COMMUNICATION IN NETWORKS WITH SPATIALLY CORRELATED FAULTS

by

Michel Paquette

M.Sc. Université du Québec en Outaouais, 2005

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Computer Science

Carleton University
Ottawa, Ontario

© 2010, Michel Paquett
Abstract

The aim of this thesis is to study communication algorithms in networks where nodes and/or communication links fail in a random dependent way. In order to capture fault dependencies, we introduce the neighborhood fault, ranged fault, and swamping communication models. In the neighborhood fault model for arbitrary networks, damaging events, called spots, occur randomly and independently with probability $p$ at nodes of a network, and cause faults in the given node and all of its neighbors. Under this model, faults at nodes within distance 2 are correlated. In the ranged fault model for geometric networks, spots occur with Poisson arrival rate $\lambda$ on a plane and cause faults in all nodes at distance at most $s$. Under this model, faults at nodes within distance $2s$ are spatially correlated. In the swamping communication model for geometric radio networks, nodes can communicate only with nodes that are at a distance greater than the swamping distance $s$ and at most at the communication range $r$. Under this model, nodes experience reception disturbances when nodes within distance $s$ transmit, hence transient message reception failures are spatially correlated. Under each of these models, we present results concerning the connectivity and the diameter of networks as well as the time of communication.

Keywords: Fault-tolerance, dependent faults, communication, crash faults, network connectivity, broadcasting, diameter, geometric radio network, topology, swamping.
Acknowledgements

My gratitude goes out to my supervisors Evangelos Kranakis and Andrzej Pelc for their constant efforts in steering me towards the right path, for their support and, most of all, for their great patience. Thanks also go out to my family and friends who gave me much support through the best and the worst years of my PhD studies. Lastly, thanks go to my love, Catherine, for her support through the final steps of my research and thesis writing.
Table of Contents

Abstract ii

Acknowledgements iii

List of Figures vii

Chapter 1  Introduction 1

1.1 Motivation 2

1.2 Problem Statement and General Background 2

1.2.1 Arbitrary Wired Networks 6

1.2.2 Wireless Communication - Geometric Networks 7

1.3 Summary of Results 9

Chapter 2  Related Work 11

2.1 Dependent Fault Models in Microelectronics 11

2.2 Communication with Random Faults 12

2.2.1 Random and Independent Faults 12

2.2.2 Random, Positively Correlated Faults 13

2.3 Wireless Data Communication 14

2.3.1 Wireless Communication with Faults 14

2.4 Communication in Networks of Unknown Topology 15

Chapter 3  The Diameter and Connectivity of Networks with Random Dependent Faults 17

3.1 Introduction 17

3.1.1 Model and Problem Definition 17

3.1.2 Our Results 18

3.2 Networks of Unbounded Degree 19

3.2.1 Connectivity Results 20
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Random Binary Thick Tree</td>
<td>21</td>
</tr>
<tr>
<td>3.2</td>
<td>Building an elementary event $\varphi_u$</td>
<td>28</td>
</tr>
<tr>
<td>3.3</td>
<td>Cycle on $\Gamma_2((x, y)) \cup \Gamma_3((x, y))$</td>
<td>35</td>
</tr>
<tr>
<td>3.4</td>
<td>Substitute subpaths</td>
<td>36</td>
</tr>
<tr>
<td>3.5</td>
<td>The Toroidal Tree</td>
<td>38</td>
</tr>
<tr>
<td>3.6</td>
<td>Torus disconnection</td>
<td>39</td>
</tr>
<tr>
<td>4.1</td>
<td>Non-Connectivity of $U(n, r, \lambda, s)$</td>
<td>55</td>
</tr>
<tr>
<td>4.2</td>
<td>Algorithm $A^*$</td>
<td>68</td>
</tr>
<tr>
<td>5.1</td>
<td>Partition $\mathcal{P}$</td>
<td>80</td>
</tr>
<tr>
<td>5.2</td>
<td>Coverage of transmissions</td>
<td>89</td>
</tr>
<tr>
<td>5.3</td>
<td>Partition $\mathcal{P}^2$</td>
<td>93</td>
</tr>
<tr>
<td>5.4</td>
<td>The spokesmen of a block</td>
<td>98</td>
</tr>
<tr>
<td>5.5</td>
<td>Proximity by sector of spokesmen</td>
<td>99</td>
</tr>
<tr>
<td>5.6</td>
<td>Overall proximity of spokesmen</td>
<td>99</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

In the past 20 years, large communication systems have become integral to all aspects of our society. From personal communications to business operations and scientific computing, data networks now support all modern data communication needs. In the current information age, data networks are at the foundation of modern life.

Large parallel and distributed information systems have become essential in activities as diverse as banking, mineral/oil prospection, and peace-keeping operations. These systems allow the efficient computation of large amounts of data, the execution of high complexity algorithms and the dissemination of information to distributed locations. To scale with the increase in task volume and complexity, more processors and communication paths are added to these systems. However, with their growth in size and complexity, distributed systems become increasingly vulnerable to component failures; the resulting failures can lead to the loss of financial assets, scientific data, or even human life.

In the past, largely parallel systems were almost exclusively built using wired communication systems; this is changing. Wireless technologies are now at the center of upcoming networking standards. For instance, vehicular ad hoc networks, (VANETs), will increase convenience and safety for automobile drivers, in the near future by providing mobile information support and, specifically, by forwarding warning messages. These networks will be composed of an infrastructured part, partly or fully wired, as well as a wireless infrastructureless (ad hoc) part. This technology combination will bring forward reliability challenges associated to each of these technologies.

The enhancements to safety brought about by such technologies are considerable. On the other hand, existing work supporting the Offsetting Behavior Hypothesis (c.f., e.g., [50]) strongly suggests, in this instance, that automobile drivers will rely on this technology to forewarn them of dangers, becoming dependent on the network
functionality. This is the case for information technologies, which are now becoming integrated both in our lives (c.f., e.g., [39]), our work habits and ultimately, our economy. Hence, it becomes essential to incorporate fault-tolerance in network and algorithm designs so that information processing and dissemination may be reliable.

1.1 Motivation

The fundamental questions of network reliability have received much attention in past research under the assumption that components fail randomly and independently (cf., e.g. [7, 11, 12, 44, 45, 46, 47, 54] and the survey [48]). On the other hand, empirical work has shown that positive correlation of faults is a more reasonable assumption for networks [20, 56, 59].

There are many modern age problems which witness to this fact. For instance, on August 14, 2003, faults cascaded on the power distribution network and deprived part of North America of electricity. More common and widespread failure causes are computer viruses and Internet worms. These software phenomena take advantage of some type of proximity - social, logical or physical - between computers to spread. Finally, lightning strikes are known to deprive entire city blocks of electricity, when hitting one component of an electrical distribution network.

As our society is increasingly dependent on information networks, it becomes essential to study questions related to their tolerance of positively correlated faults. However, little analytic work has been done for communication networks under this assumption about faults.

1.2 Problem Statement and General Background

Under probabilistic models, analytic answers to the fundamental questions of communication have been developed under the assumption of fault independence. In this thesis, we provide some answers to the following question.

**Question 1** How does removing the assumption of fault independence affect the analytic results concerning fundamental questions of fault-tolerant communication?
To answer this question, we will adhere to the conventional models for networks. A network is a collection of interconnected computational units. It is modeled as an undirected graph $G = (V, E)$ with a set of nodes $V$ connected by a set of undirected links $E$. We say that two nodes are adjacent (or neighbors) if they share a link. The geometric distance between nodes $u, v \in V$ is the length of the straight line segment between node $u$ and node $v$; it is denoted by $\text{dist}(u, v)$. The graph distance or hop count between nodes $u, v \in V$ is the minimum number of links which must be traversed from $u$ to reach $v$; it is denoted by $\text{hop}(u, v)$. The maximum hop count between any two nodes, $D$, is called the diameter of the network. For a node $u$, we denote the set of its neighbors by $\Gamma(u)$. For a set of nodes $U$, we denote by $\Gamma(U)$ the union of all sets of neighbors of nodes $u \in U$, i.e., $\Gamma(U) = \bigcup_{u \in U} \Gamma(u)$. For a node $u$, we denote by $\Gamma_d(u)$ the set of all nodes $w$ such that $\text{hop}(u, w) = d$ and we denote by $\Gamma_{\leq d}(u)$ the set of all nodes $w$ such that $\text{hop}(u, w) \leq d$; we call the latter the distance $d$ neighborhood. The notations $\Gamma(u)$ and $\Gamma_1(u)$ are equivalent, however for simplicity of notation $\Gamma(u)$ is used whenever possible.

An important type of wireless networks is obtained from a set of stations in the plane where each station $u$ has communication range $r_u$. The resulting network is modeled as a directed graph in which stations are nodes and a directed edge exists from $u$ to $v$ if $v$ is at geometric distance at most $r_u$ from $u$. I.e., a directed edge goes from a node $u$ to a node $v$ when the node $v$ can receive the information sent by the node $u$. Such networks are called geometric radio networks (GRNs). In Chapter 4, we assume that the range is the same for all nodes, and hence the considered GRNs are modeled as undirected graphs. In Chapter 5, we also assume that all nodes possess the same minimum restriction on the distance of communication, the swamping range; hence the considered wireless networks are also modeled as undirected graphs.

A node is said to be functional, or fault-free, when it executes only its predefined algorithm without any deviation, and doing so, transmits all messages correctly, in a timely manner and without any loss; a node which is not functional is said to be faulty. The same definitions apply to links, with the exception that they do not execute algorithms. Faults can be of different types: at opposite ends of the spectrum are crash faults and Byzantine faults. Faults of the crash type cause faulty entities to
stop all communication; these can neither send, receive nor relay any message. For all practical purposes, they can be considered as removed from the network. Faults of the Byzantine type cause arbitrary (even malicious) behavior in the affected entities. We say that faults are permanent when their effect lasts for the entire duration of a communication process; otherwise, the faults are said to be transient. In Chapter 3 and Chapter 4, we assume that permanent faults of the crash type occur in nodes. In Chapter 5, we assume that transient faults of the crash type occur in incoming links.

The fault-free part of a network $G$ is the graph induced on $G$ by the set of all functional nodes, i.e., it is the graph obtained by removing from $G$ all faulty nodes and their incident edges. We consider communication in the fault-free part of networks, specifically inside large connected components of functional nodes, i.e., a connected component of the network which contains a constant fraction of all nodes. We will focus on the case where communication is synchronous; nodes have synchronized clocks and the communication process is executed in fixed time steps, called rounds. We will first consider communication in arbitrary wired networks and then in geometric wireless networks.

One of the most important communication tasks is broadcasting. In this process, a source node attempts to transmit a message to all other nodes of the network. This process is successful if, upon termination, all functional nodes, connected to the source by a fault-free path, have received the source message. In the case of wired networks, we measure the time of communication under the all-port message passing model, where nodes can communicate with all their neighbors during each round, and under the 1-port message passing model, in which every node can communicate with at most one neighbor during each round. Under the all-port model, broadcast can be completed in time $D$ if the fault-free part of the network has diameter $D$. For networks of diameter $D$ and maximum degree $\Delta$, broadcast can be completed in time at most $D\Delta$ under the 1-port model.

In the case of wireless networks, we will assume either the radio communication model or the swamping communication model. In the radio communication model, in each round, each node either sends a message or listens to the channel. In the first case, we say that the node is a sender, otherwise, it is a receiver. In a fixed
round, a node \( v \) receives a message if and only if it is a receiver, and exactly one of its neighbors is a sender. If no neighbor of \( v \) is a sender, then there is no message on the channel which \( v \) can receive. If more than one neighbor of \( v \) sends a message, we say that a collision occurs at \( v \); \( v \) can only perceive noise on the channel. Nodes do not have collision detection abilities, i.e., they cannot distinguish collision noise from background noise (which is apparent when no messages are heard).

We describe the swamping communication model. The graph is built from a set of nodes, equipped with communication range \( r \) and limited by a minimum distance requirement of \( s \) (the swamping distance). Two nodes \( u, v \in V \) located at geometric distance \( \text{dist}(u, v) \) greater than \( s \) and at most \( r \) from one-another are neighbors and share an undirected link \( (u, v) \in E \) in the graph \( G \); no other links exist in \( G \). As usual for wireless networks, communication between nodes is synchronous. In each round, each node is either a sender or a receiver. A node \( u \), which is a transmitter in a given round, sends a message to the entire set of its neighbors \( \Gamma(u) \) within the same round; this transmission makes it impossible for all nodes within distance \( s \) to receive any message. That is, for each round when a node within geometric distance \( s \) of it transmits, a node \( v \) receives no message; in this case, only noise is heard by \( v \), indistinguishable from the background noise heard when no messages are sent. In a fixed round, a node \( v \) receives a message if and only if it is a receiver, exactly one of its neighbors is a sender, and no node within distance \( s \) sends a message. If no neighbor of \( v \) is a sender, then there is no message on the channel which \( v \) can receive. If more than one neighbor of \( v \) sends a message, a collision occurs at \( v \). Nodes do not have collision detection abilities, i.e., they cannot distinguish collision noise from background noise (which is apparent when no messages are heard).

The swamping communication model can be viewed as a GRN on which radio communication is implemented with additional transient reception faults on all nodes at close proximity of a transmitter, i.e., a node cannot receive messages at each round when some node within distance \( s \) from it transmits.

Let \( n \) be the parameter which represents the number of nodes in the graph (Chapters 3 and 5) or the arrival rate of nodes on a plane (Chapter 4). We say that an event occurs in the graph with high probability (w.h.p.) if its probability converges
to 1 as \( n \) grows to infinity. We say that an event occurs on the graph with constant (positive) probability if its probability \( p \) is bounded away from 0 and from 1 for all \( n \), i.e., if there exist constants \( \epsilon_1, \epsilon_2 \) such that \( 0 < \epsilon_1 < p < \epsilon_2 < 1 \) for all \( n \). Throughout the thesis, \( \log \) means logarithm with base 2 and \( \ln \) means the natural logarithm. Given an event \( A \), we denote its complement by \( \neg A \). As usual, \( \omega(f) \) denotes the set of functions \( g \) such that \( g/f \to \infty \). The asymptotic relation notation (cf., [42]) summarized in Table 1.1 is used throughout this thesis.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Intuition</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f(n) \in O(g(n)) )</td>
<td>( g ) is an asymptotic upper bound on ( f ) (up to constant factor)</td>
<td>( \exists (C &gt; 0), n_0: \forall (n &gt; n_0) ) (</td>
</tr>
<tr>
<td>( f(n) \in \Omega(g(n)) )</td>
<td>( g ) is an asymptotic lower bound on ( f ) (up to constant factor)</td>
<td>( \exists (C &gt; 0), n_0: \forall (n &gt; n_0) ) (</td>
</tr>
<tr>
<td>( f(n) \in \Theta(g(n)) )</td>
<td>( f ) is bounded tightly by ( g ) asymptotically</td>
<td>( \exists (C, C' &gt; 0), n_0: \forall (n &gt; n_0) ) (</td>
</tr>
<tr>
<td>( f(n) \in o(g(n)) )</td>
<td>( f ) is dominated by ( g ) asymptotically</td>
<td>( \forall (C &gt; 0), \exists n_0: \forall (n &gt; n_0) ) (</td>
</tr>
<tr>
<td>( f(n) \in \omega(g(n)) )</td>
<td>( f ) dominates ( g ) asymptotically</td>
<td>( \forall (C &gt; 0), \exists n_0: \forall (n &gt; n_0) ) (</td>
</tr>
</tbody>
</table>

Table 1.1: Asymptotic relation notation table (cf., [1]), annotated.

### 1.2.1 Arbitrary Wired Networks

When communication devices transmit information using guided media, such as wire pairs and coaxial cables, the existence of links in the network is not defined by the geometric distances between nodes. The distance metrics of these networks are based on the number of links standing between nodes rather than the geometric node locations; in these networks we often use the graph distance, also known as the hop count, as the distance metric. We refer to such graphs as arbitrary because of the unrestricted nature of their topology; instances of such graphs model large wired communication networks, such as telephony networks or the Internet.

For these networks, we can reasonably assume that damaging events affect nodes
randomly and independently at their locations, with some fixed probability $p$. Moreover, residual energy from affected sites can travel through the guided medium, affecting neighboring nodes with some fixed probability $q$. This occurs, for instance, when lightning strikes electrical lines, disabling power and causing damage to household electronics. We specifically address this problem in the worst case, when $q = 1$, in Chapter 3.

Problem 1 Communication with node fault clusters: Consider an arbitrary network. Let damaging events occur randomly and independently at node sites with probability $p$, causing faults in the subjected node and in all nodes with whom they share a link.

1. What are the thresholds on $p$, the probability of occurrence of a damaging event, such that the connectivity of the fault-free part of the network is preserved with high probability?

2. Can we show networks inside which efficient communication is possible within connected components of functional nodes despite the occurrence of these faults?

1.2.2 Wireless Communication - Geometric Networks

When devices communicate data using wireless technology, in the absence of collisions, incoming signals are heard when their strength is sufficient for the receiver to decode the information. For most wireless networks, communication devices use omnidirectional antennas, and thus radio signal strength becomes a function, dependent only on the environment and on the distance. Consequently, models for such networks often use the assumption that the received signal strength is a function only of the distance from a source to a receiver, i.e., nodes share a link if they are within a certain distance from each other. These graphs are called geometric.

Wireless networks are vulnerable to phenomena causing interference on the radio signals (the links) and mechanical/electrical phenomena which damage the communication equipment (the nodes) of the network. These phenomena have effects within certain distances and, in the geometric network setting, can affect nearby nodes.
Communication with Damaging Events

Electrostatic discharges are a well known cause of electronic equipment breakdown (c.f., e.g., [21]). These discharges occur when the electrical tension levels are too high for the insulating material to prevent the flow of electrons between components (c.f., e.g., [52]). As a consequence of such discharges, heat is produced in a very small area, often causing the breakdown of microscopic components (c.f., e.g., [30]). Since the discharges occur, [52], when tension levels reach or exceed the dielectric breakdown threshold, a reasonable assumption is that phenomena affect only nodes within a certain radius. The following problem is addressed in Chapter 4.

Problem 2 Communication with spatially correlated node faults: Consider network nodes and damaging phenomena (spots) that are both distributed in the plane by distinct Poisson processes. Nodes have a fixed communication range $r$ and all nodes within distance $s$ of a spot are faulty.

1. What are the thresholds on $s$, $r$, and on the arrival rates such that

   (a) the fault-free part of the graph remains connected with high probability?

   (b) a constant fraction of the fault-free nodes in the graph form a connected component?

2. Within a connected component of fault-free nodes, can we show a data dissemination algorithm which achieves communication efficiently without prior knowledge of the network topology?

Communication in the Presence of Swamping

One of the known problems commonly faced by radio transceivers is that of swamping (cf., e.g., [6, 10, 38]). When two wireless nodes are at close proximity, their receivers cannot adapt to strong incoming signals; communication becomes difficult, even impossible. In contrast to traditional radio communication models, nodes at close proximity are not able to communicate directly; intermediate nodes will be needed to relay their messages. The following problem is addressed in Chapter 5.
Problem 3 Communication in the Presence of Swamping: Consider nodes whose transceivers allow communication with other nodes within distance $r$, but at distance greater than $s$ of them. Consider further that transmissions from nodes within distance $s$ make the reception of any message impossible.

1. Can we show efficient broadcasting algorithms for communication in lattice networks? Can we further show bounds on the time of communication in these networks?

2. Can we show efficient broadcasting algorithms for communication in networks of unknown topology? Can we further show bounds on the time of communication in these networks?

1.3 Summary of Results

Throughout this thesis, analytic results will be developed with the use of combinatorial tools, probability theory and graph theory.

In Chapter 3, we provide answers to Problem 1: communication with node fault clusters. In this chapter, we detail the results of our work, contained in [34] and [37]. In order to capture fault dependencies, we introduce the neighborhood fault model, where damaging events, called spots, occur randomly and independently with probability $p$ at nodes of a network, and cause faults in the given node and all of its neighbors. Faults at distance at most 2 become dependent in this model and are positively correlated. We investigate the impact of spot probability on the feasibility and time of communication in the fault-free part of the network. We show a network which supports fast communication with high probability, if $p \leq 1/(c \log n)$. We also show that communication is not feasible with high probability in most classes of networks, for constant spot probabilities. This negative result holds, in particular, for the important case of the hypercube. For smaller spot probabilities, high probability communication is supported even by bounded degree networks. It is shown that the torus supports communication with high probability when $p$ decreases faster than $1/n^{1/2}$, and does not when $p \in 1/O(n^{1/2})$. Furthermore, a network built of tori is designed, with the same fault-tolerance properties and additionally supporting fast
communication. We show, however, that networks of degree bounded by a constant \( d \) do not support communication with high probability, if \( p \in 1/O(n^{1/d}) \). This work demonstrates that there is a gap between the fault-tolerance of networks when faults occur independently as opposed to when they occur with positive correlation.

In Chapter 4, we study Problem 2: communication with spatially correlated node faults. In this chapter, we detail the results of our work, contained in [35] and [36]. We study the feasibility and time of communication in random geometric radio networks, where nodes fail randomly with positive correlation. We consider a set of radio stations with the same communication range, obeying Poisson distribution on a unit square region. In order to capture fault dependencies, we introduce the \textit{ranged spot} model in which damaging events, called \textit{spots}, are Poisson distributed on the region, causing faults in all nodes located within distance \( s \) from them. Node faults within distance \( 2s \) become dependent in this model and are positively correlated. We investigate the impact of the spot arrival rate on the feasibility and the time of communication in the fault-free part of the network. We first show threshold functions on the spot probability, a constant factor away, above which broadcasting is not feasible and below which it is, with high probability. We then design an algorithm which broadcasts correctly with probability \( 1 - \epsilon \) in faulty random geometric radio networks of diameter \( D \) in time \( O(D + \log 1/\epsilon) \).

In Chapter 5, we study Problem 3: communication in the presence of swamping. In this chapter, we detail the results of our work, [33]. We address the problem of broadcasting in a wireless network under a novel communication model: the \textit{swamping} communication model. In this model, nodes communicate only with those nodes at distance greater than \( s \) and at most \( r \) from them. For the \( n \)-node lattice networks, we present algorithms of optimal time complexity, respectively \( O(n/r + r/(r - s)) \) for the lattice line and \( O(\sqrt{n}/r + r/(r - s)) \) for the two-dimensional lattice. We also consider networks of unknown topology of diameter \( D \) and of a parameter \( g \) (\textit{granularity}); \( g = 1/\alpha \) and \( \alpha \) is the smallest geometric distance between any two nodes of the network. We present broadcast algorithms for networks where nodes are placed in the line and in the plane with respective time complexities \( O(D/l + g^2) \) and \( O(Dg/l + g^4) \), where \( l \in \Theta(\max\{(1 - s), \alpha\}) \).
Chapter 2

Related Work

2.1 Dependent Fault Models in Microelectronics

Dependent fault models were introduced in the field of microelectronic circuit manufacturing in the mid 1980’s. These fault models were used as a basis in the study of Very Large Scale Integration (VLSI) circuit manufacturing production yields. In this research, faults were viewed as the result of impurities on silicon wafers which occurred randomly and independently and affected nearby circuit components in a dependent way. Results were proposed mainly according to two approaches: the quadrat-based and center-satellite approaches.

In [55], Stapper proposed a coarse approach based on a square mesh partition (called quadrat) of the silicon wafers. He proposed the assumption that faults resulting from a fixed impurity occurred solely in the square containing this impurity. This quadrat-based model offered provably good results and ease of use required by the industry. Then, in [41], Meyer and Pradhan introduced a detailed model based on the center-satellite concept for ecological sampling [58]. In this model, impurities and resulting faults were viewed as the result of 4 parameter functions which were considered hard to estimate by practitioners. Later on, in [8], Blough and Pelc proposed a simplified center-satellite model of VLSI manufacturing defects for the study of the memory array reconfiguration problem.

Both the center-satellite and quadrat-based approaches were recently used for System on Chip (SoC) (cf., e.g., [28, 40]) and VLSI (cf., e.g., [13, 60]) applications. In [13], an application of quadrat-based analysis was presented in the analysis of the reliability of VLSI memory for long-lived spacecraft applications. In this paper, the authors stated that it is difficult to use the center-satellite models. In [40], the center-satellite approach was used to precisely model faults on multi-core SoC ICs (integrated circuits). This approach was advantageously compared to the Square
Nyeman (quadrat based) approach. In [60], the center-satellite approach, as refined by Blough and Pelc in [8], was used to precisely model faults in certain IC applications. Throughout this field of research, the consensus was that results originating from the center-satellite approach, as opposed to quadrat-based approaches, were more difficult to apply but provided better prediction quality.

This thesis focuses on communication network failures which occur in a dependent way. We consider both networks of nodes placed in Euclidean space, and arbitrary wired networks. The geometry-dependent, quadrat-based approach to fault dependencies is not usable for part of our work. We base our fault models on the center-satellite approach, using geometric distances in the case of wireless networks and graph distances in the case of arbitrary wired networks.

2.2 Communication with Random Faults

The research from the previous section should be contrasted with work on fault-tolerant communication in networks. The problem of communication in networks where faults occur randomly and independently has been widely studied in the past. Many results concerned random link and/or node failures (cf., e.g. [7, 11, 12, 44, 45, 46, 47, 54] and the survey [48]) but, in most cases faults were assumed to be independent. On the other hand, empirical work has shown the positive correlation of faults to be a more reasonable assumption for networks [20, 56, 59]. To the best of our knowledge, prior to our work, such analytic results were proposed only for a few families of graphs, under dependent percolation models (e.g., [4, 9, 25]).

2.2.1 Random and Independent Faults

One of the first papers to address the problem of communication in networks affected by random and independent faults is [7]. In this paper, Bienstock showed the existence of networks in which $O(\log n)$-time broadcast can be done, under the 1-port message passing model, with high probability, despite links which fail randomly and independently with positive constant probability. In [11], Chlebus, Diks and Pelc designed a network of logarithmic degree which can support high probability communication in time $O(\log n)$ when faults occur randomly and independently on links
and nodes with any constant probability smaller than 1. Then, in [12], the same authors showed fast and reliable fault-tolerant broadcasting algorithms for hypercube networks. They focused on communication under the 1-port message passing model (also called whispering model). In this setting, they provided algorithms broadcasting with optimal time order (Θ(log n)) when nodes and links fail randomly and independently with probability bounded below some constant. In [46], we designed a similar network which supports communication with high probability in time O(log² n) with faults of Byzantine type, occurring with probability bounded below some constant. It remains open if a network can be designed to support communication in logarithmic time with high probability under the same assumptions for faults.

2.2.2 Random, Positively Correlated Faults

The assumption of fault independence was used in past research to solve the fundamental problems of communication with more ease. On the other hand, physical and software phenomena generally affect physical components, causing failures in a positively correlated way. For instance, on August 14, 2003, faults cascaded on the power distribution network and deprived part of North America of electricity. More commonly, lightning strikes hitting one node of an electric network cause power outages in entire city blocks. Also, software phenomena, like computer viruses and worms, cause faults which occur in locations which are related to already infected sites.

In fact, empirical work has shown that positive correlation of faults is a more reasonable assumption for networks [20, 56, 59]. In particular, in [59], empirical evidence was given supporting that data packets losses are spatially correlated in networks. In [56], the assumption of this spatial correlation was used to enhance network traffic management. Moreover, in [20], it was stated that the environment provides many phenomena that may lead to spatially correlated faults; the authors simulated failures in a sensor network using a model much like those proposed in this thesis. We observe that this previous work was of empirical and experimental nature.
2.3 Wireless Data Communication

In the late 1960’s the interest in wireless data communication became apparent with the ALOHA system research project. As part of this project ([3]), Abramson made a case for the practicality of wireless data communication over the conventional wired, point-to-point alternative. He proved the efficiency of a random channel access multiplexing scheme for the system, compensating for the issue of collisions.

Later on, large wireless switching networks were considered. In [22], directions were provided for efficient broadcasting techniques in radio packet-switching networks. Based on heuristics and simulation results, the authors proposed controlled flooding, directional flooding and hierarchical routing schemes as practical information communication techniques, compensating, through redundancy, for node unavailability and failure. This paper left open the analysis of optimization in view of message collisions.

Although the issue of fault-tolerant wireless communication had been stated already in [22], literature on this topic is not plentiful. It is reviewed in Section 2.3.1.

2.3.1 Wireless Communication with Faults

As seen in Section 2.2.1, the questions of fault-tolerant communication have been addressed in the setting of arbitrary wired networks. In contrast, few results are known about fault-tolerant communication in geometric radio networks. To the best of our knowledge, all existing analytic results examine the problem of broadcasting in networks where, either the number of faults is bounded above by a known quantity (cf., e.g., [31]), or faults occur randomly and independently (cf., e.g., [32, 54]). One of the first papers on reliable broadcast in geometric radio networks is [43]. In this work, the authors showed a reliable communication protocol on wireless mobile networks which tolerates faults and disconnections and which is more efficient than flooding.

Later, in [31], the problem of broadcasting was considered for a fault-free connected component of a radio network whose nodes are located at points of square grids and communicate within a square of size $r$. For an upper bound $t$ on the number of faulty nodes, in worst-case location, the authors proposed a $\Theta(D + t)$-time oblivious broadcast algorithm and a $\Theta(D + \log(\min(r, t)))$-time adaptive broadcast
algorithm, both operating on a connected fault-free component of diameter $D$.

