(\rho)\Big]\mathbb{P}[C(\rho)] + k\mathbb{E}\Big[h_\rho\big(D(A,B_1)\big)\Big|\overline{C}(\rho)\Big](1-\mathbb{P}[C(\rho)]),
\end{aligned}$$

and

$$\mathbb{E}\Big[h_\rho\big(D(A,B_1)\big)\Big|\overline{C}(\rho)\Big] \le \mathbb{E}\Big[h_\rho\big(D(A,B_1)\big)\Big|C(\rho)\Big] = s_\rho.$$

$\blacksquare$

Lemma 8.5 *Assume user A with coordinates (x_A, y_A) is the only user in Σ_{n,i_1j_1}, and user B with coordinates (x_B, y_B) is the only user in Σ_{n,i_2j_2} with $(i_2, j_2) \neq (i_1, j_1)$. Assume (i_2, j_2) belongs to B_{n,i_1,j_1}, the neighborhood of dependence for (i_1, j_1). Assume user C is uniformly distributed in $[0,1]^2 - \Sigma_{n,i_1j_1} - \Sigma_{n,i_2j_2}$. Let $T_{n,i_1j_1i_2j_2}$ be the indicator rv that neither user A nor user B is connected with user C. We have*

$$\mathbb{E}[T_{n,i_1j_1i_2j_2}] \leq \begin{cases} 1-(2-\kappa_\rho)s_\rho + 4\ell_n^2 & \text{if} \quad (i_1,j_1) \in R_4 \\ \\ 1-0.25s_\rho + 2\ell_n^2 & \text{if} \quad (i_1,j_1) \in R_5 \\ \\ 1-(1-0.5\kappa_\rho)s_\rho + 4\ell_n^2 & \text{if} \quad (i_1,j_1) \in R_6. \end{cases} \tag{8.8}$$

Lemma 8.5 is established in Appendix E. Note that (8.8) holds irrespective of the specific coordinates of users A and B.

8.3 The main results

First a word on the notation used in this Section: We write $a_n \sim b_n$ to indicate a_n is asymptotically equal to b_n, i.e.

$$\limsup_{n\to\infty} \frac{a_n}{b_n} = 1.$$

Throughout this Section, we assume that the boundary function $\rho : \mathbb{N}_0 \to \mathbb{R}_+$ is chosen such that the function $s_\rho : \mathbb{N}_0 \to \mathbb{R}_+$ is of the form

$$s_{\rho_n} = \frac{\log n + \alpha + o(1)}{n}$$

with α in $\mathbb{R}$. We first establish some asymptotic equivalences to facilitate the proof of our main results.

1) According to the definition of h_ρ in (2.17), $h_\rho(r) > 0$ when $r < \rho$ and $h_\rho(r) < 1$ when $r > 0$. Also, $h_\rho(r)$ being non-increasing with r, we get

$$s_{\rho_n} = 2\pi \int_0^{\rho_n} h_{\rho_n}(r) r\, dr > 2\pi \int_0^{0.5\rho_n} h_{\rho_n}(0.5\rho_n) r\, dr = \frac{h_{\rho_n}(0.5\rho_n)}{4} \pi \rho_n^2 = \Theta(\rho_n^2),$$

and

$$s_{\rho_n} = 2\pi \int_0^{\rho_n} h_{\rho_n}(r) r\, dr < 2\pi \int_0^{\rho_n} r\, dr = \pi \rho_n^2 = \Theta(\rho_n^2).$$

It follows that

$$\rho_n = \Theta(\sqrt{s_{\rho_n}}) = \Theta\left(\sqrt{\frac{\log n}{n}}\right).$$

2) We immediately have

$$t_n = \lceil \frac{\rho_n}{\ell_n} \rceil = \Theta(\sqrt{n^3 \log n}).$$

3) The number of squarelets in R_1 is $(m_n - 2t_n)^2 = \Theta(n^4)$. The number of squarelets in R_2 is $4t_n^2 = \Theta(n^3 \log n)$. The number of squarelets in R_3 is $4(m_n - 2t_n)t_n = \Theta(n^3\sqrt{n \log n})$.

4) Similarly, the number of squarelets in R_4, R_5 and R_6 are $\Theta(n^4)$, $\Theta(n^3 \log n)$ and $\Theta(n^3\sqrt{n \log n})$, respectively.

5) For any (i,j) in Γ_n, the cardinality of its neighborhood of dependence $B_{n,ij}$ is $\Theta(t_n^2) = \Theta(n^3 \log n)$.

Proposition 8.1 *Under the enforces assumptions, it holds that*

$$\lim_{n\to\infty} P_{iso,2,P}(n;\rho_n) = e^{-e^{-\alpha}} \tag{8.9}$$

Proposition 8.1 is established in Section 8.4 under the assumption that network nodes are placed according to a homogeneous Poisson point process in $[0,1]^2$ with density n. Next, we prove a similar result assuming that the n network nodes are uniformly and independently distributed in $[0,1]^2$.

Theorem 8.1 *Under the enforced assumptions, it holds that*

$$\lim_{n\to\infty} P_{iso,2}(n;\rho_n) = e^{-e^{-\alpha}}. \tag{8.10}$$

Proof. Let M_n be the number of nodes located in $[0,1]^2$. Since nodes follow Poisson point process with parameter n, M_n is a Poisson random variable with parameter n. Conditioning on M_n, we have

$$P_{iso,2,P}(n;\rho_n) = \sum_{k=0}^{\infty} P_{iso,2}(k;\rho_n)\mathbb{P}[M_n = k], \tag{8.11}$$

and Chebyshev's inequality thus yields

$$\mathbb{P}[|M_n - n| \geq \sqrt{n}\log n] \leq \frac{n}{n\log^2 n} = \frac{1}{\log^2 n}.$$

According to Proposition 8.1, with $s_{\rho_n} = \frac{\log n+\alpha+o(1)}{n}$, we have $\lim_{n\to\infty} P_{iso,2,P}(n;\rho_n) = e^{-e^{-\alpha}}$, which is strictly larger than zero. Thus

$$\begin{aligned} e^{-e^{-\alpha}} = \lim_{n\to\infty} P_{iso,2,P}(n;\rho_n) &= \lim_{n\to\infty} \sum_{k=0}^{\infty} P_{iso,2}(k;\rho_n)\mathbb{P}[M_n = k] \\ &= \lim_{n\to\infty} \sum_{k\in A_n} P_{iso,2}(k;\rho_n)\mathbb{P}[M_n = k] \qquad (8.12) \end{aligned}$$

where $A_n = \{0, 1, \ldots, : n - \sqrt{n}\log n \leq k \leq n + \sqrt{n}\log n\}$.

The basic idea in the following proof is to show that for different k belong to A_n, $P_{iso,2}(k;\rho_n)$ almost remains unchanged as n tends to infinity. This fact will immediately lead to our desired result.

Consider the scenario where users numbered $1, 2, \ldots,$ are uniformly and independently deployed in $[0, 1]^2$. Fix $n = 2, 3, \ldots$. The probability of connectivity between a pair of users is computed by the BCG model with parameter ρ_n. Let k be an integer, and denote by $Z_{n,k}$ the number of isolated users in the graph formed by the first k users. Clearly $\mathbb{P}[Z_k = 0] = P_{iso,2}(k;\rho_n)$. With $i = 1, \ldots, k$, we define

$$I_{i,k} \quad := \quad \mathbf{1}[\text{The } i^{th} \text{ user is isolated in the graph formed by the first } k \text{ users}].$$

For any k in A_n, we have

$$\begin{aligned} P_{iso,2}(k+1;\rho_n) &= P_{iso,2}(k;\rho_n)\Big(1 - \mathbb{P}[I_{k+1,k+1} = 1|Z_k = 0]\Big) \\ &\quad + \Big(1 - P_{iso,2}(k;\rho_n)\Big)\mathbb{P}[Z_{k+1} = 0|Z_k > 0]. \qquad (8.13) \end{aligned}$$

The conditional event $[Z_{k+1} = 0 | Z_k > 0]$ indicates that the $(k+1)^{rst}$ user has to connect to at least one isolated user among the first k users. The probability of this event is at most $s_{\rho_n} = \Theta(\frac{\log n}{n})$.

