University of Memphis

University of Memphis Digital Commons

Electronic Theses and Dissertations

4-19-2011

Probabilistic Problems in Graph Theory

Karen Rosemarie Johannson

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

Recommended Citation

Johannson, Karen Rosemarie, "Probabilistic Problems in Graph Theory" (2011). *Electronic Theses and Dissertations*. 179.
https://digitalcommons.memphis.edu/etd/179

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.

To the University Council:

The Dissertation Committee for Karen Rosemarie Johannson certifies that this is the final approved version of the following electronic dissertation: "Probabilistic Problems in Graph Theory."

Béla Bollobás, Ph.D.
Major Professor

We have read this dissertation and recommend its acceptance:

Paul Balister, Ph.D.

Anna Kaminska, Ph.D.

Jenő Lehel, Ph.D.

Vladimir Nikiforov, Ph.D.

Accepted for the Graduate Council:

Karen D. Weddle-West, Ph.D.
Vice Provost for Graduate Programs

PROBABILISTIC PROBLEMS IN GRAPH THEORY

by

Karen Rosemarie Johannson

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Mathematical Sciences

The University of Memphis

May, 2011

Acknowledgement

I would like to thank my advisor, Dr. Béla Bollobás, for all of his guidance and assistance, his unfailing dedication to his students and for all of the wonderful opportunities he provided me as a student.

I would also like to express my thanks to my advisory committee, Dr. Paul Balister, Dr. Anna Kaminska, Dr. Jenő Lehel, and Dr. Vladimir Nikiforov for their careful reading of this manuscript and all of their helpful comments. Their suggestions helped me make many improvements both to the proofs and in the presentation.

I am grateful for support received from the Natural Sciences and Engineering Research Council of Canada in the form of a Postgraduate Scholarship.

I would like to give my thanks to Tom Coker for his collaboration and to my fellow graduate students for their help over the years. In particular, I am indebted to Fabricio Benevides for his assistance in the final stretch. My thanks are also due to Ms. Tricia Simmons for her great kindness and for her care of forgetful and disorganized students.

I am thankful for the love and support that I have always received from my family. Most of all, I will be eternally grateful to my sweet David. Without his love and encouragement, I would not have worked as hard or smiled as much.

Abstract

Johannson, Karen Rosemarie. Ph.D. The University of Memphis. May 2011. Probabilistic Problems in Graph Theory. Major Professor: Dr. Béla Bollobás.

In this thesis, I examine two different problems in graph theory using probabilistic techniques. The first is a question on graph colourings. A proper total k-colouring of a graph $G = (V, E)$ is a map $\varphi : V \cup E \to \{1, 2, \ldots, k\}$ such that $\varphi|_V$ is a proper vertex colouring, $\varphi|_E$ is a proper edge colouring, and if $v \in V$ and $vw \in E$ then $\varphi(v) \neq \varphi(vw)$. Such a colouring is called adjacent vertex distinguishing if for every pair of adjacent vertices, u and v, the set $\{\varphi(u)\} \cup \{\varphi(uw) : uw \in E\}$, the 'colour set of u', is distinct from the colour set of v. It is shown that there is an absolute constant C such that the minimal number of colours needed for such a colouring is at most $\Delta(G) + C$.

The second problem is related to a modification of bootstrap percolation on a finite square grid. In an $n \times n$ grid, the 1×1 squares, called sites, can be in one of two states: 'uninfected' or 'infected'. Sites are initially infected independently at random and the state of each vertex is updated simultaneously by the following rule: every uninfected site that shares an edge with at least two infected sites becomes itself infected while each infected site with no infected neighbours becomes uninfected. This process is repeated and the central question is, when is it either likely or unlikely that all sites eventually become infected? Here, both upper and lower bounds are given for the probability that all sites eventually become infected and these bounds are used to determine a critical probability for the event that all sites eventually become infected.

Contents

List of Figures

Chapter 1

Preface

1.1 Introduction

The idea of applying probability to problems in combinatorics, and graph theory in particular, is a well-studied concept. Probabilistic techniques have proven their usefulness in many different ways. The notion of random graphs was introduced by Erdős and Rényi [19] who, in a 1959 paper, gave estimates on the size of the largest component in a graph when edges are chosen independently at random. The study of random graphs has been used extensively to obtain results on 'almost every graph' of a particular edge density. On the other hand, probabilistic techniques can often be used to prove the existence of some object or structure with a particular property by showing that a random selection will, with positive probability possesses the desired property. In 1947, Erdős [17] gave a lower bound on the Ramsey numbers $R(k, k)$ by showing that when n is sufficiently large in terms of k, a randomly 2-coloured K_n will contain a monochromatic K_k with positive probability.

The first stumbling block in many of these types of problems is generally an attempt to determine the probability of some combination of events that are not independent. Many probabilistic tools have been developed to take advantage of the structure of the dependencies between different events. The Lovász Local Lemma, due to Erdős and Lovász [18], is one such result that can be used to show that if certain collections of unlikely events are such that each event does not depend on too many others, then with positive probability, none occurs. A number of other results have been given that show that under certain circumstances, collections of events that are not independent behave nearly as if they were, such as Janson's inequalities [27], Talagrand's inequality [34], Harris' Lemma [22] and the van den-Berg-Kesten inequality [8]. In this thesis, I use a number of these tools to give results on two different types of problems.

The first type of problem examined is the minimum number of colours required for a particular type of graph colouring. A *proper total colouring* of a graph G is a colouring of both the edges and vertices of G so that every pair of adjacent vertices receive different colours, every pair of incident edges receive different colours and each vertex has a colour different from the colours of its incident edges. In Chapter 2, a strengthening of this condition is considered. Given a total colouring of a graph G, the colour set of a vertex u is the set of all colours on edges incident to u and the colour of u itself. A proper total colouring is said to be *adjacent vertex distinguishing* if for every pair of adjacent vertices u and v, the colour set of u is different from the colour set of v. The idea of 'vertex distinguishing colourings' was introduced independently by Aigner, Triesch and Tuza [1]; Burris and Schelp [11]; and Černý, Horňák, and Soták [12] who each looked at edge colourings that distinguish all vertices. This condition was relaxed by Zhang, Liu and Wang [40] who introduced 'adjacent vertex distinguishing edge colourings'.

The study of adjacent vertex distinguishing total colourings was introduced by Zhang, Chen, Li, Yao, Lu and Wang [39] who gave a number of results on the number of colours required for such a colouring of several classes of graphs. In general, the best-known upper bound for the number of colours required for an adjacent vertex distinguishing total colouring of an arbitrary graph G was $2\Delta(G)+1$ when the maximum degree of G is large. It was conjectured by Zhang *et al.* [39] that for any graph G, there is an adjacent vertex-distinguishing total colouring of G using $\Delta(G)+3$ colours. Here, it is shown that there is a constant C such that, for any graph G, there is an adjacent vertex-distinguishing total colouring with $\Delta(G)+C$ colours. This upper bound follows a similar result given by Hatami [23] on adjacent vertex distinguishing edge colourings.

The strategy is to start with any proper total colouring and randomly 'uncolour' edges, independently with some probability p. Depending on the value of p, it can be shown that with positive probability, after the edges are randomly uncoloured, all adjacent vertices are distinguished in the subgraph of still-coloured vertices. If the uncoloured edges are then properly coloured with a set of new colours, the resulting colouring is adjacent vertex-distinguishing. However, it is possible that many new colours would be needed. In order to avoid this, a random set of uncoloured edges is altered so that the uncoloured subgraph has maximum degree bounded above by some absolute constant. The uncoloured subgraph can then be recoloured with only a constant number of new colours. The events that two particular edges were uncoloured and remain so may no longer be independent. A careful analysis, using such tools as Talagrand's inequality and the Lovász Local Lemma, shows that this can be done is such a way that there is a positive probability that after the uncolouring and alteration, every pair of adjacent vertices are distinguished.