Then, in [54], the problem of connectivity was considered for a square grid of $n$ sensors with communication range $r$ on a unit square when faults occur at the nodes randomly and independently with probability $1 - p$. The authors showed that if $p r^2 \approx \log \frac{n}{n}$, then the functional nodes are all part of a connected component w.h.p.

More recently, in [2], the concept of region fault-tolerant spanners was introduced for point sets in the plane. These spanners remain connected despite the removal of the point set inside any one convex region. In particular, the authors showed that any set of $n$ points admits an $O(n \log n)$-size region fault-tolerant $(1 + \epsilon)$-spanner, for any constant $\epsilon > 0$.

### 2.4 Communication in Networks of Unknown Topology

The question of communication in unknown networks (wired and wireless) is related to the question of communication in networks with faults. Instead of designing networks whose fault-free part remains connected with high probability or algorithms which broadcast correctly with high probability in spite of faults, this research aims at efficient communication in connected components of a network, for which the topology is largely unknown. In fact, in [14], it was stated that broadcasting algorithms which function in unknown GRNs also function in the resulting fault-free connected components of faulty GRNs.

Distributed communication in unknown radio networks was first examined in [5]. In this work, the authors presented a randomized communication algorithm which completes broadcast with probability $1 - \epsilon$ within time $O((D + \log n/\epsilon) \log n)$.

In [15], an algorithm was shown to broadcast in arbitrary GRNs in time $O(D)$ under the assumption that nodes have a large amount of knowledge about the network, i.e., given that all nodes have a knowledge radius larger than $R$, the largest communication radius. The authors also showed that algorithms broadcasting in time $O(D + \log n)$ are asymptotically optimal, for unknown GRNs when nodes communicate spontaneously and either can detect collisions or have knowledge of node locations at some positive distance $\epsilon$, arbitrarily small.

More recently, in [16], it was shown that the time of broadcast depends on the
network diameter $D$ and the smallest geometric distance $\alpha$ (denoted $d$ in their paper) between any two nodes. Under the conditional wake-up model, where nodes start transmitting only after hearing a first message, the authors proposed an algorithm that completes broadcasting in time $O(Dg)$. They also proved that, in this context, every broadcasting algorithm requires $\Omega(D\sqrt{g})$ time. Under the spontaneous wake up model, where nodes may transmit from the beginning of the communication process, the authors combined two sub-optimal algorithms into one algorithm, which completes broadcasting in optimal time $O(\min(D + g^2, D\log g))$.

In [17], under the conditional wakeup model, $\Omega(Dg)$ was shown to be the tight lower bound on broadcasting time. However, for networks where nodes locations are restricted to the vertices of a grid of squares of size $\alpha$, the authors proposed an $O(Dg^{5/6}\log g)$-time broadcasting algorithm, thus showing that the broadcast time is not always linearly dependent on $g$.

In [19], the problem of broadcasting in unknown topology networks was proposed given that nodes do not perceive their location accurately and that they do not know the minimum distance $\alpha$ between them. Under the spontaneous wake up model, the authors showed a broadcasting algorithm maintaining optimal time complexity $O(\min(D + g^2, D\log g))$ in these conditions given an upper bound $\alpha/2$ on the inaccuracy of node location perception; beyond this upper bound on inaccuracy, the authors showed that broadcasting is impossible.
Chapter 3

The Diameter and Connectivity of Networks with Random Dependent Faults

3.1 Introduction

In this chapter, we consider the connectivity, diameter and degree of the functional part of graphs with dependent positively correlated faults. A preliminary version of this chapter has appeared in the Proceedings of the 32nd International Symposium on Mathematical Foundations of Computer Science, (MFCS 2007), under the title “Communication in Networks with Random Dependent Faults [34].” A complete version of this chapter has appeared in Networks, under the title “The Diameter and Connectivity of Networks with Random Dependent Faults” [37]. To the best of our knowledge, this is the first analytic work which provides connectivity results for arbitrary graphs under random dependent fault models.

3.1.1 Model and Problem Definition

We consider wired networks, modeled as an undirected graph $G = (V, E)$ with a set of nodes $V$ connected by a set of undirected links $E$. We consider faults which are permanent and of the crash type. For networks both of bounded and unbounded maximum degree $\Delta$, we study the connectivity and the diameter $D$ of their fault-free part. These network properties are related to the feasibility and time to complete communication processes in networks. Under different communication models, the message dissemination time decreases as either $D$, $\Delta$, or both of these parameters decrease. Hence, we seek networks both with low maximum degree $\Delta$ and low diameter $D$.

In order to capture fault dependencies, we introduce the neighborhood fault model, where damaging events, called spots, occur randomly and independently at nodes of
a network, with probability $p$, and cause permanent crash faults in the given node and all of its neighbors. Faults in nodes at distance at most 2 from one another become dependent in this model and are positively correlated. We say that a node is spot-free, when no spot occurs on it.

Throughout this chapter, whenever we discuss the connectivity and diameter of a graph, we refer to these properties of its fault-free part, unless otherwise mentioned. Let $\mathcal{F}$ be a family of graphs. We say that a property $\Phi$ of graphs in $\mathcal{F}$ holds with high probability (w.h.p.), if the probability of the event that a graph $G \in \mathcal{F}$ has property $\Phi$ converges to 1 as the number of nodes of $G$ increases. Thus we say, e.g., that a family of graphs is connected w.h.p., has logarithmic diameter or logarithmic maximum degree w.h.p. When a family $\mathcal{F}$ of graphs is connected w.h.p., we also say that almost certain connectivity holds in $\mathcal{F}$. The empty graph is considered to be disconnected. A path is a sequence of nodes for which each pair of successive nodes shares a link in the graph $G$. A path with no repeated nodes is a simple path. A cycle is a path in which the start and end nodes are the same. A simple cycle is a cycle with no repeated nodes (except the start and end nodes). In this chapter, we only consider simple paths and simple cycles, unless mentioned otherwise.

We focus on the problem of preservation of connectivity and of low diameter properties, both guaranteed with high probability, in spite of faults. We design networks both of bounded and of unbounded maximum degree which are connected and have a small diameter w.h.p., despite relatively high spot probabilities. We also prove bounds on the spot probability such that almost certain connectivity does not hold.

### 3.1.2 Our Results

All our results address the following problem: For which spot probabilities $p$ do there exist networks which remain connected with high probability, and if so, is their diameter small? In our positive results we also seek networks of low maximum degree.

In Section 3.2, we address the questions regarding networks of unbounded degree. We first show that there exists a constant $c$, such that for the values $p \leq 1/(c \log n)$ of the spot probability, there exists an $n$-node graph which, with high probability, has logarithmic degree, is connected and has logarithmic diameter. On the negative side,
we show that, for constant spot probability values $p > 0$, there exist two constants $c_1 < \frac{1}{\log(1/(p(1-p)))}$ and $c_2 > \frac{1}{\log(1/(1-p))}$ such that for any graph $G$: if the maximum degree of $G$ is at most $c_1 \log n$ then $G$ is disconnected with high probability; if the minimum degree of $G$ is at least $c_2 \log n$ then $G$ has all nodes faulty with high probability. In either case, almost certain connectivity does not hold. This leaves some very particular networks undecided. For example, this negative result does not cover the important case of the hypercube, for some constant spot probabilities. Therefore, we study the hypercube separately and prove that, for any constant spot probability $0 < p \leq 1$, almost certain connectivity does not hold. The above should be contrasted with the results from [11, 12] showing that, for independent faults, connectivity and low diameter with high probability hold for arbitrary constant fault probabilities in some graphs, and for small constant fault probability, even in the hypercube.

In Section 3.3, we investigate bounded degree networks. We show that the torus remains connected with high probability when $p \in 1/\omega(n^{1/2})$. However, the diameter of an $n$-node torus is $\Omega(\sqrt{n})$ and the fault-free part has the same large diameter. Hence we seek networks with the same fault-tolerance properties, but with small diameter. We construct a bounded degree network built of tori, whose fault-free part is connected and has diameter $O(\log n)$ with high probability, whenever $p \in 1/\omega(n^{1/2})$; for this network, both $D$ and $\Delta$ are of small magnitude. On the negative side, we show that for the torus and the above constructed network, almost certain connectivity does not hold when $p \in 1/O(n^{1/2})$. Finally, we prove that for networks whose maximum degree $\Delta$ is a constant, almost certain connectivity does not hold when $p \in 1/O(n^{1/\Delta})$.

### 3.2 Networks of Unbounded Degree

In this section, we focus on networks whose maximum degree $\Delta$ is not bounded above by a constant. We first design an $n$-node network which, with high probability, has logarithmic degree, remains connected and has diameter $D \in O(\log n)$ when the spot probability is at most $1/(c \log n)$, for some positive constant $c$. We then establish two bounds on the degree of nodes for which almost certain connectivity does not hold when the spot probability is a positive constant, for a large class of networks.
3.2.1 Connectivity Results

This section is dedicated to proving the following result.

**Theorem 3.2.1** There exist $n$-node graphs which, with high probability, have logarithmic degree and whose fault-free part remains connected with diameter $O(\log n)$, for spot probability $p \leq 1/(c \log n)$, where $c$ is some positive constant.

The network construction is based on a binary tree structure where each node of the tree represents a group of nodes and each link of the tree represents a random set of links between nodes of adjacent groups. To be more precise, for a given $m$, we define a random $n$-node graph $G(n, m, p_l)$. Let $x = \lceil n/m \rceil$ and partition the set of all nodes into subsets $S_1, \ldots, S_x$, of size $m$, called **supernodes** ($S_x$ has size at most $m$). Let $\mathcal{S} = \{S_1, \ldots, S_x\}$ be the set of all supernodes.

Let $L = \lceil \log x \rceil$. Arrange all supernodes into a binary tree $T$ with $L + 1$ levels $0, 1, \ldots, L$, placing each supernode $S_i$ on level $\lceil \log i \rceil$. Level 0 contains the root and levels $L - 1$ and $L$ contain leaves of $T$. The supernode $S_1$ is the root of $T$. For every $1 < i \leq x$, supernode $S_{\lfloor i/2 \rfloor}$ is the parent of $S_i$. Supernodes with even labels are left children and those with odd labels are right children. If a supernode is a parent or a child of another supernode, we say that these supernodes are adjacent in $T$.

The set of edges of $G(n, m, p_l)$ is defined as follows. If supernodes $S_i$ and $S_j$ are adjacent in $T$, then there is an edge in $G(n, m, p_l)$ between any node in $S_i$ and any node in $S_j$ with probability $p_l$, the link probability. Moreover, supernodes have no interior links. The graph $G(n, m, p_l)$ is called a Random Binary Thick Tree (RBTT). See Figure 3.1.

In the remainder of this section, we analyze RBTT and with this graph show that Theorem 3.2.1 holds. In order to prove Theorem 3.2.1, we need to establish an upper bound on the number of spots in each supernode, and a lower bound on the number of functional nodes in each supernode. These two results will be used in the proof of two connectivity lemmas needed in the proof of the theorem. We prove Theorem 3.2.1 by considering the case of the $n$-node RBTT with link probability $p_l = 1/18 \ln n$ and $m = \lceil 1152 \ln^2 n \rceil$ nodes per supernode, when the spot probability
is $p \leq 1/(768 \ln n)$. Hence, we take $c = 768/\ln 2$. For these $m$ and $p_l$, we obtain a graph which has logarithmic maximum degree w.h.p.

Let $C_1$ be the event that all supernodes in $RBTT$ contain fewer than $6 \ln n + 1$ spots.

**Claim 3.2.2** The event $C_1$ occurs with probability at least $1 - 1/n$.

**Proof.** Let $p = k(n)/(768 \ln n)$, with $k(n) \in O(1)$ and $0 < k(n) \leq 1$. Let $N_i$ be the random variable which counts the number of spots in the supernode $S_i$. Since $1152 \ln^2 n \leq m < 1152 \ln^2 n + 1$, we have $(3/2)k(n) \ln n \leq E[N_i] = mp < k(n)((3/2) \ln n + 1/4)$. Using Chernoff bounds ([27], inequality (5)) with parameter $(1 + \epsilon) = 4/k(n)$, under the above assumptions for $k(n)$, we can show that $N_i < 6 \ln n + 1$ with high probability. More precisely,

$$
\Pr[C_1] = \Pr[(\forall S_i \in \mathcal{S}) \ N_i < 6 \ln n + 1]
\geq 1 - n \sum_{S_i \in \mathcal{S}} \Pr[N_i \geq (4/k(n))k(n)((3/2) \ln n + 1/4)]
\geq 1 - n \left(e^{(4/k(n)-1)} \left(\frac{4}{k(n)}\right)^{(4/k(n))}k(n)((3/2) \ln n)\right)^k(n)((3/2) \ln n)
\geq 1 - n \left(e^{(4/k(n)-1)} \left(\frac{k(n)e}{4}\right)^{(4/k(n))}k(n)((3/2) \ln n)\right)
= 1 - n \left(e^{(4/k(n)-1)} \left(\frac{k(n)e}{4}\right)^{6 \ln n}\right)
= 1 - n \left(e^{(4/k(n)-1)} \left(\frac{k(n)e}{4}\right)^{6 \ln n}\right)
\geq 1 - 1/n
$$
for all \( k(n) \in O(1) \) where \( 0 < k(n) \leq 1 \). Clearly, for \( p = 0 \) the claim holds. 

For any constant \( 0 < \epsilon \leq 1 \), let \( C_2 \) be the event that each supernode in \( RBTT \) has more than \( 288(1 - \epsilon) \ln^2 n \) functional nodes.

**Claim 3.2.3** The event \( C_2 \) occurs with probability at least \( 1 - 1/n^{d\log n} \), for some constant \( d > 0 \).

**Proof.** Fix the set of spots in the graph. Pick a spot-free node \( u \). The event \( FF_u \) that the node \( u \) is fault-free occurs if no link of \( u \) has a spot as an endpoint. Since each supernode is adjacent to at most 3 other supernodes, \( u \) may have a link to at most \( 18 \ln n + 3 \) spots, if event \( C_1 \) holds. It follows, for large enough \( n \), that

\[
\Pr[FF_u] \geq \Pr[FF_u \land C_1] \\
= \Pr[FF_u | C_1] \Pr[C_1] \\
\geq \left(1 - \frac{1}{18 \ln n}\right)^{18 \ln n + 3} \Pr[C_1] \\
\geq \frac{1}{4} \left(1 - \frac{1}{n}\right).
\]

Let \( N_i \) be the random variable which counts the number of functional nodes in \( S_i \). Since under event \( C_1 \) at most \( 6 \ln n + 1 \) spots occur in any supernode, at least \( m - 6 \ln n - 1 \) nodes are spot-free, if \( C_1 \) occurs. Thus,

\[
E[N_i] \geq \Pr[C_1] \cdot (m - 6 \ln n - 1) \cdot \Pr[FF_u] \\
\geq \left(1 - \frac{1}{n}\right) \cdot (1152 \ln^2 n - 6 \ln n - 1) \cdot \frac{1}{4} \left(1 - \frac{1}{n}\right) \\
\geq 288 \left(1 - \frac{6}{1152 \ln n} - \frac{1}{1152 \ln^2 n}\right) \left(1 - \frac{1}{n}\right)^2 \ln^2 n
\]

and therefore \( E[N_i] \geq 288(1 - \epsilon') \ln^2 n \), for any constant \( 0 < \epsilon' \leq 1 \), if \( n \) is sufficiently large. For a fixed set of spots, the events \( FF_u \) are independent for all nodes, i.e., for nodes \( u \neq v \), the probability of the event \( FF_u \) is the same whether \( FF_v \) occurs or not. Thus we use a Chernoff bound ([27], inequality (7)) to show that, with high probability, \( N_i \geq 288(1 - \epsilon')(1 - \epsilon'') \ln^2 n \) for any constant \( 0 < \epsilon'' \leq 1 \). It follows that the probability of this event is

\[
\Pr[C_2] = \Pr[(\forall S_i \in S) N_i > 288(1 - \epsilon')(1 - \epsilon'') \ln^2 n]
\]
\[
\geq 1 - \sum_{S_i \in S} \Pr[N_i \leq 288(1 - \epsilon')(1 - \epsilon'')\ln^2 n]
\]
\[
\geq 1 - \sum_{S_i \in S} e^{-(\epsilon'')^2 288(1 - \epsilon')\ln^2 n/2}
\]
\[
\geq 1 - nn^{-144(1 - \epsilon')(\epsilon'')^2 \ln n}
\]
\[
\geq 1 - n^{-d \log n}
\]

for some constant \(d > 0\).

Using the previous results, we now present two connectivity lemmas in preparation for the proof of the main theorem of this section.

**Lemma 3.2.4** All functional nodes are connected to at least one functional node in each supernode adjacent to their own, with probability exceeding \(1 - 1/n^{13}\).

**Proof.** Fix a node \(u\). Let \(N(u)\) denote the set of supernodes adjacent to the supernode containing \(u\). Consider the event \(\gamma_{u,S}\) that \(u\) has a link to at least one functional node in a given supernode \(S \in N(u)\). The event \(\gamma_{u,S}\) occurs unless all links from \(u\) to functional nodes in \(S\) do not exist. It follows from the above and from Claim 3.2.3 that, for constant \(0 < \epsilon \leq 1\),

\[
\Pr[\gamma_{u,S}] \geq \Pr[\gamma_{u,S} \cap C_2] = \Pr[C_2] \Pr[\gamma_{u,S} | C_2]
\]
\[
\geq \Pr[C_2] \left(1 - (1 - 1/18 \ln n)^{288(1 - \epsilon')\ln^2 n}\right)
\]
\[
\geq \left(1 - n^{-d \ln n}\right) \left(1 - n^{-16(1 - \epsilon)}\right).
\]

This holds, in particular, for very small constant values of \(\epsilon\) and hence \(\Pr[\gamma_{u,S}] \geq 1 - n^{-15}\). Since the graph contains at most \(n\) functional nodes, which should be connected to at least one functional node in at most 3 supernodes, the estimated probability is at least

\[
\Pr[(\forall u \in V (\forall S \in N(u))) \gamma_{u,S}] \geq 1 - \sum_{u \in V} \sum_{S \in N(u)} \Pr[\neg \gamma_{u,S}]
\]
\[
\geq 1 - 3nn^{-15} > 1 - n^{-13}.
\]

**Lemma 3.2.5** All functional node pairs in supernodes at distance 3 from one another are connected by a fault-free path of length 3 with probability at least \(1 - 1/n^{1.9}\).
Proof. This lemma is proven in steps, defining connectivity probabilities and lower bounds on the number of connected nodes at distances 1, 2, and 3.

Fix 4 distinct supernodes, $S_r, S_s, S_t, S_u$, which form a simple path in $RBTT$, i.e., $S_r$ is adjacent to $S_s$, which is adjacent to $S_t$, which is adjacent to $S_u$.

Fix a node $r$ in $S_r$. Let $X_s$ be the random variable which counts the number of functional nodes $s \in \Gamma(r)$ located in $S_s$. From Claim 3.2.3, each supernode contains more than $288(1-\epsilon)\ln^2 n$ fault-free nodes, for any constant $0 < \epsilon \leq 1$ with probability $1-1/n^{d\log n}$, for some positive constant $d$. In the $RBTT$ construction, the probability of a link existing between nodes of adjacent supernodes is $p_l = 1/18 \ln n$. Hence,

$$E[X_s] \geq \Pr[C_2]\frac{288(1-\epsilon)\ln^2 n}{18 \ln n} = (1 - n^{-d\log n})16(1 - \epsilon)\ln n \geq (1 - \epsilon')\ln n,$$

with some constant $1 > \epsilon' > \epsilon$. We use a Chernoff bound ([27], inequality (7)) with parameter $\epsilon = \sqrt{3/(8(1 - \epsilon'))}$ to show that a fixed functional node has at most $16(1 - \sqrt{3/(8(1 - \epsilon'))})(1 - \epsilon')\ln n$ such neighbors with probability

$$\Pr \left[ X_s \leq 16 \left( 1 - \sqrt{\frac{3}{8(1 - \epsilon')}} \right) (1 - \epsilon') \ln n \right] \leq e^{-\left(\frac{3}{8(1 - \epsilon')}\right)^2 \frac{16(1 - \epsilon')\ln n}{2}} = n^{-3}.$$
use a Chernoff bound ([27], inequality (7)) to show that, if event $A$ occurs, event $B$ occurs with probability $1 - 1/n^{k \log n}$, for some positive constant $k$.

Assume event $A \cap B$ occurs. Fix a functional node $u$ in $S_u$. Fix a subset $S' \subseteq S$ of functional nodes, each of which is a neighbor of some element of $S$, with size $72 (1 - \epsilon'' \ln^2 n)$. Denote by $P_{S'u}$ the event that there exists a link between node $u$ and some node in $S'$. This event occurs unless $u$ has no link to any node in $S'$. Hence,

$$
\Pr[P_{S'u} | B \cap A] = \left(1 - \left(1 - 1/18 \ln n \right)^{72 (1 - \epsilon'' \ln^2 n)}\right) \\
\geq 1 - e^{-72 (1 - \epsilon'' \ln^2 n)/(18 \ln n)} \geq 1 - n^{-4(1 - \epsilon')}.
$$

Consider the event $P_{rstu}$ that there exists a fault-free path of the form $rstu$ from a fixed node $r$ to a fixed node $u$. The event $P_{rstu}$ may occur even if $A$ or $B$ do not occur. Hence,

$$
\Pr[P_{uijk}] \geq \Pr[P_{S'u} \cap B \cap A] = \Pr[A] \Pr[B | A] \Pr[P_{S'u} | B \cap A] \\
\geq \left(1 - n^{-3}\right) \left(1 - 1/n^{k \log n}\right) \left(1 - n^{-4(1 - \epsilon')}\right) \geq \left(1 - n^{-3 + \epsilon''} \right),
$$

for some small constant $\epsilon'' > 0$, and in particular for $\epsilon'' = 0.1$.

There are at most $n$ functional nodes in $RBT T$, each with $O(\log^2 n)$ other functional nodes in supernodes at distance 3. Hence, there are $O(n \log^2 n)$ functional node pairs in supernodes at distance 3. It follows that all node pairs in supernodes at distance 3 are connected by a fault-free path of length 3 with probability at least $1 - n^{-1.9}$.

Combining the previous lemmas, we now prove Theorem 3.2.1.

**Proof of Theorem 3.2.1.** The $RBT T$ contains $\Theta(n/ \log^2 n)$ supernodes connected in a binary-tree structure. Hence, the diameter $D$ of the $RBT T$ (before the spots appear) is $\Theta(\log n)$. It follows from the construction that the maximum degree of the $RBT T$ is $O(\log n)$, with high probability. By Lemma 3.2.5, all functional node pairs in supernodes at distance 3 are connected by at least one fault-free path of length 3 with probability greater than $1 - 1/n^{1.9}$. Therefore, all functional nodes in the subgraph $RBT T'$ composed of the root supernode $S_1$ and of all supernodes at distances a multiple of 3 from $S_1$ are connected with this probability. Moreover, functional nodes not in $RBT T'$ are in supernodes adjacent to supernodes in $RBT T'$. Thus, by Lemma
3.2.4, all these functional nodes are connected to at least one functional node in \( RBTT' \) with probability exceeding \( 1 - 1/n^{1.8} \). Hence, with probability exceeding \( 1 - 1/n^{1.8} \), the fault-free part of \( RBTT \) is connected.

We now investigate the diameter of the fault-free part of \( RBTT \). From the above argument, we first observe that nodes \( u \) and \( v \) located in supernodes of \( RBTT' \) at distance \( d_{u,v} \) in the binary tree structure \( T \) are connected with high probability by a path of length \( d_{u,v} \) in \( RBTT \). Moreover, functional nodes \( s \) and \( t \) in all other supernodes at distance \( d_{s,t} \) in the binary tree structure \( T \) are connected with high probability by a path of length at most \( d_{s,t} + 4 \). Since the diameter of the binary tree structure \( T \) is \( \Theta(\log n) \) and since all supernodes of the \( RBTT \) contain functional nodes with high probability, it follows from the above argument that the diameter of the fault-free part of \( RBTT \) is also \( \Theta(\log n) \), with high probability.

3.2.2 Non-Connectivity Results

We have shown that it is possible to build logarithmic-degree graphs which remain connected with high probability in spite of spot probabilities \( p \leq 1/(c \log n) \), for some positive constant \( c \). The natural question then is whether it is possible to build arbitrarily large networks which remain connected with high probability despite larger spot probabilities. In what follows, we show that for constant spot probabilities, a large family of networks does not have this property. More formally, the following theorem holds.

**Theorem 3.2.6** For any constant spot probability \( p > 0 \), there exist constants \( c_1 \) and \( c_2 \) such that in an \( n \)-node graph \( G \): if the maximum degree of \( G \) is at most \( c_1 \log n \) then the fault-free part of the graph is disconnected with high probability; if the minimum degree of \( G \) is at least \( c_2 \log n \) then the graph has all nodes faulty with high probability. In either case, almost certain connectivity does not hold.

In order to prove Theorem 3.2.6, we need a few lemmas. First, we describe an event which implies that a small fault-free part of the graph is surrounded by faulty nodes, and bound its probability. We then determine the size of a proper set of nodes for which these events are independent in the graph. With these lemmas, we bound
the maximum degree of graphs for which the fault-free part is not connected, with high probability. We then bound the minimum degree of graphs such that all nodes are faulty, with high probability.

We say that a node $u$ is *insular* if it is not faulty and all its neighbors who have neighbors at distance 2 from $u$ are faulty. Denote by $I_u$ the event that $u$ is insular.

**Lemma 3.2.7** In a graph of maximum degree $\Delta$, $\Pr[I_u] \geq (1 - p)^{\Delta+1}p^\Delta$, for spot probability $p$.

**Proof.** An equivalent definition of event $I_u$ is as follows: No node in $\Gamma_{\leq 1}(u)$ is a spot and, for each node $i \in \Gamma(u)$, either there exists a node $g \in \Gamma(i) \cap \Gamma_2(u)$ which is a spot, or $\Gamma(i) \cap \Gamma_2(u) = \emptyset$. In order to obtain a lower bound on the probability of event $I_u$, we now describe the construction of an elementary event $\varphi_u \subseteq I_u$: Sequentially, for each node $i \in \Gamma(u)$, mark one unmarked neighbor node which is not in $\Gamma_{\leq 1}(u)$ and label it $m_i$. If no such node exists, then mark nothing. Denote by $M$ the set of nodes marked by this process. Clearly, $|M| \leq |\Gamma(u)| \leq \Delta$. The event $\varphi_u$ occurs if all nodes in $M$ are spots and all nodes in $\Gamma_{\leq 1}(u)$ are not.

By definition, if all nodes in $M$ are spots, then each node $i \in \Gamma(u)$ for which there exists a node $m_i \in M$, is faulty. Still by definition, all nodes $i \in \Gamma(u)$ for which there is no $m_i \in M$ are adjacent to a $m_u \in M$, $i \neq u$ or have no neighbor at distance 2 from $u$. Therefore, if all marked nodes are spots and all nodes in $\Gamma_{\leq 1}(u)$ are not spots, the event $I_u$ is implied. See Figure 3.2.

It follows from the preceding discussion that

$$\Pr[I_u] \geq \Pr[\varphi_u] = (1 - p)^{|\Gamma_{\leq 1}(u)|}p^{|M|} \geq (1 - p)^{\Delta+1}p^\Delta.$$

**Lemma 3.2.8** For all $n$-node graphs with maximum degree at most $\Delta$, there exist more than $n/\Delta^{2h+1}$ nodes $c_i$ with disjoint $h$-hop neighborhoods $\Gamma_{\leq h}(c_i)$.

**Proof.** Take any graph whose maximum degree is at most $\Delta$. Algorithm $C$ constructs a sequence $\{c_0, c_1, \ldots\}$ of nodes with disjoint $h$-hop neighborhoods $\Gamma_{\leq h}(c_i)$. 


Figure 3.2: Building an elementary event \( \varphi_u \): \( \Gamma(u) = \{a, b, c\} \); for node \( a \), no node is marked because its neighbors are in \( \Gamma_{\leq 1}(u) \); for node \( b \), node \( e \) is marked and labeled \( m_b \); for node \( c \) no node is marked because its neighbor \( e \) outside \( \Gamma_{\leq 1}(u) \) is already marked and labeled \( m_b \). If \( m_b \) is a spot and \( u, a, b, c \) are not spots, then \( u \) is insular.

Algorithm \( C \)

\[
i \leftarrow 0
\]

\[
c_0 \leftarrow \text{pick a node } v \in V
\]

mark nodes in \( \Gamma_{\leq 2h}(c_0) \)

repeat

\[
i \leftarrow i + 1
\]

\[
c_i \leftarrow \text{pick an unmarked node adjacent to a marked node}
\]

mark nodes in \( \Gamma_{\leq 2h}(c_i) \)

until all \( v \in V \) are marked

We now prove by induction that the sequence \( \{c_0, c_1, \ldots\} \) satisfies the statement of the lemma. For the sequence \( \{c_0\} \), this is trivial. By inductive hypothesis, assume \( \{c_0, c_1, \ldots, c_i\} \) has been constructed such that there is no pair \( c_j \neq c_k \) with intersecting sets \( \Gamma_{\leq h}(c_j) \) and \( \Gamma_{\leq h}(c_k) \). Since, by construction, all nodes fewer than \( 2h + 1 \) hops away from any node \( c_j, j \leq i \), are marked and \( c_{i+1} \) is chosen from the unmarked node set, it is at least \( 2h + 1 \) hops away from any \( c_j, j < i + 1 \). Furthermore, for any two nodes \( u, v \) with distance at least \( 2h + 1 \), \( \Gamma_{\leq h}(u) \) and \( \Gamma_{\leq h}(v) \) are disjoint. Therefore, the inductive hypothesis holds for \( \{c_0, c_1, \ldots, c_{i+1}\} \).