Moreover,

$$\begin{aligned}
\mathbb{P}[I_{k+1,k+1} = 1 | Z_k = 0] &= \frac{\mathbb{P}[I_{k+1,k+1} = 1, Z_k = 0]}{\mathbb{P}[Z_k = 0]} \\
&< \frac{\mathbb{P}[I_{k+1,k+1} = 1]}{\mathbb{P}[Z_k = 0]} \\
&= \frac{\mathbb{P}[I_{1,k+1} = 1]}{P_{iso,2}(k; \rho_n)}.
\end{aligned}$$

From (8.6), it can be shown that for any k belongs to A_n,

$$\mathbb{P}[I_{1,k+1} = 1] = \Theta\left(\frac{1}{n}\right),$$

because

$$\mathbb{P}[I_{1,k+1} = 1] \geq (1 - s_{\rho_n})^k \geq (1 - s_{\rho_n})^{n+\sqrt{n}\log n} = \Theta\left(\frac{1}{n}\right),$$

and

$$\begin{aligned}
\mathbb{P}[I_{1,k+1} = 1] \leq\ & \frac{(m_n - 2t_n)^2}{m_n^2}\Big(1 - s_{\rho_n}\Big)^k + \frac{4t_n^2}{m_n^2}\Big(1 - 0.25 s_{\rho_n}\Big)^k \\
& + \frac{4(m_n - 2t_n)}{m_n^2} \sum_{w_{n,ij}=1}^{t_n} \left(1 - 0.5 s_{\rho_n} - \frac{(w_{n,ij} - 1)\ell_n}{\rho_n \pi} s_{\rho_n}\right)^k \\
\leq\ & \frac{(m_n - 2t_n)^2}{m_n^2}\Big(1 - s_{\rho_n}\Big)^{n-\sqrt{n}\log n} + \frac{4t_n^2}{m_n^2}\Big(1 - 0.25 s_{\rho_n}\Big)^{n-\sqrt{n}\log n} \\
& + \frac{4(m_n - 2t_n)}{m_n^2} \sum_{w_{n,ij}=1}^{t_n} \left(1 - 0.5 s_{\rho_n} - \frac{(w_{n,ij} - 1)\ell_n}{\rho_n \pi} s_{\rho_n}\right)^{n-\sqrt{n}\log n}
\end{aligned}$$

where

$$\frac{(m_n - 2t_n)^2}{m_n^2}\Big(1 - s_{\rho_n}\Big)^{n-\sqrt{n}\log n} = \Theta\left(\frac{1}{n}\right), \quad \frac{4t_n^2}{m_n^2}\Big(1 - 0.25 s_{\rho_n}\Big)^{n-\sqrt{n}\log n} = o\left(\frac{1}{n}\right)$$

and

$$
\begin{aligned}
& \frac{4(m_n-2t_n)}{m_n^2}\sum_{w_{n,ij}=1}^{t_n}\left(1-0.5s_{\rho_n}-\frac{(w_{n,ij}-1)\ell_n}{\rho_n\pi}s_{\rho_n}\right)^{n-\sqrt{n}\log n} \\
< \quad & \frac{4(m_n-2t_n)}{m_n^2}e^{-0.5s_{\rho_n}(n-\sqrt{n}\log n)}\sum_{w_{n,ij}=1}^{t_n}e^{-\frac{(n-\sqrt{n}\log n)(w_{n,ij}-1)\ell_n}{\rho_n\pi}s_{\rho_n}}=o\left(\frac{1}{n}\right).
\end{aligned} \tag{8.14}
$$

The expression (8.14) follows from the following fact: With $f_n \sim n$,

$$
\begin{aligned}
\ell_n\sum_{w_{n,ij}=1}^{t_n}e^{-\frac{f_n(w_{n,ij}-1)\ell_n}{\rho_n\pi}s_{\rho_n}} &< \int_0^{\rho_n}e^{-\frac{f_ns_{\rho_n}}{\rho_n\pi}x}\,dx+\ell_n \\
&= \frac{\rho_n\pi}{f_ns_{\rho_n}}(1-e^{-\frac{f_ns_{\rho_n}}{\pi}})+\ell_n \\
&= \frac{\rho_n\pi}{f_ns_{\rho_n}}-\Theta\left(\sqrt{\frac{1}{n\log n}}\frac{1}{n^{\frac{1}{\pi}}}\right)+\frac{1}{n^2} \\
&= O\left(\frac{1}{\sqrt{n\log n}}\right).
\end{aligned} \tag{8.15}
$$

Thus from (8.13),

$$
|P_{iso,2}(k+1;\rho_n)-P_{iso,2}(k;\rho_n)|<\mathbb{P}[I_{1,k+1}=1]+\mathbb{P}[Z_{k+1}=0|Z_k>0]=O\left(\frac{\log n}{n}\right).
$$

Since the cardinality of A_n is $\Theta(\sqrt{n}\log n)$, it follows that

$$
\max_{k\in A_n}P_{iso,2}(k;\rho_n)-\min_{k\in A_n}P_{iso,2}(k;\rho_n)=O\left(\frac{\log^2 n}{\sqrt{n}}\right),
$$

and

$$
\lim_{n\to\infty}\left(\max_{k\in A_n}P_{iso,2}(k;\rho_n)-\min_{k\in A_n}P_{iso,2}(k;\rho_n)\right)\sum_{k\in A_n}\mathbb{P}[M_n=k]=0. \tag{8.16}
$$

Finally for any given n, we have

$$
\max_{k\in A_n}P_{iso,2}(k;\rho_n)\sum_{k\in A_n}\mathbb{P}[M_n=k]\geq\sum_{k\in A_n}P_{iso,2}(k;\rho_n)\mathbb{P}[M_n=k], \tag{8.17}
$$

and

$$\min_{k\in A_n} P_{iso,2}(k;\rho_n)\sum_{k\in A_n}\mathbb{P}[M_n=k] \le \sum_{k\in A_n} P_{iso,2}(k;\rho_n)\mathbb{P}[M_n=k]. \tag{8.18}$$

According to (8.16), (8.17) (8.18) and (8.12), we get

$$\begin{aligned} e^{-e^{-\alpha}} &= \lim_{n\to\infty}\max_{k\in A_n} P_{iso,2}(k+1;\rho_n)\sum_{k\in A_n}\mathbb{P}[M_n=k] \\ &= \lim_{n\to\infty}\min_{k\in A_n} P_{iso,2}(k;\rho_n)\sum_{k\in A_n}\mathbb{P}[M_n=k] \end{aligned}$$

Thus for any k in A_n, we have

$$\lim_{n\to\infty} P_{iso,2}(k;\rho_n)\sum_{k\in A_n}\mathbb{P}[M_n=k] = e^{-e^{-\alpha}}, \tag{8.19}$$

and taking $k=n-1$, we conclude that

$$\lim_{n\to\infty} P_{iso,2}(n;\rho_n) = e^{-e^{-\alpha}}. \tag{8.20}$$

■

8.4 A proof of Proposition 8.1

First a word on the notation used in this Section: We write $a_n \lesssim b_n$ to indicate a_n is asymptotically smaller than b_n, i.e.

$$\limsup_{n\to\infty}\frac{a_n}{b_n}\le 1.$$

The main effort in establishing Proposition 8.1 consists in using Corollary 3.3 in order to show that the distribution of

$$C'(n;\rho_n) = \sum_{(i,j)\in\Gamma_n} J'_{ij}(n;\rho_n)$$

converges to a Poisson distribution with parameter $e^{-\alpha}$. Our proof is composed of two steps: First, we prove that

$$\lim_{n\to\infty} \mathbb{E}[C'(n;\rho_n)] = e^{-\alpha}. \tag{8.21}$$

Next we will show that the upper bound in (3.14) converges to zero, i.e., we prove that

$$\lim_{n\to\infty} \sum_{(i_1,j_1)\in\Gamma_n} \sum_{(i_2,j_2)\in B_{i1j1}} \mathbb{E}[J'_{i1j1}(n;\rho_n)]\mathbb{E}[J'_{i2j2}(n;\rho_n)] = 0, \tag{8.22}$$

and

$$\lim_{n\to\infty} \sum_{(i_1,j_1)\in\Gamma_n} \sum_{(i_2,j_2)\in B_{i1j1}\setminus(i_1,j_1)} \mathbb{E}[J'_{i1j1}(n;\rho_n)J'_{i2j2}(n;\rho_n)] = 0. \tag{8.23}$$

Step 1 In order to prove (8.21), we need to evaluate $\mathbb{E}[J'_{ij}(n;\rho_n)]$, $i,j = 1,\ldots,m_n$. Note that $J'_{ij}(n;\rho_n) = 1$ if and only if $J_{ij}(n;\rho_n) = 1$ and $N_{ij}(n) = 1$, which indicates that there is only one node in $\Sigma_{n,ij}$ and it is isolated. Thus

$$\begin{aligned}\mathbb{E}[J'_{ij}(n;\rho_n)] &= \mathbb{E}[J_{ij}(n;\rho_n)N_{ij}(n)] \\ &= \mathbb{P}[\text{The squarelet } \Sigma_{n,ij} \text{ only contains one isolated node }].\end{aligned}$$