The second type of question examined here is the 'average behaviour' over time of a deterministic process that begins with a random subset of the vertices of a graph. Given a graph G, with vertices in one of two states: 'infected' or ‘uninfected’, and $r \in \mathbb{Z}^+$, the bootstrap process for G with parameter r is an update rule for the states of vertices, defined as follows. Infected vertices remain infected forever and every uninfected vertex with at least r infected neighbours becomes itself infected. This process is repeated and a set of initially infected vertices is said to ‘percolate’ if eventually all vertices become infected. The bootstrap process was introduced by Chalupa, Leath and Reich [13] who studied the process on infinite regular trees and related bootstrap percolation to the study of ferromagnetism.

Bootstrap percolation has been well-studied in the case where the graph G is a finite $n \times n$ integer lattice and the parameter for infection is $r = 2$. The vertices are thought of as 1×1 squares, called sites, in the $n \times n$ grid with two sites adjacent if they share an edge. An uninfected site that shares an edge with at least two infected sites becomes infected. The bootstrap process is an example of a cellular automaton, as introduced by von Neumann [32] and suggested by Ulam [36]. When the set of initially infected vertices are chosen independently at random, one would like to know when percolation is either likely or unlikely. Aizenmann and Lebowitz [2] first gave bounds on the critical probability of percolation and, later, a sharp threshold was given by Holroyd [24].

In Chapters 3–5, a modification of the usual bootstrap process is considered where infected sites can potentially recover from their infection, although they remain susceptible to future re-infection. The update rule for the infection status of sites is modified so that each infected site with no infected neighbours becomes uninfected when the modified update rule is applied. As before, uninfected sites with at least two infected neighbours become infected. These two rules are applied simultaneously to every site in the grid. As in the study of the usual bootstrap process, sites in a finite integer lattice are initially infected independently at random. The process is repeated and one asks, again, when it is either likely or unlikely that all sites eventually become infected. With this modified update rule, it is significant that the rule is applied simultaneously to all sites. The process would be quite different if sites were chosen at random, one at a time, to have their state updated.

The difficulty, compared to the analysis of the usual bootstrap percolation, is that the event that a particular site is initially infected and remains infected depends on the infection status of other, nearby, sites. If there is an infected site with no other infected sites within distance 2, then after the modified bootstrap update rule is applied, this site is uninfected and does not affect the infection of any of its neighbours.

Instead of considering the infection status of single sites, the infection status of certain pairs of sites is considered. These events are not, in general, independent since two pairs may depend on overlapping sets of sites. It is shown, though, that this modified bootstrap process can be reasonably compared to another model, where pairs of sites are infected independently. Using this approximation and modifications of the tools used by Holroyd [24] and others to analyze usual bootstrap percolation, both upper and lower bounds are given on the probability for percolation in the modified bootstrap process. These bounds on the likelihood of percolation are used to determine an asymptotic value for the critical probability of percolation. A slightly different analysis is used in the two cases where percolation is shown to be likely and when it is shown to be unlikely.

In Section 1.2, a few of the probabilistic tools that are used repeatedly are given, as well as some notation.

1.2 Probability tools and notation

A few results are used repeatedly throughout and those are stated here for reference. In a probability space, the probability of an event E is denoted $\mathbb{P}(E)$, or $\mathbb{P}_p(E)$ if the probability measure depends on a parameter p. The expected value of a random variable X is denoted $\mathbb{E}(X)$.

The following lemma, a Chernoff-type bound, gives estimates for the unlikelihood of a binomial random variable being either much larger or much smaller than its mean. This result can be found, in the following form, in Alon and Spencer [3, pp 267–268], for example.

Lemma 1.2.1. Let X be a binomial random variable with parameters $n \in \mathbb{Z}^+$ and $p \in (0,1)$. For $pn < m < n$,

$$\mathbb{P}(X \geq m) \leq e^{m-np} \left(\frac{np}{m}\right)^m$$

and for $0 < m < pn$,

$$\mathbb{P}(X < m) \leq e^{-(m-np)^2/2pn}.$$

The following two results are useful for examining the probabilities of 'increasing events' in a cube $\{0,1\}^n$. For every $n \in \mathbb{Z}^+$, let Q^n be the cube $\{0,1\}^n$ and for $\mathbf{p} = (p_1, \ldots, p_n) \in (0,1)^n$, let $Q^n_{\mathbf{p}}$ denote the cube with probability measure given by product measure with respect to $\mathbf{p}$. That is, if $\mathbf{x} = (x_1, \ldots, x_n) \in Q^n_{\mathbf{p}}$ is taken at random, then for every $i = 1, \ldots, n$, $\mathbb{P}(x_i = 1) = p_i$ and for each $i \neq j$, the events $\{x_i = 1\}$ and $\{x_j = 1\}$ are independent.

The cube Q^n is endowed with a coordinate-wise partial ordering as follows. For each $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in Q^n$ write $(a_1, \ldots, a_n) \leq (b_1, \ldots, b_n)$ iff for every $i = 1, \ldots, n$, $a_i \leq b_i$. An event $E \subseteq Q^n$ is called *increasing*, or an 'up-set', if for every $\mathbf{e} \in E$ and $\mathbf{x} \in Q^n$ with $\mathbf{e} \leq \mathbf{x}$, then $\mathbf{x} \in E$. Similarly, an event E is called *decreasing* or a 'down-set' if for every $\mathbf{e} \in E$ and $\mathbf{x} \in Q^n$ with $\mathbf{x} \leq \mathbf{e}$, then $\mathbf{x} \in E$.

The following two results on increasing and decreasing events can be found, for example, in Bollobás and Riordan [10, pp. 39–44]. The first lemma is due to Harris [22].

Lemma 1.2.2. Let A and B be subsets of Q_p^n. If A and B are both increasing events or if both A and B are decreasing events, then

$$\mathbb{P}(A \cap B) \geq \mathbb{P}(A)\,\mathbb{P}(B).$$

For sets $A, B \subseteq Q^n$, define an event $A \,\square\, B \subseteq A \cap B$ as follows. For each $\mathbf{x} \in Q^n$, then $\mathbf{x} \in A \,\square\, B$ iff there exist $I = I(\mathbf{x}), J = J(\mathbf{x}) \subseteq \{1, 2, \ldots, n\}$ with $I \cap J = \emptyset$ and such that for any $\mathbf{y} \in Q^n$, if for every $i \in I$, $y_i = x_i$ then $\mathbf{y} \in A$ and for any $\mathbf{z} \in Q^n$, if for every $j \in J$, $z_j = x_j$, then $\mathbf{z} \in B$. The set $A \,\square\, B$ is the set of elements in $A \cap B$ whose membership in A and B can be certified by two disjoint sets of indices and is called the event that A and B *occur disjointly.*

The following lemma on the probability of two events occurring disjointly is called the van den Berg-Kesten inequality [8].

Lemma 1.2.3. Let A and B be subsets of Q_p^n. If A and B are both up-sets or if both are down-sets, then

$$\mathbb{P}(A \,\square\, B) \leq \mathbb{P}(A)\mathbb{P}(B).$$

Something much stronger than Lemma 1.2.3 is in fact true, although the stronger result is not used here. In 2000, Reimer [33] showed that if A and B are any events in Q_p^n, then $\mathbb{P}(A \,\square\, B) \leq \mathbb{P}(A)\mathbb{P}(B)$.

Talagrand's inequality [34] is a useful tool for the analysis of certain random variables that depend on a finite number of Bernoulli random variables. The original formulation of Talagrand's inequality gives concentration results for such random variables about their median value. The following form of Talagrand's inequality, due to McDiarmid and Reed [30], is a variation giving bounds on the probability that a random variable of this type is far from its mean. For further details on the Talagrand inequality see, for example, Talagrand [35].

Theorem 1.2.4. Fix $c > 0$, $r \geq 0$ and $d \geq 0$. Suppose that g is a non-negative random variable with mean μ and $g = g(X_1, \ldots, X_n)$ where $X_1, \ldots, X_n$ are independent Bernoulli 0-1 random variables satisfying

(a) if $\mathbf{x}, \mathbf{x}' \in \{0,1\}^n$ differ in exactly one coordinate, then $|g(\mathbf{x}) - g(\mathbf{x}')| \leq c$ and

(b) for any $s \geq 0$, if $g(\mathbf{y}) \geq s$, there is a set $I \subseteq [1, n]$ with $|I| \leq rs + d$ such that if $\mathbf{y}' \in \{0,1\}^n$ agrees with $\mathbf{y}$ on the coordinates in I, then $g(\mathbf{y}') \geq s$.