Finally, for a graph whose maximum degree is at most \( \Delta \), there are fewer than \( \sum_{k=0}^{2h} \Delta^k = \Delta^{2h+1} \) nodes which can be marked at every step of the construction. Therefore, the above construction defines a sequence \( C \) of more than \( \frac{n}{\Delta^{2h+1}} \) nodes.
With Lemma 3.2.7 and Lemma 3.2.8, we can show the following result.

**Lemma 3.2.9** Let $0 < p < 1$ be a constant spot probability. For any constant $c_1 < \frac{1}{\log(1/(p(1-p)))}$, any graph with maximum degree at most $c_1 \log n$ has the fault-free part disconnected, with high probability.

**Proof.** In Lemma 3.2.7, we bound the probability of event $I_u$, that a fixed node $u$ is insular. With an upper bound $c_1 \log n$ on the degree of the nodes,

$$
\Pr[I_u] \geq (1-p)^{c_1 \log n + 1} p^{c_1 \log n} = (1-p)^{c_1 \log(1-p)} n^{c_1 \log p} = \frac{1-p}{n^{c_1 \log(1/(p(1-p)))}}.
$$

Let $\mathcal{S}$ be a set of nodes in the graph for which the events $I_u$ are independent. Since event $I_u$ occurs if spots are located at distance 2 from $u$, the occurrence of this event is independent for nodes whose neighborhoods of radius at most 2 are disjoint. Furthermore, since all nodes have degree at most $c_1 \log n$, by Lemma 3.2.8, we have $|\mathcal{S}| \geq n/(c_1 \log n)^5$. Partition $\mathcal{S}$ in two sets $\mathcal{S}_1$ and $\mathcal{S}_2$, each containing at least $n/(3(c_1 \log n)^5)$ nodes. By showing that inside each of these sets there is at least one insular node, we show the existence of two functional nodes with no fault-free path joining them, i.e., we show that the graph is disconnected. The probability estimate of this event is the same for sets $\mathcal{S}_1$ and $\mathcal{S}_2$. We have, for the set $\mathcal{S}_1$,

$$
\Pr[\exists u \in \mathcal{S}_1 I_u] = 1 - \Pr[\forall u \in \mathcal{S}_1 \neg I_u] \\
\geq 1 - \Pr[\neg I_u]^{\mathcal{S}_1} \\
\geq 1 - \left(1 - \frac{1-p}{n^{c_1 \log(1/(p(1-p)))}}\right)^{n/(3(c_1 \log n)^5)} \\
\geq 1 - e^{-\frac{(1-p)n^{1-c_1 \log(1/(p(1-p)))}}{3(c_1 \log n)^5}}.
$$

Take $c_1 < \frac{1}{\log(1/(p(1-p)))}$. Then we have $1 - c_1 \log(1/(p(1-p))) > 0$ and $\Pr[\exists u \in \mathcal{S}_1 I_u]$ converges to 1 as $n$ grows. It follows that $\Pr[\exists u \in \mathcal{S}_1 I_u] \Pr[\exists v \in \mathcal{S}_2 I_v]$ also converges to 1 as $n$ grows.

**Lemma 3.2.10** Let $0 < p < 1$ be a constant spot probability. For any constant $c_2 > \frac{1}{\log(1/(1-p))}$, any graph with minimum degree at least $c_2 \log n$ has all nodes faulty with high probability.
Proof. Consider a node $u$ with at least $c_2 \log n$ neighbors. Let $F_u$ denote the event that $u$ is faulty. $F_u$ does not occur if none of $u$ and its neighbors is a spot. Let $F_V$ denote the event that all nodes in the graph are faulty. We can bound the probability of this event as

$$\Pr[F_V] = \Pr[\forall u \in V F_u] = 1 - \Pr[\exists u \in V \neg F_u] \geq 1 - \sum_{u \in V} \Pr[\neg F_u] = 1 - n \cdot (1 - p)^{c_2 \log n + 1} = 1 - n \cdot (1 - p) \cdot n^{c_2 \log(1 - p)} = 1 - \frac{(1 - p)}{n^{c_2 \log(1/(1 - p))} - 1}.$$ 

Take $c_2 > \frac{1}{\log(1/(1 - p))}$, then we have $c_2 \log(1/(1 - p)) - 1 > 0$ and $\Pr[F_V]$ converges to 1 as $n$ grows. 

Notice that if $p = 1$ then all nodes are faulty with probability 1. Hence, Lemmas 3.2.9 and 3.2.10 imply Theorem 3.2.6.

Theorem 3.2.6 indicates that almost certain connectivity does not hold for a large class of graphs, when the spot probability is a positive constant. However, the bounds $c_1 \log n$ and $c_2 \log n$ do not coincide. Since $c_1 < \frac{1}{\log(1/(1 - p))}$ and $c_2 > \frac{1}{\log(1/(1 - p))}$, we have $c_1 < c_2$ for all positive values of $p$. It remains open whether or not there exists an arbitrarily large graph which remains connected with high probability despite constant spot probabilities.

Remark 3.2.11 Choosing $c_2 = \frac{1}{\log(1/(1 - p))}$ and $p = 1/2$ in the argument of Lemma 3.2.10 yields $\Pr[F_V] = 1/2$ and $c_2 = 1$. Moreover, the probability of this event increases with $p$. This covers the important case of the $n$-node hypercube for $p \geq 1/2$.

We will now attempt to provide insight into the question of what happens when node degrees lie between these bounds. For example, when $p = 1/2$, we have $c_1 < 1/2$ and $c_2 > 1$. From Remark 3.2.11, the $n$-node hypercube is disconnected at least with some constant probability when $p \geq 1/2$, however Theorem 3.2.6 does not cover this graph for values of $p < 1/2$. We will investigate this important case in the following section.
3.2.3 The Hypercube

The hypercube $H_k$ of dimension $k$ is a $2^k$-node graph with the set of nodes having labels from $\{0, 1\}^k$ and the set of links between nodes whose labels have a Hamming distance of 1. Hence, the hypercube $H_k$ of dimension $k = \log n$ contains $n$ nodes and has diameter $\log n$. We now show the following theorem.

**Theorem 3.2.12** Let the spot probability $p$ be a positive constant. Then the probability that the fault-free part of $H_k$ is connected is bounded above by a constant strictly less than 1 as $n \to \infty$.

From Remark 3.2.11, for constant $p \geq 1/2$, the fault-free part of the hypercube is empty with at least a positive constant probability, which implies that it is connected with a probability bounded above by a constant less than 1. We now show that for constant $0 < p < 1/2$, the fault-free part of the hypercube is disconnected with high probability. This will prove Theorem 3.2.12.

In order to prove that the fault-free part of the hypercube is disconnected for constant $0 < p < 1/2$, we first show for these values of $p$ that there exists at least one pair of functional nodes with some minimum constant distance in the hypercube. We will then show that there exists no fault-free path of sufficient length to connect this pair of nodes.

**Claim 3.2.13** Assume constant spot probability $0 < p < 1/2$. For any positive integer constant $c$, the $n$-node hypercube $H_k$ contains at least one pair of fault-free nodes at Hamming distance at least $c$, with high probability.

**Proof.** Consider event $FF_u$ that a node $u$ is fault-free. Event $FF_u$ occurs if and only if there is no spot on nodes in $\Gamma_{\leq 1}(u)$. Furthermore, if two nodes $u$ and $v$ are at distance at least 3, then $FF_u$ and $FF_v$ are independent. For the hypercube, an improvement of the Gilbert-Varshamov bound for binary codes due to Varshamov [57] implies the existence of a set $S_3$ of nodes at distance at least 3 from each other with size $|S_3| \in \Omega(\frac{2^k}{k})$. Let $|S_3| \geq \frac{a2^k}{k}$, for some constant $\alpha > 0$. (Since the distance between nodes in $H_k$ is the Hamming distance between their labels, the result directly
transfers from binary codes to hypercubes.) It follows from the preceding discussion that

\[
\Pr[(\exists u \in V) \text{ } FF_u] = 1 - \Pr[\forall u \in V \neg FF_u] \\
\geq 1 - (1 - \Pr[FF_u])^{\lvert S_3 \rvert} \\
\geq 1 - \left(1 - (1 - p)^{k+1}\right)^{\frac{\alpha n}{\log n}} \\
= 1 - \left(1 - (1 - p)(1 - p)^{\log n}\right)^{\frac{\alpha n}{\log n}} \\
= 1 - \left(1 - \frac{(1 - p)^n}{\alpha n}\right)^{\frac{\alpha n}{\log n}}.
\]

For \( p < 1/2 \), \( \log 1/(1 - p) < 1 \) and therefore, we have at least one functional node with high probability.

Pick one functional node \( u \) and, without loss of generality, label the nodes of \( H_k \) so that \( u \) has the label \( 0^k \). Pick a positive integer \( c \) and consider the subset of nodes with labels \( 1^c \{0, 1\}^{k-c} \). These nodes induce a sub-hypercube \( H'_{k-c} \) of \( H_k \) with nodes at Hamming distance at least \( c \) from \( 0^k \). Varshamov’s result applied to \( H'_{k-c} \) shows a set of nodes \( S'_3 \) at Hamming distance at least 3 from each other with size \( \lvert S'_3 \rvert \in \Omega(\frac{2k-c}{k-c}) \in \Omega(\frac{n}{\log n}) \). The above probabilistic argument applied to \( H'_{k-c} \) proves the lemma.

**Lemma 3.2.14** The fault-free part of the \( n \)-node hypercube \( H_k \) is disconnected with high probability, for constant spot probability \( 0 < p < \frac{1}{2} \).

**Proof.** Let \( \epsilon \) be an arbitrary positive constant and let \( c = (1 + \epsilon)/\log(1/(1 - p)) \). For any constant \( 0 < p < 1/2 \), \( c \) is a positive constant. Since \( 0 < p < 1/2 \), by Claim 3.2.13, there is a pair \((u, v)\) of functional nodes at some Hamming distance \( L \) bounded below by a constant \( l = [c] \geq c \), with high probability. For nodes \( u \) and \( v \) to be connected, there must exist a fault-free path of length at least \( L \) between them. Since each path of length at least \( l \) includes a subpath of length \( l \), the event \( E \) that all paths of length \( l \) are faulty implies the event that all paths of length at least \( L \) are faulty. In particular, \( E \) implies that all paths of length at least \( L \) are faulty. In what follows, we show that \( E \) occurs with high probability.

Consider any path \( x = x_1x_2\ldots x_l \) of length \( l \) in \( H_k \). For the path \( x \) to be fault-free, the set \( S \) of all nodes \( x_1, x_2, \ldots, x_l \) and all their neighbors must be spot-free. We now
determine the size of the set $S$. For any node $x_i$ in the path, let $Z(x_i)$ be the set of neighbors of $x_i$ outside the path. We have $|Z(x_i)| \geq k - l - 1$ for any $i = 1, \ldots, l$. Furthermore, in view of the properties of the hypercube, we have $|Z(x_i) \cap Z(x_j)| \leq 2$ for any $i \neq j$. There are fewer than $l^2$ pairs $\{i, j\}$ with $i \neq j$. Hence, the size of $S$ is at least

\[ l + l(k - l - 1) - 2l^2 = lk - 3l^2. \]

Denote $\delta = 3l^2$. Since $l$ is constant, so is $\delta$. Furthermore, each node in $H_k$ is an endpoint to, at most, $\log^{l-1} n$ paths of length $l$; there are at most $n \log^{l-1} n$ paths of length $l$ in $H_k$. Let $P_l$ be the set of all paths of length $l$. Let $P_x$ be the event that a path $x$ is fault-free. Since $l \geq \lceil c \rceil \geq 1 + \frac{1+ \epsilon}{\log(1/(1-p))}$, it follows from the preceding discussion that

\[
\Pr[(\exists x \in P_l) \ P_x] \leq \sum_{x \in P_l} \Pr[P_x] \\
\leq n \log^{l-1} n \cdot (1 - p)^{lk - \delta} \\
= \frac{\log^{l-1} n}{(1 - p)^{\delta}} \cdot n(1 - p)^{l \log n} \\
= \frac{\log^{l-1} n}{(1 - p)^{\delta}} \cdot n^{1 - l \log(1/(1-p))} \\
\leq \frac{\log^{l-1} n}{(1 - p)^{\delta} n^c} \\
\to 0.
\]

From Remark 3.2.11, the fault-free part of the hypercube is empty with at least some constant probability for $p \geq 1/2$, hence, for these values of $p$, the probability of existence of a single non-empty connected component is at most some constant strictly less than 1. From Lemma 3.2.14, the fault-free part of the hypercube is disconnected with probability converging to 1 as $n \to \infty$ when $p < 1/2$, thus, for these values of $p$, the fault-free part is connected with a probability bounded above by a constant strictly less than 1. Together, Remark 3.2.11 and Lemma 3.2.14 show that the probability that the fault-free part of the hypercube is connected is bounded above by a constant strictly less than 1 for constant $p > 0$. Hence, Remark 3.2.11 and Lemma 3.2.14 imply Theorem 3.2.12.
3.3 Networks of Bounded Degree

The RBTT presented in Section 3.2 remains connected despite relatively high spot probabilities. However, its degree is unbounded. Therefore, it is natural to ask if bounded-degree networks can also remain connected with high probability with comparable spot probabilities.

In this section we construct bounded-degree networks which remain connected despite inverse polynomial spot probabilities and for which the diameter of the fault-free part is of smallest possible order. On the negative side, we prove that bounded-degree networks can tolerate no more than inverse polynomial spot probabilities.

3.3.1 Connectivity Results

We now study the properties of two networks: the torus and a torus-based tree-like network that we call the Toroidal Tree. We show that the $n$-node torus remains connected with high probability despite spot probability $1/\omega(n^{1/2})$. However, the diameter of the torus is quite large. Thus we design a tree-like structure based on the torus which provides the same fault-tolerance properties and has diameter $\Theta(\log n)$, even with faults.

The Torus

In this section, we show an upper bound on the spot probability such that the fault-free part of the torus remains connected. Denote by $T_{m\times k}$ the $m \times k$ torus with $m, k \geq 4$. The torus has the set of nodes $\{u = (u_x, u_y) : u_x \in \{0, 1, \ldots, m - 1\}, u_y \in \{0, 1, \ldots, k - 1\}\}$ and the set of links $\{(u, v) : |u_x - v_x| \mod m + |u_y - v_y| \mod k = 1\}$.

**Theorem 3.3.1** The fault-free part of the $n$-node torus $T_{m\times k}$, $m, k \geq 4$, is connected with high probability for $p \in 1/\omega(n^{1/2})$.

To prove Theorem 3.3.1, we show that, if spots are far enough from each other, fault-free node pairs are connected by paths which circumvent all faulty nodes.

**Claim 3.3.2** For each node $u$ in $T_{m\times k}$, $m, k \geq 4$, there exists a cycle in the set $\Gamma_2(u) \cup \Gamma_3(u)$ which includes all nodes in $\Gamma_2(u)$.
Proof. Since the torus is vertex-symmetric, we need only prove the statement for a fixed node \((x, y)\). The cycle illustrated in Figure 3.3 lies in the set \(\Gamma_2((x, y)) \cup \Gamma_3((x, y))\) and includes all nodes of \(\Gamma_2((x, y))\).

![Figure 3.3: Cycle on \(\Gamma_2((x, y)) \cup \Gamma_3((x, y))\)](image)

Lemma 3.3.3 The fault-free part of \(T_{m \times k}\), \(m, k \geq 4\), is connected if all spots are at distance at least 5 from each other.

Proof. Consider any pair of spots \(i \neq j\) at distance at least 5. For the spot \(i\), all nodes within \(\Gamma_2(i)\) and \(\Gamma_3(i)\) are respectively at least at distance 3 and 2 from \(j\). It follows that if all spots are at least at distance 5 from each other, then each spot \(s\) is surrounded by a cycle \(C_s\) of functional nodes as described in Claim 3.3.2.

Consider a path \(P\) between two functional nodes \(u\) and \(v\) for which at least one node is faulty. The faulty nodes in \(P\) are in a subpath located at distance less than 2 from a fixed spot \(S\). The nodes \(w\) and \(x\) adjacent to this subpath in \(P\) are at distance 2 from \(S\). Since all nodes in \(\Gamma_2(S)\) are included in a cycle \(C_s\) of functional nodes around \(S\), there exists a path of functional nodes between \(w\) and \(x\). It follows from the above argument that for each faulty subpath of a faulty path between functional nodes \(u\) and \(v\), there exists a functional alternative subpath and hence, all functional nodes are connected. See Figure 3.4.
We now prove Theorem 3.3.1.

**Proof of Theorem 3.3.1.** Let $D_5$ be the event that all spots are at distance at least 5 from each other. By Lemma 3.3.3, the event $D_5$ implies that the fault-free part of the torus is connected. Alternatively, $D_5$ is the event that, for all spots $i$, no set $\Gamma_{\leq 5}(i)$ contains a spot $j \neq i$. We note that $|\Gamma_{\leq 5}(i)| \leq 61$.

Let $S$ be the set of spots. Let $D_{5,i}$ be the event that, for spot $i$, the set $\Gamma_{\leq 5}(i)$ contains no spot $j \neq i$. The event $D_5$ is the intersection of all $D_{5,i}$, $i \in S$. For a fixed spot $i$, the probability of the event $D_{5,i}$ is $Pr[D_{5,i}] = (1 - p)^{|\Gamma_{\leq 5}(i)| - 1} > 1 - 60p$.

Let the spot probability be $p \in \frac{1}{\omega\left(\frac{n}{2}\right)}$. Hence $p = \frac{1}{\omega(n^{1/2})}$, where $f(n) \to \infty$ as $n \to \infty$. Let $A$ be the event that there are at most $2n^{1/2}$ spots. We have $Pr[D_5] \geq Pr[D_5 \cap A] = Pr[\bigcap_{i \in S} D_{5,i} \cap A] = Pr[A] Pr[\bigcap_{i \in S} D_{5,i} | A]$. We first prove that $Pr[\bigcap_{i \in S} D_{5,i} | A] \to 1$, as $n \to \infty$. Suppose that $A$ holds.

$$Pr\left[\bigcap_{i \in S} D_{5,i}\right] = 1 - Pr\left[\bigcup_{i \in S} \neg D_{5,i}\right]$$

$$\geq 1 - \sum_{i \in S} Pr[\neg D_{5,i}]$$

$$\geq 1 - 2n^{1/2}(1 - (1 - 60p))$$

$$= 1 - 2n^{1/2} \frac{60}{n^{1/2} f(n)}$$

$$= 1 - \frac{2 \cdot 60}{f(n)} \to 1 \text{ as } n \to \infty.$$
It follows that $\Pr[\bigcap_{i \in S} D_{5,i} | A] \to 1$ as $n \to \infty$. Moreover, using a Chernoff bound ([27], inequality (5)), with parameter $(1 + \epsilon) = 2f(n)$ we can show that $\Pr[A] \geq 1 - ((e/(2f(n)))^{2n})$. Hence, $\Pr[A] \to 1$ as $n \to \infty$. The two above facts imply that $\Pr[A] \Pr[\bigcap_{i \in S} D_{5,i} | A] \to 1$ as $n \to \infty$, and hence $\Pr[D_5] \to 1$ as well.

The Toroidal Tree

We now design a bounded-degree network which provides the same fault-tolerance as the torus, while also having logarithmic diameter. Since the diameter of a bounded-degree graph is at least logarithmic, this is optimal. The network construction is based on two binary trees, $T$ and $T'$, connected by a link between their root nodes. Each node of $T, T'$ represents a group of nodes, and groups adjacent in $T, T'$ have a subset of nodes in common. More precisely, for constant $k \geq 4$, we define the $n$-node graph $G(n, k)$. Assume that the set of nodes can be partitioned exactly as described below; this is easy to modify in the general case, by adding nodes to a leaf group.

Let the sets $T_1, \ldots, T_x$ and $T'_1, \ldots, T'_x$ be tori with $2k$ rows $\{0, 1, \ldots, 2k-1\}$ and $k$ columns $\{0, 1, \ldots, k-1\}$; $|x - x'| \leq 1$. We describe the construction for the tree $T$; the same construction is applied for the tree $T'$. Arrange all $T_i$ as the nodes of $T$, with $L + 1$ levels $0, 1, 2, \ldots, L$, placing each $T_i$ on level $\lceil \log i \rceil$ of $T$. Level 0 contains the root of $T$ and levels $L - 1$ and $L$ contain the leaves. For every $1 < i \leq x$, $T_{\lfloor i/2 \rfloor}$ is the parent of $T_i$. Supernodes with even labels are left children and those with odd labels are right children. Identify row 0 of each child torus with row $k$ of its parent torus to connect them. Identify the row 0 of both roots in $T$ and $T'$ to connect the trees $T$ and $T'$. Note that there are three types of nodes in this graph. Nodes that belong to one torus $T_i$ have degree 4, nodes that belong to two tori $T_i$ (at row 0 of $T_i$ and $T'_i$) have degree 6, and nodes that belong to three tori $T_i$ have degree 8.

It follows from the above description that $x + x' = \lfloor (n - 2k^2)/(2k - 1)k) \rfloor + 1$ tori are located on $L = \lfloor \log(x + x' + 1) \rfloor$ levels in $G(n, k)$. The graph $G(n, k)$ is called a Toroidal Tree. See Figure 3.5.

Lemma 3.3.4 The fault-free part of the Toroidal Tree is connected if all spots are at distance at least 5.
Proof. Since the Toroidal Tree is build of $k \times 2k$ tori $\mathcal{T}_i$, $k \geq 4$, and since faults are at least at distance 5 from each other, by Lemma 3.3.3 the fault-free part of each $\mathcal{T}_i$ is connected. Since $k \geq 4$, at least one common node is functional for each pair of adjacent tori. Therefore, the fault-free part of each adjacent pair of tori is connected and the lemma follows.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{toroidal_tree.png}
\caption{The Toroidal Tree: each planar region is a constant degree torus identical to $\mathcal{T}_1$. The line where planar regions intersect represents the set of nodes shared by adjacent tori.}
\end{figure}

Theorem 3.3.5 For $p \in 1/\omega(n^{1/2})$, the fault-free part of the $n$-node Toroidal Tree remains connected and has diameter $O(\log n)$, with high probability.

Proof. To prove connectivity of the fault-free part of the graph, we repeat the proof of Theorem 3.3.1, substituting Lemma 3.3.3 with Lemma 3.3.4. More precisely, let $D_5$ be the event that all spots are at distance at least 5 from each other. By Lemma 3.3.4, the event $D_5$ implies connectivity of the Toroidal Tree. Moreover, we observe that, for this graph, $|\Gamma_{\leq 5}(u)| \leq 161$. Substituting 161 for 61 in the probability estimation for $\Pr[D_5]$ in Theorem 3.3.1, we find that, for $p \in 1/\omega(n^{1/2})$, the Toroidal Tree is connected with high probability.

Since every torus $\mathcal{T}_i$ forming the Toroidal Tree has a constant number of nodes, for each of them, the diameter of its fault-free part is bounded. Moreover, since the tori are arranged in a binary tree, the diameter of the fault-free part of the Toroidal Tree is $O(\log n)$. By construction, the Toroidal Tree has maximum degree 8.
3.3.2 Non-Connectivity Results

In this section, we show that almost certain connectivity of bounded-degree graphs does not hold for relatively small spot probabilities. We first show that the bounds on the spot probability provided in Theorem 3.3.1 and Theorem 3.3.5 are tight both for the torus and for the Toroidal Tree. We then show that for arbitrary bounded-degree networks, if spot probability is the inverse of some polynomial, then almost certain connectivity does not hold.

The Torus and Toroidal Tree

We first show a lower bound on the spot probability, such that almost certain connectivity of the torus does not hold. Then, we extend the proof to show a similar lower bound for the Toroidal Tree. The following lower bounds match the upper bounds from Theorem 3.3.1 and Theorem 3.3.5, thus showing that the results are tight.

Theorem 3.3.6 For spot probability $p \in 1/O(n^{1/2})$, almost certain connectivity of the fault-free part of the $n$-node torus $T_{m \times k}$ does not hold.

Proof. Let $J_u$ be the event that a node $u$ is functional and has 2 spots close to it, such that all neighbors of node $u$ are faulty (see Figure 3.6). The probability of this event is $\Pr[J_u] \geq (1 - p)^7 p^2$. If two events $J_u$, $J_v$ occur, then there exist two functional nodes with no fault-free path connecting them in the torus, i.e., the graph is disconnected. We will now show that two such nodes exist with at least some positive constant probability, when $p \in 1/O(n^{1/2})$.

![Figure 3.6: Torus disconnection: two spots disconnect a torus.](image)

We observe that the events $J_u$ are independent for nodes which are at distance 5 from each other. By Lemma 3.2.8, there exists a set $\mathcal{S}$ of nodes at distance at least 5...
from each other, such that $|S| = kn$, for some constant $k$. Partition $S$ into two sets $S_1$ and $S_2$, each containing at least $kn/3$ nodes. Let $p \geq c/n^{1/2}$, with $c$ a positive constant. The probability of the event that at least one node in the set is insular is the same for sets $S_1$ and $S_2$. We have, for the set $S_1$,

$$
\Pr[\exists u \in S_1 J_u] = 1 - \Pr[\forall u \in S_1 \neg J_u]
= 1 - (1 - \Pr[J_u])^{|S_1|}
\geq 1 - (1 - (1 - p)^2p^2)^{kn/3}
\geq 1 - (1 - (1/2)c^2/n)^{kn/3}
\geq 1 - e^{-kc^2/6}, a positive constant.
$$

Therefore, $\Pr[\exists u \in S_1 J_u] \Pr[\exists v \in S_2 J_v]$ is also a positive constant and almost certain connectivity of the graph does not hold.

Consider now the Toroidal Tree. Inside the set of constant size component tori forming the Toroidal Tree, there is a set $S$ of linear size with the same property as for the torus. Hence, using the same argument we can show the following result.

**Theorem 3.3.7** For spot probability $p \in 1/O(n^{1/2})$, almost certain connectivity of the fault-free part of the $n$-node Toroidal Tree does not hold.

**Arbitrary Graphs of Bounded Degree**

We showed in the preceding section that in the case of the $n$-node torus and Toroidal Tree, spot probabilities $p \in 1/O(n^{1/2})$ cannot be tolerated for these graphs to remain connected with high probability. In the following theorem, we show that a similar phenomenon occurs for all graphs whose degree is bounded by a constant.

**Theorem 3.3.8** Let $G$ be an $n$-node graph with maximum degree $\Delta$, for $\Delta$ a fixed positive integer. Let the spot probability $p$ be $\in \Omega(n^{-1/\Delta})$. Then the probability that the fault-free part of $G$ is connected is bounded above by a constant strictly less than 1.

**Proof.** Consider the event $I_u$ that a node $u$ is insular, as defined in Section 3.2. The graph $G$ is disconnected, in particular, if the event $I_u$ occurs for two different nodes. We estimate the probability of this event.
Let $p \geq c/n^{1/\Delta}$, with $c$ a positive constant. By Lemma 3.2.7, a fixed node $u$ is insular with probability $Pr[I_u] \geq (1 - p)^{\Delta + 1} p^\Delta$. From Lemma 3.2.8 we know that any bounded degree graph has a set $S$ of at least $n/\Delta^{2h+1}$ nodes with disjoint $h$-radius neighborhoods. Since event $I_u$ depends only on the status of nodes at distance 2 from $u$, the occurrence of these events is independent for nodes $u$ which have radius 2 disjoint neighborhoods, i.e., for $h = 2$. Partition $S$ into two sets $S_1$ and $S_2$, each containing at least $n/(3\Delta^5)$ nodes. The probability that one node in the set is insular is the same for both sets $S_1$ and $S_2$. We estimate, for the set $S_1$,
\[
Pr[\exists u \in S_1 \ I_u] = 1 - Pr [\forall u \in S_1 \neg I_u]
\geq 1 - (1 - (1 - p)^{\Delta + 1} p^\Delta)^{|S_1|/\Delta^5}
\geq 1 - \left(1 - \frac{1}{2} \left(\frac{c}{n^{1/\Delta}}\right)^{\Delta}\right)^{\frac{3n^5}{\Delta^5}}
= 1 - \left(1 - \frac{1}{2} \cdot \frac{\Delta}{n}\right)^{\frac{3n^5}{\Delta^5}},
\]
which is larger than some positive constant for all values of $n$.

Therefore, $Pr[\exists u \in S_1 \ I_u] Pr[\exists v \in S_2 \ I_v]$ is also larger than some positive constant and almost certain connectivity of $G$ does not hold.

3.4 Conclusion

We provided analytic results on connectivity and diameter of the fault-free part of arbitrary networks in the presence of dependent, positively correlated faults. We introduced the neighborhood fault model where damaging events, called spots, occur randomly and independently at nodes of a network with probability $p$, and cause faults in the affected node and its neighbors.