Conditioning on the coordinates (X,Y) of this only node in $\Sigma_{n,ij}$ and the number of nodes $K_{ij}(n)$ outside $\Sigma_{n,ij}$, we have

$$\begin{aligned}
& \mathbb{E}[J'_{ij}(n;\rho_n)] \\
= \; & \mathbb{E}[J_{ij}(n;\rho_n)N_{ij}(n)] \\
= \; & \mathbb{E}\big[\mathbb{E}[J_{ij}(n;\rho_n)N_{ij}(n)|X,Y,K_{ij}(n)]\big] \\
= \; & \mathbb{P}[N_{ij}(n)=1]\mathbb{E}\big[\mathbb{E}[J'_{ij}(n;\rho_n)|N_{ij}(n)=1,X,Y,K_{ij}(n)]\big] \\
= \; & \mathbb{P}[N_{ij}(n)=1]\sum_{k_{ij}=0}^{\infty}\Big(\mathbb{P}[K_{ij}(n)=k_{ij}] * \\
& \mathbb{E}\big[\mathbb{E}[J'_{ij}(n;\rho_n)|N_{ij}(n)=1,X,Y,K_{ij}(n)=k_{ij}]\big]\Big) \qquad (8.24)
\end{aligned}$$

where

$$\mathbb{P}[N_{ij}(n)=1] = n\ell_n^2 e^{-n\ell_n^2},$$

and

$$\mathbb{P}[K_{ij}(n)=k_{ij}] = \frac{e^{-n(1-\ell_n^2)}(n(1-\ell_n^2))^{k_{ij}}}{k_{ij}!}.$$

Since nodes are placed according to a homogeneous Poisson point process, given the number of nodes k_{ij} located outside $\Sigma_{n,ij}$, these nodes are uniformly and independently distributed in $[0,1]^2 - \Sigma_{n,ij}$. Thus for any given coordinates (x,y) in $\Sigma_{n,ij}$, it is not difficult to see that

$$\begin{aligned}
& \mathbb{E}[J'_{ij}(n;\rho_n)|N_{ij}(n)=1,X=x,Y=y,K_{ij}(n)=k_{ij}] \\
= \; & \Big(\mathbb{E}[J'_{ij}(n;\rho_n)|N_{ij}(n)=1,X=x,Y=y,K_{ij}(n)=1]\Big)^{k_{ij}}. \qquad (8.25)
\end{aligned}$$

Moreover,

$$\mathbb{E}[J'_{ij}(n;\rho_n)|N_{ij}(n)=1,X=x,Y=y,K_{ij}(n)=1] = \mathbb{E}[T_{n,ij}]$$

introduced in Lemma 8.3. Thus for any (x, y) in $\Sigma_{n,ij}$,

$$\mathbb{E}[J'_{ij}(n;\rho_n)|N_{ij}(n)=1, X=x, Y=y, K_{ij}(n)=1] = \mathbb{E}[T_{n,ij}]$$

$$\begin{cases} \in (1-s_{\rho_n}, 1-s_{\rho_n}+\ell_n^2] & \text{if} \quad (i,j)\in R_1 \\ \\ \leq 1-0.25s_{\rho_n}+\ell_n^2 & \text{if} \quad (i,j)\in R_2 \\ \\ \leq 1-0.5s_{\rho_n}-\frac{(w_{n,ij}-1)\ell_n}{\rho_n\pi}s_{\rho_n}+\ell_n^2 & \text{if} \quad (i,j)\in R_3. \end{cases}$$

Since the above inequalities are independent of the specific coordinates (x, y), we have

$$\mathbb{E}[\mathbb{E}[J'_{ij}(n;\rho_n)|N_{ij}(n)=1, X, Y, K_{ij}(n)=1]]$$

$$\begin{cases} \in (1-s_{\rho_n}, 1-s_{\rho_n}+\ell_n^2] & \text{if} \quad (i,j)\in R_1 \\ \\ \leq 1-0.25s_{\rho_n}+\ell_n^2 & \text{if} \quad (i,j)\in R_2 \\ \\ \leq 1-0.5s_{\rho_n}-\frac{(w_{n,ij}-1)\ell_n}{\rho_n\pi}s_{\rho_n}+\ell_n^2 & \text{if} \quad (i,j)\in R_3. \end{cases} \tag{8.26}$$

According to (8.24), (8.25) and (8.26), we can get

$$\mathbb{E}[J'_{ij}(n;\rho_n)] \begin{cases} \sim \frac{e^{-\alpha}}{n^4} & \text{if} \quad (i,j) \in R_1 \\ \\ \lesssim \frac{e^{-0.25\alpha}}{n^{3.25}} & \text{if} \quad (i,j) \in R_2 \\ \\ \lesssim n^{-3} e^{\frac{-ns_{\rho_n}}{2}} e^{-\frac{ns_{\rho_n}(w_{n,ij}-1)\ell_n}{\rho_n \pi}} & \text{if} \quad (i,j) \in R_3. \end{cases} \tag{8.27}$$

We immediately have

$$\lim_{n\to\infty} \sum_{(i,j)\in R_1} \mathbb{E}[J'_{ij}(n;\rho_n)] = \lim_{n\to\infty} (m_n - 2t_n)^2 \frac{e^{-\alpha}}{n^4} = e^{-\alpha}, \tag{8.28}$$

and

$$\lim_{n\to\infty} \sum_{(i,j)\in R_2} \mathbb{E}[J'_{ij}(n;\rho_n)] = 0 \tag{8.29}$$

since

$$\sum_{(i,j)\in R_2} \mathbb{E}[J'_{ij}(n;\rho_n)] \lesssim \Theta\left(n^3 \log n\right) \frac{e^{-0.25\alpha}}{n^{3.25}}.$$

It is little bit tricky to evaluate $\sum_{(i,j)\in R_3} \mathbb{E}[J'_{ij}(n;\rho_n)]$. We divide R_3 into $4(m_n - 2t_n)$ groups. Each of them contains t_n squarelets whose $w_{n,ij}$ ranging from 1 to t_n. According to (8.27),

$$\begin{aligned} \sum_{(i,j)\in R_3} \mathbb{E}[J'_{ij}(n;\rho_n)] &\lesssim 4(m_n - 2t_n) n^{-1} e^{-n\frac{s_{\rho_n}}{2}} * \ell_n \sum_{w_{n,ij}=1}^{t_n} e^{-\frac{n(w_{n,ij}-1)\ell_n}{\rho_n \pi} s_{\rho_n}} \\ &< 4n e^{-n\frac{s_{\rho_n}}{2}} \frac{\rho_n \pi}{n s_{\rho_n}} \\ &= \Theta\left(\frac{1}{\sqrt{n}}\right) \Theta\left(\sqrt{\frac{n}{\log n}}\right) \\ &= \Theta\left(\frac{1}{\sqrt{\log n}}\right) \end{aligned} \tag{8.30}$$

where the second inequality holds due to (8.15).

Finally

$$\lim_{n\to\infty} \sum_{(i,j)\in R_3} \mathbb{E}[J'_{ij}(n;\rho_n)] = 0 \tag{8.31}$$

From (8.28), (8.29) and (8.31), we prove that

$$\lim_{n\to\infty} \mathbb{E}[C'(n;\rho_n)] = e^{-\alpha}. \tag{8.32}$$

Step 2 First a word on the notation used in Step 2: To simplify expressions, we will write $J'_{i_1j_1}$ $J'_{i_2j_2}$, $N_{i_1j_1}$ and $N_{i_2j_2}$ instead of $J'_{i_1j_1}(n;\rho_n)$ $J'_{i_2j_2}(n;\rho_n)$, $N_{i_1j_1}(n)$ and $N_{i_2j_2}(n)$.

The main target in step 2 is to evaluate

$$\mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] \quad \text{with}(i_1,j_1)\in\Gamma_n \quad \text{and} \quad (i_2,j_2)\in B_{n,i_1j_1}$$

and

$$\mathbb{E}[J'_{i_1j_1}J'_{i_2j_2}] \quad \text{with}(i_1,j_1)\in\Gamma_n \quad \text{and} \quad (i_2,j_2)\in B_{n,i_1j_1}\backslash(i_1,j_1).$$

Note that $J'_{i_1j_1}J'_{i_2j_2} = 1$ if and only if $J_{i_1j_1}$, $J_{i_2j_2}$, $N_{i_1j_1}$ and $N_{i_2j_2}$ are all equal to 1, which indicates that Σ_{n,i_1j_1} and Σ_{n,i_2j_2} each contains only one node and both nodes are isolated. Thus

$$\begin{aligned}\mathbb{E}[J'_{i_1j_1}J'_{i_2j_2}] &= \mathbb{E}[J_{i_1j_1}N_{i_1j_1}J_{i_2j_2}N_{i_2j_2}] \\ &= \mathbb{P}[\text{Squarelets } \Sigma_{n,i_1j_1} \text{ and } \Sigma_{n,i_2j_2} \text{ each contains one isolate node}].\end{aligned}$$