For any $t \geq 0$,

$$\mathbb{P}(g - \mu \geq t) \leq \exp\left(-\frac{t^2}{2c^2(r\mu + d + rt)}\right),$$
$$\mathbb{P}(g - \mu \leq -t) \leq \exp\left(-\frac{t^2}{2c^2(r\mu + d + t/3c)}\right).$$

Very often, in the applications that follow, there are large collections of events that are not mutually independent, but for which each event does not depend on too many others in the collection. There are many results that measure how far the probabilities of combinations of these events are from what they would be if the events were mutually independent. The two that are used here are one of the Janson inequalities (Theorem 1.2.5 below) and the Lovász Local Lemma (Theorem 1.2.6 below) which give upper and lower bounds, respectively, on the probability that none of the events in such a collection occur.

The following is an inequality due to Janson [27] and can be found in this form, for example, in [28, p.33].

Theorem 1.2.5. Fix $n \in \mathbb{Z}^+$, let $\mathbf{p} \in (0,1)^n$, and let $\{A_i : i \in I\}$ be a finite collection of sets so that for each $i \in I$, $A_i \subseteq \{1, 2, \ldots, n\}$. Choose $\mathbf{x} \in Q^n_{\mathbf{p}}$ at random, according to the product measure given by $\mathbf{p}$ and for each $i \in I$, let $B_i \subseteq Q^n_{\mathbf{p}}$ be the event that for every $j \in A_i$, $x_j = 1$. Set

$$\Delta = \sum_{\substack{B_i, B_j \\ A_i \cap A_j \neq \emptyset}} \mathbb{P}_{\mathbf{p}}(B_i \wedge B_j)$$

and let $\mu = \sum_{i \in I} \mathbb{P}_{\mathbf{p}}(B_i)$. Then

$$\mathbb{P}_{\mathbf{p}}(\wedge_{i \in I} \overline{B}_i) \leq e^{-\mu + \Delta}.$$

The next theorem, due to Erdős and Lovász [18], is known as the Lovász Local Lemma. For the form below, see for example [9, pp 21–22].

Theorem 1.2.6. Let $A_1, \ldots, A_n$ be events in a probability space and for each $i \in \{1, \ldots, n\}$, let $\Gamma(i) \subseteq \{1, \ldots, n\}$ be such that A_i is mutually independent of the events $\{A_j : \{1, \ldots, n\} \setminus (\Gamma(i) \cup \{i\})\}$. If there are $x_1, \ldots, x_n \in [0, 1)$ such that for all $i \in [1, n]$,

$$\mathbb{P}(A_i) < x_i \prod_{j \in \Gamma(i)} (1 - x_j)$$

then

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \bar{A}_i\right) \geq \prod_{i=1}^{n} (1 - x_i)$$

and in particular, $\mathbb{P}(\cap_{i=1}^n \bar{A}_i) > 0$.

As a special case, let $p \in (0, 1)$ and $d \in \mathbb{Z}^+$ be such that $ep(d + 1) \leq 1$ and suppose that $A_1, A_2, \ldots, A_n$ are events such that for every $i \in \{1, 2, \ldots, n\}$,

$\mathbb{P}(A_i) \le p$ and $|\Gamma(i)| \le d$. Taking $x_1 = x_2 = \ldots, x_n = \frac{1}{d+1}$, since

$$p \le \frac{1}{e(d+1)} \le \frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^d < x_i \prod_{j \in \Gamma(i)} (1 - x_j),$$

then $\mathbb{P}(\cap_{i=1}^n \bar{A}_i) \ge \prod_{i=1}^n (1 - x_i) > 0$.

Throughout, some standard notation is used and recorded here for reference. For any $n \in \mathbb{Z}^+$ and $p \in (0,1)$ and X a binomial random variable with parameters n and p, write $X \sim \text{Bin}(n,p)$. Similarly, in a slight abuse of notation, for any set A and $p \in (0,1)$ denote by $X \sim \text{Bin}(A,p)$ a random subset $X \subseteq A$ where each element of A is included in X independently with probability p.

For any $a, b \in \mathbb{Z}$ with $a < b$, set $[a,b] = \{a \le n \le b : \ n \in \mathbb{Z}\}$ and for any $a \ge 1$, define $[a] = [1,a]$, unless otherwise stated. The symbol log is used to denote the natural logarithm.

Some standard graph theory notation is used in the following chapters. For any $n \ge 3$, C_n is used to denote a cycle on n vertices and K_n for a complete graph on n vertices. In any graph $G = (V,E)$ and $v \in V$, the neighbourhood of v is denoted by $N(v) = \{u \in V : \ uv \in E\}$.

Chapter 2

Vertex-distinguishing total colourings

2.1 Introduction

A well-studied concept is that of the total chromatic number. A *proper total colouring* of a graph is a colouring of both vertices and edges so that every pair of adjacent vertices receive different colours, every pair of adjacent edges receive different colours and every vertex receives a colour different from the colour of each of its incident edges. In this chapter, proper total colourings are considered that have the additional property that for any adjacent vertices u and v, the set of colours incident to u is different from the set of colours incident to v. It is shown that there is a constant C so that for any graph G, there exists such a colouring using at most $\Delta(G) + C$ colours.

This type of question is a natural extension of the study of 'vertex-distinguishing edge colourings'. Before proceeding with the details regarding total colourings, some background is given on the related results on edge colourings.

2.1.1 Vertex distinguishing edge colourings

The study of proper colourings that induce different colour sets on different vertices was introduced independently by Aigner, Triesch and Tuza [1]; Burris and Schelp [11]; and Černý, Horňák and Soták [12]. Each of these teams examined the number of colours needed to properly edge colour a graph so that every vertex has a colour set different from that of every other vertex.

Zhang, Liu, and Wang [40] relaxed this condition, examining proper edge colourings that distinguish pairs of adjacent vertices.

Definition 2.1.1. Given a graph $G = (V, E)$, the *adjacent vertex distinguishing edge chromatic number*, denoted $\chi'_a(G)$, is the least k such that there exists φ, a proper edge k-colouring of G, with the property that if $u, v \in V$ with $uv \in E$, then $\{\varphi(uw) : w \in N(u)\} \neq \{\varphi(vz) : z \in N(v)\}$.

In their paper, Zhang *et al.* determine the exact value of $\chi'_a(G)$ for several classes of graphs and conjecture that if G is a connected graph with $V(G) \geq 6$, then $\chi'_a(G) \leq \Delta(G) + 2$.

Balister, Győri, Lehel and Schelp [4] showed that if G is a graph with $\Delta(G) = 3$ then $\chi'_a(G) \leq 5$. They also showed that if G is a bipartite graph then $\chi'_a(G) \leq \Delta + 2$ and for G any graph, $\chi'_a(G) \leq \Delta(G) + O(\log_2 \chi(G))$. The upper bound on $\chi'_a(G)$ for arbitrary graphs was sharpened by Hatami [23] who, using probabilistic techniques, showed that if G is a graph with $\Delta(G) \geq 10^{20}$, then

$$\chi'_a(G) \leq \Delta(G) + 300. \tag{2.1}$$

2.1.2 Total colourings

Let $G = (V, E)$ be a simple graph with no loops or multiple edges. For $k \in \mathbb{Z}^+$, a map $\varphi : V \cup E \to \{1, 2, \ldots, k\} = [k]$ is called a *proper total k-colouring of G* iff

- for every $u, v \in V$, if $uv \in E$, then $\varphi(u) \neq \varphi(v)$ and $\varphi(u) \neq \varphi(uv)$, and
- for every pair $uv, uw \in E$ of adjacent edges, $\varphi(uv) \neq \varphi(uw)$.