Our results show clear differences between the assumption of independent faults and that of the neighborhood fault model. For example, while under independent faults with small constant fault probability $p > 0$ the fault-free part of the hypercube remains connected with high probability [12], this is not the case under the neighborhood fault model with any positive constant spot probability. Likewise, the fault-free part of the torus is connected with high probability for fault probability $p \in 1/\omega(n^{1/4})$.
when faults are independent, but this is not the case for such spot probabilities under the neighborhood fault model.

In the following chapter, we will focus on wireless networks and study the problem of communication with spatially correlated faults.
Chapter 4

Communication in Random Geometric Radio Networks with Positively Correlated Random Faults

4.1 Introduction

In this chapter, we consider the feasibility and time of communication in random geometric radio networks, where nodes fail randomly with positive correlation. A preliminary version of this chapter appeared in the Proceedings of the ADHOC-NOW 2008 [35] and a complete version was published in Ad Hoc and Sensor Wireless Networks [36]. Both versions have the title "Communication in random geometric radio networks with positively correlated random faults". To the best of our knowledge, this is the first analytic work which provides connectivity results for GRNs under random dependent fault models.

4.1.1 Model and Problem Definitions

We seek to model a network composed of mobile stations moving under the Random Waypoint mobility model [29] inside an open region, e.g., a train station or a plaza. Under this mobility model, mobile stations alternately move and pause for random amounts of time, choosing a direction, distance and speed at random at every movement phase. Here, we assume that the mobile stations move at low, pedestrian-like speeds, making the network appear static for the short duration of communication processes; the distances and directions of movement are chosen in some uniform way. We further assume that neither the boundaries of the open region nor the other mobile stations have any effect on the mobile station movements. Hence, any snapshot of the graph is a set of stations distributed on a plane by a Poisson Process. Due to the short duration of the communication processes, we consider that the faults are
permanent. The proposed static model also applies to networks of sensors spread randomly in hostile environments where individual placement and replacement of units is not possible.

We focus attention on a unit square region of the plane. Node locations occur with Poisson arrival rate \( n \). We fix a parameter \( r \), called the communication range; any two nodes at geometric distance at most \( r \) from one another can exchange messages directly. We define the \textit{ranged spot} fault model. Damaging phenomena, called \textit{spots}, occur on the plane with Poisson arrival rate \( \lambda \). Some examples of damaging phenomena are lightning strikes (and other electrostatic discharges), electromagnetic pulses and explosions. We fix a parameter \( s > 0 \), called the \textit{spot range}. Each spot causes permanent crash faults in all nodes within distance \( s \) of it, i.e., inside the disk of radius \( s \) centered at it, which we call the \textit{spot disk}. For a fixed spot \( i \), we denote the corresponding spot disk by \( D_i \). Faulty nodes can neither send nor receive messages for the entire communication process. More formally, consider the undirected fault-free graph \( G(V, E) \), where \( V \) is the random set of nodes occurring on the unit square with Poisson arrival rate \( n \), and \( E \) is the set of all node pairs \( \{u, v\} \) for which the Euclidean distance is at most \( r \). Let \( S \) be the set of spots which occur on the unit square with Poisson arrival rate \( \lambda \). Let \( F \) be the set of faulty nodes, i.e., all the nodes in \( V \) whose location is within distance \( s \) from at least one spot in \( S \). We consider the graph \( G[V'] \) induced on \( G \) by the set \( V' = V \setminus F \) of all functional nodes. To remind the reader how it is built, throughout this chapter, we will denote the graph \( G[V'] \) by \( U(n, r, \lambda, s) \).

As usual in wireless network algorithms, communication is assumed synchronous in the graph \( U(n, r, \lambda, s) \). All communication is done using the same base frequency, modulation, and encoding, hence using a single channel; we use the radio communication model. Furthermore, we work under the assumption that nodes cannot detect collisions.

In this chapter, we study the question of feasibility and efficiency of communication in the fault-free graph \( U(n, r, \lambda, s) \). Specifically, we say that a graph is connected w.h.p. when the event that it is connected occurs w.h.p. On the other hand, we say that a graph is not connected w.h.p. when the event that it is disconnected occurs
with constant probability greater than 0.

4.1.2 Our Results

In Section 4.3, we study the question for which spot arrival rate \( \lambda \) is the graph \( U(n, r, \lambda, s) \) non-empty w.h.p. We give answers to the question for which parameters \( s = s(n) \) and \( \lambda = \lambda(n, s) \) there exist any fault-free nodes in the unit square, i.e., when the fault-free graph \( U(n, r, \lambda, s) \) is non-empty, w.h.p. For \( s \in o(1) \), we find a threshold function \( l(n, s) \) and constants \( L_1, L_2 \) such that, for \( \lambda \geq L_1 \cdot l(n, s) \) fault-free nodes do not exist, w.h.p., while for \( \lambda \leq L_2 \cdot l(n, s) \) they do exist w.h.p. For \( s \in \Omega(1) \), we show that, for \( \lambda \in o(1/s^2) \) fault-free nodes do not exist, w.h.p., and for \( \lambda \in o(1/s^2) \) they do exist, w.h.p.

Later, in Section 4.4, we explore the question for which \( \lambda \) is the graph \( U(n, r, \lambda, s) \) connected w.h.p. We give answers to the question for which parameters \( s = s(n) \), \( r = r(n) \) and \( \lambda = \lambda(n, s, r) \), the fault-free graph \( U(n, r, \lambda, s) \) is connected, w.h.p. Connectivity is equivalent to feasibility of communication in our setting. We restrict attention to the case of small spot range, more precisely, we work under the assumption \( s \in o(r) \). In the case \( r \in o(1) \), we find a threshold function \( c(n, s, r) \) and constants \( C_1, C_2 \) such that, for \( \lambda \geq C_1 \cdot c(n, s, r) \) the graph \( U(n, r, \lambda, s) \) is not connected w.h.p., and for \( \lambda \leq C_2 \cdot c(n, s, r) \) it is connected, w.h.p. Then, in the case \( r \in \Omega(1) \), and for \( \lambda \in o(1/s^2) \), we show that for the values of \( \lambda \) for which the graph \( U(n, r, \lambda, s) \) contains at least one node w.h.p., it is also connected w.h.p.

Finally, in Section 4.5, we propose a fast and reliable broadcasting algorithm. Under the additional restriction on the spot range, when \( s \in o(1/\sqrt{n}) \), we show an algorithm which accomplishes broadcast with probability at least \( 1 - \epsilon \) in time \( O(D + \log 1/\epsilon) \) in the graph \( U(n, r, \lambda, s) \) of diameter \( D \).

Before we show these results, we present, in Section 4.2, some background information and preparatory lemmas.

4.2 Background and Preliminaries

The following theorem is referred to as Chernoff bounds (cf., e.g., [27]).
Theorem 4.2.1 Let $X$ be the random variable which counts the number of successes in a series of independent Bernoulli trials. The probability that the value of $X$ will deviate from $E[X]$ is bounded as follows:

\[
\begin{align*}
\Pr[X \geq (1 + \epsilon)E[X]] &\leq \left(\frac{e^\epsilon}{(1+\epsilon)^{1+\epsilon}}\right)^{E[X]}, \ \epsilon > 0 \\
\Pr[X \geq (1 + \epsilon)E[X]] &\leq e^{-\epsilon^2 E[X]/3}, \ 0 \leq \epsilon \leq 1 \\
\Pr[X \leq (1 - \epsilon)E[X]] &\leq e^{-\epsilon^2 E[X]/2} \leq \left(\frac{e^\epsilon}{(1+\epsilon)^{1+\epsilon}}\right)^{E[X]}, \ 0 \leq \epsilon \leq 1
\end{align*}
\]

In the following theorem, Chernoff bounds are applied to Poisson processes in a plane.

Theorem 4.2.2 Let $X$ be the random variable which counts Poisson arrivals, with rate $\lambda$, on a region $R$ with area $|R| = A$. Then,

\[
\begin{align*}
\Pr[X \geq (1 + \epsilon)\lambda A] &\leq \left(\frac{e^\epsilon}{(1+\epsilon)^{1+\epsilon}}\right)^{\lambda A}, \ \epsilon > 0 \\
\Pr[X \geq (1 + \epsilon)\lambda A] &\leq e^{-\epsilon^2 \lambda A/3}, \ 0 \leq \epsilon \leq 1 \\
\Pr[X \leq (1 - \epsilon)\lambda A] &\leq e^{-\epsilon^2 \lambda A/2} \leq \left(\frac{e^\epsilon}{(1+\epsilon)^{1+\epsilon}}\right)^{\lambda A}, \ 0 \leq \epsilon \leq 1
\end{align*}
\]

Proof. Consider the Poisson process on the region $R$ with arrival rate $\lambda$. Partition the region $R$ into a mesh of squares of dimensions $1/\sqrt{\lambda h} \times 1/\sqrt{\lambda h}$, with $h$ a variable parameter. In the region $R$, there are $A\lambda h$ such squares. For a fixed square $i$ of the partition, let $arrival_i$ be the event that at least one arrival occurs inside this square. We have $\Pr[arrival_i] = 1 - e^{-\lambda/(\lambda h)} = 1 - e^{-1/h}$. Let $X$ be the random variable which counts the number of times that the events $arrival_i$ occur, for $i = 1, 2, \ldots, A\lambda h$. It follows from the above that $E[X] = A\lambda h(1 - e^{-1/h})$. By expanding the exponential, the expression becomes

\[
E[X] = A\lambda h(1 - e^{-1/h}) = A\lambda h(1 - (1 - 1/h + (1/h)^2/2 - \ldots)) = A\lambda h(1/h - (1/h)^2/2 + \ldots) = A\lambda(1 - (1/h)/2 + (1/h)^3/3! - \ldots)
\]

and hence, for $h = h(\lambda) \in \omega(1)$, we have

\[
A\lambda(1 - (1/h)/2) \leq E[X] \leq A\lambda.
\]
Since \( \lim_{h \to \infty} A \lambda \left(1 - (1/h)/2\right) = A \lambda \), it follows that \( \lim_{h \to \infty} E[X] = A \lambda \) as well.

Since each square has the same area and since events of the Poisson process occur randomly and independently inside the squares, we have a series of Bernoulli trials with mean arbitrarily close to \( A \lambda \). Hence, Chernoff bounds may be used, as in Theorem 4.2.1, to bound the value of \( X \).

Theorem 4.2.2 implies the following remark.

**Remark 4.2.3** For any \( \lambda A \in \Omega(1) \), there exists a constant \( k \) such that \( X < k \lambda A \) w.h.p. For any \( \lambda A \in o(1) \), there exists a constant \( k \) such that \( X < k \) w.h.p.

**Proof.** Let \( f(n) \in \omega(1) \), then

\[
\Pr[X \geq f(n) \lambda A] \leq \left( \frac{e^{f(n)-1}}{(f(n))^{f(n)}} \right)^{\lambda A} \leq \left( \frac{e}{f(n)} \right)^{\lambda A f(n)} \to 0 \text{ as } n \to \infty.
\]

for \( \lambda A \in \Omega(1) \) and \( f(n) \in \omega(1) \) and for \( \lambda A \in o(1) \) and \( f(n) \in \omega(\lambda A) \). Hence, in the first case, there must exist a function \( k(n) \in o(\omega(1)) = O(1) \) such that the probability converges, at least, to a constant, i.e., the function \( k(n) \) must be smaller than all the functions which are larger than all constants (hence \( k(n) \in O(1) \)). In the second case, must exist a function \( k(n) \in o(\omega(\lambda A)) = O(\lambda A) \) such that the probability converges, at least, to a constant, i.e., the function \( k(n) \) must be smaller than all the functions which grow faster than \( \lambda A \) (hence \( k(n) \in O(\lambda A) \)).

In the sequel, let \( F_v \) be the event that a fixed node \( v \) is faulty, i.e., that there exists at least one spot within distance \( s \) of it. Then, for spot arrival rate \( \lambda \) we have

\[
\Pr[F_v] = 1 - e^{-\lambda \pi s^2}.
\]

While distant faults are independent, the presence of a faulty node within distance \( 2s \) from a fixed node \( v \) implies that there is a spot which might be close enough to \( v \) to make it faulty, i.e., the occurrence of a fault at a node can never decrease the probability of a fault on another node. This is why faults are positively correlated. Hence, the following fact applies to the events \( F_v \).

**Fact 1** For any set \( Z \) of nodes,

\[
\Pr[\bigcap_{v \in Z} F_v] \geq \prod_{v \in Z} \Pr[F_v].
\]
In our considerations, we use various partitions of the unit square into meshes of subsquares of size $\frac{1}{\sigma} \times \frac{1}{\sigma}$. This approach is widely used to solve problems related to communication, and in particular, to build communication algorithms. For simplicity of presentation, we always assume that $\sigma$ is an integer and thus we ignore the issue of incomplete squares. Since our results are of asymptotic nature, this extra assumption can be easily removed. Moreover, we assume throughout the chapter that squares of the partition include their top and left edges, but not their right and bottom edges.

4.3 Liveness of the Graph

In this section, we show bounds on the spot arrival rate $\lambda$ for which functional nodes exist in the unit square, i.e., the graph $U(n, r, \lambda, s)$ contains at least one node, w.h.p. We say that the graph $U(n, r, \lambda, s)$ is alive if it contains at least one node; otherwise, we say that it is dead. In the sequel, let $alive_R$ be the event that a region $R$ of the unit square contains at least one functional node; let $dead_R$ be the event that the region $R$ contains no functional node; the events alive and dead written without subscripts denote the respective events for the unit square. In the following sections, we will prove Theorems 4.3.1 and 4.3.2.

**Theorem 4.3.1** For $s \in o(1)$, there exist two positive constants, $L_1$ and $L_2$, such that if the spot arrival rate $\lambda \geq L_1 \frac{\ln(\min\{n, 1/s^2\})}{s^2}$, then the graph $U(n, r, \lambda, s)$ is dead, w.h.p., and if $\lambda \leq L_2 \frac{\ln(\min\{n, 1/s^2\})}{s^2}$, then $U(n, r, \lambda, s)$ is alive, w.h.p.

**Theorem 4.3.2** For $s \in \Omega(1)$, the graph $U(n, r, \lambda, s)$ is dead w.h.p. if $\lambda \in \omega(1/s^2)$ and alive w.h.p. if $\lambda \in o(1/s^2)$.

Theorem 4.3.1 will follow from Claims 4.3.3, 4.3.4 and 4.3.7. Theorem 4.3.2 will follow from Claims 4.3.5 and 4.3.9.

4.3.1 Event dead

**Claim 4.3.3** For all $s$, the graph $U(n, r, \lambda, s)$ is dead w.h.p. when $\lambda = \alpha \frac{\ln n}{\pi s^2}$, for any constant $\alpha > 1$. 
Proof. Recall that $V$ is the set of all nodes inside the unit square. Consider the event $\text{dead}$. We have

$$\Pr[\text{dead}] = \Pr\left[\bigcap_{v \in V} F_v\right] \geq \prod_{v \in V} \Pr[F_v],$$

by Fact 1. By Theorem 4.2.2, for any constant $k > 1$, there are at most $kn$ nodes in the graph, w.h.p. For $\lambda = \ln(nf(n))/(\pi s^2)$, $f(n) \in \omega(1)$ we have

$$\Pr[\text{dead}] \geq (\Pr[F_v])^kn = (1 - e^{-\lambda \pi s^2})^kn = (1 - 1/(nf(n)))^kn \to 1 \text{ as } n \to \infty.$$ 

In particular, for $\alpha > 1$ and $f(n) = n^{\alpha - 1} \in \omega(1)$, the spot arrival rate is

$$\lambda = \ln(n \cdot n^{\alpha - 1})/(\pi s^2) = \alpha \ln n/(\pi s^2)$$

and thus the claim holds.

Partition the unit square into a mesh of $2/s^2$ squares of dimensions $s/\sqrt{2} \times s/\sqrt{2}$. Let $Q$ be the set of all these squares. Each square $q \in Q$ has area $s^2/2$ and has all its points within distance $s$ from all points in the concentric disk $d_q$ of diameter $s$. For a fixed square $q$, the occurrence of a spot in $d_q$ implies the event $\text{dead}_q$. In turn, the conjunction of $\text{dead}_q$ for all squares $q$ implies the event $\text{dead}$. We have, by Fact 1,

$$\Pr[\text{dead}] = \Pr\left[\bigcap_{q \in Q} \text{dead}_q\right] \geq \prod_{q \in Q} \Pr[\text{dead}_q] = (\Pr[\text{dead}_q])^{2/s^2}.$$ 

We prove that the event $\text{dead}$ occurs w.h.p. respectively for the cases $s \in O(1)$ and $s \in \omega(1)$ in Claims 4.3.4 and 4.3.5.

Claim 4.3.4 For $s \in o(1)$ and any constant $\alpha > 1$, the graph $U(n, r, \lambda, s)$ is dead w.h.p. when $\lambda = \frac{4\alpha \ln(1/s^2)}{\pi s^2}$.

Proof. For $s \in o(1)$ and $\lambda = \frac{4\ln(f(n)/s^2)}{\pi s^2}$, $f(n) \in \omega(1)$ we have

$$\Pr[\text{dead}] \geq (\Pr[\text{dead}_q])^{2/s^2} = (1 - e^{-\lambda \pi s^2/4})^{2/s^2} \geq (1 - e^{-\ln(f(n)/s^2)})^{2/s^2}$$

$$= (1 - s^2/f(n))^{2/s^2} \to 1 \text{ as } n \to \infty.$$ 

In particular, for $\alpha > 1$ and $f(n) = (1/s^2)^{\alpha - 1} \in \omega(1)$, the spot arrival rate is

$$\lambda = 4 \ln((1/s^2) \cdot (1/s^2)^{\alpha - 1})/(\pi s^2) = 4\alpha \ln(1/s^2)/(\pi s^2)$$

and thus the claim holds.
Claim 4.3.5 For $s \in \Omega(1)$, the graph $U(n, r, \lambda, s)$ is dead, w.h.p., when $\lambda \in \omega(1/s^2)$.

Proof. When $s \in \Omega(1)$, there is a constant $c$ such that the partition of the unit square contains $c$ squares. Hence, for $\lambda = f(n)/s^2$, $f(n) \in \omega(1)$ we have

$$\Pr[\text{dead}] \geq (1 - e^{-\lambda\pi s^2/4})^c \geq 1 - ce^{-f(n)\pi/4} \to 1 \text{ as } n \to \infty.$$ 

\[\square\]

4.3.2 Event alive

Recall the event $F_v$ that a fixed node $v$ is faulty. In our model, these events are positively correlated for all nodes within distance $2s$ from one another. A set $S$ of nodes for which the events $F_u$ are independent is a set whose elements have a distance greater than $2s$ from one another. Call such sets sparse. In the following lemma, we show that there exist large sparse sets.

Lemma 4.3.6 A square $A$ with area $|A|$ contains a sparse set $S$ of size at least $k|A| \min\{n, 1/s^2\}$, for some positive constant $k$, w.h.p., if $|A| \min\{n, 1/s^2\} \to \infty$ as $n \to \infty$.

Proof. Partition the square $A$ into a mesh of $s \times s$ squares. Group the squares in $3 \times 3$ matrices of squares, called blocks. Let $B$ be the set of all blocks. For each block $b \in B$ label the central square $c_b$. For any two blocks $b, b' \in B$, $b \neq b'$, the squares $c_b$ and $c_{b'}$ are at distance greater than $2s$ from one-other, and hence, the events $F_u$ and $F_v$ for the nodes $u \in c_b$ and $v \in c_{b'}$ are independent. Hence, a sparse set of nodes can be constructed by choosing one node inside each $c_b$, $b \in B$, which contains at least one node. Thus the size of the sparse set $S$ is at least as large as the number of squares $c_b$ which contain at least one node. For each block $b \in B$ let $C_b$ be the event that the central square $c_b$ contains at least one node. Since nodes occur on the unit square with Poisson arrival rate $n$, the number of nodes in a $s \times s$ square is a Poisson random variable with parameter $ns^2$. For a fixed block $b$, the event $C_b$ has probability $1 - e^{-ns^2}$. Let $X$ be the random variable which counts the number of times that events $C_b$ occur, for $b \in B$. The expected value of $X$ is

$$E[X] = \frac{|A|}{9s^2}(1 - e^{-ns^2}).$$
We consider 2 cases: 1) $s^2 \in o(1/n)$ and 2) $s^2 \in \Omega(1/n)$.

Case 1: When $s^2 \in o(1/n)$, we expand the exponential and obtain

$$E[X] = \frac{|A|}{9s^2}(1 - (1 - ns^2 + (ns^2)^2/2 - \ldots)) \geq \frac{|A|}{9s^2}(0.9ns^2) = \frac{|A|n}{10}.$$ 

Case 2: When $s^2 \in \Omega(1/n)$, $\Pr[C_b]$ converges to a positive constant and we have

$$E[X] = \frac{|A|}{9s^2}(1 - e^{-ns^2}) \geq \frac{|A|k}{9s^2}, \text{ for some constant } k > 0.$$ 

Since the random variable $X$ counts the number of successes (events $C_b$) in a series of independent Bernoulli trials, we use Theorem 4.2.1 to show, whenever $E[X] \to \infty$, that, w.h.p., $(0.9)E[X] < X < (1.1)E[X]$. Since $|S| \geq X$, the lemma follows.

Using Lemma 4.3.6, we now prove Claim 4.3.7.

**Claim 4.3.7** For $s \in o(1)$, the graph $U(n, r, \lambda, s)$ is alive with high probability when $\lambda = \frac{\beta \ln(\min\{n, 1/s^2\})}{\pi s^2}$, for any constant $\beta < 1$.

**Proof.** We consider the cases $s \leq 1/\sqrt{n}$ and $s \geq 1/\sqrt{n}$ (with $s \in o(1)$). In both cases, we define the event *sparse* that there exists a sparse set $S$ of size, respectively at least $k'n$ (in the first case) and at least $k'/s^2$ (in the second case), for some constant $k'$. In both cases, the event *alive* is implied if at least one node in $S$ is not faulty. Assume that the event *sparse* holds.

**Case 1:** $s \leq 1/\sqrt{n}$.

Let $\lambda = \ln(nf(n))/(\pi s^2)$, $f(n) \in o(1)$. Then we have

$$\Pr[F_v] = 1 - e^{-\lambda s^2} = 1 - e^{-\ln(f(n))}$$

and

$$\Pr[\text{alive}] = 1 - \Pr[\bigcap_{v \in V} F_v] \geq 1 - \Pr[\bigcap_{v \in S} F_v] = 1 - (\Pr[F_v])^{|S|}$$

$$\geq 1 - \left(1 - e^{-\ln(f(n))}\right)^{k'n} = 1 - \left(1 - \frac{1}{nf(n)}\right)^{k'n}$$

$$\to 1 \text{ as } n \to \infty.$$ 

**Case 2:** $s \geq 1/\sqrt{n}$ (with $s \in o(1)$).

Let $\lambda = \ln(f(n)/s^2)/(\pi s^2)$, $f(n) \in o(1)$. Then we have

$$\Pr[F_v] = 1 - e^{-\lambda s^2} = 1 - e^{-\ln(f(n)/s^2)}$$
and
\[ \Pr[\text{alive}] \geq 1 - \left(1 - e^{-\ln(f(n)/s^2)}\right)^{k'/s^2} = 1 - \left(1 - s^2/f(n)\right)^{k'/s^2} \to 1 \]
as \( n \to \infty \). In particular, for \( \beta < 1 \) and \( f(n) = \min\{n, 1/s^2\}^{\beta-1} \in o(1) \), the spot arrival rate is
\[ \lambda = \ln(\min\{n, 1/s^2\} \cdot \min\{n, 1/s^2\}^{\beta-1})/(\pi s^2) = \beta \ln(\min\{n, 1/s^2\})/(\pi s^2). \]
Moreover, since \( \min\{n, 1/s^2\} \to \infty \) for \( s \in o(1) \), by Lemma 4.3.6, the event \textit{sparse} occurs w.h.p. in both cases. Hence the claim follows. \( \blacksquare \)

**Remark 4.3.8** For \( s \in o(1/\sqrt{n}) \) and \( \lambda = \frac{\ln(cn)}{\pi s^2} \), where \( c \) is a positive constant, the graph \( U(n, r, \lambda, s) \) is dead with constant probability.

**Proof.** In the proofs of Claims 4.3.3 and 4.3.7 the probability estimates are both bounded away from 0 as \( n \to \infty \) for any \( f(n) \in \Theta(1) \). \( \blacksquare \)

**Claim 4.3.9** For \( s \in \Omega(1) \), the graph \( U(n, r, \lambda, s) \) is alive w.h.p. when \( \lambda \in o(1/s^2) \).

**Proof.** Consider the event \textit{free} that no spot occurs in the \( 3s \times 3s \) square concentric with the unit square. Under this event, there exists a square of constant area \( s^2 \leq 1 \) where any node will be fault-free and inside the unit square. Thus, for \( \lambda = \frac{f(n)}{s^2} \), \( f(n) \in o(1) \), we have
\[ \Pr[\text{alive}] \geq \Pr[\text{alive}|\text{free}] \Pr[\text{free}] \geq (1 - e^{-ns^2})e^{-\lambda \cdot 9s^2} \]
\[ = (1 - e^{-s^2n})e^{-9f(n)} = (1 - e^{-s^2n})(1 - 9f(n) + (9f(n))^2/2 - \ldots) \]
\[ \geq (1 - e^{-s^2n})(1 - 9f(n)) \to 1 \text{ as } n \to \infty. \]

\( \blacksquare \)

### 4.4 Connectivity of \( U(n, r, \lambda, s) \)

In the preceding section, we have shown a threshold for the spot arrival rate for which the graph \( U(n, r, \lambda, s) \) is non-empty w.h.p. We now answer the question for which spot arrival rate the graph \( U(n, r, \lambda, s) \) is connected w.h.p.
It has been shown, in [49], that for any real number \( c \), if \( r \geq \sqrt{\frac{\ln n + c}{\pi n}} \), then the probability that the graph \( U(n, r, \lambda, s) \), with \( \lambda = 0 \), is connected is at least \( e^{-e^{-c}} \), as \( n \to \infty \). If we substitute \( e^{-c} = f(n) \), assume that \( f(n) \in o(1) \) and recall that \( e^{-f(n)} = 1 - f(n) + f(n)^2/2 - \ldots \geq 1 - f(n) \), then we see that if

\[
r \geq \sqrt{\frac{\ln n + \ln 1/f(n)}{\pi n}}
\]

then

\[
\Pr[U(n, r, 0, s) \text{ is connected}] \geq 1 - f(n).
\]

Hence, it is natural to investigate the connectivity of the graph \( U(n, r, \lambda, s) \) under the assumption \( r^2 \geq \frac{\ln n + \ln 1/f(n)}{\pi n} \), for some \( f(n) \in o(1) \), when we know that connectivity is guaranteed w.h.p. without faults. In what follows we make this assumption.

The main results of this section are Theorems 4.4.1 and 4.4.2. In Theorem 4.4.1, for spot range \( s \) of lower order of magnitude than the communication range \( r \) and for \( r \in o(1) \), we show a threshold for the spot arrival rate \( \lambda \) below which the graph \( U(n, r, \lambda, s) \) is connected w.h.p. and above which it is not. For the case \( r \in \Omega(1) \), the separation is different: in Theorem 4.4.2, we show thresholds for the spot arrival rate \( \lambda \) below which the graph \( U(n, r, \lambda, s) \) is connected w.h.p. and above which it is dead w.h.p.

**Theorem 4.4.1** For \( s \in o(r) \) and \( r \in o(1) \), there exist two positive constants, \( C_1 \) and \( C_2 \), such that if \( \lambda \geq C_1 \ln\left(\frac{\ln\min\{n, 1/s^2\}}{\ln(1/r^2)}\right)/s^2 \), then the graph \( U(n, r, \lambda, s) \) is not connected w.h.p., and if \( \lambda \leq C_2 \ln\left(\frac{\ln\min\{n, 1/s^2\}}{\ln(1/r^2)}\right)/s^2 \), then the graph \( U(n, r, \lambda, s) \) is connected, w.h.p.

Theorem 4.4.1 will follow from Lemmas 4.4.3, 4.4.4, 4.4.5, and 4.4.6.

**Theorem 4.4.2** For \( s \in o(r) \) and \( r \in \Omega(1) \),

1. if \( s \in o(1) \), then there exist two positive constants, \( C_3 \) and \( C_4 \), such that
   
   (a) for \( \lambda \leq C_3 \ln(\min\{n, 1/s^2\})/s^2 \), the graph \( U(n, r, \lambda, s) \) is connected, w.h.p.
   
   (b) for \( \lambda \geq C_4 \ln(\min\{n, 1/s^2\})/s^2 \), the graph \( U(n, r, \lambda, s) \) is dead w.h.p.
2. if $s \in \Omega(1)$, then

(a) for $\lambda \in o(1/s^2)$, the graph $U(n, r, \lambda, s)$ is connected, w.h.p.

(b) for $\lambda \in \omega(1/s^2)$, the graph $U(n, r, \lambda, s)$ is dead w.h.p.,

Theorem 4.4.2 will follow from Theorems 4.3.1 and 4.3.2 and from Lemmas 4.4.7 and 4.4.8.

4.4.1 Non-Connectivity Results

In this section, we show conditions on the spot arrival rate implying, with high probability, non-connectivity of the graph $U(n, r, \lambda, s)$ by the existence of two functional nodes which cannot communicate with one another in the unit square.