Conditioning on the coordinates (X_1, Y_1) of the only node in Σ_{n,i_1j_1}, the coordinates (X_2, Y_2) of the only node in Σ_{n,i_2j_2} and the number of nodes $K_{i_1j_1i_2j_2}$ outside Σ_{n,i_1j_1} and Σ_{n,i_2j_2}, we have

$$\begin{aligned}
& \mathbb{E}[J'_{i_1j_1}J'_{i_2j_2}] \\
= \; & \mathbb{E}[J_{i_1j_1}J_{i_2j_2}N_{i_1j_1}N_{i_2j_2}] \\
= \; & \mathbb{E}[\mathbb{E}[J_{i_1j_1}J_{i_2j_2}N_{i_1j_1}N_{i_2j_2}|X_1, Y_1, X_2, Y_2, K_{i_1j_1i_2j_2}]] \\
= \; & \mathbb{P}[N_{i_1j_1} = 1]\mathbb{P}[N_{i_2j_2} = 1] \\
& \mathbb{E}\big[\mathbb{E}[J'_{i_1j_1}J'_{i_2j_2}|N_{i_1j_1} = 1, N_{i_2j_2} = 1, X_1, Y_1, X_2, Y_2, K_{i_1j_1i_2j_2}]\big] \\
= \; & \mathbb{P}[N_{i_1j_1} = 1]\mathbb{P}[N_{i_2j_2} = 1] \sum_{k_{i_1j_1i_2j_2}=0}^{\infty} \Big(\mathbb{P}[K_{i_1j_1i_2j_2} = k_{i_1j_1i_2j_2}] \\
& \mathbb{E}\big[\mathbb{E}[J'_{i_1j_1}J'_{i_2j_2}|N_{i_1j_1} = 1, N_{i_2j_2} = 1, X_1, Y_1, X_2, Y_2, K_{i_1j_1i_2j_2} = k_{i_1j_1i_2j_2}]\big]\Big),
\end{aligned} \tag{8.33}$$

where

$$\mathbb{P}[N_{i_1j_1} = 1] = \mathbb{P}[N_{i_2j_2} = 1] = n\ell_n^2 e^{-n\ell_n^2}$$

and

$$\mathbb{P}[K_{i_1j_1i_2j_2} = k_{i_1j_1i_2j_2}] = \frac{e^{-n(1-2\ell_n^2)}(n(1-2\ell_n^2))^{k_{i_1j_1i_2j_2}}}{k_{i_1j_1i_2j_2}!}.$$

Since nodes are placed according to a homogeneous Poisson point process, given the number of nodes $k_{i_1j_1i_2j_2}$ located outside Σ_{n,i_1j_1} and Σ_{n,i_2j_2}, these nodes are uniformly and independently distributed in $[0, 1]^2 - \Sigma_{n,i_1j_1} - \Sigma_{n,i_2j_2}$.

Thus for any given coordinates (x_1, y_1) and (x_2, y_2) in Σ_{n,i_1j_1} and Σ_{n,i_2j_12} respectively, it is not difficult to see that

$$\mathbb{E}[J'_{i_1j_1}J'_{i_2j_2}|N_{i_1j_1} = N_{i_2j_2} = 1, (X_1, Y_1, X_2, Y_2) = (x_1, y_1, x_2, y_2), K_{i_1j_1i_2j_2} = k_{i_1j_1i_2j_2}] =$$
$$\Big(\mathbb{E}[J'_{i_1j_1}J'_{i_2j_2}|N_{i_1j_1} = N_{i_2j_2} = 1, (X_1, Y_1, X_2, Y_2) = (x_1, y_1, x_2, y_2), K_{i_1j_1i_2j_2} = 1]\Big)^{k_{i_1j_1i_2j_2}}. \tag{8.34}$$

Moreover, according to Lemma 8.5, we have

$$\mathbb{E}[\mathbb{E}[J'_{i_1j_1}J'_{i_2j_2}|N_{i_1j_1} = 1, N_{i_2j_2} = 1, X_1, Y_1, X_2, Y_2, K_{i_1j_1i_2j_2} = 1]] = \mathbb{E}[T_{n,i_1j_1i_2j_2}]$$

$$< \begin{cases} 1 - (2 - \kappa_{\rho_n})s_{\rho_n} + 4\ell_n^2 & \text{if} \quad (i_1, j_1) \in R_4 \quad \text{and} \quad (i_2, j_2) \in B_{n,i_1j_1} \backslash (i_1, j_1) \\ 1 - 0.25s_{\rho_n} + 2\ell_n^2 & \text{if} \quad (i_1, j_1) \in R_5 \quad \text{and} \quad (i_2, j_2) \in B_{n,i_1j_1} \backslash (i_1, j_1) \\ 1 - (1 - 0.5\kappa_{\rho_n})s_{\rho_n} + 4\ell_n^2 & \text{if} \quad (i_1, j_1) \in R_6 \quad \text{and} \quad (i_2, j_2) \in B_{n,i_1j_1} \backslash (i_1, j_1). \end{cases} \tag{8.35}$$

According to (8.33), (8.34) and (8.35), we can get

$$\mathbb{E}[J'_{i_1j_1}J'_{i_2j_2}] = \begin{cases} O\left(n^{\kappa_{\rho_n} - 8}\right) & \text{if} \quad (i_1, j_1) \in R_4 \quad \text{and} \quad (i_2, j_2) \in B_{n,i_1j_1} \backslash (i_1, j_1) \\ O\left(n^{-6.25}\right) & \text{if} \quad (i_1, j_1) \in R_5 \quad \text{and} \quad (i_2, j_2) \in B_{n,i_1j_1} \backslash (i_1, j_1) \\ O\left(n^{0.5\kappa_{\rho_n} - 7}\right) & \text{if} \quad (i_1, j_1) \in R_6 \quad \text{and} \quad (i_2, j_2) \in B_{n,i_1j_1} \backslash (i_1, j_1). \end{cases} \tag{8.36}$$

According to the asymptotic equivalences established in Section 8.3, it is plain to obtain the following results:

$$\lim_{n\to\infty} \sum_{(i_1,j_1)\in R_4} \sum_{(i_2,j_2)\in B_{n,i_1j_1}\setminus(i_1,j_1)} \mathbb{E}[J'_{i_1j_1} J'_{i_2j_2}] = 0, \tag{8.37}$$

$$\lim_{n\to\infty} \sum_{(i_1,j_1)\in R_5} \sum_{(i_2,j_2)\in B_{n,i_1j_1}\setminus(i_1,j_1)} \mathbb{E}[J'_{i_1j_1} J'_{i_2j_2}] = 0, \tag{8.38}$$

and

$$\lim_{n\to\infty} \sum_{(i_1,j_1)\in R_6} \sum_{(i_2,j_2)\in B_{n,i_1j_1}\setminus(i_1,j_1)} \mathbb{E}[J'_{i_1j_1} J'_{i_2j_2}] = 0. \tag{8.39}$$

Next we are going to evaluate $\mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}]$ when (i_1, j_1) belongs to R_4, R_5 and R_6, respectively.

1) Assume that (i_1, j_1) is in R_4 and (i_2, j_2) is in B_{n,i_1j_1}.

   According to the definitions of R_4, R_1 and B_{n,i_1j_1}, both (i_1, j_1) and (i_2, j_2) belong to R_1, it is then clear that

$$\mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] = \Theta(n^{-8}).$$

   It follows that

$$\begin{aligned} \sum_{(i_1,j_1)\in R_4} \sum_{(i_2,j_2)\in B_{n,i_1j_1}} \mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] &< \Theta(n^4)\Theta(n^3 \log n)\Theta\left(n^{-8}\right) \\ &= \Theta\left(\frac{\log n}{n}\right). \end{aligned}$$

   Thus

$$\lim_{n\to\infty} \sum_{(i_1,j_1)\in R_4} \sum_{(i_2,j_2)\in B_{n,i_1j_1}} \mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] = 0. \tag{8.40}$$

2) Assume that (i_1, j_1) is in R_5 and (i_2, j_2) is in B_{n,i_1j_1}.

According to our analysis in Step 1, $\mathbb{E}[J'_{ij}] = O(n^{-3.25})$ for any (i,j) belongs to Γ_n, it is clear that

$$\mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] = O(n^{-6.5}).$$

It follows that

$$\begin{aligned}\sum_{(i_1,j_1)\in R_5}\sum_{(i_2,j_2)\in B_{n,i_1j_1}} \mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] &= \Theta(n^3\log n)\Theta(t_n^2)O\left(n^{-6.5}\right) \\ &= O\left(\frac{\log n}{\sqrt{n}}\right).\end{aligned}$$

Thus

$$\lim_{n\to\infty}\sum_{(i_1,j_1)\in R_5}\sum_{(i_2,j_2)\in B_{n,i_1j_1}} \mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] = 0. \tag{8.41}$$

3) Assume that (i_1, j_1) is in R_6 and (i_2, j_2) is in B_{n,i_1j_1}.