In other words, $\varphi|_V$ is a proper vertex colouring, $\varphi|_E$ is a proper edge colouring and every vertex receives a colour different from the colour of each of its incident edges.

The *total chromatic number of* G, denoted $\chi''(G)$, is the least k for which there exists a proper total k-colouring of G.

The maximum degree of a graph G is denoted, as usual, by $\Delta(G)$. Under any proper total colouring a vertex of maximum degree in G receives a colour different from that of any of its edges and thus $\chi''(G) \geq \Delta(G) + 1$.

Definition 2.1.2. Let $G = (V, E)$ be a graph and φ be a proper total colouring of G. For each $v \in V$ the *colour set of* v *(with respect to* φ*)* is the set

$$C_\varphi(v) = \{\varphi(v)\} \cup \{\varphi(vw) : w \in N(v)\}.$$

A vertex $v \in V$ is said to be *distinguished from* u *by* φ iff $C_\varphi(u) \neq C_\varphi(v)$ and φ is said to be *adjacent vertex distinguishing* iff every pair of adjacent vertices in G are distinguished from each other by φ.

The least k for which G has an adjacent vertex distinguishing total k-colouring is called the *adjacent vertex distinguishing total chromatic number*, denoted $\chi_{at}(G)$.

The study of adjacent vertex distinguishing total colourings was first introduced by Zhang, Chen, Li, Yao, Lu and Wang [39] who gave the following precise values of χ_{at} several classes of graphs. They showed that for the n-cycle

C_n on at least 4 vertices, $\chi_{at}(C_n) = 4$, and that for complete graphs

$$\chi_{at}(K_n) = \begin{cases} n+1 & \text{if } n \text{ is even, and} \\ n+2 & \text{if } n \text{ is odd.} \end{cases} \tag{2.2}$$

In the same paper, Zhang *et al.* [39] made the following conjecture.

Conjecture 2.1.3. For every graph G,

$$\chi_{at}(G) \leq \Delta(G) + 3.$$

There are graphs that attain the upper bound in Conjecture 2.1.3. For example, when n is odd, $\chi_{at}(K_n) = n + 2 = \Delta(K_n) + 3$, as above. Note that the adjacent vertex-distinguishing total chromatic number is not, in general, monotone with respect to subgraphs. In Figure 2.1 there is an example of a graph G containing K_3 as a subgraph but with $\chi_{at}(G) = 4 < 5 = \chi_{at}(K_3)$.

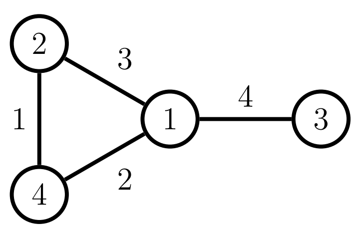

Figure 2.1: Example of a subgraph with $\chi_{at}(H) > \chi_{at}(G)$

Since an adjacent vertex distinguishing total colouring is also a proper total colouring, for any graph G, $\chi''(G) \leq \chi_{at}(G)$. While it has been conjectured, independently by both Behzad [7] and Vizing [37], that $\chi''(G) \leq \Delta(G) + 2$, currently, the best-known upper bound for graphs with sufficiently large maximum degree was given by Molloy and Reed [31] who showed that there exists a constant Δ_0 such that if G is any graph with $\Delta(G) \geq \Delta_0$, then

$$\chi''(G) \leq \Delta(G) + 10^{26}. \tag{2.3}$$

While a proof of Conjecture 2.1.3 would require a significant improvement on the known upper bound for the total chromatic number of an arbitrary graph, in the case $\Delta(G) = 3$, the conjecture has been verified, independently by Wang [38], Chen [14] and Hulgan [25]. Hulgan, in fact, showed the following stronger result.

Theorem 2.1.4. For every graph $G = (V, E)$ with $\Delta(G) = 3$ and $G \neq K_4$, there is a proper total 6-colouring, φ, of G such that $|\varphi[V] \cap \varphi[E]| \leq 1$.

This theorem implies that if G is a graph with $\Delta(G) = 3$ then $\chi_{at}(G) \leq 6$. Indeed, suppose φ is a proper total colouring of $G = (V, E)$ with $|\varphi[V] \cap \varphi[E]| \leq 1$ and u and v are adjacent vertices. Since φ is a proper colouring, $\varphi(u) \neq \varphi(v)$ and at most one of the colours $\varphi(u)$ or $\varphi(v)$ can appear as an edges colour and hence $C_\varphi(u) \neq C_\varphi(v)$. Thus, φ is an adjacent vertex-distinguishing total colouring.

Hulgan [26] further conjectured that if G is a graph with maximum degree 3, then $\chi_{at}(G) \leq 5$.

For graphs of larger maximum degree, Liu, An, and Gao [29] showed that if G is a graph with $\Delta(G) = \Delta$ sufficiently large and $\delta(G) \geq 32\sqrt{\Delta \ln \Delta}$, then $\chi_{at}(G) \leq \Delta + 10^{26} + 2\sqrt{\Delta \ln \Delta}$.

2.1.3 Results

Following an argument similar to that used by Hatami [23] to prove the upper bound given in equation (2.1), in joint work with T. Coker [15], the following upper bound for $\chi_{at}(G)$ was found.

Theorem 2.1.5. There exists $C_0 > 0$ such that for every graph G,

$$\chi_{at}(G) \leq \chi''(G) + C_0.$$

Applying Molloy and Reed's [31] upper bound on $\chi''(G)$, yields an upper bound on $\chi_{at}(G)$ in terms of $\Delta(G)$.

Theorem 2.1.6. There exists $C' > 0$ such that for every graph G,

$$\chi_{at}(G) \leq \Delta(G) + C'.$$

The idea of the proof of Theorem 2.1.5 is to begin with a proper total colouring and recolour of some of the vertices and edges so that the resulting colouring remains a proper total colouring and becomes adjacent vertex distinguishing, but in such a way that the number of new colour added is bounded by an absolute constant. While the process of recolouring vertices is deterministic, the edges to be recoloured are chosen at random and probabilistic techniques are used to show that there is a 'good' choice of edges for recolouring in a way to obtain an adjacent vertex distinguishing total colouring.

Different techniques are used to deal with vertices of relatively small degree and those with high degree that are not distinguished from some neighbour by the initial proper total colouring. In Section 2.2 it is shown that any proper total colouring can be redefined on the vertices to obtain a proper total colouring that distinguishes vertices of degree at most $\Delta(G)/2$ from each of their neighbours. Probabilistic techniques come into play in Section 2.3, where it is shown that given any proper total colouring, there is a subset of the edges that can be recoloured with no more than a constant number of new colours so that vertices of degree more than $\Delta(G)/2$ are distinguished from their neighbours. Finally, in Section 2.4, these previous two results are combined to prove Theorem 2.1.5.

The proof that the edges of a graph G can be recoloured appropriately requires an assumption that the maximum degree of G is at least as large as a fixed constant. However, once Theorem 2.1.5 is proved for graphs with sufficiently large maximum degree, it immediately holds for all graphs, potentially with a larger constant C_0.

Since different techniques are applied to the subgraph induced by the vertices of 'low degree' and to that induced by the vertices of 'high degree', it will be convenient to use the following notation.

Definition 2.1.7. For any graph $G = (V, E)$, set

$$V_\ell = \{v \in V : \deg(v) \leq \Delta(G)/2\} \text{ and}$$
$$V_h = \{v \in V : \deg(v) > \Delta(G)/2\}.$$

Vertices in the set V_ℓ are said to be *vertices of low degree*, while vertices in the set V_h are said to be *vertices of high degree*.

In this chapter, the following notation for graphs is used. For any graph $G = (V, E)$ and for sets of vertices $A, B \subseteq V$, let the set of edges between A and B be denoted $E(A, B) = \{uv \in E : u \in A \text{ and } v \in B\}$. This notation is used even if the sets A and B are not disjoint.

Given $F \subseteq E$, let $G[F]$ be the subgraph of G induced by the edges in F. For $v \in V$, let the degree of v in F be $\deg_F(v) = |\{f \in F : v \in f\}|$ and denote by $F(v) = \{vw \in F : w \in N(v)\}$, the edges of F that are incident to v.