Denote by $P_{left}$ and $P_{right}$ the two rectangular halves of the unit square. Partition $P_{left}$ and $P_{right}$ respectively into meshes of $r \times r$ squares. Group these squares in matrices of $5 \times 5$ squares, called blocks; let $B_{left}$ and $B_{right}$ be the sets of these blocks. For each block $b$, denote by $c_b$ the central square in this block and by $p_b$ the union of 8 squares adjacent to $c_b$. Let $alive_b$ be the event that $c_b$ contains at least one functional node. Let $surround_b$ be the event that $p_b$ contains no functional node. Let $isolation_b$ be the intersection of events $surround_b$ and $alive_b$. If $isolation_b$ occurs, and there is at least one functional node outside $b$, then nodes in $c_b$ have no functional path to this external functional node, and then, the graph $U(n, r, \lambda, s)$ is disconnected. In particular, we show non-connectivity w.h.p. by proving that events $isolation_{b_1}$ and $isolation_{b_2}$, $b_1 \in B_{left}$ and $b_2 \in B_{right}$, occur w.h.p. Note that, for distinct blocks $b_1$ and $b_2$, events $surround_{b_1}$ and $surround_{b_2}$ are independent. See Figure 4.1.

We first examine non-connectivity in the case when $r \in o(1)$ and $s \in o(1/\sqrt{n})$, in Lemma 4.4.3. Non-connectivity for $r \in o(1)$ and for larger values of $s \in o(r)$ will be addressed in Lemma 4.4.4. The case $s \in \Omega(1)$ is treated in the next section. We show that for these values of $r$, the graph is connected w.h.p. for those spot arrival rates for which it is alive w.h.p.

**Lemma 4.4.3** Fix any constants $\alpha > 8$ and $\beta > 1$. For $s \in o(1/\sqrt{n})$ and $r \in o(1)$, the graph $U(n, r, \lambda, s)$ is not connected w.h.p. when $\lambda = \beta \ln \left( \frac{\alpha nr^2}{\ln(1/\sqrt{n})} \right) / \pi s^2$. 
Figure 4.1: Non-Connectivity of $U(n, r, \lambda, s)$: (a) If the black region, $p_b$, contains no fault-free node and the central square $c_b$ in the block $b$ contains at least one fault-free node $u$, then $u$ is disconnected from any node outside $b$. (b) In particular, if two disjoint events $isolation_{b_1}$ and $isolation_{b_2}$ occur, $b_1 \in B_{left}$ and $b_2 \in B_{right}$, then the graph $U(n, r, \lambda, s)$ is disconnected.

Proof. For $f(n) \in \omega(1)$, consider the set $\Lambda$ of spot arrival rates of the form $\lambda = \ln \left( \frac{\alpha r^2 n}{\ln (1/(r^2f(n)))} \right) / (\pi s^2)$. Consider two subsets of $\Lambda$: $\Lambda_1$ consisting of these $\lambda$ of the form $\lambda = \ln(g(n)r^2n)/(\pi s^2)$, with $g(n) \in O(1)$ and $\Lambda_2$ consisting of these $\lambda$ of the same form with $g(n) \in \Omega(1)$. In each case, we show that there exists at least one occurrence of the event $isolation_b$ in each set $B_{left}$ and $B_{right}$ and thus, that the graph $U(n, r, \lambda, s)$ is disconnected.

Case 1: $\lambda = \ln(g(n)r^2n)/(\pi s^2)$, with $g(n) \in O(1)$

Fix a block $b$ and consider the event $alive_b$. Consider the subsquare $c'_b \subset c_b$ whose points are at distance greater than $2s$ from $p_b$, i.e., for which the contained nodes become faulty independently from nodes in $p_b$. For $s/r \to 0$ as $n \to \infty$, $|c'_b| > 0.9r^2$, for large $n$. From Lemma 4.3.6, since $0.9mr^2 > 0.9 \log n \in \omega(1)$, it follows that, w.h.p., there is a sparse set of nodes $S_b$, in $c'_b$, of size at least $kmr^2$, for some positive constant $k$. Events $F_v$, $v \in S_b$, occur independently. Let $A$ be the event that the above lower bound on the size of the sparse set $S_b$ holds. Assume $A$. Then,

$$
\Pr[alive_b] = 1 - \Pr[(\forall v \in c_b) F_v] \geq 1 - \Pr[(\forall v \in c'_b) F_v] \geq 1 - \Pr[(\forall v \in S_b) F_v]
$$
\[
\begin{align*}
\geq & \ 1 - (\Pr[F_v])^{knr^2} = 1 - (1 - e^{-\lambda\pi s^2})^{knr^2} \\
& = 1 - (1 - e^{-\ln(g(n))r^2n})^{knr^2} = 1 - (1 - 1/g(n))r^2n)^{knr^2} \\
& \geq c' \in \Theta(1)
\end{align*}
\]

since \(g(n) \in O(1)\). Since \(\Pr[A] \to 1\) for large \(n\), this implies that the probability of event \(\text{alive}_b\) is at least a positive constant \(c\). Let \(\mathcal{A}_{left}\) be the set of all blocks \(b \in B_{left}\) for which the event \(\text{alive}_b\) occurs. Since the probability of the event \(\text{alive}_b\) is a constant, the expected size of the set \(\mathcal{A}_{left}\) is a constant fraction of \(|B_{left}|\). The number of blocks in \(B_{left}\) is \(|B_{left}| = 1/50r^2\). For \(r \in o(1)\), \(|B_{left}|\) grows to infinity as \(n \to \infty\) and thus, under the preceding assumptions, we use Theorem 4.2.1 to show that \(|\mathcal{A}_{left}| \geq (0.9)c/(50r^2)\) w.h.p. Assume this bound on \(|\mathcal{A}_{left}|\) and let \(k/r^2 = (0.9)c/(50r^2)\) for the remainder of the proof.

Fix a block \(b\) and consider the event \(\text{surround}_b\). Using Theorem 4.2.2, we can show that, w.h.p., at most \(\alpha nr^2\) nodes are in \(p_b\); let \(E\) be the event that this bound holds. Assume \(E\). Then, we have, by Fact 1,

\[
\Pr[\text{surround}_b] = \Pr[\bigcap_{v \in p_b} F_v] \geq \prod_{v \in p_b} \Pr[F_v] \geq (1 - e^{-\lambda\pi s^2})^{\alpha nr^2}
\]

and since \(\Pr[E] \to 1\) for large \(n\), we have that \(\Pr[\text{surround}_b] \geq (0.9)(1 - e^{-\lambda\pi s^2})^{\alpha nr^2}\), for large \(n\). Then, the probability that there exists a block \(b \in B_{left}\) for which event \(\text{isolation}_b\) occurs is

\[
\Pr[\exists b \in B_{left} \text{ isolation}_b] = \Pr[\exists b \in \mathcal{A}_{left} \text{ surround}_b] \\
= 1 - \Pr[\forall b \in \mathcal{A}_{left} \neg \text{surround}_b] \\
= 1 - (\Pr[\neg \text{surround}_b])^{\mathcal{A}_{left}} \\
\geq 1 - \left(1 - (0.9) \left(1 - \frac{\ln(1/(r^2f(n))))}{\alpha r^2n}\right)\right)^{k/r^2} \\
= 1 - (1 - (0.9)r^2f(n))^{k/r^2} \to 1 \text{ as } n \to \infty.
\]

The same calculations apply to the second half of the unit square, thus showing the occurrence of at least 2 events \(\text{isolation}_b\) w.h.p. This concludes the argument in the first case.
Case 2: $\lambda = \ln(g(n)r^2n)/(\pi s^2)$, with $g(n) \in \Omega(1)$

Consider again the event $\text{surround}_b$. For $\lambda = \ln(g(n)r^2n)/(\pi s^2)$, with $g(n) \in \Omega(1)$, the same argument as in case 1 implies

$$\Pr[\text{surround}_b] \geq (0.9)(1 - e^{-\lambda\pi s^2}nr^2) = (0.9)(1 - 1/(g(n)r^2n))nr^2 \geq c' \in \Theta(1).$$

Let $\mathcal{S}_{left}$ be the set of all blocks in $B_{left}$ for which the event $\text{surround}_b$ occurs. Since the probability of $\text{surround}_b$ is constant, the expected size of the set $\mathcal{S}_{left}$ is a constant fraction of $|B_{left}|$. The number of blocks in $B_{left}$ is $|B_{left}| = 1/50r^2$. For $r \in o(1)$, $|B_{left}|$ grows to infinity as $n \to \infty$ and thus, under the preceding assumptions, we use Theorem 4.2.1 to show that $|\mathcal{S}_{left}| \geq (0.9)c'/(50r^2)$ w.h.p. Assume this bound on $|\mathcal{S}_{left}|$ and let $k/r^2 = (0.9)c'/(50r^2)$ for the remainder of the proof.

From Remark 4.3.8, if the spot arrival rate is $\lambda = \ln(nh(n))/(\pi s^2)$, $h(n) \in \Omega(1)$, we find a positive constant probability that the graph $U(n, r, \lambda, s)$ is dead. Hence, consider the subset of spot arrival rates of the form $\lambda = \ln(nh(n))/(\pi s^2)$, $h(n) \in o(1)$. Then, for these values of $\lambda$, the probability that there exists a block $b \in B_{left}$ for which event $\text{isolation}_b$ occurs is

$$\Pr[\exists b \in B_{left} \text{ isolation}_b] = \Pr[\exists b \in \mathcal{S}_{left} \text{ alive}_b] = 1 - \Pr[\forall b \in \mathcal{S}_{left} \neg \text{ alive}_b] = 1 - (\Pr[\neg \text{ alive}_b]|_{\mathcal{S}_{left}}) \geq 1 - (1 - (1 - (1 - e^{-\lambda\pi s^2}kr^2))^k/r^2) \geq 1 - (1 - e^{-\ln(nh(n))}kr^2)^k/r^2 = 1 - (1 - 1/(nh(n)))k'kn \to 1 \text{ as } n \to \infty.$$

The same calculations apply to the second half of the unit square, thus showing the occurrence of at least 2 events $\text{isolation}_b$ w.h.p. This concludes the argument in the second case.

To conclude the proof, fix the function $f(n) = 1/r$. Since $r \in o(1)$, we have $f(n) \in \omega(1)$. Hence the corresponding

$$\tilde{\lambda} = \ln \left( \frac{\alpha r^2n}{\ln(1/(r^2f(n)))} \right) / (\pi s^2) = \ln \left( \frac{\alpha r^2n}{\ln(1/r)} \right) / (\pi s^2)$$
Let $\lambda < \beta \ln \left( \frac{\alpha r^2 n}{\ln(1/r^2)} \right) / (\pi s^2)$, for any constant $\beta > 1$. Indeed,

$$\tilde{\lambda} = \ln \left( \frac{\alpha r^2 n}{\ln(1/r)} \right) / (\pi s^2) = \ln \left( \frac{\alpha r^2 n}{0.5 \ln(1/r^2)} \right) / (\pi s^2)$$

$$= \left( \ln \left( \frac{\alpha r^2 n}{\ln(1/r^2)} \right) + \ln 2 \right) / (\pi s^2) \leq \beta \ln \left( \frac{\alpha r^2 n}{\ln(1/r^2)} \right) / (\pi s^2).$$

It follows that all $\lambda = \beta \ln \left( \frac{\alpha r^2 n}{\ln(1/r^2)} \right) / (\pi s^2)$, for any constant $\beta > 1$, are also in $\Lambda$ which proves the lemma. Note that, under the assumption $s = o(1/\sqrt{n})$, we have $\min\{n, 1/s^2\} = n$.

**Lemma 4.4.4** Fix any constant $\beta > 1$. For $s \in o(r)$ and $r \in o(1)$, the graph $U(n, r, \lambda, s)$ is disconnected w.h.p. when $\lambda = 4\beta \ln \left( \frac{8r^2/s^2}{\ln(1/r^2)} \right) / (\pi s^2)$.

**Proof.** For $f(n) \in \omega(1)$, consider the set $\Lambda$ of spot arrival rates of the form $\lambda = 4\ln \left( \frac{8r^2/s^2}{\ln(1/r^2)} \right) / (\pi s^2)$. Consider two subsets of $\Lambda$: $\Lambda_1$ consisting of these $\lambda$ of the form $\lambda = \ln(g(n)r^2/s^2)/(\pi s^2)$, with $g(n) \in O(1)$ and $\Lambda_2$ consisting of these $\lambda$ of the form $\lambda = 4\ln(g(n)r^2/s^2)/(\pi s^2)$, with $g(n) \in \Omega(1)$. In each case, we show that there exists at least one occurrence of the event $\text{isolation}_b$ in each set $B_{left}$ and $B_{right}$ and thus, that the graph $U(n, r, \lambda, s)$ is disconnected.

Case 1: $\lambda = \ln(g(n)r^2/s^2)/(\pi s^2)$, with $g(n) \in O(1)$

Fix a block $b$ and consider the event $\text{alive}_b$. By a reasoning similar to Lemma 4.4.3, w.h.p., there is a sparse set of nodes in $c'_b$, $S$, of size at least $kr^2/s^2$, for some positive constant $k$, for which the events $F_v$, $v \in S$, occur independently; let $A$ be the event that the lower bound $kr^2/s^2$ on the size of $S$ holds. Assuming the event $A$ we have

$$\Pr[\text{alive}_b] \geq (1 - \Pr[\bigcap_{v \in S} F_v]) \geq (1 - (\Pr[F_v])^{kr^2/s^2})$$

$$= (1 - (1 - e^{-\lambda s^2})^{kr^2/s^2}).$$

Let $\lambda = \ln(g(n)r^2/s^2)/(\pi s^2)$, with $g(n) \in O(1)$. Then we have

$$\Pr[\text{alive}_b] \geq (1 - (1 - e^{-\lambda s^2})^{kr^2/s^2}) = \left( 1 - \left( 1 - \frac{s^2}{g(n)r^2} \right)^{kr^2/s^2} \right)$$

$$\geq (1 - e^{-k/g(n)/2}) \geq c' \in \Theta(1).$$

Since $\Pr[A] \to 1$ as $n \to \infty$, this implies that the event $\text{alive}_b$ occurs at least with some positive constant probability $c$. Let $A_{left}$ be the set of all blocks in $B_{left}$ for which
the event $alive_b$ occurs. Since the probability of $alive_b$ is a constant, the expected size of the set $A_{left}$ is a constant fraction of $|B_{left}|$. The number of blocks in $B_{left}$ is $|B_{left}| = \frac{1}{50r^2}$. For $r \in o(1)$, $|B_{left}|$ grows to infinity as $n \to \infty$ and thus, under the preceding assumptions, we use Theorem 4.2.1 to show that $|A_{left}| \geq (0.9)c/(50r^2)$ w.h.p. Assume this bound on $|A_{left}|$ and let $k/r^2 = (0.9)c/(50r^2)$ for the remainder of the proof.

Fix a block $b$ and consider the event $surround_b$ that all nodes in $p_b$ are faulty. Partition $p_b$ into a mesh of $s/\sqrt{2} \times s/\sqrt{2}$ squares. Let $Q_b$ be the set of these $8r^2/s^2$ squares. Let $dead_q$ be the event that the disk of diameter $s$ concentric with $q \in Q_b$ contain at least one spot. Since all points in the square are within distance $s$ from any point in this disk, $dead_q$ implies that no node is functional in the square $q$. Hence, $surround_b$ is implied by the intersection of all $dead_q$, $q \in Q_b$. Then, we have, by Fact 1,

$$Pr[surround_b] = Pr[\bigcap_{q \in Q_b} dead_q] \geq \prod_{q \in Q_b} Pr[dead_q] = (Pr[dead_q])^{\left|Q_b\right|} = \left(1 - e^{-\lambda s^2/4}\right)^{8r^2/s^2}.$$ 

Then, the probability that there exists a block in $B_{left}$ for which event $isolation_b$ occurs is

$$Pr[\exists b \in B_{left} isolation_b] = Pr[\exists b \in A_{left} surround_b]$$

$$= 1 - Pr[\forall b \in A_{left} \neg surround_b]$$

$$= 1 - (Pr[\neg surround_b])^{|A_{left}|}$$

$$\geq 1 - \left(1 - \left(1 - e^{-\lambda s^2/4}\right)^{8r^2/s^2}\right)^{k/r^2}$$

$$= 1 - \left(1 - \left(1 - e^{-\ln\left(\frac{s^2/s^2}{\ln(1/(r^2f(n)))}\right)}\right)^{8r^2/s^2}\right)^{k/r^2}$$

$$= 1 - \left(1 - \frac{s^2}{8r^2} \ln\left(\frac{1}{r^2f(n)}\right)\right)^{k/r^2}$$

$$\geq 1 - \left(1 - e^{-\ln\left(\frac{1}{r^2f(n)}\right)}\right)^{k/r^2}$$

$$= 1 - (1 - r^2f(n))^{k/r^2}.$$
\[ \rightarrow 1 \text{ as } n \rightarrow \infty. \]

The same calculations apply to the second half of the unit square, thus showing the occurrence of at least 2 events \( \text{isolation}_b \) w.h.p. This concludes the argument in the first case.

**Case 2:** \( \lambda = 4 \ln(g(n)r^2/s^2)/(\pi s^2) \), with \( g(n) \in \Omega(1) \)

From above, we have \( \Pr[\text{surround}] \geq (1 - e^{-\lambda \pi s^2/4}s^2/(3\sqrt{3} s^2) \). With \( g(n) \in \Omega(1) \) and \( \lambda = 4 \ln(g(n)r^2/s^2)/(\pi s^2) \) we have

\[
\Pr[\text{surround}_b] \geq (1 - e^{-\ln(g(n)r^2/s^2)})s^2/s^2 = (1 - s^2/g(n)r^2)s^2/s^2
\]

\[ \geq e^{-8/g(n)} \geq c' \in \Theta(1). \]

Let \( S_{\text{left}} \) be the set of all blocks in \( B_{\text{left}} \) for which the event \( \text{surround}_b \) occurs.

Since the probability of \( \text{surround}_b \) is a constant, the expected size of the set \( S_{\text{left}} \) is a constant fraction of \( |B_{\text{left}}| \). The number of blocks in \( B_{\text{left}} \) is \( |B_{\text{left}}| = \frac{1}{\delta r^2} \).

For \( r \in o(1) \), \( |B_{\text{left}}| \) grows to infinity as \( n \to \infty \) and thus, under the preceding assumptions, we use Theorem 4.2.1 to show that \( |S_{\text{left}}| \geq (0.9)c'/(50r^2) \) w.h.p. Let \( k/r^2 = (0.9)c'/(50r^2) \) in for the remainder of the proof.

From Remark 4.3.8, if the spot arrival rate is \( \lambda = 4 \ln(r^2h(n))/(\pi s^2) \), \( h(n) \in \Omega(1) \), we find a positive constant probability that the graph \( U(n, r, \lambda, s) \) is dead. Consider the subset of spot arrival rates a constant factor away from this bound, of the form \( \lambda = \ln(nh(n))/(\pi s^2) \), \( h(n) \in o(1) \). Then, the probability that there is a block in \( B_{\text{left}} \) for which event \( \text{isolation}_b \) occurs is

\[
\Pr[\exists b \in B_{\text{left} \text{ isolation}_b}] = \Pr[\exists b \in S_{\text{left} \text{ alive}_b}] = 1 - \Pr[\forall b \in S_{\text{left} \text{ -alive}_b}]
\]

\[
= 1 - (\Pr[\neg \text{alive}_b]^{\text{S_{left}}})
\]

\[
\geq 1 - (1 - (1 - (1 - e^{-\lambda \pi s^2/k^2/s^2})^{k/r^2})
\]

\[
= 1 - ((1 - e^{-\ln(h(n)/s^2)})^{k^2/s^2})^{k/r^2}
\]

\[
= 1 - (1 - s^2/h(n))^{k^2/s^2} \rightarrow 1 \text{ as } n \to \infty.
\]

The same calculations apply to the second half of the unit square, thus showing the occurrence of at least 2 events \( \text{isolation}_b \) w.h.p. This concludes the argument in the second case.
A gap remains between the two spot arrival rate ranges. In fact, as the spot arrival rate increases, the number of \( r \times r \) squares of the partition which contain no functional node steadily increases. These dead squares appear randomly and uniformly over the unit square. Hence, the probability that a fixed square containing functional nodes becomes surrounded by dead squares is monotonically increasing. It follows that the graph is disconnected also for those spot arrival rates between the ranges in parts 1 and 2.

To conclude the proof, fix the function \( f(n) = 1/r \). Since \( r \in o(1) \), we have \( f(n) \in \omega(1) \). Hence, the corresponding

\[
\hat{\lambda} = \ln \left( \frac{8r^2/s^2}{\ln(1/(r^2(f(n))))} \right) / (\pi s^2) = \ln \left( \frac{8r^2/s^2}{\ln(1/r)} \right) / (\pi s^2)
\]

is in \( \Lambda \). We show that \( \hat{\lambda} < \beta \ln \left( \frac{8r^2/s^2}{\ln(1/r^2)} \right) / (\pi s^2) \), for any constant \( \beta > 1 \). Indeed,

\[
\hat{\lambda} = \ln \left( \frac{8r^2/s^2}{\ln(1/r)} \right) / (\pi s^2) = \ln \left( \frac{8r^2/s^2}{0.5 \ln(1/r^2)} \right) / (\pi s^2)
\]

\[
= \left( \ln \left( \frac{8r^2/s^2}{\ln(1/r^2)} \right) + \ln 2 \right) / (\pi s^2) \leq \beta \ln \left( \frac{8r^2/s^2}{\ln(1/r^2)} \right) / (\pi s^2).
\]

It follows that all \( \lambda = \beta \ln \left( \frac{8r^2/s^2}{\ln(1/r^2)} \right) / (\pi s^2) \), for any constant \( \beta > 1 \), are also in \( \Lambda \) which proves the lemma. Note that, under the assumption \( s \in \Omega(1/\sqrt{n}) \), we have \( \Theta(\min\{n, 1/s^2\}) = \Theta(1/s^2) \).

The preceding lemmas concern only the case when \( r \in o(1) \). As stated in Theorem 4.4.2, for \( r \in \Omega(1) \), thresholds on the spot arrival rate separate the case of connected \( U(n, r, \lambda, s) \) from the case when it is dead. Hence, we do not provide any non-connectivity result for \( r \in \Omega(1) \) and defer this case to the next section.

### 4.4.2 Connectivity Results

In this section, we show conditions on the spot arrival rate guaranteeing connectivity of the graph \( U(n, r, \lambda, s) \) w.h.p. We show connectivity of \( U(n, r, \lambda, s) \) by proving the existence of a fault-free node in each square of a sufficiently fine partition of the unit square w.h.p. This implies the existence of a fault-free path between any pair of nodes of the graph \( U(n, r, \lambda, s) \) and hence this graph is connected.
Partition the unit square into a mesh of $r/\sqrt{5} \times r/\sqrt{5}$ squares, called blocks. Let $B$ be the set of all blocks. The distance between any two points in the blocks which are adjacent by an edge (edge-adjacent) is at most $r$. Hence, functional nodes in edge-adjacent blocks can communicate with each other.

Partition each block $b \in B$ into a mesh of $3s \times 3s$ squares called tiles. Let $T_b$ be the set of all these $r^2/(45s^2)$ tiles for the block $b$. For a fixed tile $t \in T_b$, let $\text{free}_t$ be the event that it contains no spot. Under the event $\text{free}_t$, the central $s \times s$ square $c_t \subset t$ is at distance greater than $s$ from all spots. Let $a_t$ be the event that $c_t$ contains at least one node. Since node arrivals and spot arrivals are independent, the events $\text{free}_t$ and $a_t$ are independent. Moreover, for all $t \neq t' \in T_b$, the events $a_t, a_{t'} (\text{free}_t, \text{free}_{t'})$ are independent since they are respectively the result of arrivals inside non-overlapping tiles $t$ and $t'$.

Consider the event $\text{alive}_b$ that a fixed block $b$ contains at least one functional node. The event $\text{alive}_b$ is implied by the existence of a tile $t \in T_b$ where both the events $\text{free}_t$ and $a_t$ occur. Let $\text{alive}_b' = \{\exists t \in T_b \text{ s.t. } \text{free}_t \cap a_t\}$ be this sub-event of $\text{alive}_b$. Hence, the probability of event $\text{alive}_b$ that a fixed block $b$ contains at least one functional node is

$$\Pr[\text{alive}_b] \geq \Pr[\text{alive}_b'] = \Pr[\exists t \in T_b \text{ free}_t \cap a_t]$$

$$= 1 - \Pr[\forall t \in T_b \neg \text{free}_t \cup \neg a_t] = 1 - (\Pr[\neg \text{free}_t \cup \neg a_t])^{|T_b|}$$

$$= 1 - (1 - \Pr[\text{free}_t \cap a_t])^{|T_b|} = 1 - (1 - \Pr[\text{free}_t] \Pr[a_t])^{|T_b|}$$

$$= 1 - (1 - e^{-9s^2}(1 - e^{-ns^2}))^{r^2/(45s^2)}.$$ 

Let $\text{connect}$ be the event that each block $b \in B$ contains at least one functional node. We have $\Pr[\text{connect}] \geq \Pr[\forall b \in B \text{ alive}_b']$. In the next two lemmas we establish upper bounds on $\lambda$ that guarantee connectivity w.h.p. of $U(n, r, \lambda, s)$ for $s \in o(1/\sqrt{n})$ and $s \in \Omega(1/\sqrt{n})$ respectively.

**Lemma 4.4.5** For $s \in o(1/\sqrt{n})$ and any constant $\alpha < 1$, the graph $U(n, r, \lambda, s)$ is connected, w.h.p., when the spot arrival rate is $\lambda = \alpha \ln \left(\frac{nr^2}{45\ln(1/\sqrt{r^2})}\right)/9s^2$.

**Proof.** For $s \in o(1/\sqrt{n})$, we have that $ns^2$ is in $o(1)$. Hence,

$$\Pr[\text{alive}_b'] = 1 - (1 - e^{-9s^2}(1 - e^{-ns^2}))^{r^2/(45s^2)}$$
\[ = 1 - (1 - e^{-9s^2\lambda}(1 - (1 - ns^2 + (ns^2)^2/2 - \ldots))r^2/(45s^2)) \geq 1 - (1 - ns^2 e^{-9s^2\lambda}r^2/(45s^2)). \]

We observe that the events alive' \(i\) are independent for all \(b \in B\) since they concern arrivals inside disjoint blocks. For \(\lambda = \ln\left(\frac{nr^2}{45 \ln(1/(r^2f(n)))}\right)/9s^2\), \(f(n) \in o(1)\), we have

\[
\Pr[\text{connect}] \geq \Pr[\forall b \in B \text{ alive}'_b] = (\Pr[\text{alive}'_b])^{5/r^2} \\
\geq (1 - (1 - ns^2 e^{-9s^2\lambda}r^2/(45s^2)))^{5/r^2} \\
= \left(1 - \left(1 - ns^2 e^{-\ln\left(\frac{nr^2}{45 \ln(1/(r^2f(n)))}\right)}\right)r^2/(45s^2)\right)^{5/r^2} \\
= \left(1 - \left(1 - ns^2 45 \ln\left(\frac{1/(r^2f(n))}{nr^2}\right)r^2/(45s^2)\right)\right)^{5/r^2} \\
= \left(1 - \left(1 - 45s^2 \ln\left(\frac{1/(r^2f(n))}{r^2}\right)r^2/(45s^2)\right)\right)^{5/r^2} \\
\geq \left(1 - e^{-\ln(1/(r^2f(n)))}\right)^{5/r^2} = \left(1 - r^2f(n)\right)^{5/r^2} \to 1 \text{ as } n \to \infty. 
\]

Let \(\Lambda\) be the set of spot arrival rates of the form \(\lambda = \ln\left(\frac{nr^2}{45 \ln(1/(r^2f(n)))}\right)/9s^2\), \(f(n) \in o(1)\). We have shown that, for all \(\lambda \in \Lambda\), the graph \(U(n, r, \lambda, s)\) is connected w.h.p. To conclude the proof, fix the function \(f(n) = r\). Since \(r \in o(1)\), we have \(f(n) \in o(1)\). Hence, the corresponding

\[ \tilde{\lambda} = \ln\left(\frac{nr^2}{45 \ln(1/(r^2f(n)))}\right)/9s^2 = \ln\left(\frac{nr^2}{45 \ln(1/(r^3))}\right)/9s^2 \]

is in \(\Lambda\). We show that \(\tilde{\lambda} > \alpha \ln\left(\frac{nr^2}{45 \ln(1/r^2)}\right)/9s^2\), for any constant \(\alpha < 1\). Indeed,

\[ \lambda = \ln\left(\frac{nr^2}{45 \ln(1/(r^2f(n)))}\right)/9s^2 = \ln\left(\frac{nr^2}{45 \ln(1/(r^3))}\right)/9s^2 \\
= \ln\left(\frac{nr^2}{45(1.5) \ln(1/r^2)}\right)/9s^2 = \left(\ln\left(\frac{nr^2}{45 \ln(1/r^2)}\right) - \ln(1.5)\right)/9s^2 \\
\geq \alpha \ln\left(\frac{nr^2}{45 \ln(1/r^2)}\right)/9s^2. \]

It follows that all \(\lambda = \alpha \ln\left(\frac{nr^2}{45 \ln(1/r^2)}\right)/9s^2\), for any constant \(\alpha < 1\), are also in \(\Lambda\) which proves the lemma.
Lemma 4.4.6 For \( s \in \Omega(1/\sqrt{n}) \cap o(r) \) and any constant \( \alpha < 1 \), \( U(n, r, \lambda, s) \) is connected w.h.p. when the spot arrival rate is \( \lambda = \alpha \ln \left( \frac{r^2/s^2}{45 \ln(1/r^2)} \right) / 9s^2 \).