Since neither (i_1, j_1) nor (i_2, j_2) belongs to R_2, and according to our analysis in Step 1, $\mathbb{E}[J'_{ij}] = O(n^{-3.5})$ for any (i,j) not belong to R_2, it is then clear that

$$\mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] = O(n^{-7}).$$

It follows that

$$\begin{aligned}\sum_{(i_1,j_1)\in R_6}\sum_{(i_2,j_2)\in B_{n,i_1j_1}} \mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] &= \Theta(n^3\sqrt{n\log n})\Theta(n^3\log n)O\left(n^{-7}\right) \\ &= \Theta\left(\frac{\log n\sqrt{\log n}}{n^{1.5}}\right).\end{aligned}$$

Thus

$$\lim_{n\to\infty}\sum_{(i_1,j_1)\in R_6}\sum_{(i_2,j_2)\in B_{n,i_1j_1}} \mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] = 0. \tag{8.42}$$

Finally we have

$$\lim_{n\to\infty} \sum_{(i_1,j_1)\in\Gamma_n} \sum_{(i_2,j_2)\in B_{n,i_1j_1}\setminus(i_1,j_1)} \mathbb{E}[J'_{i_1j_1}J'_{i_2j_2}] = 0 \tag{8.43}$$

and

$$\lim_{n\to\infty} \sum_{(i_1,j_1)\in\Gamma_n} \sum_{(i_2,j_2)\in B_{n,i_1j_1}} \mathbb{E}[J'_{i_1j_1}]\mathbb{E}[J'_{i_2j_2}] = 0. \tag{8.44}$$

It is well-known [40, p. 58] that

$$d_{TV}(\Pi(\mathbb{E}[C'(n;\rho_n)]),\Pi(e^{-\alpha})) \leq |\mathbb{E}[C'(n;\rho_n)] - e^{-\alpha}|$$

Thus, according to (8.32) obtained in Step 1, we conclude that

$$\lim_{n\to\infty} d_{TV}(\Pi(\mathbb{E}[C'(n;\rho_n)]),\Pi(e^{-\alpha})) = 0.$$

Moreover, by Corollary 3.3, we find

$$\lim_{n\to\infty} d_{TV}(C'(n;\rho_n),\Pi(\mathbb{E}[C'(n;\rho_n)])) = 0$$

upon using (8.43) and (8.44) obtained in Step 2, and (8.32) obtained in Step 1. Finally the triangular inequality yields

$$\lim_{n\to\infty} d_{TV}(C'(n;\rho_n),\Pi(e^{-\alpha})) = 0.$$

Thus, $C'(n;\rho_n)$ converges to a Poisson distribution with parameter $e^{-\alpha}$, and we have

$$\lim_{n\to\infty} \mathbb{P}[C'(n;\rho_n) = 0] = e^{-e^{-\alpha}}.$$

Finally, the desired result (8.9) follows upon combining (8.3), (8.4) and the two inequalities

$$\mathbb{P}[C(n;\rho_n) = 0] \leq \mathbb{P}[C'(n;\rho_n) = 0]$$

and

$$\mathbb{P}[C'(n;\rho_n)=0] \leq \mathbb{P}[C(n;\rho_n)=0] + \mathbb{P}[C''(n;\rho_n)=0].$$

8.5 Discussion

According to (8.7) established in Lemma 8.4, for each node in $\mathbb{G}_2(n;\rho_n)$, its expected degree D_n satisfies the following inequality

$$(n-1)(1-2\rho_n)^2 s_{\rho_n} < D_n < (n-1)s_{\rho_n}.$$

It is then plain that $D_n = \log n + \alpha + o(1)$ if and only if $ns_{\rho_n} = \log n + \alpha + o(1)$. Thus we establish the following result

Corollary 8.1 *The boundary function $\rho : \mathbb{N}_0 \to \mathbb{R}_+$ is selected such that the expected node degree function $D : \mathbb{N}_0 \to \mathbb{R}_+$ admits a form*

$$D_n = \log n + \alpha + o(1)$$

with α in $\mathbb{R}$, then it holds that

$$\lim_{n\to\infty} P_{iso,2}(n;\rho_n) = e^{-e^{-\alpha}}. \tag{8.45}$$

We learn from Corollary 8.1 that $\log n$ is a critical scaling for the expected node degree: $\mathbb{G}_2(n;\rho_n)$ is very unlikely (resp. likely) to contain isolated nodes if D_n is suitably larger (resp. smaller) than $\log n$. Indeed, if the link probability function h_ρ has bounded support, we really do not need to care about its specific form, the only useful information is the expected node degree, which is related to h_ρ through (8.1).

Please note that we only estimate $P_{iso,2}(n;\rho_n)$ rather than $P_{con,2}(n;\rho_n)$ in this Chapter. We conjecture the following asymptotic equivalence

$$\lim_{n\to\infty} P_{iso,2}(n;\rho_n) = \lim_{n\to\infty} P_{con,2}(n;\rho_n).$$

This asymptotic equivalence suggests that when there are a large number of nodes in the network, the absence of isolated nodes is not only a necessary condition, but also an almost sufficient condition for network connectivity. Similar results exist for the Erdös Rényi graph (ERG) model and the geometric random graph (GRG) model.

Appendix A

A proof of Proposition 6.4

We begin with some easy bounds to be used repeatedly in the proofs. With $0 \leq x < 1$, it is a simple matter to check that

$$\log(1-x) = -\int_0^x \frac{1}{1-t}dt = -x - \Psi(x) \tag{A.1}$$

where we have set

$$\Psi(x) := \int_0^x \frac{t}{1-t}dt, \quad 0 \leq x < 1.$$

The mapping $x \to \Psi(x)$ is increasing and convex on the interval $[0,1)$ with

$$0 < \Psi(x) \leq \frac{x^2}{2(1-x)}, \quad 0 \leq x < 1. \tag{A.2}$$

The standard bound

$$1 - x \leq e^{-x}, \quad x \in [0,1] \tag{A.3}$$

is now a simple consequence of the decomposition (A.1) and of the non-negativity of Ψ.

A proof of (6.30) – Fix $n = 1, 2, \ldots$ and τ in $(0,1)$. The bound (A.3) readily yields

$$\begin{aligned}\tilde{K}(n;\tau) &\leq \int_0^{1-\tau} ne^{-nb(x;\tau)}e^{b(x;\tau)}f(x)dx \\ &\leq f^\star e^{f^\star\tau}K(n;\tau). \end{aligned} \tag{A.4}$$

The assumption $\lim_{n\to\infty}\tau_n = 0$ on the range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ implies the existence of a positive integer $n_\star$ such that $\tau_n < 1$ for all $n \geq n_\star$. Reporting this fact into (A.4) (with τ replaced by τ_n for $n \geq n^\star$) gives the desired conclusion (6.30). ■

A proof of (6.32) – Fix $n = 1, 2, \ldots$ and τ in the unit interval $(0, 1)$. Since $0 \leq b(x;\tau) < 1$ for all x on the interval $(0, 1-\tau)$, we find from (6.27) that

$$\begin{aligned}\tilde{K}(n;\tau) &\geq n\int_0^{1-\tau}(1-b(x;\tau))^n f(x)dx \\ &\geq f_\star n\int_0^{1-\tau}(1-b(x;\tau))^n dx \\ &= f_\star n\int_0^{1-\tau} e^{-nb(x;\tau)}e^{-n\Psi(b(x;\tau))}dx\end{aligned}$$

with the help of the decomposition (A.1). Next, the bound (A.2) gives

$$\Psi(b(x;\tau)) \leq \frac{b(x;\tau)^2}{2\,(1-b(x;\tau))}, \quad x \in (0,1)$$

with $0 < b(x;\tau) \leq f^\star\tau$. Therefore, whenever $f^\star\tau < 1$, the uniform bound

$$\sup_{x\in(0,1)} \Psi(b(x;\tau)) \leq \frac{(f^\star\tau)^2}{2\,(1-f^\star\tau)} \tag{A.5}$$

holds.

Now pick a range function $\tau : \mathbb{N}_0 \to \mathbb{R}_+$ which satisfies (6.31) (hence also (6.29)). The latter convergence implies both $\tau_n < 1$ and $f^\star\tau_n < 1$ for large enough

n, while the former yields

$$\lim_{n\to\infty} \sup_{x\in(0,1)} \left(n\Psi(b(x;\tau_n)) \right) = 0$$

as we make use of (A.5) (where τ_n is substituted to τ). By continuity of the exponential mapping, for each ε in $(0,1)$, there exists a positive integer $n^\star(\varepsilon)$ such that

$$\inf_{x\in(0,1)} e^{-n\Psi(b(x;\tau_n))} \geq 1-\varepsilon, \quad n \geq n^\star(\varepsilon)$$

and the bound (6.32) follows. ∎

Appendix B

A proof of Lemma 6.4

Pick ε in the interval $(0, a)$. Under (6.2), there exists $\delta = \delta(\varepsilon) > 0$ such that

$$-\varepsilon|x - x_\star|^r \le h(x) \le \varepsilon|x - x_\star|^r$$

whenever $|x - x_\star| \le \delta$ in $[0, 1]$. On this range, the representation (6.1) yields

$$c + (a - \varepsilon)|x - x_\star|^r \le f(x) \le c + (a + \varepsilon)|x - x_\star|^r. \tag{B.1}$$