If φ is a total colouring of G and $D \subseteq V \cup E$, then let the set of colours appearing in D be $\varphi[D] = \{\varphi(v) : v \in D \cap V\} \cup \{\varphi(uv) : uv \in E \cap D\}$.

2.2 Vertices of low degree

The vertices of low degree in a graph G, as in Definition 2.1.7, are those with degree at most $\Delta(G)/2$. Since a vertex of low degree in G has relatively few neighbours compared to $\Delta(G)$, any total colouring of G with more than $\Delta(G)$ colours can be adjusted by recolouring some vertices so that every vertex of V_ℓ is distinguished from all of its neighbours. Recall that since $\chi''(G) \geq \Delta(G) + 1$, if φ is a proper total k-colouring of G, then $k \geq \Delta(G) + 1$.

Proposition 2.2.1. Let $G = (V, E)$ be a graph and let φ be a proper total colouring of G. There exists a proper total colouring φ' of G with $\varphi'|_{E \cup V_h} = \varphi|_{E \cup V_h}$ such that for every $v \in V_\ell$, the colouring φ' distinguishes v from each of its neighbours.

Proof. Fix a graph G and φ, a proper total k-colouring of G. Let ψ_0 be a proper total k-colouring of G with the property that among the proper total k-colourings of G that agree with φ on $E \cup V_h$, the map ψ_0 has the fewest vertices in V_ℓ not distinguished from one of its neighbours. More precisely, among the total colourings

$$\{\psi : \psi \text{ is a proper total } k\text{-colouring of } G \text{ with } \psi|_{E \cup V_h} = \varphi|_{E \cup V_h}\}$$

ψ_0 is such that the quantity

$$|\{u \in V_\ell : \exists\ v \in N(u) \cap V_\ell \text{ with } C_{\psi_0}(u) = C_{\psi_0}(v)\}|$$

is minimised. It will be shown that, in fact, every vertex in V_ℓ is distinguished from each of its neighbours with respect to ψ_0. It suffices to show that every vertex in V_ℓ is distinguished from each of its neighbours in V_ℓ. Note that every vertex $v \in V_\ell$ is distinguished from every $u \in N(v) \cap V_h$ since $|C_{\psi_0}(v)| < |C_{\psi_0}(u)|$.

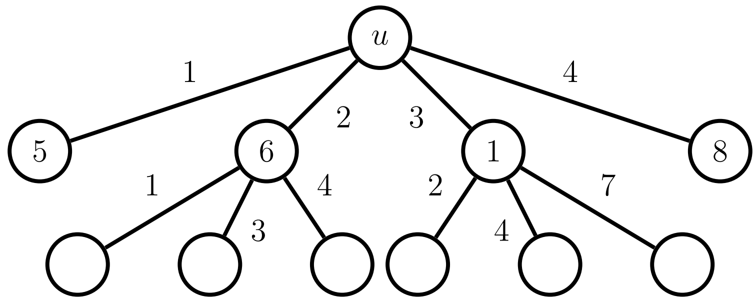

Figure 2.2: Example of colours unavailable for vertex u

Suppose that there is a vertex $u \in V_\ell$ not distinguished by ψ_0 from one of its neighbours. The vertex u will be recoloured so that the resulting total colouring is both proper and distinguishes u from all of its neighbours.

Consider which colours in $[1, k]$ should not be used to recolour u if the resulting colouring is to be proper and distinguish u from its neighbours. These colours are $\psi_0[N(u)] \cup \psi_0[\{uv : \ v \in N(u)\}]$ and also all $i \in [1, k]$ such that there is a vertex $v \in N(u)$ with $C_{\psi_0}(v) = (C_{\psi_0}(u) \setminus \{\psi_0(u)\}) \cup \{u\}$. If $i \in [k]$ is such that there is $v \in N(u)$ with $C_{\psi_0}(v) = \{i\} \cup C_{\psi_0}(u) \setminus \{\psi_0(u)\}$ and $i \neq \psi_0(v)$ then $\psi_0(v) \in \{\psi_0(uw) : w \in N(u)\}$. So, for every $v \in N(u)$, there is at most one colour $i_v \in [k] \setminus \{\psi_0(uw) : w \in N(u)\}$ such that either $i_v = \psi_0(v)$ or $C_{\psi_0}(v) = \{i_v\} \cup C_{\psi_0}(u) \setminus \{\psi_0(u)\}$.

Thus

$$
\begin{aligned}
&|(\cup_{w \in N(u)} \{\psi_0(w), \psi_0(uw)\}) \\
&\qquad \cup \{i \in [k] : \exists\ v \in N(u) \text{ with } C_{\psi_0}(v) = \{i\} \cup C_{\psi_0}(u) \setminus \{\psi_0(u)\}\}| \\
&\qquad\quad \leq 2 \deg(u) \leq 2\Delta/2 < k.
\end{aligned}
$$

Therefore, there is at least one colour $i_u \in [k]$ such that if $v \in N(u)$, then $i_u \neq \psi_0(v)$, $i_u \neq \psi_0(uv)$ and $C_{\psi_0}(v) \neq \{i_u\} \cup \{\psi_0(uw) : w \in N(u)\}$.

Define $\psi_1 : V \cup E \to [k]$ for each $x \in V \cup E$ by

$$\psi_1(x) = \begin{cases} i_u & \text{if } x = u, \\ \psi_0(x) & \text{otherwise.} \end{cases}$$

Then, ψ_1 is a proper total colouring of G with k colours, $\psi_1|_{E \cup V_h} = \psi_0|_{E \cup V_h} = \varphi|_{E \cup V_h}$ and ψ_1 has fewer vertices in V_ℓ that are not distinguished from some neighbour than ψ_0 does. This contradicts the choice of ψ_0.

Thus, ψ_0 is a proper total k-colouring of G with $\psi_0|_{E \cup V_h} = \varphi|_{E \cup V_h}$ and for every $u \in V_\ell$ and $v \in N(u)$, $C_{\psi_0}(u) \neq C_{\psi_0}(v)$. □

2.3 Vertices of high degree

In the previous section, Section 2.2, it is shown that vertices can be deterministically recoloured so that all vertices of low degree are distinguished from their neighbours. In order to alter a proper total colouring so that vertices of high degree are distinguished from their neighbours, a random approach is used. Instead of recolouring vertices, edges are chosen at random to be recoloured with a constant number of new colours.

Proposition 2.3.1. There exists $\Delta_1 > 0$ and $C_1 > 0$ such that for every graph G with $\Delta(G) \geq \Delta_1$ and φ, a proper total k-colouring of G, there is a proper total $(k + C_1)$-colouring, φ', of G such that for every $u, v \in V_h$, if $uv \in E$, then $C_{\varphi'}(u) \neq C_{\varphi'}(v)$.

Proposition 2.3.1 is proved in two steps. First it is shown that there is a set of edges in G that can be deleted (or uncoloured) so that, in the resulting subgraph, most vertices are distinguished from their neighbours and so that every vertex has relatively few neighbours that remain undistinguished from some neighbour. The second step consists of randomly deleting a few more edges incident to those vertices that were potentially not distinguished from a neighbour.

Lemma 2.3.2. For every $m, d \in \mathbb{Z}^+$ with $m \geq d + 6$, and $\varepsilon > 0$, there exists $M = M(m, d, \varepsilon) > 0$ and $\Delta_2 = \Delta_2(m, d, \varepsilon) > 0$ such that for every graph G with $\Delta(G) \geq \Delta_2$ and φ, a proper total k-colouring of G, there is a set $E_1 \subseteq E(V_h, V)$ such that for each $v \in V$, $\deg_{E_1}(v) \leq M$, and setting $\varphi_1 = \varphi|_{V \cup E \setminus E_1}$,

(a) for $u, v \in V_h$ with $uv \in E$ and $\deg(u) = \deg(v)$, if $\deg_{E_1}(u) \geq m$, then $|C_{\varphi_1}(u) \triangle C_{\varphi_1}(v)| \geq d$ and

(b) if $v \in V_h$, then $|\{u \in N_G(v) : \deg_{E_1}(u) < m\}| \leq \varepsilon\Delta(G)$.