Proof. Recall that

\[
\Pr[\text{alive}_i'] = 1 - (1 - e^{-9s^2\lambda}(1 - e^{-ns^2}))^{r^2/(45s^2)}.
\]

For \( \lambda = \ln \left( \frac{(1-e^{-ns^2})r^2/s^2}{45 \ln(1/(r^2f(n)))} \right) / 9s^2 \), \( f(n) \in o(1) \), we have

\[
\Pr[\text{connect}] \geq \Pr[\forall i \in B \text{ alive}_i'] = (\Pr[\text{alive}_i'])^{5/r^2} \geq (1 - (1 - (1 - e^{-ns^2})e^{-9s^2\lambda})^{r^2/(45s^2)})^{5/r^2} = \left(1 - (1 - (1 - e^{-ns^2})e^{-9s^2\lambda})^{r^2/(45s^2)}\right)^{5/r^2} \geq (1 - e^{-\ln(1/(r^2f(n)))})^{5/r^2} = (1 - r^2f(n))^{5/r^2} \rightarrow 1 \text{ as } n \rightarrow \infty.
\]

Let \( \Lambda \) be the set of spot arrival rates of the form \( \lambda = \ln \left( \frac{(1-e^{-ns^2})r^2/s^2}{45 \ln(1/(r^2f(n)))} \right) / 9s^2 \), \( f(n) \in o(1) \). We have shown that, for all \( \lambda \in \Lambda \), the graph \( U(n, r, \lambda, s) \) is connected w.h.p. To conclude the proof, fix the function \( f(n) = r \). Since \( r \in o(1) \), we have \( f(n) \in o(1) \). Hence the corresponding

\[
\hat{\lambda} = \ln \left( \frac{(1-e^{-ns^2})r^2/s^2}{45 \ln(1/(r^2f(n)))} \right) / 9s^2 = \ln \left( \frac{(1-e^{-ns^2})r^2/s^2}{45 \ln(1/(r^3))} \right) / 9s^2
\]

is in \( \Lambda \). We show that \( \hat{\lambda} > \alpha \ln \left( \frac{r^2/s^2}{45 \ln(1/r^2)} \right) / 9s^2 \), for any constant \( \alpha < 1 \). Indeed, for \( s \in \Omega(1/\sqrt{n}) \), there exists a constant \( c = (1 - e^{-ns^2}) \) and

\[
\hat{\lambda} = \ln \left( \frac{(1-e^{-ns^2})r^2/s^2}{45 \ln(1/(r^2f(n)))} \right) / 9s^2 = \ln \left( \frac{cr^2/s^2}{45 \ln(1/(r^3))} \right) / 9s^2 \geq \alpha \ln \left( \frac{r^2/s^2}{45 \ln(1/r^2)} \right) / 9s^2.
\]

It follows that all \( \lambda = \alpha \ln \left( \frac{r^2/s^2}{45 \ln(1/r^2)} \right) / 9s^2 \), for any constant \( \alpha < 1 \), are also in \( \Lambda \) which proves the lemma.
For large values of $r$, we show connectivity for the same range of $\lambda$ for which we have shown the graph $U(n, r, \lambda, s)$ to be alive w.h.p.

**Lemma 4.4.7** For $r \in \Omega(1)$ and $s \in o(1)$, the graph $U(n, r, \lambda, s)$ is connected, w.h.p., when the spot arrival rate is $\lambda = \alpha \frac{\ln(\min\{n, 1/s^2\})}{\pi s^2}$, for any constant $\alpha < 1$.

**Proof.** We will prove the lemma in 2 cases: $s \leq 1/\sqrt{n}$ and $s \geq 1/\sqrt{n}$ ($s \in o(1)$).

For positive constant $c = \lceil 5/r^2 \rceil$, partition the plane into a mesh of $c$ squares of area $1/c$. Let $Q$ be the set of all these squares. In such a partition, all nodes inside a square $q \in Q$ are within distance $r$ of nodes in the squares adjacent to $q$ by an edge.

We show that the event $alive_q$ that a square $q$ contains at least one functional node occurs w.h.p. for all squares $q \in Q$. The event $\bigcap_{q \in Q} alive_q$ implies that the graph $U(n, r, \lambda, s)$ is connected.

For each square $q$ of the partition, let $q'$ be the subsquare of $q$ at distance greater than $2s$ from other squares of the partition. We define the event $sparse_q$ that there exists a sparse set $S_q$ of nodes with size, respectively at least $k'n/c$ ($s \leq 1/\sqrt{n}$) and at least $k'(cs^2)$ ($s \geq 1/\sqrt{n}$), for some constant $k'$. Moreover, $\Pr[sparse_q] \to 1$ as $n \to \infty$, by Lemma 4.3.6, since the sizes of these sets grow to infinity as $n \to \infty$. Hence, since $|Q| = c$, a constant,

$$\Pr[\bigcap_{q \in Q} sparse_q] = \Pr[sparse_q]^{|Q|} \to 1 \text{ as } n \to \infty.$$  

Assume the event $sparse_q$ holds for each square $q \in Q$.

**Case 1:** $s \leq 1/\sqrt{n}$.

Let $\lambda = \ln(n f(n))/\pi s^2$, $f(n) \in o(1)$. Then we have

$$\Pr[F_v] = 1 - e^{-\lambda \pi s^2} = 1 - e^{-\ln(n f(n))}$$

and

$$\Pr[alive_q] = 1 - \Pr[\bigcap_{v \in S_q} F_v] \geq 1 - \Pr[\bigcap_{v \in S_q} F_v] = 1 - (\Pr[F_v])^{|S_q|} \geq 1 - (1 - e^{-\ln(n f(n))})^{k'n/c} \to 1 \text{ as } n \to \infty.$$
Case 2: \( s \geq 1/\sqrt{n} \) (with \( s \in o(1) \)).

Let \( \lambda = \ln(f(n)/s^2)/(\pi s^2) \), \( f(n) \in o(1) \). Then we have

\[
\Pr[F_v] = 1 - e^{-\lambda s^2} = 1 - e^{-\ln(f(n)/s^2)}
\]

and

\[
\Pr[\text{alive}_q] \geq 1 - \left(1 - e^{-\ln(f(n)/s^2)}\right)^{k/(cs^2)} = 1 - \left(1 - s^2/f(n)\right)^{k/(cs^2)} \to 1 \text{ as } n \to \infty.
\]

Since \( |Q| = c \) is a constant, we have in both cases

\[
\Pr[\bigcap_{q \in Q} \text{alive}_q] = \Pr[\text{alive}_q]^{\mid Q\mid} \to 1 \text{ as } n \to \infty.
\]

Let \( \Lambda \) be the set of spot arrival rates of the form \( \lambda = \ln\left(\min\left\{n, 1/s^2\right\} f(n)/\pi s^2\right) \), for \( f(n) \in o(1) \). We have shown that, for all \( \lambda \in \Lambda \), the graph \( U(n, r, \lambda, s) \) is connected w.h.p. To conclude the proof, fix the function \( f(n) = 1/\ln(\min\{n, 1/s^2\}) \). Since \( r \in O(1) \), we have \( f(n) \in o(1) \). Hence, the corresponding

\[
\tilde{\lambda} = \frac{\ln(\min\{n, 1/s^2\} f(n))}{\pi s^2} = \frac{\ln(\min\{n, 1/s^2\} \ln(\min\{n, 1/s^2\}))}{\pi s^2}
\]

is in \( \Lambda \). We show that \( \tilde{\lambda} > \alpha \frac{\ln(\min\{n, 1/s^2\})}{\pi s^2} \), for any constant \( \alpha < 1 \). Indeed,

\[
\tilde{\lambda} = \frac{\ln(\min\{n, 1/s^2\} f(n))}{\pi s^2} = \frac{\ln(\min\{n, 1/s^2\} \ln(\min\{n, 1/s^2\}))}{\pi s^2}
\]

\[
= \frac{\ln(\min\{n, 1/s^2\}) - \ln(\ln(\min\{n, 1/s^2\}))}{\pi s^2} \geq \alpha \frac{\ln(\min\{n, 1/s^2\})}{\pi s^2}.
\]

It follows that all \( \lambda = \alpha \ln\left(\frac{r^2/s^2}{45\ln(1/r^2)}\right)/9s^2 \), for any constant \( \alpha < 1 \), are also in \( \Lambda \) which proves the lemma.

For \( r \in O(1) \) and \( s \in O(1) \cap o(r) \), we observe that if \( r \in \Theta(1) \), then the condition \( s \in o(r) \) is impossible. Hence, necessarily, \( r \in \omega(1) \). Since the unit square has a diameter of \( \sqrt{2} \), if it is alive, then it is also connected for \( r \in \omega(1) \) and sufficiently large \( n \). Hence Lemma 4.4.8 follows from from Theorem 4.3.2.

**Lemma 4.4.8** For \( r \in \omega(1) \) and \( s \in O(1) \cap o(r) \), the graph \( U(n, r, \lambda, s) \) is connected, w.h.p., when the spot arrival rate is \( \lambda \in o(1/s^2) \).
4.5 Broadcasting Algorithm

We propose a deterministic algorithm which completes broadcast with probability $1 - \epsilon$ in time $O(D + \log 1/\epsilon)$, in the fault-free graph $U(n, r, \lambda, s)$ for $s \in o(1/\sqrt{n})$. The algorithm consists of two parts: a preprocessing part called \textit{spokesman election}, and a message transmission part. In the spokesman election part, a unique node called the \textit{spokesman} is selected in each square of a partition defined below. Only the spokesman of a square relays messages in the transmission part of the algorithm.

Partition the unit square into a mesh of $r/\sqrt{5} \times r/\sqrt{5}$ squares, called boxes, and let $S$ be the set of these boxes. Group the boxes in $5 \times 5$ matrices, called blocks, and let $B$ be the set of all these blocks. For all blocks, label its boxes 1 through 25 row by row. Further partition each box into a mesh of $1/\sqrt{n} \times 1/\sqrt{n}$ squares, called tiles. For a box $i$, let $T_i$ be the set of all tiles in this box. For all boxes, label the tiles 1 through $t = r^2n/5$ row by row.

**Algorithm $A^*$**

**Spokesman Election Part**

Nodes know their location and hence, they can compute the labels $i, j$ of their box, and tile, respectively. Nodes label themselves $(i, j)$ accordingly.

In parallel for all blocks, the algorithm executes steps $i = 1, 2, \ldots, 25$. In a step $i$, the algorithm sequentially goes through rounds $j = 1, 2, \ldots, t$. In a step $i$, at round $j$, all nodes with label $(i, j)$ (in the box $i$ and tile $j$) transmit their label and the list of labels heard from adjacent boxes. At any given round $j'$, when only one node transmits its label $(i, j')$, the message is heard by all other nodes in the box $i$ and all edge-adjacent boxes; The first node whose message has been heard is chosen as the spokesman for box $i$ by all other nodes in the box $i$ (the node itself does not know this yet) and in edge-adjacent boxes. In subsequent rounds of step $i$, nodes in the box $i$ containing the elected node $(i, j')$ are silent. The node $(i, j')$ will learn that it is the spokesman for box $i$ when, in an edge adjacent box, a unique node transmits its own label and the list of labels heard from adjacent boxes. Since all boxes, except box 25, are edge-adjacent to a box with a larger label, by the end of round 25, if a spokesman is chosen for each box then all spokesmen, with the exception of the spokesman in the box 25 are confirmed, i.e., they know that they are spokesmen. Hence, after round
25, a single transmission from the spokesman in the box 24 is sufficient to confirm the spokesman of box 25. This transmission is done in parallel by all spokesmen in the boxes labeled 24, right after the end of round 25.

Hence, in the spokesman election part of the algorithm, one spokesman is chosen and confirmed in every box if there is, in every box, a tile which contains exactly one functional node.

**Message Transmission Part**

In the first round of this part, the source transmits its message. Then, in parallel for all blocks, the algorithm is executed in identical phases \( \rho = 1, 2, \ldots \). In phase \( \rho \), rounds \( j = 1, 2, \ldots, 25 \) are executed sequentially. In a round \( j \), a spokesman of box \( j \) which has received the source message but has not relayed it yet, transmits the source message. This completes the description of algorithm \( A^* \). See Figure 4.2 and the detailed pseudocode of algorithm \( A^* \) below.

![Figure 4.2: Algorithm \( A^* \): (a) Spokesman Election part. (b) Message Transmission part.](image)

Let \( \epsilon \) be the tolerated error probability for the algorithm, i.e., we wish to broadcast with probability at least \( 1 - \epsilon \). Let \( \mathcal{A} \) be the algorithm \( A^* \) modified so that the spokesman election part uses only the first \( \frac{\ln(2D^2/\epsilon)}{\ln(1/(1-0.9e^{-(c+1)}))} \) tiles of each square.

**Theorem 4.5.1** Let \( c \) be a positive constant and \( d = \ln(1/(1 - (0.9)e^{-(c+1)})) \). For \( s \in o(1/\sqrt{n}) \), \( r^2 \geq \frac{5\ln(5D^2/\epsilon)}{dn} \), and \( \lambda \leq c/(\pi s^2) \), the algorithm \( \mathcal{A} \) broadcasts a message in time \( O(D + \log 1/\epsilon) \), with probability at least \( 1 - \epsilon \).

**Proof.** Consider a tile \( t \). There exists a subsquare \( a \) of \( t \) of area \( (1/\sqrt{n} - s)^2 = \)
Algorithm $A^*$, for all nodes in parallel.

$KnownSpokesmen_v \leftarrow \emptyset \{ \text{the set of spokesmen known to } v \}$

$Elected_v \leftarrow \text{false} \{ \text{a flag set to } \text{true} \text{ if } v \text{ is elected a spokesman} \}$

$Label_v \leftarrow \text{(square label, tile label)} \{ \text{the label of } v \}$

for $i = 1$ to 25 do

for $j = 1$ to $t$ do

if $Label_v = (i, j)$ AND $\neg \exists j', (i, j') \in KnownSpokesmen_v$ then

Send “$(i, j)$” and the contents of $KnownSpokesmen_v$

else

Listen()

Call Procedure Terminate()

Call Procedure Transmit()

Procedure Listen()

if a message “$(x, y)$, Spokesmen” is heard then

$KnownSpokesmen_v \leftarrow KnownSpokesmen_v \cup (x, y)$

if $Label_v \in Spokesmen$ then

$KnownSpokesmen_v \leftarrow KnownSpokesmen_v \cup Label_v$

$Elected_v \leftarrow \text{true}$

Procedure Terminate()

if $\exists j \text{ s.t. } Label_v = (24, j)$ AND $Spokesman_v = \text{true}$ then

Send the contents of $KnownSpokesmen_v$

else

Listen()

Procedure Transmit()

$MessageQueue_v \leftarrow ”\text{if } v \text{ is the source, the source message, else } \emptyset”$

loop

for square $i = 1 \ldots 25$ do

if $Label_v = (i, j)$ AND $Spokesman_v = \text{true}$ then

if $MessageQueue_v \neq \emptyset$ then

Send the contents of $MessageQueue_v$

$MessageQueue_v \leftarrow \emptyset$

else if a message “$x$” is heard then

$MessageQueue_v \leftarrow MessageQueue_v \cup x$
$1/n - 2s/\sqrt{n} + s^2$ whose nodes are not affected by spots in other tiles; the remaining subset $a'$ of the tile has area $2s/\sqrt{n} - s^2$. Let $good_t$ be the event that there exists exactly one node in $a$, no node in $a'$, and that the node in $a$ is not within distance $s$ of a spot. We have

$$\Pr[good_t] = e^{-(1/n-2s/\sqrt{n}+s^2)n} \cdot e^{-(2s/\sqrt{n}-s^2)n} \cdot e^{-\lambda \pi s^2} \geq (1 - 2s/\sqrt{n} + s^2) e^{-1+(2s/\sqrt{n}-s^2)-(2s/\sqrt{n}-s^2)-\frac{\epsilon}{\pi} \pi s^2} \geq 0.9e^{-(c+1)}$$

for large $n$. Let $spokesman_q$ be the event that the spokesman election part is successful in a fixed box $q$. Since $r^2 \geq \frac{5\ln(5D^2/\epsilon)}{dn}$, there are at least $nr^2/5 = \frac{\ln(5D^2/\epsilon)}{d}$ tiles in each box. Hence, the algorithm $\mathcal{A}$ can execute its spokesman election part. Then, we have

$$\Pr[spokesman_q] = 1 - (1 - \Pr[good_t])^{\ln(5D^2/\epsilon)/d}$$

$$\geq 1 - (1 - 0.9e^{-(c+1)})^{\ln(5D^2/\epsilon)/d}$$

$$= 1 - \left(\frac{5D^2}{\epsilon}\right)^{\frac{\ln(1-0.9e^{-(c+1)})}{\ln(1-0.9e^{-(c+1)})}}$$

$$= 1 - \frac{\epsilon}{5D^2}.$$ 

There are at most $5D^2$ boxes in the unit square. Hence, the event $spokesmen$ that each box contains one spokesman occurs with probability

$$\Pr[spokesmen] \geq \left(1 - \frac{\epsilon}{5D^2}\right)^{5D^2} \geq 1 - \epsilon.$$ 

We now show that, assuming the event $spokesmen$, all functional nodes are informed and we estimate the total running time of the algorithm. For each step, 25 rounds are elapsed. We say that a square with label $j$ is active if the algorithm round is $j$, i.e., when its spokesman may transmit. All squares with the same labels are located at distance at least $4r/\sqrt{5}$ from each other. Only spokesmen in active squares (with the same label $j$, at a round $j$) transmit. Hence all nodes in squares adjacent to active squares will receive the message correctly at every round when a spokesman transmits in this active square (due to large distances between squares
with the same label, there is no collision in adjacent squares). It follows that if a message is received by any square in a block $i$ at time $t$, then there exists a positive constant $\delta$ such that at time $t + \delta$ all nodes in the block $i$ will know the message. Moreover, at time $t + \delta$, the nodes in the squares outside the block $i$, but adjacent to the squares in the block $i$ also have received the message. Consider two nodes in different blocks $i$ and $j$ such that there is a sequence of edge-adjacent blocks of length $k - 1$ between them. If all nodes in the block $i$ have received the message by time $t$, it follows from the above that, at the time $t + k\delta$, the message will also be received by all nodes in the block $j$. Since the unit square is partitioned in rows and columns of $\sqrt{5}/(5r)$ blocks, there is a sequence of, at most, $2\sqrt{5}/(5r)$ blocks between any two blocks $i$ and $j$, so that consecutive blocks are edge-adjacent. Hence, the total broadcast time is at most $2\delta\sqrt{5}/(5r)$. Since the diameter of the graph is at least $1/r$, the message transmission part is completed in time $O(D)$. The spokesman election part of the algorithm terminates in time $O(\log(5D^2/\epsilon)) = O(1 + \log D + \log 1/\epsilon)$ rounds. Hence, the total execution time of the algorithm is $O(D + \log 1/\epsilon)$, with probability at least $1 - \epsilon$.

4.6 Conclusion

In this chapter, we have shown upper and lower bounds on the spot arrival rate $\lambda$ such that the random graph $U(n, r, \lambda, s)$ is non-empty, is connected and can be used for efficient broadcasting. We have presented a broadcasting algorithm which successfully terminates with probability $1 - \epsilon$, in time $O(D + \log 1/\epsilon)$. These are the first analytic results concerned with communication in GRNs in the presence of positively correlated faults.

In this chapter, we have studied a scenario where faults result from phenomena external to the nodes. In the following chapter, we study the known phenomenon of swamping, when communication faults result from causes internal to the nodes.
Chapter 5

Broadcasting in Networks of Unknown Topology in the Presence of Swamping

5.1 Introduction

One of the known problems commonly faced by radio transceivers is that of swamping (cf., e.g., [6, 10, 38]). When two wireless nodes are at close proximity, their receivers cannot adapt to strong incoming signals; communication becomes difficult, even impossible. In contrast to traditional radio communication models, nodes at close proximity are not able to communicate directly; intermediate nodes are needed to relay their messages. In this chapter, we consider a wireless network where nodes suffer from the problem of swamping: nodes cannot receive messages from nodes within distance $s$ of them (the swamping distance) and may correctly receive messages only if no node within distance $s$ from them transmits.

We study analytically the problem of broadcasting in networks where nodes may be suffering from swamping. We propose broadcasting algorithms for this novel communication model which successfully broadcast in networks of unknown topology. Moreover, we propose algorithms to broadcast in optimal time complexity in the lattice line and in the two-dimensional lattice.

We would like to thank to Dr. Ioannis Lambadaris for suggesting the swamping paradigm used in this chapter.

5.1.1 The Model and Problem Definition

Typical wireless receivers are built from a radio-frequency amplifier, a demodulator and a decoder. The amplifier adapts the strength of the received signal such that it becomes usable for the demodulator stage. However, this amplifier is not ideal.

When the received signal strength is too low its output is either too weak or
too noisy to be usable; the first situation occurs when the communication range of a receiver is exceeded, for instance. When the received signal strength is too high, its input stage becomes saturated leading to a distorted signal (cf., e.g., [53]); in this case, we say that the receiver is swamped (cf., e.g., [6, 10, 38]). This occurs when there is a radio transmitter which is too close to a receiver. We now propose our model for this fault phenomenon. In what follows, whenever we speak of the distance, it is meant in its geometric sense, unless otherwise mentioned.

We work in the swamping communication model. Our graphs are built from a set $V$ of $|V| = n$ nodes, placed on the line (Sections 5.2 and 5.3) or on the plane (Sections 5.4 and 5.5). Nodes are equipped with communication range $r$ and limited by a minimum distance requirement of $s$ (the swamping distance). Two nodes $u, v \in V$ located at distance $\text{dist}(u, v)$ greater than $s$ and at most $r$ from one-another are neighbors and share an undirected link $(u, v) \in E$ in the graph $G$; no other links exist in $G$. In each round, each node is either a sender or a receiver. A node $u$ which is a transmitter in a given round sends a message to the entire set of its neighbors $\Gamma(u)$ within the same round; this transmission also makes the receiving of messages impossible for all nodes within distance $s$. More formally, for each round when a node within distance $s$ of it transmits, a node $v$ receives no message; in this case, only noise is heard by $v$, indistinguishable from the background noise heard when no messages are sent. In a fixed round, a node $v$ receives a message if and only if it is a receiver, exactly one of its neighbors is a sender, and no node within distance $s$ sends a message. If no neighbor of $v$ is a sender, then there is no message on the channel which $v$ can receive. If more than one neighbor of $v$ sends a message, we say that a collision occurs at $v$ and $v$ can only perceive noise on the channel. Nodes do not have collision detection abilities, i.e., they cannot distinguish collision noise from background noise (which is apparent when no messages are heard).

The swamping communication model can be viewed as a GRN on which radio communication is implemented with additional transient reception faults on all nodes at close proximity of a transmitter, i.e., a node cannot receive messages at each round when some node within distance $s$ from it transmits. Alternately, we can say that all incoming links of nodes at close proximity to a transmitter fail.
Throughout this chapter, we study networks of nodes placed on the line and on the plane which are either designed (sections 5.2 and 5.4) or of unknown topology (sections 5.3 and 5.5). Nodes are location-aware, i.e., each node knows its own location with respect to some global reference, but all nodes are unaware of the location of any other node. In the cases where the topology is unknown, we restrict attention to connected networks where nodes are positioned with some minimum distance $\alpha$ from each other. Let the parameter $g = 1/\alpha$ be called granularity. Nodes are also aware of the parameter $\alpha$, the swamping distance $s$, and the communication distance $1$.

We consider the process of broadcasting under the spontaneous wake up model in which all nodes are considered to be awake when the source begins transmission. Under this model, nodes may contribute to the broadcasting process even before receiving the source message, by exchanging control messages. In the sequel, we consider that nodes execute algorithms in a synchronous way.

We consider deterministic algorithms without global knowledge (Sections 5.2, 5.3, and 5.4) and with some knowledge about messages received by nodes close by (Section 5.5). In general, the algorithm is known to all nodes and its execution is based solely on the location of nodes in the network, the history known to each node, and the parameter $g$. In Section 5.5, the algorithm execution is based on the above-mentioned information augmented by the information about messages received by nodes surrounding each node.

5.1.2 Our Results

In Section 5.2, we address the problem of broadcasting on the lattice line and show a broadcasting algorithm, $A$, which correctly broadcasts the message $m$ on the lattice line of length $n$, in time $\lceil n/r \rceil + 3(\lceil r/(r - (s + 1)) \rceil + 1)$. This order of magnitude for the time complexity is optimal.

In Section 5.3, we provide an algorithm, $B$, to correctly broadcasts a message $m$ in a network of unknown topology in the line. Given a network diameter $D$, a minimum distance between nodes $\alpha$, a granularity parameter $g = 1/\alpha$ and $l = \max\{(1 - s), \alpha\}$, Algorithm $B$ completes broadcasting in time $O(D/l + g^2)$.

In Section 5.4, we address the problem of broadcasting on the two-dimensional
lattice and show a broadcasting algorithm, $A^2$, which correctly broadcasts the message $m$ in the two-dimensional lattice line of length $n$, in time $4\lceil\sqrt{n}/r\rceil + 12(\lceil r/(r - (s + 1))\rceil) + 1$. This order of magnitude for the time complexity is optimal.

In Section 5.5, we provide an algorithm, $B^2$, to correctly broadcasts a message $m$ in a network of unknown topology in the plane. Given a network diameter $D$, a minimum distance between nodes $\alpha$, a granularity parameter $g = 1/\alpha$ and $l = \max\{(1 - s)/(3\sqrt{2}), \alpha/\sqrt{2}\}$, Algorithm $B^2$ completes broadcasting in time $O(Dg/l + g^4)$.

### 5.2 Lattice Line

Throughout this section, we assume that $r$ and $s$ are positive integers. Consider a set of $n$ nodes placed at points $0, 1, \ldots, n - 1$ in one-dimensional Euclidean space; nodes are labeled according to their location. We call this placement of the nodes the lattice line. For simplicity, the communication and swamping ranges $r$ and $s$ are integer values; each node may reach nodes which are located on points at distance at least $s + 1$ from it, and at most $r$ from it. In this section, we consider the broadcasting of a message $m$ from the node 0 to all other nodes of the line.

In the sequel, we will present an algorithm $A$ and then prove the following result:

**Theorem 5.2.1** Algorithm $A$ correctly broadcasts the message $m$ on the lattice line of length $n$, in time $\lceil n/r \rceil + 3(\lceil r/(r - (s + 1))\rceil) + 1$. This order of magnitude is optimal.

#### 5.2.1 Non-Connectivity

Consider the case when $s > 0$ and $r - s = 1$. In this case, completing the broadcasting process in the lattice line is impossible.

**Lemma 5.2.2** If $s > 0$ and $r - s = 1$, then broadcast is impossible on the lattice line.

**Proof.** The source node has label 0 and possesses the source message $m$. Since $r - s = 1$, each node $u$ has a link to a node $v$ only if it is exactly at distance $r$ from it. We may thus model any path of this network by a sequence of additions and subtractions of $r$ on the source node label. Hence, the node 0 has paths only to nodes
whose labels are multiples of \( r \). For \( r \) and \( s > 0 \) integers, we have \( r > 1 \). In this case, the network is disconnected; broadcasting is impossible.

5.2.2 Fast Broadcast

In the previous section, we have shown conditions under which broadcast is impossible. We now show that, when these conditions are not met, broadcast is possible. We further show an algorithm for broadcasting in optimal time on the line.

Consider two sets of nodes on the line, \( A_k \) and \( B_k \), where \( A_k \) (resp. \( B_k \)) is the set of all nodes whose labels are in the interval \( [k, r - 1 + k] \) (\( [r + k, 2r - 1 + k] \)). We now describe the communication scheme \( \text{Local}_k \) for disseminating a message inside these sets. The scheme \( \text{Local}_k \) consists of \( x = \lceil r/(r - (s + 1)) \rceil \) steps \( i = 0, 1, \ldots, x - 1 \), each taking two rounds. For each step \( i \), in the first round the node with label \( a_{i,k} = i \cdot (r - (s + 1)) + k \) from the set \( A_k \) transmits the message \( m \); in the second round the node in set \( B_k \) with label \( b_{i,k} = a_{i,k} + r \) transmits the message \( m \). In the following lemma, we assume the absence of collision with nodes external to the communication scheme.

**Lemma 5.2.3** Given that the node \( k \) has previously received the message \( m \), all nodes in \( A_k \) will have received the message \( m \) at the end of scheme \( \text{Local}_k \).

**Proof.** Without loss of generality, let \( k = 0 \). We show by induction that the scheme \( \text{Local}_0 \) sends the message \( m \) to all nodes in the set \( A_0 \).

**Base step:**
In step 0, first node \( a_{0,0} = 0 \) transmits, the message \( m \); the nodes with labels \( s + 1, s + 2, \ldots, r \) receive the message. In the second round of step 0, the node \( b_{0,0} = r \) transmits; the nodes with labels \( 0, 1, \ldots, r - (s + 1) \) receive the message.