The minimum $x_\star$ being unique, it follows that

$$\inf\{f(x) : \ x \in [0, 1], \ |x - x_\star| \ge \delta\} = c + r \tag{B.2}$$

for some $r > 0$. Therefore, whenever $|x - x_\star| \ge \delta$ in $[0, 1]$,

$$\begin{aligned} f(x) &\ge c + r \\ &\ge c + r|x - x_\star|^r \end{aligned} \tag{B.3}$$

since $0 \le |x - x_\star| \le 1$ on that range. On the other hand, it is also the case that

$$\begin{aligned} f(x) &\le c + (f^\star - c) \\ &\le c + \frac{f^\star - c}{\delta^r} \cdot |x - x_\star|^r \end{aligned} \tag{B.4}$$

whenever $|x - x_\star| \geq \delta$ in $[0,1]$. The desired conclusion follows by combining (B.1) with (B.3) and (B.4), in which case we can take $a_- = \min(r, a - \varepsilon)$ and

$$a_+ = \max\left\{\frac{f^\star - c}{\delta^r}, a + \varepsilon\right\}.$$

■

Appendix C

A proof of Lemma C.1

Throughout this appendix let $p \geq 3$ denote a constant. The proof of Proposition 6.3 relies on our ability to determine the validity of the inequality

$$(1-(u+v))^p \leq ((1-u)(1-v))^{p+1} \tag{C.1}$$

on the range $0 \leq u, v \leq 1$ under the constraint $u+v \leq 1$. The next technical lemma provides a simple characterization of a large region where this inequality holds.

Lemma C.1 *Fix $p \geq 3$. For $0 \leq u, v \leq 1$ with $u+v \leq 1$, the inequality (C.1) holds provided*

$$\frac{2}{p} \leq \min(u, v). \tag{C.2}$$

Proof. Fix u, v in the interval $[0, 1]$ such that $u+v \leq 1$. If this pair satisfies (C.1), then it also satisfies

$$\frac{1}{(1-u)(1-v)} \leq \left(1 + \frac{uv}{1-(u+v)}\right)^p. \tag{C.3}$$

Since $1-(u+v) \leq (1-u)(1-v)$, we see that (C.1) holds if we can show that

$$\frac{1}{1-(u+v)} \leq 1 + p\frac{uv}{1-(u+v)} \tag{C.4}$$

as we make use of the standard inequality $(1+t)^p \geq 1 + pt$ valid for all $t \geq 0$. The inequality (C.4) is equivalent to

$$\frac{1-puv}{1-(u+v)} \leq 1, \tag{C.5}$$

which can be rewritten as

$$u + v \leq puv. \tag{C.6}$$

In short, the pair u, v satisfies (C.1) if

$$\frac{1}{u} + \frac{1}{v} \leq p. \tag{C.7}$$

This last inequality is clearly satisfied if we select u and v according to (C.2). ■

Appendix D

A proof of Lemma 8.3

Since the expectation of $T_{n,ij}$ is the probability that users A and B are not connected, we can write

$$\mathbb{E}[T_{n,ij}] = 1 - \iint_{S_{n,ij}} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B,$$

where $S_{n,ij} = \overline{D}_\rho(A) \cap \big([0,1]^2 - \Sigma_{ij}\big)$ with $\overline{D}_\rho(A)$ being a closed radius-ρ disk around user A. Note that Σ_{ij} is completely contained in $\overline{D}_\rho(A)$, thus $S_{n,ij} = \overline{D}_\rho(A) \cap [0,1]^2 - \Sigma_{ij}$. Moreover, based on the definition of s_ρ in (8.1), we can see that

$$s_\rho = \iint_{\overline{D}_\rho(A)} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B.$$

1) When (i,j) belongs to R_1, $\overline{D}_\rho(A)$ is completely contained in $[0,1]^2$, thus

$$S_{n,ij} = \overline{D}_\rho(A) \cap [0,1]^2 - \Sigma_{n,ij} = \overline{D}_\rho(A) - \Sigma_{n,ij}.$$

It then follows that

$$\begin{aligned}
\mathbb{E}[T_{n,ij}] &= 1 - \iint_{S_{n,ij}} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B \\
&= 1 - \iint_{\overline{D}_\rho(A)} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B + \iint_{\Sigma_{n,ij}} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B \\
&= 1 - s_\rho + \iint_{\Sigma_{n,ij}} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B.
\end{aligned}$$

Since

$$0 < \iint_{\Sigma_{n,ij}} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B < \ell_n^2,$$

it is plain that

$$1 - s_\rho < \mathbb{E}[T_{n,ij}] < 1 - s_\rho + \ell_n^2.$$

2) When (i,j) belongs to R_2, it is without loss of generality to only consider the case that (i,j) belongs to R_{21}. We have

$$\begin{aligned}
\mathbb{E}[T_{n,ij}] &= 1 - \iint_{S_{n,ij}} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B \\
&\leq 1 - \iint_{S_{n,ij} \cap \big([x_A,1]\times[y_A,1]\big)} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B \qquad \text{(D.1)}
\end{aligned}$$

The equality in (D.1) holds if and only if $x_A = y_A = 0$. The integration region $S_{n,ij} \cap \big([x_A,1]\times[y_A,1]\big)$ is displayed as the shaded region in Fig. D.1. According to Fig. D.1, we have

$$\iint_{S_{n,ij} \cap \big([x_A,1]\times[y_A,1]\big)} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B \geq 0.25 s_\rho - l^2,$$

where the equality holds if and only if $x_A = \frac{i-1}{m_n}$ and $y_A = \frac{j-1}{m_n}$.

Thus

$$\mathbb{E}[T_{n,ij}] \leq 1 - 0.25 s_\rho + \ell_n^2.$$

3) When (i,j) belongs to R_3, we similarly have

$$\begin{aligned}
\mathbb{E}[T_{n,ij}] &= 1 - \iint_{S_{n,ij}} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B \\
&\leq 1 - \iint_{S^*} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B, \qquad \text{(D.2)}
\end{aligned}$$

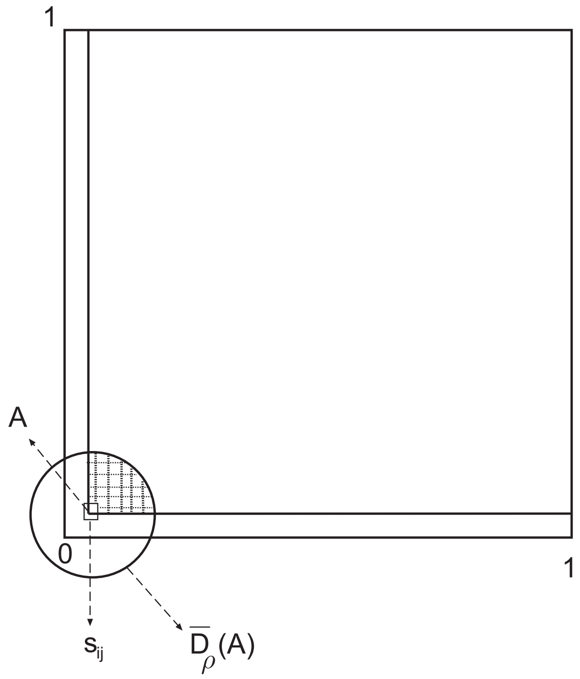


Figure D.1: Integration region when user A belongs to R_{21}.

where S^* is the shaded region displayed in Fig. D.2. In Fig. D.2, we denote the minimum distance of user A to the border of $[0,1]^2$ by d, which is equal

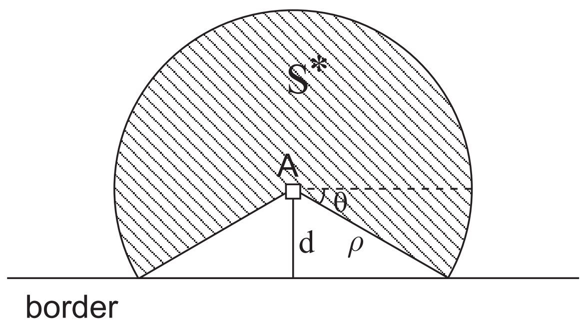


Figure D.2: Integration region when user A belongs to R_3.

to $\min(x_A, y_A, 1-x_A, 1-y_A)$. According to Fig. D.2, we have

$$\begin{aligned}
\mathbb{E}[T_{n,ij}] &\le 1 - \iint_{S^*} h_\rho\big(d(A,B)\big)\, dx_B\, dy_B \\
&< 1 - \frac{s_\rho}{2} - \frac{\theta}{\pi} s_\rho + \ell_n^2 \\
&= 1 - \frac{s_\rho}{2} - \frac{\arcsin(\frac{d}{\rho})}{\pi} s_\rho + \ell_n^2 \\
&\le 1 - \frac{s_\rho}{2} - \frac{d}{\rho\pi} s_\rho + \ell_n^2 \\
&\le 1 - \frac{s_\rho}{2} - \frac{(w_{n,ij}-1)l}{\rho\pi} s_\rho + \ell_n^2,
\end{aligned}$$

where $w_{n,ij} = \lceil \frac{d}{\ell_n} \rceil$. It is clear that w is determined by i, j and n, thus $w_{n,ij} = \min(i, j, m_n - i, m_n - j)$.