Proof. Let G be a graph and set $\Delta(G) = \Delta$. Set $\lambda = 4(m + \ln(3/\varepsilon))$ and $M = 2e\lambda$.

Set $p = \lambda/\Delta$ and select $X \subseteq E(V_h, V)$ randomly, with each edge in $E(V_h, V)$ included in X independently with probability p.

Set $E_1 = E_1(X) = X \setminus \{uv \in E : \deg_X(u) > M\}$ so that every vertex is contained in at most M edges from E_1, as in Figure 2.3.

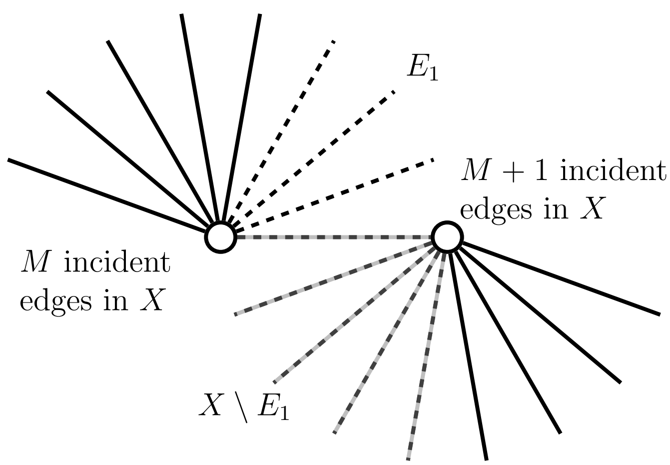

Figure 2.3: Edges contained in X and E_1.

For every $v \in V_h$ and $u \in N(v) \cap V_h$ with $\deg(u) = \deg(v)$, define the following events, depending on the randomly chosen set of edges X:

$$A_{uv} = \left\{\deg_{E_1}(u) \geq m \text{ and } |C_{\varphi_1}(u) \triangle C_{\varphi_1}(v)| < d\right\}$$
$$B_v = \left\{|\{u \in N_G(v) : \deg_{E_1}(u) < m\}| > \Delta\varepsilon\right\}.$$

If E_1 is such that none of these events occur, then E_1 is a set of edges that satisfy conditions (a) and (b) of the statement of the theorem. The Lovász Local Lemma (Theorem 1.2.6) is used to show that

$$\mathbb{P}\left(\bigcap_{uv \in E(V_h, V_h)} \overline{A_{uv}} \cap \bigcap_{v \in V_h} \overline{B_v}\right) > 0.$$

That is, with positive probability, the set E_1 satisfies both conditions (a) and (b). To prove this, estimates on $\mathbb{P}(A_{uv})$ and $\mathbb{P}(B_v)$ are required.

Claim: For each $v \in V_h$ and $u \in N(v) \cap V_h$, if $\deg(v) = \deg(u)$, then $\mathbb{P}(A_{uv}) \leq 2^{M+d} p^{m-d+1}$.

Proof of Claim: Fix $u, v \in V_h$ with $uv \in E$ and $\deg(u) = \deg(v)$. In order to estimate $\mathbb{P}(A_{uv})$, it is convenient to condition on the following event. For any $D \subseteq C_\varphi(u) \setminus \{\varphi(u)\}$ with $m \leq |D| \leq M$, let Z_D be the event that $\varphi[\{uw \in E_1 : w \in N(u)\}] = D$. That is, Z_D is the event that D is the set of colours contributed to $C_\varphi(u)$ by E_1. For each such set D, define $t = t(D) = |D \setminus C_\varphi(v)|$ and $\ell = \ell(D) = |D \cap C_\varphi(v)|$.

Fix such a set D, let $t = t(D)$ and $\ell = \ell(D)$ and suppose that Z_D holds. Set $s = |C_\varphi(u) \setminus C_\varphi(v)|$. Since $\deg(u) = \deg(v)$, then $|C_\varphi(v) \setminus C_\varphi(u)| = s$ also.

Since $s - t \leq |C_{\varphi_1}(u) \setminus C_\varphi(v)|$, if $s - t \geq d$, then

$$\mathbb{P}(|C_{\varphi_1}(u) \triangle C_{\varphi_1}(v)| < d \mid Z_D) = 0.$$

From now on, assume that $s - t < d$ and so $s < t + d \leq t + \ell + d \leq M + d$. Suppose that $E_1(v)$ is such that $|C_{\varphi_1}(u) \triangle C_{\varphi_1}(v)| \leq d - 1 < d$. Note that $C_{\varphi_1}(u) \triangle C_{\varphi_1}(v)$ is the disjoint union of the four sets $C_{\varphi_1}(v) \setminus C_{\varphi}(u)$, $C_{\varphi_1}(u) \setminus C_{\varphi}(v)$, $C_{\varphi_1}(v) \cap D$, and $C_{\varphi_1}(u) \cap \varphi[E_1(v)]$. Note that, by the definition of s, t and ℓ,

$$|C_{\varphi_1}(u) \setminus C_{\varphi}(v)| = s - t,$$
$$|C_{\varphi_1}(v) \cap D| = \ell - |D \cap \varphi[E_1(v)]|, \text{ and}$$
$$|C_{\varphi_1}(v) \setminus C_{\varphi}(u)| = s - |\varphi[E_1(v)] \setminus C_{\varphi}(u)|.$$

Thus,

$$\begin{aligned} d - 1 &\geq |C_{\varphi_1}(u) \triangle C_{\varphi_1}(v)| \\ &= |C_{\varphi_1}(u) \setminus C_{\varphi}(v)| + |C_{\varphi_1}(u) \cap \varphi[E_1(v)]| + |C_{\varphi_1}(v) \setminus C_{\varphi}(u)| + |D \cap C_{\varphi_1}(v)| \\ &\geq (s - t) + (s - |\varphi[E_1(v)] \setminus C_{\varphi}(u)|) + (\ell - |D \cap \varphi[E_1(v)]|). \end{aligned}$$

Therefore,

$$\begin{aligned} |\varphi[E_1(v)] \setminus C_{\varphi}(u)| + |\varphi[E_1(v)] \cap D| &\geq (s - t) + s + \ell - d + 1 \\ &= 2(s - t) + t + \ell - d + 1 \\ &\geq m - d + 1. \end{aligned}$$

That is, considering the colours deleted from the set $C_{\varphi}(v)$ when the edges in E_1 are removed from the graph, at least $m - d + 1$ colours are either contained in the set $C_{\varphi}(v) \setminus C_{\varphi}(u)$ or else from the set $D \cap C_{\varphi}(v)$ since these two sets are disjoint. Recall that, by definition of s and ℓ, $|C_{\varphi}(v) \setminus C_{\varphi}(u)| = s$ and

$|D \cap C_\varphi(v)| = \ell$. Thus, since $s + \ell \leq d + t + \ell \leq M + d$,

$$\begin{aligned}
\mathbb{P}(|C_{\varphi_1}(u) \triangle C_{\varphi_1}(v)| < d \mid Z_D) \\
&\leq \mathbb{P}(|\varphi[E_1(v)] \setminus C_\varphi(u)| + |\varphi[E_1(v)] \cap D| \geq m - d + 1 \mid Z_D) \\
&\leq \mathbb{P}(|\varphi[X(v)] \setminus C_\varphi(u)| + |\varphi[X(v)] \cap D| \geq m - d + 1 \mid Z_D) \\
&\leq 2^s 2^\ell p^{m-d+1} \\
&\leq 2^{M+d} p^{m-d+1}
\end{aligned}$$

uniformly, for all choices of D. Thus, for each $uv \in E(V_h)$,

$$\mathbb{P}(A_{uv}) \leq 2^{M+d} p^{m-d+1}.$$

Claim: There exists a constant c_0 such that if $v \in V_h$ then $\mathbb{P}(B_v) \leq 3e^{-c_0 \Delta}$.