**Inductive hypothesis:**
Assume that the nodes with labels in \( [0, i \cdot (r - (s + 1))] \) have received the message by the end of of step \( i - 1 \). Then, by the end of step \( i \), the nodes with labels in \( [0, (i + 1) \cdot (r - (s + 1))] \) will have received the message.

**Proof of the inductive hypothesis:**
In the first round of step \( i \), the node \( a_{i,0} = i(r - (s + 1)) \) sends the message and the
nodes with labels in

\[ [i \cdot (r - (s + 1)) + (s + 1), i \cdot (r - (s + 1)) + r] \]

receive \( m \) from the node \( a_{i,0} = i \cdot (r - (s + 1)) \). In the second round of step \( i \), the node \( b_{i,0} = i(r - (s + 1)) + r \) sends the message and the nodes with labels in

\[ [i \cdot (r - (s + 1)), i \cdot (r - (s + 1)) + r - (s + 1)] \]

receive \( m \) from the node \( b_{i,0} = r + i \cdot (r - (s + 1)) \). The latter interval overlaps the set of nodes which had previously received the message \( m \). The largest label of the nodes in the receiving interval may be rewritten as \( (i + 1) \cdot (r - (s + 1)) \). This proves the inductive hypothesis.

Hence, after \( \lceil r/(r - (s + 1)) \rceil \) steps of scheme \( \text{Local}^*_k \), the nodes with labels in the interval \( [0, (\lceil r/(r - (s + 1)) \rceil) \cdot (r - (s + 1))] \) will have received the message. Since \( \lceil r/(r - (s + 1)) \rceil (r - (s + 1)) \geq r \), this proves the lemma.

Consider now the scheme \( \text{Local}^*_k \) consisting of the scheme \( \text{Local}_k \) to which one step is added: the step \( \lceil r/(r - (s + 1)) \rceil + 1 \).

**Lemma 5.2.4** Given that the node \( k \) has previously received the message \( m \), all nodes in \( A_k \) and \( B_k \) will have received the message \( m \) at the end of scheme \( \text{Local}^*_k \).

**Proof.** To obtain the set of labels of the nodes in \( B_k \) we need only add \( r \) to each label contained in the set \( A_k \). Hence, to prove the lemma, we need only show that if the nodes with labels in \([a, b] \cap A_k\) are informed at step \( i \), then, at step \( i + 1 \), the nodes with labels in \([a + r, b + r] \cap B_k\) are informed. For simplicity and without loss of generality, we set \( k = 0 \) in the following proof.

Observe that in step \( i \), the node with label \( i \cdot (r - (s + 1)) + r \) in interval \( B_k \) transmits the message to the nodes in subinterval

\[ [i \cdot (r - (s + 1)), i \cdot (r - (s + 1)) + r - (s + 1)] \]

of interval \( A_k \). Also observe than in step \( i + 1 \), the node with label \((i + 1) \cdot (r - (s + 1))\) in interval \( A \) transmits the message to the nodes in the subinterval

\[ [(i + 1) \cdot (r - (s + 1)) + (s + 1), (i + 1) \cdot (r - (s + 1)) + r] \]
of interval \( B_k \). To see that these intervals differ exactly by \( r \), we subtract \( r \) from the latter and obtain

\[
[(i + 1) \cdot (r - (s + 1)) + (s + 1) - r, (i + 1) \cdot (r - (s + 1)) + r - r] \\
[i \cdot (r - (s + 1)), (i + 1) \cdot (r - (s + 1))] 
\]

which corresponds to the former. The lemma follows.

Consider algorithm \( A \) for broadcasting on a line which consists of 2 parts. In the first part, the message is sparsely transmitted throughout the line by transmitting the message \( m \) sequentially by nodes \( 0, r, \ldots, \lfloor n/r \rfloor r \). In the second part, the scheme \( \text{Local}_k \) is executed, in parallel, for \( k = 0, 4r, \ldots, 4(\lfloor n/(4r) \rfloor - 1)r \) and then for \( k = 2r, 6r, \ldots, (4(\lfloor n/(4r) \rfloor - 1) - 2)r \). The algorithm is terminated by executing scheme \( \text{Local}_k \) for \( k = n - 2r \).

We extend the validity of Algorithm \( A \) to any source node by re-labeling nodes sequentially left to right such that the source has label 0. The sparse transmission part of the algorithm is then executed from node 0 through the positive node labels and then through the negative node labels. The remainder of the algorithm is identical.

We now prove the main theorem of this section.

**Proof of Theorem 5.2.1.** We begin by showing the algorithm correctness and then its execution time.

From Lemma 5.2.4, the scheme \( \text{Local}_k \) will successfully broadcast the message for each interval of size \( 2r \), given that the node with label \( k \) has received the message, and assuming no collision caused by transmissions by nodes external to the communication intervals. Since all transmissions that are executed in parallel originate at nodes whose labels differ by \( 4r \) in our algorithm, no collision occurs. Hence, if the sparse transmission part of the algorithm is successful, then the second part of the algorithm will also be successful. Moreover, if the second part of the algorithm is successful, the node whose label is \( n - 2r \) will have received the message before the termination phase of the second part. Hence, the termination phase of the algorithm will also be successful. Thus the algorithm correctly broadcasts the message \( m \).

We now count the number of rounds necessary to execute the algorithm. The number of rounds to complete the first part of \( A \) is \( \lfloor n/r \rfloor \). Each phase of the second
part and the termination phase of the algorithm are completed in time \([r/(r-(s+1))] + 1\). Hence, the algorithm is completed in \([n/r] + 3([r/(r-(s+1))] + 1)\) rounds.

We will now prove optimality. In order to disseminate the message from one end of the line to the other, \(n/r\) constitutes a trivial lower bound. For one node to receive a message, only one node within distance \(r\) must send a message. Therefore, avoiding all collisions, at most \(2(r-(s+1))\) nodes learn the message within each interval of size \(2r\), at each turn. It follows that, for all nodes of any interval of such size to know the message \(m\), it takes at least \(r/(r-(s+1))\) steps. Hence, the lower bound on transmission time is \(\Omega(n/r + r/(r-(s+1)))\).

### 5.3 Highway Model

In this section, we analyze the problem of broadcasting along a line segment of length \(L\) where nodes are placed by an adversary. Each node \(u\) is equipped to communicate with all nodes that are both within distance \(1\) and at distance greater than \(s\) from it. Hence, in this section, we assume that \(r = 1\) for simplicity. More formally, we describe the highway model. The communication range of a node \(u\) is the interval within distances \((s, 1]\) from \(u\). The size of the communication range is the length of this interval, i.e., \(1 - s\). The adversary designs the network such that it is connected and the distance between any pair of nodes \(u, v\) is at least \(\alpha\). We say that a network is connected if, for any node pair \(u, v\), there exists a path in the network from node \(u\) to node \(v\). Observe that the network is connected only if \(\alpha \leq 1\).

Message collisions result in noise indistinguishable from background noise and nodes are not equipped to detect these collisions. However, in this section we use the apparent silence from collisions to discover the presence of nodes through a collision-causing algorithm.

Nodes are aware of the parameter \(\alpha\) (and \(g = 1/\alpha\)) and the coordinate system of the line segment of length \(L\). Each node also knows the parameter \(s\), its swamping distance, and its communication distance \(1\).

In this section, we present a broadcasting algorithm \(B\) and show the following result.
Theorem 5.3.1 Algorithm $B$ broadcasts a message $m$ in a network of diameter $D$ in time $O(D/l + g^2)$, where $l = \max\{(1 - s), \alpha\}$.

In order to prove the main result of this section, we need several preparatory lemmas. The following fact requires no proof.

**Fact 2** For any node $u$ in a connected network, there is at least one node $v$ within the set $\Gamma(u)$.

### 5.3.1 Partition $\mathcal{P}$ of the Line

We now define a partition, called $\mathcal{P}$, on which our communication algorithm will operate. For each line segment in the partition below, the segment includes its leftmost point and excludes its rightmost point so that there will be no intersection between adjacent segments. We provide a graphical representation of the partition in Figure 5.1 and describe it below in detail.

![Partition P](image)

Figure 5.1: Partition $\mathcal{P}$

Partition the line into line segments of length 3, called *regions*. The line contains $\lceil L/3 \rceil$ regions, where $\lfloor L/3 \rfloor$ are of length 3 and at most one (the rightmost) is shorter, and even may consist of a single point.

Further, partition each region into smaller line segments, called *blocks*, of length $l = \max\{(1 - s), \alpha\}$. Here, $l \leq 1$ since both $\alpha \leq 1$ and $1 - s \leq 1$. Each region
contains \( \mu = \lceil 3/l \rceil \) blocks, where \( \lceil 3/(1-s) \rceil \) are of length \( l \) and at most one (the rightmost) is shorter, and even may consist of a single point. For each region, label blocks 1, 2, \ldots, \mu, from left to right.

Partition also each block into line segments of length \( \alpha \), called homes. Each block contains \( \nu = \lceil l/\alpha \rceil \) homes, where \( \lceil l/\alpha \rceil \) are of length \( \alpha \) and at most one (the rightmost) is shorter, and even may consist of a single point. For each block, label homes 1, 2, \ldots, \nu, from left to right.

**Partition Properties**

We now show communication properties related to the partition defined above. We first show that transmissions in distinct regions do not collide, if properly scheduled. We then show that transmissions by a few distinguished nodes in a part or all of a block can reach all neighbors of nodes on this block or part of a block. However, before showing these properties, we observe that since homes are of length at most \( \alpha \), at most one node can occupy each home. Hence, we have the following lemma.

**Lemma 5.3.2** Each home contains at most one node.

**Lemma 5.3.3** Transmissions from unique nodes inside identically labeled blocks in distinct regions do not collide.

**Proof.** Consider nodes \( u, v \) in different regions and identically labeled blocks. Each region has length 3 and each block has length \( l \leq 1 \). Because the block labels are identical in each region, the minimum distance between two identically-labeled blocks (that contain the nodes \( u, v \)) is \( 3-l \geq 2 \). Since each line segment of the partition excludes its rightmost point, there is no point within distance 1 of both \( u \) and \( v \).

**Lemma 5.3.4** Consider any pair \( u, v \) of nodes within distance \( 1-s \). Also consider the set \( \mathcal{U} \) of all nodes inclusively located between \( u \) and \( v \). We have that \( \Gamma(\mathcal{U}) = \Gamma(u) \cup \Gamma(v) \).

**Proof.** Consider two nodes \( u, v \) at distance \( d \leq 1-s \) from one-another; \( u \) is to the left of \( v \) and \( u \) is at coordinate 0. Consider the right part of the range of \( u \) and \( v \).
Then, the range of \( u \) to the right covers the interval \((s, 1]\). Similarly, the range of \( v \) to the right covers the interval \((s + d, 1 + d]\). Since \( 0 < d \leq 1 - s \), we have that \( s + d \leq s + 1 - s = 1 \). Hence, the ranges of \( u \) and \( v \) overlap and cover the interval \((s, 1 + d]\).

Consider any node \( w \) between \( u \) and \( v \), i.e., at distance \( \text{dist}(u, w) \) from \( u \), with \( 0 < \text{dist}(u, w) < d \). The range of \( w \) to the right covers the interval \((s + \text{dist}(u, w), 1 + \text{dist}(u, w)]\). Since \( 0 < \text{dist}(u, w) < d \), the range of \( w \) to the right is completely included in the ranges of \( u \) and \( v \) to the right. The argument is symmetric for the left.

In the following sections, we describe communication procedures that will enable nodes to broadcast messages to all nodes of their networks.

5.3.2 Procedure \( D^* \) for Neighborhood Discovery

We now define procedures used to communicate once from each node to all other nodes within distance 1 of them. We refer to this process as Neighborhood Discovery.

Procedure \( D \)

We now present Procedure \( D \), in which nodes in distinct homes inside a region sequentially send a message while other nodes listen. This procedure is executed in parallel over all regions and for all \((\text{block}, \text{home})\) labels sequentially. All homes with some \((\text{block}, \text{home})\) label transmit a message while all other nodes listen for incoming messages. More formally, refer to the code for Procedure \( D \). By Lemma 5.3.3, no collision occurs in this procedure. Hence we claim that each node will gain knowledge of all nodes located within its communication range as a result of Procedure \( D \). We have the following lemma:

**Lemma 5.3.5** After one execution of Procedure \( D \), nodes know of all nodes within distance 1 and greater than \( s \) of them.

**Proof.** Fix any node \( u \) and the set \( \Gamma(u) \) of all nodes within its communication range. By Lemma 5.3.3, no message collision can occur during Procedure \( D \). Since
Procedure $\mathcal{D}$

In parallel for all regions

- $N_u \leftarrow \emptyset$ // the set of nodes known to $u$
- $H_u \leftarrow$ the (block, home) label of $u$

for block $= 1..\mu$ do

  for home $= 1..\nu$ do

    if $H_u = (\text{block}, \text{home})$ then
      Transmit hello
    else if a hello is heard then
      $N_u \leftarrow N_u \cup (\text{block}, \text{home})$

the procedure makes nodes in all the (block, home) couples transmit, it then follows that $u$ must receive messages from all the nodes in the set $\Gamma(u)$.

By the previous lemma, since the graph is connected, each node will discover at least one node within its communication range by the end of procedure $\mathcal{D}$. We will use this fact in the following subsection to allow all nodes to discover all other nodes within distance $s$ of them.

Procedure $\mathcal{D}^*$

Recall that, in our model, it is not possible for any node to hear messages from nodes within distance $s$ of them. Observe that the length of the path between two nodes within distance $s$ is only bounded above by the diameter $D$, in many cases. We now concentrate on a time-guaranteed procedure for discovery of nodes within distance 1.

Consider Procedure $\mathcal{D}_{(b,h)}$ which uses the absence of distinguishable messages from collisions to discover nodes within distance $s$.

**Lemma 5.3.6** By Procedure $\mathcal{D}_{(b,h)}$, nodes neighbor to $(b,h)$ know all other nodes within distance 1 of them in time $\Theta(g)$.

**Proof.** The time complexity of Procedure $\mathcal{D}_{(b,h)}$ is in $\Theta(\mu \nu)$. With $\alpha \leq l \leq 1$, we have that $\mu = \lceil 3/l \rceil \in \Theta(1/l)$ and $\nu = \lceil \sqrt{2l}/\alpha \rceil \in \Theta(l/\alpha)$. Hence,

$$\mu \nu \in \Theta((1/l)(l/\alpha)) = \Theta(1/\alpha) = \Theta(g).$$
We now prove correctness. Fix a node $u$ which shares a link with the node $(b, h)$. During the execution of Procedure $\mathcal{D}_{(b,h)}$, the node $(b, h)$ will transmit messages at every round. A message from $(b, h)$ will be heard by $u$ at every round when no collision occurs at $u$. Furthermore, when no message can be distinguished, another node within distance 1 of $u$ must be transmitting from the home with label $(\text{block, home})$ (as defined in the procedure). Since Procedure $\mathcal{D}_{(b,h)}$ schedules all nodes to transmit in pairs with $(b, h)$, upon completion of this procedure, the node $u$ will have discovered all nodes $w$ for which the distance $\text{dist}(u, w)$ from $u$ is at most 1.

Now consider Procedure $\mathcal{D}^*$ consisting of one execution of Procedure $\mathcal{D}$ followed by the execution of Procedure $\mathcal{D}_{(b,h)}$ for all $(b, h) \in \{1, 2, \ldots, \mu\} \times \{1, 2, \ldots, \nu\}$. In plain words, Procedure $\mathcal{D}^*$ schedules colliding transmissions for all $(\text{block, home})$ couple pairs

$$((b, h), (b', h')) \in \{(1, 2, \ldots, \mu) \times (1, 2, \ldots, \nu)\}^2.$$

More formally, refer to the pseudo code for Procedure $\mathcal{D}^*$.

**Procedure $\mathcal{D}^*$**

Call Procedure $\mathcal{D}$

for $b = 1..\mu$ do
  for $h = 1..\nu$ do
    Call Procedure $\mathcal{D}_{(b,h)}$
Lemma 5.3.7 By Procedure $D^*$, nodes know all other nodes within distance 1 of them in time $\Theta(g^2)$.

Proof. The time complexity of Procedure $D_{(b,h)}$ is in $\Theta(\mu \nu)$. Hence, the time complexity of Procedure $D^*$ is in $\Theta(\mu^2 \nu^2)$. By the above and by Lemma 5.3.6, the time complexity of Procedure $D^*$ is therefore in $\Theta(g^2)$.

We now prove correctness. By Fact 2, for any node $u$, since the graph is connected, there exists a node $(b, h)$ such that Lemma 5.3.6 will hold. By the above and by Lemma 5.3.6, all nodes know all other nodes that are within distance 1 of them.

With knowledge of all nodes within distance 1, nodes have the basic tools to select distinguished nodes to relay messages for all nodes of a block. We discuss such a procedure in the following subsection.

5.3.3 Selection of Spokesman Nodes

We now describe a procedure for selection of distinguished nodes for each block known as spokesmen. We wish to select these spokesmen in order to avoid collisions and speed up the broadcasting process. Before we concentrate on the different cases, we present the following fact.

Fact 3 Given location-awareness, if a sender includes its location inside a message, then a receiver can determine all points where the message may be received.

Proof. The sender knows its own location and therefore can incorporate this as part of his message. The receiver then knows the origin of the received message and hence can determine the covered region.

Consider the spokesman selection procedure that elects, for each block,

1. right (left) boundary spokesmen: the node in the rightmost (leftmost) home known to be completely contained within the transmission range of a sender, if this home is the rightmost (leftmost) home on the block;

2. right (left) range spokesmen: the node in the rightmost (leftmost) home known to be completely contained in the transmission range of a sender, if this home is not the rightmost (leftmost) home on the block;
3. right (left) potential spokesmen: the node in the rightmost (leftmost) home known to be partially contained within the transmission range of a sender.

We now show that the spokesman selection procedure making the above selections selects unique spokesmen for each type.

**Lemma 5.3.8** The spokesman selection procedure selects at most one node for each spokesman type.

**Proof.** Given that right and left boundary spokesmen are unique by definition (those nodes in the home that is closest to the block boundaries), we prove the lemma for right and left range and potential spokesmen. In the case when \( l = \alpha \), there is only one home per block, hence the lemma holds in this case. We now prove the lemma for the case when \( l = 1 - s \).

Given that the range of a node is of size \( 1 - s = l \), the range of a transmitter always encloses at least one of the homes that is closest to the block boundaries; call this home a boundary home. For any set \( S \) of transmitters whose ranges enclose a same boundary home, the intersection of their communication ranges with the block \( t \) defines a set \( I \) of intervals for which one is the largest. This largest interval is the communication range of a node \( u \in S \) that includes all other communication ranges inside of the set \( I \). By Fact 3, all nodes located inside this interval know the limits of the communication range of \( u \). It follows that the potential and range spokesmen for the set of nodes \( S \) are unique. These spokesmen are right (left) potential and range spokesmen if the leftmost (rightmost) home of \( t \) is completely included in the range of \( u \) and not the rightmost (leftmost) home of \( t \).

It also follows from the above discussion that for any pair of transmitters \( u \) and \( v \) whose ranges do not enclose a same boundary home, the spokesmen types defined will be different (right vs. left spokesmen).

5.3.4 Broadcasting Algorithm \( B \)

In order to complete the broadcasting algorithm, we need a final procedure to transmit the message \( m \) from the source to all other nodes of the network. We now describe Procedure \( T \).
Procedure $\mathcal{T}$

$S_u \leftarrow$ the label of the block, home containing $u$

**In parallel for all regions**

for $block = 1..\mu$ do

if $S_u = block$ then

update spokesman status

Spokesmen transmit the message $m$ in the following order:

1) left boundary spokesman,
2) right boundary spokesman,
3) left range spokesman,
4) right range spokesman,
5) left potential spokesman,
6) right potential spokesman

else

Listen to incoming messages for 6 turns

---

**Lemma 5.3.9** Procedure $\mathcal{T}$ broadcasts the message correctly through the network in time $O(D/l)$.

**Proof.** Consider a network $G$ of diameter $D$ built by the adversary under the swamping model. Consider also the network $G'$ with the same nodes and links as $G$, but where nodes may receive messages from multiple neighbors in one round without collisions. Let the broadcasting algorithm $\mathcal{F}$ execute such that, when a node receives a message $m$ the first time, it transmits this message to all its neighbors the next turn. The algorithm $\mathcal{F}$ executes in $\Theta(D)$ rounds on the network $G'$. We prove the lemma statement by comparing the execution of Procedure $\mathcal{T}$ on $G$ to the execution of Algorithm $\mathcal{F}$ on $G'$.

Consider $G$ and the partition $\mathcal{P}$. Since each region has $\lceil 3/l \rceil$ blocks, where $l = \max\{\alpha, (1 - s)\}$ and since each block has a constant number of spokesmen, the broadcast algorithm sequentially makes all spokesmen of a region communicate every $\Theta(1/l)$ turns. From Lemma 5.3.3, the process is collision-free. From Lemma 5.3.4, the spokesmen of a block reach all the nodes that can be reached by any node on
their block that do know the message $m$. It then follows that the message $m$ being relayed through the network may be slowed down by a factor $O(1/l)$ with respect to the execution of Algorithm $\mathcal{F}$ in $G'$. Hence, for any network $G$ of diameter $D$, the total transmission time is in $O(D/l)$.

**Algorithm $B$**

- **In parallel** for all nodes
- Call Procedure $D^*$
- Call Procedure $T$

**Proof of Theorem 5.3.1.** From Lemma 5.3.9, the time of execution of Procedure $T$ is $O(D/l)$. From Lemma 5.3.7, the time of execution of Procedure $D^*$ is $\Theta(g^2)$. Adding these times together, we get a total time of $O(D/l + g^2)$.

**5.4 Two-Dimensional Lattice**

In Section 5.2, we have shown an optimal time broadcast algorithm for the lattice line. We now extend this result to multi-dimensional lattices. Hence, we consider the set $V$ of $n$ nodes placed at Euclidean coordinates $(i, j)$ for $i = 0, 1, \ldots, \sqrt{n} - 1$ and $j = 0, 1, \ldots, \sqrt{n} - 1$. We call this placement of the nodes the **two-dimensional lattice**.

Consider that each node has a communication range $r$ and a swamping range $s$ such that $r - s > 1$. Throughout this section, we assume that $r$ and $s$ are positive integers. We call transmission annulus of $u$ the region at distance greater than $s$ and at most $r$ from a node $u$ and denote it by $A_u$. Each node $u$ shares a link with each node $v$ located within $A_u$. The set of links $E$ is the union of all these shared links. We will present Algorithm $A^2$, an extension of Algorithm $A$, to broadcast a message in this two-dimensional lattice. In this section we will prove the following result:

**Theorem 5.4.1** Algorithm $A^2$ broadcasts in time $4[\sqrt{n}/r] + 12([r/(r-(s+1))] + 1)$. This order of magnitude is optimal.

In order to present Algorithm $A^2$ and prove the main theorem of this section, we need a preparatory lemma. Fix one row $l$ of nodes in the square lattice and consider
the region covered by all the transmission annuli in the execution of algorithm $A$ on this line.

**Lemma 5.4.2** Algorithm $A$ broadcasts the message to all nodes on $l$ and to all nodes within distance $\lfloor \sqrt{3}r/2 \rfloor$ from $l$.

**Proof.** In Algorithm $A$, nodes broadcast a message along a line. To do so, nodes at distance at most $r-(s+1)$ from one to the next transmit the message. The algorithm is successful because each node sends the message symmetrically to intervals of length $r-s$, resulting in complete coverage of the line by the set of transmitting nodes. It follows that, if each transmitter on a line also covers a length $r-s$ on a parallel lattice line then, complete coverage of this line would be achieved by the end of this Algorithm $A$. For a line $l_d$ parallel to $l$ and at distance $d$ from $l$ and for a fixed node $u$ on the line $l$, let the segment $l_{sd}$ be a line segment resulting from the intersection of the communication annulus of $u$ and the line $l_d$.

We now evaluate the length of $l_{sd}$. Let $L$ denote the length of some line segment $l_{sd}$. For $d \leq s$, there are 2 line segments on each lattice line. By the law of cosines, we have that the length $L$ is

$$L^2 = r^2 + s^2 - 2rs\cos(\theta) \geq r^2 - 2rs + s^2 = (r - s)^2.$$  

Hence, for all line segments with one endpoint at distance $s$ from $u$ and another at distance $r$ we have that $L \geq r - s$ for all segments $l_{sd}$. See Figure 5.2.

![Figure 5.2](image-url)

On the other hand, for $s < d \leq r$, there is a single line segment on each lattice
line with length \( l_d = 2\sqrt{r^2 - d^2} \). In this case, for \( d \leq \sqrt{3r^2/4 + rs/2 - s^2/4} \) we have

\[
l_d \geq 2\sqrt{r^2 - (3r^2/4 + rs/2 - s^2/4)} = 2\sqrt{r^2/4 - rs/2 + s^2/4} = 2(r - s)/2.
\]

Since \( r > s \), this is also true if \( d \leq \sqrt{3r}/2 \).

Hence, Algorithm \( \mathcal{A} \) completes message dissemination on all lines within distance \( \lfloor \sqrt{3r}/2 \rfloor \) from \( l \).

Building on Algorithm \( \mathcal{A} \) and on the fact that it broadcasts the message \( m \) to all nodes on a line \( l \) and to all nodes within distance \( d = \lfloor \sqrt{3r}/2 \rfloor \) from \( l \), Algorithm \( \mathcal{A}^2 \) operates in 4 phases:

1. Execution of Algorithm \( \mathcal{A} \) on the horizontal line \( l \) of the source node.
2. Execution of Algorithm \( \mathcal{A} \) on the vertical lines at coordinates \( d, 7d, 13d, \ldots \)
3. Execution of Algorithm \( \mathcal{A} \) on the vertical lines at coordinates \( 3d, 9d, 15d, \ldots \)
4. Execution of Algorithm \( \mathcal{A} \) on the vertical lines at coordinates \( 5d, 11d, 17d, \ldots \)

As it was the case for Algorithm \( \mathcal{A} \), algorithm \( \mathcal{A}^2 \) is valid for any source node.

We now prove the main theorem of this section.

**Proof of Theorem 5.4.1.** By Lemma 5.4.2, and considering the set of lines at distance \( 2d = 2\lfloor \sqrt{3r}/2 \rfloor \), Algorithm \( \mathcal{A}^2 \) will achieve complete coverage of the lattice within its execution, given that no collision occurs.

We now demonstrate that no collision occurs from simultaneous transmissions. To show this, we show that any two nodes transmitting in one round are at distance at least \( 2r \) from one-another. Any simultaneous transmission occurs from nodes at distance \( 6\lfloor \sqrt{3r}/2 \rfloor \) from one-another. We have that the distance between any two nodes transmitting in the same round is

\[
6\lfloor \sqrt{3r}/2 \rfloor > 3\sqrt{3r} - 6 > 5.19r - 6 = 2r + (3.19r - 6) > 2r \text{ for } r \geq 2.
\]

An implication of Lemma 5.2.2 is that broadcasting in lattice networks with swamping is impossible for \( r < 2 \). Hence the assumption that \( r \geq 2 \) is true in all cases when broadcast is possible.
Since Algorithm $\mathcal{A}^2$ is a sequence of 4 executions of Algorithm $\mathcal{A}$, it runs in the time of 4 execution of algorithm $\mathcal{A}$ on lines of length $\sqrt{n}$. Hence, the execution time of Algorithm $\mathcal{A}^2$ is $4[\sqrt{n}/r] + 12([r/(r - (s + 1))] + 1).

We now prove that our algorithm is of optimal time complexity. Nodes within the swamping radius of a transmitting node cannot receive any message; the maximum number of nodes which can receive a message in one round within the communication radius of a node is then the nodes within its annulus. Consider a set of nodes $N$, at distance $L$ from the source node, where $L$ is some multiple of $r$. Consider further that this set $N$ is the intersection of a disk of diameter $r$ and the square lattice. At least $L/r$ rounds are needed for the message to reach the set $N$. At the following round, broadcasting within $N$ may begin. From the communication model, at most $\pi(r^2 - s^2)$ nodes inside $N$ may receive the message within each round. Hence, because $2\pi r(r - s) > \pi(r^2 - s^2)$, less than $2\pi r(r - s)$ nodes receive the message in each round. Therefore, the total broadcasting time is at least

$$\left\lceil \frac{L}{r} \right\rceil + \left\lceil \frac{\pi r^2}{2\pi r(r - s)} \right\rceil = \left\lceil \frac{L}{r} + \frac{r}{2(r-s)} \right\rceil.$$ 

For communication in the $n$-node square lattice, we obtain a lower bound on the broadcasting time which is

$$\left\lceil \frac{\sqrt{n}}{2r} \right\rceil + \left\lceil \frac{r}{2(r-s)} \right\rceil \in \Omega \left( \frac{\sqrt{n}}{2r} + \frac{r}{2(r-s)} \right),$$

for $r > 1$ and $s$ integers and for $r - s > 1$. This lower bound matches the time complexity of Algorithm $\mathcal{A}^2$.

By the same technique used to extend Algorithm $\mathcal{A}$ to Algorithm $\mathcal{A}^2$, we may extend the Algorithm $\mathcal{A}$ to an algorithm $\mathcal{A}^d$, broadcasting in the $d$-dimensional lattice, when $d \in \Theta(1)$. By a proof similar to that of Theorem 5.4.1, we may show the following result.

**Lemma 5.4.3** Algorithm $\mathcal{A}^d$ broadcasts in the $d$-dimensional lattice, $d \in \Theta(1)$, with time complexity in $\Theta(d\sqrt{n} + dr/(r - (s + 1)))$. 