Appendix E

A proof of Lemma 8.5

It is plain that

$$\mathbb{E}[T_{n,i_1j_1i_2j_2}] = \iint_{\Sigma'_{n,i_1j_1i_2j_2}} \big(1 - h_\rho\big(d(A,C)\big)\big)\,\big(1 - h_\rho\big(d(B,C)\big)\big)\; dx_C\, dy_C,$$

where $\Sigma'_{n,i_1j_1i_2j_2} = [0,1]^2 - \Sigma_{n,i_1j_1} - \Sigma_{n,i_2j_2}$.

Since user C is uniformly distributed in $\Sigma'_{n,i_1j_1i_2j_2}$,

$$\begin{aligned} \mathbb{E}[T_{n,i_1j_1i_2j_2}] &= 1 - \iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big)\, dx_C\, dy_C \\ &\quad - \iint_{\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(B,C)\big)\, dx_C\, dy_C \\ &\quad + \iint_{\overline{D}_\rho(A)\cap\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big) h_\rho\big(d(B,C)\big)\, dx_C\, dy_C \end{aligned} \tag{E.1}$$

1) When (i_1, j_1) belongs to R_4, both $\overline{D}_\rho(A)$ and $\overline{D}_\rho(B)$ are completely contained in $[0,1]^2$, thus we have

$$\begin{aligned} \iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big)\, dx_C\, dy_C &= \iint_{\overline{D}_\rho(A)-\Sigma_{n,i_1j_1}-\Sigma_{n,i_2j_2}} h_\rho\big(d(A,C)\big)\, dx_C\, dy_C \\ &\geq s_\rho - 2\ell_n^2, \end{aligned}$$

and similarly

$$\iint_{\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(B,C)\big)\,dx_C\,dy_C \geq s_\rho - 2\ell_n^2.$$

Moreover,

$$\begin{aligned}
&\iint_{\overline{D}_\rho(A)\cap\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big)h_\rho\big(d(B,C)\big)\,dx_C\,dy_C \\
\leq\ & \max\Big\{\iint_{\overline{D}_\rho(A)\cap\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho^2\big(d(A,C)\big)\,dx_C\,dy_C, \\
&\quad \iint_{\overline{D}_\rho(A)\cap\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho^2\big(d(B,C)\big)\,dx_C\,dy_C\Big\} \\
<\ & \max\Big\{\iint_{\overline{D}_\rho(A)} h_\rho^2\big(d(A,C)\big)\,dx_C\,dy_C, \iint_{\overline{D}_\rho(B)} h_\rho^2\big(d(B,C)\big)\,dx_C\,dy_C\Big\} \\
=\ & \kappa_\rho s_\rho,
\end{aligned}$$

where the last equality holds according to the definitions of s_ρ and κ_ρ in (8.1) and (8.2), respectively.

Thus the desired result when (i_1, j_1) belongs to R_4 follows from (E.1).

2) When (i_1, j_1) belongs to R_5, we have

$$\iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big)\,dx_C\,dy_C \geq 0.25 s_\rho - 2\ell_n^2.$$

Also it is clear that

$$\begin{aligned}
&\iint_{\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(B,C)\big)\,dx_C\,dy_C \\
\geq\ & \iint_{\overline{D}_\rho(A)\cap\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(B,C)\big)\,dx_C\,dy_C \\
\geq\ & \iint_{\overline{D}_\rho(A)\cap\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big)h_\rho\big(d(B,C)\big)\,dx_C\,dy_C.
\end{aligned}$$

Thus the desired result follows from (E.1).

3) When (i_1, j_1) belongs to R_6, it is without loss of generality to only consider the case that (i_1, j_1) belongs to R_{61}. According to Fig. E.1[1], since $y_B \geq 0$, we have

$$\iint_{\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(B,C)\big)\, dx_C\, dy_C \geq 0.5s_\rho - 2\ell_n^2,$$

and

$$\begin{aligned}
&\iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big)\, dx_C\, dy_C \\
&\quad - \iint_{\overline{D}_\rho(A)\cap\overline{D}_\rho(B)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big) h_\rho\big(d(B,C)\big)\, dx_C\, dy_C \\
\geq\ & \iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big)\, dx_C\, dy_C \\
&\quad - \iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big) h_\rho\big(d(B,C)\big)\, dx_C\, dy_C \\
=\ & \iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}} h_\rho\big(d(A,C)\big) \big(1 - h_\rho\big(d(B,C)\big)\big)\, dx_C\, dy_C \\
\geq\ & \iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}\cap([0,1]\times[y_A,1])} h_\rho\big(d(A,C)\big) \big(1 - h_\rho\big(d(B,C)\big)\big)\, dx_C\, dy_C
\end{aligned}$$

The integration region $\overline{D}_\rho(A) \cap \Sigma'_{n,i_1j_1i_2j_2} \cap ([0,1] \times [y_A, 1])$ is displayed as the shaded region in Fig. E.1. It is plain from the figure that

$$\begin{aligned}
&\iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}\cap([0,1]\times[y_A,1])} h_\rho\big(d(A,C)\big)\, dx_C\, dy_C \\
\geq\ & \iint_{\overline{D}_\rho(A)\cap([0,1]\times[y_A,1])} h_\rho\big(d(A,C)\big)\, dx_C\, dy_C \\
&\quad - \iint_{\Sigma_{n,i_1j_1}\cup\Sigma_{n,i_2j_2}} h_\rho\big(d(A,C)\big)\, dx_C\, dy_C \\
\geq\ & 0.5s_\rho - 2\ell_n^2
\end{aligned}$$

[1]Our analysis is based on the assumption that $y_A \geq y_B$, the case that $y_A < y_B$ can be similarly analyzed.

and

$$
\begin{aligned}
&\iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}\cap([0,1]\times[y_A,1])} h_\rho\big(d(A,C)\big)h_\rho\big(d(B,C)\big)\,dx_C\,dy_C \\
\leq\;& \max\Big\{\iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}\cap([0,1]\times[y_A,1])} h_\rho^2\big(d(A,C)\big)\,dx_C\,dy_C, \\
&\quad \iint_{\overline{D}_\rho(A)\cap\Sigma'_{n,i_1j_1i_2j_2}\cap([0,1]\times[y_A,1])} h_\rho^2\big(d(B,C)\big)\,dx_C\,dy_C\Big\} \\
\leq\;& \max\Big\{\iint_{\overline{D}_\rho(A)\cap([0,1]\times[y_A,1])} h_\rho^2\big(d(A,C)\big)\,dx_C\,dy_C, \\
&\quad \iint_{\overline{D}_\rho(A)\cap([0,1]\times[y_A,1])} h_\rho^2\big(d(B,C)\big)\,dx_C\,dy_C\Big\} \\
\leq\;& 0.5\kappa_\rho s_\rho
\end{aligned}
$$

Thus the desired result follows from (E.1).

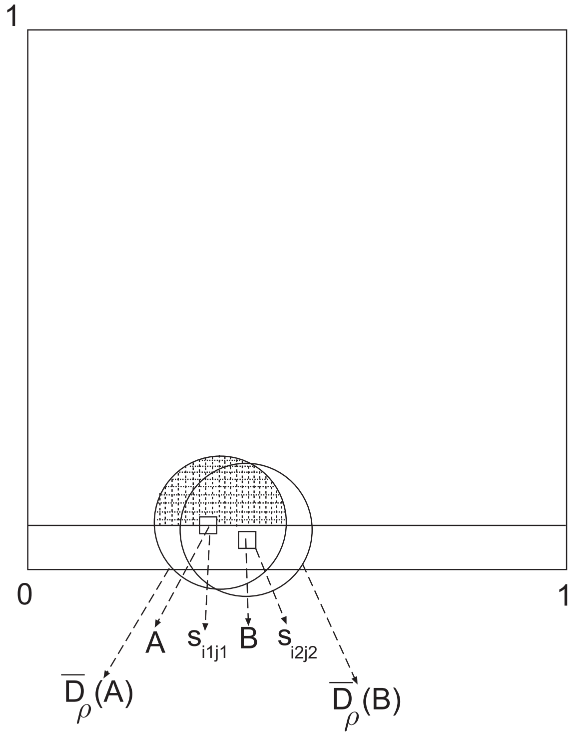


Figure E.1: Integration region when user A belongs to R_{61}.

BIBLIOGRAPHY

[1] N. Alon and J. H. Spencer, "The Probabilistic Method," (Second Edition), *Wiley-Science Series in Discrete Mathematics and Optimization*, John Wiley & Sons, New York (NY) 2000.

[2] M.J.B. Appel and R.P. Russo, "The connectivity of a graph on uniform points on $[0,1]^d$," *Statisitcs & Probability Letters* **60** (2002), pp. 351-357.