Proof of Claim: Given $v \in V_h$, consider the event B_v that $|\{u \in N(v) : \deg_{E_1}(u) < m\}| > \varepsilon\Delta$. Set

$$\begin{aligned}
V_m &= \{u \in V_h : \ \deg_X(u) < m\}, \\
V_M &= \{u \in V_h : \deg_X(u) > M\}, \text{and} \\
V_N &= \{u \in V_h : \exists\ w \in V_M \cap N(u) \text{ with } uw \in X\}.
\end{aligned}$$

Then, by the definition of E_1,

$$\{u \in V_h : \deg_{E_1}(u) < m\} \subseteq V_m \cup V_M \cup V_N$$

and so by the pigeonhole principle, if $|\{u \in N(v) : \deg_{E_1}(u) < m\}| > \varepsilon\Delta$, then one of the following three events occurs: either $|N(v) \cap V_m| > \varepsilon\Delta/3$, $|N(v) \cap V_M| > \varepsilon\Delta/3$, or $|N(v) \cap V_N| \geq \varepsilon\Delta/3$. The probability of each of these three events is considered separately, although the calculations are similar.

Case 1: Consider the event $|N(v) \cap V_M| > \Delta\varepsilon/3$. Note that for each $w \in V_M$, the quantity $\deg_X(w)$ is a binomial random variable with parameters $\deg_G(w)$ and p.

$$\begin{aligned}
\mathbb{E}(|N(v) \cap V_M|) &= \sum_{w \in N(v)} \mathbb{P}(w \in V_M) \\
&= \sum_{w \in N(v)} \mathbb{P}(\deg_X(w) > M) \\
&\le \sum_{w \in N(v)} \mathbb{P}(\mathrm{Bin}(\Delta, p) > M) \\
&\le \sum_{w \in N(v)} e^{M - p\Delta} \left(\frac{p\Delta}{M}\right)^M && \text{(by Lemma 1.2.1)} \\
&\le \Delta e^{M-\lambda} \left(\frac{\lambda}{M}\right)^M . && \text{(since } \lambda = p\Delta\text{)}
\end{aligned}$$

Changing the status of any one edge in X changes the size of the set V_M by at most 2 and if $|N(v) \cap V_M| \ge a$, this event can be certified by the status of a collection of at most Ma edges. Since $M = 2e\lambda$ and $\lambda \ge \ln(3/\varepsilon)$, it follows that $e^{M-\lambda}(\lambda/M)^M \le e^{-\lambda}(1/2^M) < \varepsilon/3$. Thus, by Theorem 1.2.4,

$$\begin{aligned}
\mathbb{P}(|N(v) \cap V_M| \ge \Delta\varepsilon/3) &\le \exp\left(-\frac{(\Delta\varepsilon/3 - \Delta e^{-\lambda}(e\lambda/M)^M)^2}{2 \cdot 2^2 M \Delta\varepsilon/3}\right) \\
&= \exp\left(-3\Delta \frac{(\varepsilon/3 - e^{-\lambda}(e\lambda/M)^M)^2}{8M\varepsilon}\right).
\end{aligned}$$

Case 2: Now, consider the event $|N(v) \cap V_N| \ge \varepsilon\Delta/3$. Fix $u \in N(v)$. Then,

$$\begin{aligned}
&\mathbb{P}(u \notin (V_M \cup V_m) \text{ and } \exists w \in N(u) \cap V_M \text{ with } uw \in X) \\
&\qquad \le \mathbb{P}(\exists\, w \in N(u) \cap V_M \text{ with } uw \in X) \\
&\qquad \le \sum_{w \in N(u)} \mathbb{P}(w \in N_M \text{ and } uw \in X) \\
&\qquad = \sum_{w \in N(u)} \mathbb{P}(w \in N_M \mid uw \in X)\mathbb{P}(uw \in X)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{w \in N(u)} e^{M-\lambda} \left(\frac{\lambda}{M}\right)^M p \qquad \text{(as above)} \\
&\leq \Delta p e^{M-\lambda} \left(\frac{\lambda}{M}\right)^M \\
&= \lambda e^{-\lambda} \left(\frac{\lambda e}{M}\right)^M .
\end{aligned}
$$

Thus,

$$
\mathbb{E}(|\{u \in N(v) : \exists\ uw \in X \text{ with } w \in V_M\}|) \leq \Delta \lambda e^{-\lambda} \left(\frac{\lambda e}{M}\right)^M .
$$

Changing the status of one edge, say wz, with respect to X to obtain X' changes the status in V_m of at most 2 vertices (w and z) and only if either w or z had exactly M or $M+1$ incident edges in X. For example, suppose $wz \notin X$ and X is changed to $X' = X \cup \{wz\}$. If $\deg_X(w) = M$, then $\deg_{X'}(w) = M+1$ and at most $M+1$ new vertices $u \in N(v)$ are such that $uw \in X'$. Similarly in the case where $\deg_X(w) = M+1$ and X' is obtained from X by removing the edge wz. Thus, changing the status of one edge in X changes the value of $|N(v) \cap V_N|$ by at most $2M+2$, taking into account the two endpoints of the vertices of the affected edge. As before, the event that $|N(v) \cap V_N| \geq a$ can be certified by the status of a collection of at most Ma edges. By the choice of M and λ,
$\varepsilon/3 > \lambda e^{-\lambda} \left(\frac{\lambda e}{M}\right)^M$ and so by Theorem 1.2.4,

$$
\begin{aligned}
\mathbb{P}(|N(v) \cap V_N| > \Delta\varepsilon/3) &\leq \exp\left(-\frac{(\Delta\varepsilon/3 - \Delta\lambda e^{-\lambda}\left(\frac{\lambda e}{M}\right)^M)^2}{2(2M+2)^2 M \Delta\varepsilon/3}\right) \\
&= \exp\left(-3\Delta \frac{(\varepsilon/3 - \lambda e^{-\lambda}(\lambda e/M)^M)^2}{8M(M+1)^2\varepsilon}\right).
\end{aligned}
$$

Case 3: Finally, consider the event $|N(v) \cap V_m| > \Delta\varepsilon/3$.

$$\begin{aligned}
\mathbb{E}(|N(v) \cap V_m|) &= \sum_{w \in N(v)} \mathbb{P}(w \in V_m) \\
&= \sum_{w \in N(v)} \mathbb{P}(\deg_X(w) < m) \\
&\leq \deg(v)\mathbb{P}(\text{Bin}(\Delta/2, p) < m) \\
&\leq \deg(v)e^{-(m-p\Delta/2)^2/(2p\Delta/2)} \qquad \text{(by Lemma 1.2.1)} \\
&= \deg(v)e^{-(m-\lambda/2)^2/\lambda}.
\end{aligned}$$

Talagrand's inequality (Theorem 1.2.4) is applied to the random variable $Y(v) = \deg(v) - |N(v) \cap V_m| = |\{u \in N(v) : \deg_X(u) \geq m\}|$. Note that $\mathbb{E}(Y(v)) = \deg(v) - \mathbb{E}(|N(v) \cap V_m|)$. As in Case 1, changing the status of any edge changes the value of $\deg(v) - |N(v) \cap V_m|$ by at most 2 and the the event $\deg(v) - |N(v) \cap V_m| \geq a$ can be certified by a collection of at most ma edges. By the choice of λ, $(m - \lambda/2)^2/\lambda \geq \ln(3/\varepsilon)$ and therefore, since $e^{-(m-\lambda/2)/\lambda} < \varepsilon/3$, by Theorem 1.2.4,

$$\begin{aligned}
&\mathbb{P}(|N(v) \cap V_m| > \Delta\varepsilon/3) \\
&\quad= \mathbb{P}(\deg(v) - |N(v) \cap V_m| < \deg(v) - \Delta\varepsilon/3) \\
&\quad= \mathbb{P}(\deg(v) - |N(v) \cap V_m| - \mathbb{E}(Y(v)) < \mathbb{E}(|N(v) \cap V_m| - \Delta\varepsilon/3) \\
&\quad\leq \mathbb{P}(\deg(v) - |N(v) \cap V_m| - \mathbb{E}(Y(v)) < \Delta e^{-(m-\lambda/2)^2/\lambda} - \Delta\varepsilon/3) \\
&\quad\leq \exp\left(-\frac{(\Delta\varepsilon/3 - \Delta e^{-(m-\lambda/2)^2/\lambda})^2}{2 \cdot 2^2(m(\Delta/2 - \Delta e^{-(m-\lambda/2)^2/\lambda}) + (\Delta\varepsilon/3 - \Delta e^{-(m-\lambda/2)^2/\lambda})/6)}\right) \\
&\quad= \exp\left(-\Delta\frac{(\varepsilon/3 - e^{-(m-\lambda/2)^2/\lambda})^2}{8(m(\frac{1}{2} - e^{-(m-\lambda/2)^2/\lambda}) + (\varepsilon/18 - e^{-(m-\lambda/2)^2/\lambda}/6))}\right).
\end{aligned}$$