5.5 City Model

We now consider the task of broadcasting in a connected network of unknown topology. In particular, we consider networks with nodes placed at points on the plane, located at least at some geometric distance $\alpha$ from each other. Each node $u$ is equipped to communicate with all nodes that are both within distance 1 and at distance greater than $s$ from it. Hence, in this section, we assume that $r = 1$ for simplicity. More formally, we describe the city model. The communication range of a node $u$ is the annulus centered at $u$ with radii $s$ and 1. The size of the communication range is the width of this annulus, i.e., $1 - s$. The adversary designs the network such that it is connected and the distance between any pair of nodes $u, v$ is at least $\alpha$. We say that a network is connected if, for any node pair $u, v$, there exists a path in the network from node $u$ to node $v$. Observe that the network is connected only if $\alpha \leq 1$.

Nodes are aware of the parameter $\alpha$ (and $g = 1/\alpha$) and the coordinate system of the plane. Each node also knows the parameter $s$, its swamping distance, and its communication distance 1.

We wish to complete broadcasting in a collision avoidance scheme. We will use the assumption of spontaneous wake-up of the nodes. In this section we use the apparent silence from collisions to discover the presence of nodes through a collision-causing process, used before the transmission part of the broadcasting algorithm. Moreover, we will assume that nodes know about the transmissions made within close proximity. We will show the following result.

**Theorem 5.5.1** Algorithm $B^2$ broadcasts a message $m$ in a network of diameter $D$ in time $O(Dg/l + g^4)$, where $l = \max\{(1 - s)/(3\sqrt{2}), \alpha/\sqrt{2}\}$.

5.5.1 Partition $P^2$ of the Plane

We now define a partition, called $P^2$, on which our communication algorithm will operate.

Each square in the partition below includes its North border, its West border, and both its North vertices; it excludes its East border, its South border and both its South vertices. We provide a graphical representation of the partition in Figure
5.3 and now describe it below.

Partition the plane into a mesh of $3 \times 3$ squares called **regions**.

Further partition each region into a mesh of $l \times l$ squares, called **blocks**, with length $l = \max\{(1 - s)/(3\sqrt{2}), \alpha/\sqrt{2}\}$. Here, $l \leq 1/\sqrt{2}$ since both $\alpha \leq 1$ and $(1 - s)/3 \leq 1$. Each region contains $\mu = \lceil 3/l \rceil^2$ blocks, where $\lceil 3/l \rceil^2$ are of area $l^2$ and at most $2\lceil 3/l \rceil + 1$ are smaller, and even may consist of a single line or point. For each region, label blocks $1, 2, \ldots, \mu$, from West-East row by row, North to South.

Partition also each block into a mesh of $\alpha/\sqrt{2} \times \alpha/\sqrt{2}$ squares, called **homes**. Each block contains $\nu = \lceil \sqrt{2}l/\alpha \rceil^2$ homes, where $\lceil \sqrt{2}l/\alpha \rceil^2$ are of area $\alpha^2/2$ and at most $2\lceil \sqrt{2}l/\alpha \rceil + 1$ are smaller, and even may consist of a single line or point. For each block, label homes sequentially $1, 2, \ldots, \nu$, from West-East row by row, North to South.

**Partition Properties**

In section 5.3, we showed properties for the partition $\mathcal{P}$. We now show the validity of lemmas 5.3.2 and 5.3.3 for the partition $\mathcal{P}^2$. For ease of reading, we now repeat these lemmas.

**Lemma 5.3.2** Each home contains at most one node.

Observe that since homes have diameter at most $\alpha$, at most one node can occupy each home. Hence, Lemma 5.3.2 holds for partition $\mathcal{P}^2$. 
Lemma 5.3.3  *Transmissions from unique nodes inside identically labeled blocks in distinct regions do not collide.*

Consider nodes $u, v$ in different regions and identically labeled blocks. Since each region has side length 3 and each block has side length $l \leq 1$, Lemma 5.3.3 holds for $P^2$.

In the following sections, we describe communication procedures that will enable nodes to broadcast messages to all nodes of their networks.

### 5.5.2 Procedure $\mathcal{D}$ for Nodes in Range

Recall Procedure $\mathcal{D}$ in which nodes send a message sequentially based to their home label. Since Lemma 5.3.2 and Lemma 5.3.3 both hold for $P^2$, we have that Lemma 5.3.5 also still holds for $P^2$.

**Lemma 5.3.5**  *Upon completion of Procedure $\mathcal{D}$, nodes know of all nodes within distance 1 and greater than $s$ of them.*

By repeating Procedure $\mathcal{D}$ $i$ times (augmenting the *hello* message with the location of known nodes), a node $u$ can learn about other nodes within hop distance $i$ ($\Gamma_{\leq i}(u)$). However, the hop distance from a node $u$ to a node $v$ may be arbitrarily large, even if $v$ is within geometric distance 1 of $u$. 

Procedure $\mathcal{D}$

In parallel for all nodes $u \in V$

$N_u \leftarrow \emptyset$  // the set of nodes known to $u$

$H_u \leftarrow$ the (block, home) label of $u$

for block = 1..$\mu$ do

for home = 1..$\nu$ do

if $H_u = (\text{block}, \text{home})$ then

Transmit *hello*

else if a *hello* is heard then

$N_u \leftarrow N_u \cup (\text{block}, \text{home})$
Hence, for diameter $D$ graphs, the use of Procedure $D$ alone could take as many as $Dg^2$ rounds to discover the existence of all nodes within distance 1. In this case, the message could be transmitted without the assumption of spontaneous wake up from the source to the nodes. We have the following lemma:

**Lemma 5.5.2** In all networks where nodes are placed on the plane, of diameter $D$ and granularity $g$, the broadcast time is in $O(Dg^2)$.

Hence, under our communication model, Procedure $D$ is insufficient to speed up broadcast in the spontaneous wake-up model as opposed to the conditional wake-up model, in the worst case.

**5.5.3 Procedure $D^*$ for Neighborhood Discovery**

Recall Procedure $D_{(b,h)}$ using collisions to discover nodes within distance $s$.

```
Procedure $D_{(b,h)}$

// $N_u$ is the set of nodes known to $u$
// $H_u$ is the (block, home) label of $u$

In parallel for all nodes $u \in V$
for block = 1..$\mu$ do
  for home = 1..$\nu$ do
    if $H_u = (\text{block, home})$ OR $H_u = (b, h)$ then
      Transmit hello
    else if no hello is heard AND $(b, h) \in N_u$ then
      $N_u \leftarrow N_u \cup (\text{block, home})$
```

**Lemma 5.5.3** By Procedure $D_{(b,h)}$, nodes neighbor to $(b, h)$ know all other nodes within geometric distance 1 of them in time $\Theta(g^2)$.

**Proof.** The time complexity of Procedure $D_{(b,h)}$ is in $\Theta(\mu\nu)$. With $\alpha \leq l \leq 1$, we have that $\mu = \lceil 3/l \rceil \in \Theta((1/l)^2)$ and $\nu = \lceil \sqrt{2l}/\alpha \rceil \in \Theta((l/\alpha)^2)$. Hence,

$$\mu\nu \in \Theta((1/l)^2(l/\alpha)^2) = \Theta((1/\alpha)^2) = \Theta(g^2).$$
We now prove correctness. Consider the execution of Procedure $D_{(b,h)}$, during which the node $(b,h)$ will transmit messages at every round. A message from $(b,h)$ will be heard by $u$ at every round when no collision occurs at $u$. Furthermore, when no message can be distinguished, another node within distance 1 of $u$ must be transmitting from the home with label $(block, home)$ (as defined in the procedure). Since Procedure $D_{(b,h)}$ schedules all nodes to transmit in pairs with $(b,h)$, upon completion of this procedure, the node $u$ will have discovered all nodes $w$ for which the geometric distance $dist(u,w)$ from $u$ is at most 1.

Recall Procedure $D^*$ consisting of one execution of Procedure $D$ followed by the execution of Procedure $D_{(b,h)}$ for all $(b,h) \in \{1,2,\ldots,\mu\} \times \{1,2,\ldots,\nu\}$. For the plane, Procedure $D^*$ allows the discovery of nodes within distance 1. More formally, refer to the pseudo code for Procedure $D^*$.

**Procedure $D^*$**

Call Procedure $D$

for $b = 1..\mu$ do

for $h = 1..\nu$ do

Call Procedure $D_{(b,h)}$

Procedure $D^*$ accomplishes the same function in the plane as it does in the line however, with increased time complexity.

**Lemma 5.5.4** By Procedure $D^*$, nodes know all other nodes within distance 1 of them in time $\Theta(g^4)$.

**Proof.** The time complexity of Procedure $D_{(b,h)}$ is in $\Theta(\mu \nu)$. Hence, the time complexity of Procedure $D^*$ is in $\Theta(\mu^2 \nu^2)$. By the above and by Lemma 5.5.3, the time complexity of Procedure $D^*$ is therefore in $\Theta(g^4)$.

We now prove correctness. For any node $u$, since the graph is connected, by Fact 2 there exists a node $(b,h)$ such that Procedure $D_{(b,h)}$ will be executed. By the above and by Lemma 5.5.3, all nodes know all other nodes that are within distance 1 of them. ■
With knowledge of all nodes within distance 1, nodes have the basic tools to select distinguished nodes to relay messages from within a bounded-diameter region of the plane. We discuss such a procedure in the following section.

5.5.4 Selection of Spokesman Nodes

In this section, we assume that nodes know which nodes of their own block possess the source message $m$. The spokesmen nodes are those nodes in each row, column and diagonal which possess the message and which are located in the home which is closest to either end of that row, column or diagonal. We now state the following lemma.

**Lemma 5.5.5** If all spokesmen of a block $b$ transmit in a collision-avoidance scheme, then all nodes neighbor to any node in $b$ will receive the source message.

The proof will be given following some preliminary facts and discussion. More formally, the rules for deciding which nodes are spokesmen are as follows: For a row (column) of homes of partition $\mathcal{P}^2$, among nodes possessing the message, those two nodes in homes closest to the West and East (North and South) borders of a block in $\mathcal{P}^2$ are spokesmen. For a diagonal of homes of partition $\mathcal{P}^2$, among nodes possessing the message, those two nodes in homes closest to the borders of a block in $\mathcal{P}^2$ are spokesmen. See Figure 5.4.

If a spokesman is chosen in column (row) $i$ because of its proximity to the North or South (West or East) border, then it has the label $N_i$ and/or $S_i$, resp. ($W_i$ and/or $E_i$, resp.). If a spokesman is chosen in Southeast-Northwest (Southwest-Northeast) diagonal $i$ because of its proximity to the Southeastern or Northwestern (Southwestern or Northeastern) border, then it has the label $SE_i$ and/or $NW_i$, resp. ($SW_i$ and/or $NE_i$, resp.). Spokesmen can hold many labels.

Observe that there are $O(l^2g^2)$ homes inside a block; there are $O(lg)$ rows of homes, $O(lg)$ columns of homes and $O(lg)$ diagonals of homes inside a block; there are at most 2 spokesmen elected for each row, each column and each diagonal. Hence, each block contains $O(lg)$ spokesmen. We now claim that only these spokesmen are necessary to broadcast.

Before presenting the proof, we recall the following fact.
Fact 4 Consider two vertices $A$ and $B$ and the line $AB$ joining them. The line $l$ perpendicular to $AB$ and through its center defines two halfplanes $H_{A,B}$ and $H_{B,A}$. The halfplane $H_{A,B}$ (resp. $H_{B,A}$) contains $A$ ($B$) and has all points closer to $A$ ($B$) than to $B$ ($A$).

We now proceed to the presentation of two preparatory lemmas: Lemma 5.5.6 and Lemma 5.5.7. Using these lemmas, we will then prove Lemma 5.5.5.

Lemma 5.5.6 The set of spokesmen of a block is closer to any point $p$ outside the block than any non-spokesman node.

Proof. Consider the sector $S$ of a plane defined by the angle $ACB$ of a triangle. We first show that if the angle $\theta$ at $C$ is at most $\pi/2$, then all points in the sector outside the triangle $ACB$ are closer to $A$ and $B$ than they are to $C$.

Consider the halfplanes defined by the vertex pairs $A, C$ and $B, C$ as described in Fact 4. If the node $C$ is closer than $A$ and $B$ to a point $p$, then $p$ is in the intersection of $H_{C,A}$ and $H_{C,B}$. Moreover, if $\theta = \pi/2$, then $S \cap H_{C,A} \cap H_{C,B}$ is a rectangle contained within the triangle $ACB$. As $\theta$ decreases, the region $S \cap H_{C,A} \cap H_{C,B}$ remains contained within the triangle $ACB$. See Figure 5.5. It follows that all other points of $S$ are closer to either $A$ or $B$. 
Now consider a non-spokesman node $u$ and the set of all spokesmen in its row, column and diagonals. For $u$ not to be a spokesman, it must have one spokesman on each side of itself for its row, column and diagonals. Let these spokesmen be labeled sequentially $u_1, u_2, \ldots, u_8$ in a clockwise order around the node $u$. Consider a partitioning of the plane around $u$ by the set of half-lines starting at $u$ and going through $u_1, u_2, \ldots, u_8$. Call these plane regions the sectors $u_i u u_{(i+1) \mod 8}$.

Since the distance between nodes is at least $\alpha$ and because of the geometry of the partition, we have that the angle of each sector $u_i u u_{(i+1) \mod 8}$ is less than $\pi/2$. By the first part of the argument, the node $u$ is farther from any point in a sector $u_i u u_{(i+1) \mod 8}$ and outside the triangle $u_i u u_{(i+1) \mod 8}$, than the spokesmen $u_i$ and $u_{(i+1) \mod 8}$. See Figure 5.6.

**Lemma 5.5.7** For any point $p$ and any non-spokesman node $u$, there is always a spokesman node $v$ that is farther from $p$ than $u$. 

Figure 5.5: Proximity by sector of spokesmen: $A$ and $B$ are in the halfplanes containing all points of the sector not in the triangle $ACB$.

Figure 5.6: Overall proximity of spokesmen: for all points outside of the gray region, there is always a spokesman that is closer than $u$. For all points, there is always a spokesman that is farther than $u$. 

Figure 5.7: Proximity by sector of spokesmen: $A$ and $B$ are in the halfplanes containing all points of the sector not in the triangle $ACB$. 

Figure 5.8: Overall proximity of spokesmen: for all points outside of the gray region, there is always a spokesman that is closer than $u$. For all points, there is always a spokesman that is farther than $u$. 

**Lemma 5.5.7** For any point $p$ and any non-spokesman node $u$, there is always a spokesman node $v$ that is farther from $p$ than $u$. 

**Proof.** Consider a non-spokesman node \( u \) and the set of all spokesmen in its row, column and diagonals. For \( u \) not to be a spokesman, it must have one spokesman on each side of itself for its row, column and diagonals. Let these spokesmen be labeled sequentially \( u_1, u_2, \ldots, u_8 \) in a clockwise order around the node \( u \). Recall Fact 4. Consider all halfplanes \( H_{u,u_i}, i = 1, 2, \ldots, 8 \). These halfplanes are those points to which \( u \) is closer than \( u_i \), or those from which \( u_i \) is farther than \( u \). Since the distance between nodes is at least \( \alpha \) and because of the geometry of the partition, we have that each angle \( u_i u u_{(i+1) \mod 8} \) is less than \( \pi/2 \). Therefore, the union of these halfplanes covers the entire plane. See Figure 5.6.

We now prove the main lemma of this section.

**Proof of Lemma 5.5.5.** Fix a node \( u \) inside the block \( b \). Fix a node \( v \) inside the set of neighbors of \( u \). If \( u \) is a spokesman, we are done. Otherwise, we must show that there is a spokesman \( w \) that shares a link with \( v \).

If \( u \) is not a spokesman, then from Lemma 5.5.6 and from Lemma 5.5.7, there is a spokesman \( w \) that is closer to \( v \) than \( u \) and there is a spokesman \( w' \) that is farther. For some \( \delta \), \( w \) is at distance \( \text{dist}(u,v) - \delta < \text{dist}(v,w) < \text{dist}(u,v) \) of \( v \) and \( w' \) is at distance \( \text{dist}(u,v) < \text{dist}(v,w') < \text{dist}(u,v) + \delta \) from \( v \). Since \( u \) shares a link with \( v \), we know that \( s < \text{dist}(u,v) < 1 \). Moreover, for \( \delta < (1 - s)/2 \), either \( s < \text{dist}(u,v) - \delta < 1 \) or \( s < \text{dist}(u,v) + \delta < 1 \). Since the diameter of a block is \((1 - s)/3 < (1 - s)/2\), at least one of \( w \) and \( w' \) shares a link with \( v \).

**5.5.5 Broadcasting Algorithm \( B^2 \)**

In order to complete the broadcasting algorithm, we need a final procedure to transmit the message \( m \) from the source to all other nodes of the network. We now describe Procedure \( T^2 \). Procedure \( T^2 \) is executed in parallel for all regions. Sequentially for all blocks, we have the set of spokesmen transmit the message \( m \) on a turn basis. Spokesmen send the message only once each and the procedure ends when no more node sends the message. More formally, refer to the pseudo code for Procedure \( T^2 \).

**Lemma 5.5.8** Procedure \( T^2 \) broadcasts the message correctly through the network in time \( O(Dg/l) \).
Procedure $T^2$

\[ S_u \leftarrow \text{the label of the } (\text{block, home}) \text{ containing } u \]

In parallel for all regions

repeat

\textbf{for} block = 1..$\mu$ \textbf{do}

update spokesman status

\textbf{if} \( u \) is a spokesman AND \( S_u = \text{block} \) AND \( u \) has not sent the message \textbf{then}

\textbf{for} valid row indices \( i = 1, \ldots \) \textbf{do}

Spokesmen \( E_i, W_i \) transmit the message \( m \) in order

\textbf{for} valid column indices \( i = 1, \ldots \) \textbf{do}

Spokesmen \( N_i, S_i \) transmit the message \( m \) in order

\textbf{for} valid diagonal indices \( i = 1, \ldots \) \textbf{do}

Spokesmen \( NE_i, SE_i, NW_i, SW_i \) transmit the message \( m \) in order

\textbf{else}

Listen to incoming messages until all spokesmen have transmitted

\textbf{until} no node has transmitted in an iteration
Proof. Consider a network $G$ of diameter $D$ built by the adversary under the swamping model. Consider also the network $G'$ with the same nodes and links as $G$, but where nodes may receive messages from multiple neighbors in one round without collision. Let the nodes of $G'$ execute the broadcasting algorithm $\mathcal{F}$: when a node receives a message $m$ the first time, it transmits this message to all its neighbors the next turn. For the network $G'$, the algorithm $\mathcal{F}$ executes in $\Theta(D)$ rounds. We prove the lemma statement by comparing the execution of Procedure $T^2$ on $G$ to the execution of Algorithm $\mathcal{F}$ on $G'$.

Since each region has $\lceil 3/l \rceil^2$ blocks, where $l = \max\{\alpha/\sqrt{2}, (1 - s)/(3\sqrt{2})\}$ and since each block has $O(lg)$ spokesmen, the broadcast algorithm sequentially makes all spokesmen of a region communicate every $O(g/l)$ turn. From Lemma 5.3.3, the process is collision-free. From Lemma 5.5.5, the spokesmen of a block reach all the nodes that can be reached by any node on their block that do know the message $m$. It then follows that the message $m$ being relayed through the network may be slowed down by a factor $O(g/l)$ with respect to the broadcast time of Algorithm $\mathcal{F}$. Hence, for any network $G$ of diameter $D$, the total transmission time is in $O(Dg/l)$.

Algorithm $B^2$

\begin{itemize}
  \item In parallel for all nodes
  \item Call Procedure $D^*$
  \item Call Procedure $T^2$
\end{itemize}

Proof of Theorem 5.5.1. From Lemma 5.5.8, the time of execution of Procedure $T^2$ is $O(Dg/l)$. From Lemma 5.3.7, the time of execution of Procedure $D^*$ is $\Theta(g^4)$. Adding these times together, we get a total time of $O(Dg/l + g^4)$.

5.6 Conclusion

In this chapter, we have shown algorithms for broadcasting under a novel communication model, the swamping communication model. We have shown algorithms of optimal time complexity for the line and the grid. We have also shown algorithms for broadcasting in networks of unknown topology, with nodes placed on the line, and in the plane.
Chapter 6

Conclusion and Future Research

6.1 General Remarks

In this thesis, we have explored the topic of communication in the presence of positively correlated faults in three contexts: networks of known topology, random networks, and networks given by the adversary. We also considered three types of faults. The first type of faults was studied for wired networks: faults originating from inside the nodes, affecting the nodes and their neighbors; this fault type could occur as the result of power outages. The second type of faults was studied for wireless networks: faults were considered to affect nodes, but to originate from events outside of the nodes, modeling situations such as electrostatic damage. Finally, also for wireless networks, we considered faults originating at nodes and causing transient reception disturbances at close nodes through the phenomenon of swamping.

Our starting hypothesis was that positively correlated faults would be harder to tolerate in networks than random independent faults. We first examined the effects of positively correlated random faults in wired networks; damaging events, spots, occur randomly and independently at nodes of a network, causing permanent crash faults in the given node and its neighbors. Indeed, removing the assumption of fault independence has profound repercussions on the reliability of these networks. Our results include thresholds on fault probabilities indicating that the fault-free part of networks cannot be connected with high probability given constant fault probabilities. This validated our initial hypothesis in the case of wired networks. To give more insight into the effects of positively correlated faults, we needed to investigate if this finding also held for wireless networks.

We continued with the study of wireless networks in the presence of Poisson distributed damaging phenomena causing permanent crash faults in the nodes within distance $s$, the spot radius. Our results include an algorithm which successfully
broadcasts in many cases when the network is connected. Moreover, we calculated thresholds on fault probability such that the network would be connected with high probability. It turns out that these thresholds indicate that the fault-free part of networks could not be connected with high probability given constant fault arrival rate. We conjecture that this connectivity with high probability is impossible given constant fault probabilities or constant fault arrival rates in the dependent fault context for nodes.

Given our previous findings on network connectivity, we shifted focus to networks of unknown topology as an abstraction of the connected parts of networks resulting from faults. It is well known ([14]) that broadcasting algorithms valid for networks of unknown topology are valid for all networks, including the connected fault-free part of geometric radio networks. In this context, for geometric radio networks, we studied the problem of communication with positively correlated reception faults at nodes, caused by the swamping phenomenon. In our model, transmissions from nodes cause transient faults in all links of nodes within distance \( s \), the swamping distance. We found broadcasting algorithms which perform in optimal time complexity for lattice networks. We also found broadcasting algorithms successfully broadcasting in all connected geometric radio networks of unknown topology, despite the problem of swamping.

By substituting the parameters in our developments, the results from this thesis may be used by network and system designers to evaluate the reliability of distributed systems and to guide them in the choice of network properties. In the following sections, we detail our conclusions on each part of our research and we give a direction for future research.

6.2 The Diameter and Connectivity of Networks with Random Dependent Faults

In Chapter 3, we provided analytic results on connectivity and diameter of the fault-free part of wired networks in the presence of dependent, positively correlated faults. We introduced the neighborhood fault model where damaging events, called spots, occur randomly and independently at nodes of a network with probability \( p \), and
cause permanent faults of the crash type in the affected node and its neighbors.

Our results show clear differences between the assumption of independent faults and that of the neighborhood fault model. For example, while under independent faults with small constant fault probability \( p > 0 \) the fault-free part of the hypercube remains connected with high probability [12], this is not the case under the neighborhood fault model with any positive constant spot probability. Likewise, the fault-free part of the torus is connected with high probability for fault probability \( p \in 1/\omega(n^{1/4}) \) when faults are independent, but this is not the case for such spot probabilities under the neighborhood fault model.

It remains open whether or not there exists a network, which, under the neighborhood fault model, has the fault-free part connected with high probability despite constant spot probabilities. We conjecture that this is not the case.

The neighborhood fault model is a first step in graph modeling of dependent positively correlated faults in arbitrary networks. It would be interesting to analyze more precise center-satellite based models in which independent spots cause faults in nodes with probability decreasing with the distance of the node from the spot.

6.3 Communication in Random Geometric Radio Networks with Positively Correlated Random Faults

In Chapter 4, we provided analytic results on connectivity and diameter of the fault-free part of GRNs in the presence of dependent, positively correlated faults. We introduce the ranged spot model in which damaging events, called spots, are Poisson distributed on a region, causing permanent faults of the crash type in all nodes located within distance \( s \) of them.

We have shown upper and lower bounds on the spot arrival rate \( \lambda \) such that the random graph \( U(n, r, \lambda, s) \) is non-empty, is connected and can be used for efficient broadcasting. We have presented a broadcasting algorithm which successfully terminates with probability \( 1 - \epsilon \), in time \( O(D + \log 1/\epsilon) \).

We propose a few related topics for future research.

1. Some of the bounds leave a gap to be explored in future research.
2. Another question which was left open and may be addressed in future research is that of network design versus randomness. More precisely, is it possible to place nodes on the plane in order to obtain better performance than with random placement? If so, can we find the optimal design to provide higher reliability and broadcast speed? Finally, how does the performance of these networks compare to the cases when damaging events are placed by an adversary, when it knows the network topology and when it does not?

6.4 Broadcasting in Networks of Unknown Topology in the Presence of Swamping

In Chapter 5, we have shown algorithms for broadcasting under a novel communication model, the swamping communication model. The swamping communication model can be viewed as a GRN on which radio communication is implemented with additional transient reception faults on all nodes at close proximity of a transmitter, i.e., a node cannot receive messages at each round when some node within distance \( s \) from it transmits. We have shown algorithms of optimal time complexity for the lattice line and the two-dimensional lattice. We have also shown algorithms for broadcasting in networks of unknown topology, where nodes are placed on the line, and on the plane. In the latter case, there remains an interesting consideration in our research for which we propose the following question:

**Question 1** Given the properties of the partition and of the spokesmen selection algorithm, is it possible to execute a valid spokesman selection algorithm in each node \( u \) using only the node locations and the history of messages known to \( u \), without any exchange of messages between nodes of a same box?

Furthermore, the effects of the swamping phenomenon are still unknown for random graphs. We propose the following question:

**Question 2** Consider a network built of Poisson distributed nodes with communication range 1 causing, at every transmission, transient message reception errors to all
other nodes within distance $s$. Under which conditions on $s$ is it possible to broadcast a message in the entire network or in a large part of it, w.h.p.? How does the parameter $s$ affect the diameter and the speed of broadcast in such networks?

6.5 Communication in Geometric Radio Networks with Interference Sources

The most common cause of radio communication interference is electromagnetic noise. This noise emanates from noise-causing objects and propagates as any other electromagnetic signal. When the power of the noise received at a radio receiver is strong enough, it becomes difficult to decode pertinent incoming radio signals.

A well known performance parameter of communication equipment is the bit error rate with respect to the noise level [26]. This characteristic changes depending on the material, the encoding and on the modulation/demodulation type. As a consequence, there is a threshold signal-to-noise ratio ($S_{NR}$) below which communication equipments may not establish reliable communication with any reasonable probability. As the noise level reaches a certain threshold, the probability of bits being misinterpreted becomes very high; then communication becomes difficult.

Radio transmission power decreases approximately as $1/d^\alpha$ where $d$ is the distance to the source and $\alpha$ is an environment-dependent constant of value at least 2 [18]. In fact, the Friis transmission equation [18] for calculating the received radio power $P_r$ as a function of distance $d$, transmitted power $P_t$, transmitter/receiver antenna gains $G_t$ and $G_r$, and wavelength $\lambda$ is

$$ P_r = P_t G_t G_r \left( \frac{\lambda}{4\pi d} \right)^\alpha, $$

where $\alpha$ is a constant dependent on the environment; $\alpha = 2$ in free space and typically $3 \leq \alpha \leq 5$ for practical applications. Refer to [23, 24] for references on the related SINR model and to [51] for references on the power assignment problem.

**Question 3** Consider sources of Gaussian noise with power $P$ distributed on the plane by a Poisson Process with arrival rate $\lambda$. Consider further nodes with transmission power $T$ and threshold signal to noise ratio $SNR$ distributed on the plane by a Poisson Process with arrival rate $n$. Under which conditions is it possible to broadcast a message in the entire network or in a large part of it, w.h.p.?
Index

broadcasting, 4
success of, 4

communication
collision, 5
round, 4
synchronous, 4
communication model
radio, 4, 44
swamping, 5, 73

fault
Byzantine, 4
crash, 3, 17, 44
neighborhood fault model, 9, 17
of a link, 3
of a node, 3
permanent, 4, 17, 44
ranged spot model, 10, 44
transient, 4, 73

IC, integrated circuit, 11

message passing model
all-port, 4
1-port, 4

network
almost certain connectivity, 18
$\alpha$, 10, 74
city model, 92
conditional wake up model, 16

$D$, diameter, 3
dist$(u, v)$, geometric distance between $u$ and $v$, 3
g, granularity, 10, 74
highway model, 79
hop$(u, v)$, hop count between $u$ and $v$, 3
lattice line, 75
spontaneous wake up model, 16, 74
2-dimensional lattice, 88

node
fault-free, 3
faulty, 3
functional, 3
receiver, 4
sender, 4

SoC, system on chip, 11

VLSI, very large scale integration, 11

w.h.p., with high probability, 5
Bibliography