[3] R. Arratia, R. Goldstein, L. Gordon, "Two moments suffice for Poisson approximations: the Chen-Stein method," *Annals of Probability*, **17** (1989), pp. 9-25.

[4] Chen Avin, "Distance Graphs: From Random Geometric Graphs to Bernoulli Graphs and Between, " Under review.

[5] A. D. Barbour, L. Holst, S. Janson, "Poisson Approximation," Oxford Studies in Probability, 1992.

[6] A. D. Barbour and L. Holst, "Some applications of the Stein-Chen method for proving Poisson convergence," *Advances in Applied Probability*, **21** (1989), pp. 74-90.

[7] C. Bettstetter et al. "The Node Distribution of the Random Waypoint Mobility Model for Wireless Ad Hoc Networks." *IEEE Transactions on Mobile Computing*, **2** (2003), pp.257-269.

[8] C. Bettstetter and C. Hartmann, "Connectivity of Wireless Multihop Networks in a Shadow Fading Environment." *ACM.Springer Wireless Networks*, **11** (2005), pp. 571-579.

[9] B. Bollobás, "Random Graphs," Cambridge University Press, New York, 2001.

[10] J. Bourgain and G. Kalai, "Influences of variables and threshold intervals under group symmetries," *Geometric Functional Analysis*, 1997.

[11] I. D. Chakeres and E. M. Belding-Royer, "The utility of Hello messages for determining link connectivity," *Wireless Personal Multimedia Communications*, 2:504–508, October 27 2003.

[12] H. A. David and H. N. Nagaraja, "Order Statistics," Wiley Series in Probability and Statistics, John Wiley & Sons, Hoboken (NJ), 2003.

[13] P. Deheuvels, "Strong limit theorems for maximal spacings from a general univariate distribution," *The Annals of Probability* **12** (1984), pp. 1181-1193.

[14] M. P. Desai and D. Manjunath, "On the connectivity in finite ad hoc networks," *IEEE Communication Letters*, **10** (2002), pp. 437-490.

[15] L. Devroye, "Laws of iterated logarithm for order statistics of uniform spacings," *The Annals of Probability* **9** (1981), pp. 860-867.

[16] A. Ephremides, "Energy concerns in wireless networks" *Wireless Communications*, **9** (2002), pp. 48 - 59.

[17] P. Erdös and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci.* **5** (1960), pp. 17-61.

[18] D. Estrin et al, "Embedded, Everywhere: A Research Agenda for Networked Systems of Embedded Computers," National Research Council Report, 2001.

[19] W. Feller, "An Introduction to Probability Theory and Its Applications," Vol. 1, 3rd ed. New York, Wiley, 1968.

[20] C. H. Foh and B. S. Lee "A Closed Form Network Connectivity Formula for One-Dimensional MANETs," *Proc. IEEE ICC* 2004

[21] C. H. Foh, G. Liu, B. S. Lee, B. -C. Seet, K. -J. Wong and C. P. Fu, "Network connectivity of one-dimensional MANETs with random waypoint movement," *IEEE Communications Letters* **9** (2005), pp. 31-33.

[22] E. Friedgut and G. Kalai, "Every monotone graph property has a sharp threshold," *Proceedings of the American mathematical Society* **124** (1996), pp. 2993-3002.

[23] A. Ghasemi and S. Nader-Esfahani, "Exact probability of connectivity in one-dimensional ad hoc wireless networks", *IEEE Communications Letters* **10** (2006), pp. 251-253.

[24] E. N. Gilbert, "Random plane networks," *SIAM Journal* **9** (1961), pp. 533-543.

[25] E. Godehardt and J. Jaworski, “On the connectivity of a random interval graph,” *Random Structures and Algorithms* **9** (1996), pp. 137-161.

[26] A. Goel, S. Rai, and B. Krishnamachari, “Sharp thresholds for monotone properties in random geometric graphs,” *Annals of Applied Probability* **15** (2005).

[27] A.D. Gore, “Comments on “On the connectivity in finite ad hoc networks”,” *IEEE Communications Letters* **10** (2006), pp. 88-90.

[28] P. Gupta and P. R. Kumar. “Critical Power for Asymptotic Connectivity in Wireless Networks,” pp. 547-566, in *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*. Edited by W.M. McEneany, G. Yin, and Q. Zhang, Birkhauser, Boston, 1998. ISBN 0-8176-4078-9.

[29] G. Han and A. M. Makowski, “Poisson convergence can yield very sharp transitions in geometric random graphs,” Information Theory & Applications Workshop, La Jolla, CA, February 2006.

[30] G. Han and A. M. Makowski, “Very sharp transitions in one-dimensional MANETs,” in Proceedings of the International Conference on Communications (ICC 2006), Istanbul (Turkey), June 2006.

[31] G. Han and A.M. Makowski, “A very strong zero-one laws for connectivity in one-dimensional geometric random graphs,” *IEEE Communications Letters*, **11**, pp 55-57, January, 2007.

[32] G. Han and A.M. Makowski, "On the critical communication range under node placement with vanishing densities," in IEEE International Symposium on Information Theory, June, 2007.

[33] G. Han and A.M. Makowski, "Connectivity in one-dimensional geometric random graphs: Poisson approximations, zero-one laws and phase transitions," submitted to *IEEE Transactions on Information Theory* (2007).

[34] G. Han and A.M. Makowski, "A strong zero-one law for connectivity in one-dimensional geometric random graphs with non-vanishing densities," submitted to *IEEE Transactions on Information Theory* (2007).

[35] G. Han and A.M. Makowski, "A very strong zero-one law for connectivity in one-dimensional geometric random graphs with non-vanishing densities," submitted to *IEEE Transactions on Information Theory* (2007).

[36] R. Hekmat and P. Van Mieghem. "Connectivity in Wireless Ad Hoc Networks with a Log-normal Radio Model," *Mobile Networks and Applications, The Journal of Special Issues on Mobility of Systems, Users, Data and Computing*, 2005.

[37] J. Hüsler, "Maximal, non-uniform spacings and the covering problem," *Journal of Applied Probability* **25** (1988), pp. 519-528.

[38] B. Krishnamachari, S. Wicker, R. Bejar, and M. Pearlman, "Critical Density Thresholds in Distributed Wireless Networks," *book chapter in Communications, Information and Network Security*, Eds. H. Bhargava, H.V. Poor, V. Tarokh, and S. Yoon, Kluwer Publishers, December 2002.

[39] P. Lévy, "Sur la division d'un segment par des points choisis au hasard," *Comptes Rendus de l' Académie des Sciences de Paris* **208** (1939), pp. 147-149.

[40] T. Lindvall, "Lectures on the Coupling Method," John Wiley & Sons, New York (NY), 1992

[41] H. Maehara, "On the intersection graph of random arcs on a circle," *Random Graphs'87*, pp.159-173.

[42] G.L. McColm, "Threshold functions for random graphs on a line segment," *Combinatorics, Probability and Computing* **13** (2004), pp. 373-387.

[43] R. Meester and R. Roy, "Continuum Percolation", Campbridge University Press, 1996.

[44] D. Miorandi and E. Altman, "Coverage and connectivity of ad-hoc networks in presence of channel randomness," IEEE Infocom, Miami, March 13-17, 2005.

[45] S. Muthukrishnan and G. Pandurangan, "The bin-covering technique for thresholding random geometric graph properties," in Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms (SODA 2005), Vancouver (BC), 2005.

[46] M. Penrose, "Random Geometric Graphs," Oxford Studies in Probability, 2003.

[47] M.D. Penrose, "The longest edge of the random minimal spanning tree," *Annals of Applied Probability* **7** (1997), pp. 340-361.

[48] R. Pyke, "Spacings," *Journal of the Royal Statistical Society*, Series B (Methodological) **27** (1965), pp. 395-449.

[49] T. Rappaport, "Wireless communications: Principles & Practice," Upper Saddle River Prentice-Hall PTR, 2002.

[50] P. Santi, D. Blough and F. Vainstein, "A probabilistic analysis for the range assignment problem in ad hoc networks," in the Proceedings of the 2nd ACM International Symposium on Mobile Ad hoc Networking & Computing (MobiHoc 2001), Long Beach (CA), 2001.

[51] P. Santi, "The Critical Transmitting Range for Connectivity in Mobile Ad Hoc Networks," *IEEE Trans. on Mobile Computing*, **4** (2004), pp. 310-317.

[52] E. V. Slud, "Entropy and maximal spacing for random partitions," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebeiete* **41** (1978), pp. 341-352.

[53] D. Stoyan, W. S. Kendall, J. Mecke, "Stochastic Geometry and Its Applications," 2nd Edition, Wiley series in probability and statistics, 1996.

[54] S. S. Venkatesh, "Connectivity of metric random graphs." *Random Structures and Algorithms*, submitted 2004.

[55] M. Weiser, "The Computer for the Twenty-First Century," *Scientific American*, pp. 94-104, September 1991