Set

$$c_0 = \min\left\{\frac{3(\varepsilon/3 - e^{-\lambda}(e\lambda/M)^M)^2}{8\varepsilon M}, \frac{3(\varepsilon/3 - \lambda e^{-\lambda}(e\lambda/M)^M)^2}{8\varepsilon M(M+1)^2},\right.$$
$$\left.\frac{(\varepsilon/3 - e^{-(m-\lambda/2)^2/\lambda})^2}{8(m(\frac{1}{2} - e^{-(m-\lambda/2)^2/\lambda}) + (\varepsilon/18 - e^{-(m-\lambda/2)^2/\lambda}/6))}\right\}.$$

Then, $\mathbb{P}(|N(v) \cap A| > \varepsilon\Delta) \leq 3e^{-c_0\Delta}$. Thus, for each $v \in V_h$,

$$\mathbb{P}(B_v) \leq 3e^{-c_0\Delta}.$$

In order to apply the Lovász Local Lemma, it remains to determine how many of the events A_{uv} and B_w are not mutually independent. Let $u, v, w, z \in V_h$. The events A_{uv} and A_{wz} are independent if $d(\{u,v\},\{w,z\}) \geq 4$, the event A_{uv} is independent of B_w if $d(\{u,v\},w) \geq 5$ and B_u is independent of B_w if $d(u,w) \geq 6$. Fix u and v. For every vertex w with $d(\{u,v\},w) \leq 4$, there are at most Δ vertices $z \in N(w)$ and so, for $\Delta \geq 2$, the event A_{uv} is independent of all but at most

$$(1 + \Delta + \Delta(\Delta-1) + \Delta(\Delta-1)^2 + \Delta(\Delta-1)^3 + \Delta(\Delta-1)^4)\Delta \leq \Delta^6$$

other events A_{wz} and all but at most

$$1 + \Delta + \Delta(\Delta-1) + \Delta(\Delta-1)^2 + \Delta(\Delta-1)^3 + \Delta(\Delta-1)^4 \leq \Delta^5$$

events B_w. Meanwhile, the event B_v is independent of all but at most

$$(1 + \Delta + \Delta(\Delta-1) + \Delta(\Delta-1)^2 + \Delta(\Delta-1)^3 + \Delta(\Delta-1)^4)\Delta \leq \Delta^6$$

events A_{wz} and all but at most

$$1+\Delta+\Delta(\Delta-1)+\Delta(\Delta-1)^2+\Delta(\Delta-1)^3+\Delta(\Delta-1)^4 \le \Delta^5$$

events B_w.

Set $\gamma_1 = \ln\Delta/\Delta^7$ and $\gamma_2 = 1/\Delta^5$ and let $uv \in E(V_h)$. Recall that λ and M are constants that depend only on m and ε. Using the inequality $(1-t) \ge e^{-t-t^2}$, which is valid for $t < 0.68$,

$$\begin{aligned}
\gamma_1(1-\gamma_1)^{\Delta^6}(1-\gamma_2)^{\Delta^5} &= \frac{\ln\Delta}{\Delta^7}\left(1-\frac{\ln\Delta}{\Delta^7}\right)^{\Delta^6}\left(1-\frac{1}{\Delta^5}\right)^{\Delta^5} \\
&\ge \frac{\ln\Delta}{\Delta^7}e^{-\ln\Delta/\Delta(1+\ln\Delta/\Delta^7)}e^{-(1+1/\Delta^5)} \\
&\ge \frac{\ln\Delta}{5\Delta^7} && \text{(for } \Delta \ge 4) \\
&\ge \frac{2^{M+d}\lambda^{m-d+1}}{\Delta^{m-d+1}} && \text{(for } \Delta \text{ large, } m-d+1 \ge 7) \\
&\ge \mathbb{P}(A_{u,v}).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\gamma_2(1-\gamma_1)^{\Delta^6}(1-\gamma_2)^{\Delta^5} &= \frac{1}{\Delta^5}e^{-\ln\Delta/\Delta(1+\ln\Delta/\Delta^7)}e^{-(1+1/\Delta^5)} \\
&\ge \frac{1}{\Delta^5}e^{-\ln\Delta(1+\ln\Delta/\Delta)}e^{-1+1/\Delta^5} \\
&\ge \frac{1}{5\Delta^5} && \text{(as above)} \\
&\ge 3e^{-c_0\Delta} && \text{(for } \Delta \text{ large)} \\
&\ge \mathbb{P}(B_v).
\end{aligned}$$

Therefore, since $\gamma_1, \gamma_2 \in (0,1)$, by the Lovász Local Lemma (Theorem 1.2.6), the probability that none of the events

$$\{A_{uv} : uv \in E(V_h, V_h)\} \cup \{B_v : v \in V_h\}$$

occurs is positive and hence there is a set E_1 of edges satisfying conditions (a) and (b). $\square$

Next, by deleting a few more edges from G, the vertices in V_h that might not have been distinguished from one of their neighbours in $G[E \setminus E_1]$ can be made to have colour sets different from their neighbours. For simplicity of presentation, for each $\alpha \in (0, 1)$, define

$$V_\alpha = V_\alpha(G) = \{v \in V(G) : \deg(v) > \alpha\Delta(G)\}.$$

In the proof of Proposition 2.3.1, only $\alpha = 1/2$ is used, in which case $V_\alpha = V_h$. The following lemma is stated in a general form with α arbitrary.

Lemma 2.3.3. For each $M > 0$, $B \geq 2$, and $\alpha, \beta > 0$ with $\alpha > \beta$, there is a $\Delta_3 > 0$ so that the following holds. Let $G = (V, E)$ be a graph with $\Delta(G) \geq \Delta_3$, let $E_1 \subseteq E$ be such that $\Delta(G[E_1]) \leq M$, let φ be proper total colouring of G, and let $L \subseteq V_\alpha$ be such that if $v \in V_\alpha$ then $|N(v) \cap L| \leq \beta\Delta(G)$. There exists a set of edges, $E_2 \subseteq E \setminus E_1$ so that, setting $\varphi_2 = \varphi|_{V \cup E \setminus (E_1 \cup E_2)}$,

(a) if $u \in L$ then $\deg_{E_2}(u) = B$,

(b) if $v \notin L$ and $\deg(v) > \alpha\Delta(G)$, then $|E_2 \cap E(v)| \leq B - 1$, and

(c) if $u, v \in L$ with $uv \in E$, then $C_{\varphi_2}(u) \neq C_{\varphi_2}(v)$.

Proof. Fix $M > 0$, $B \geq 2$, and $\alpha, \beta > 0$ with $\alpha > \beta$. Fix a graph $G = (V, E)$ and $E_1 \subseteq E$ with $\Delta(G[E_1]) \leq M$. Set $\Delta = \Delta(G)$. Let φ be a proper total colouring of G and let $L \subseteq V_\alpha$ be such that if $v \in V_\alpha$, then $|N(v) \cap L| \leq \beta\Delta$.

Note that for each $u \in L$, $|N(u) \setminus L| \geq \alpha\Delta - \beta\Delta \geq B + M$ as long as $\Delta_3 \geq \frac{B+M}{\alpha-\beta}$. Select E_2 at random as follows: for each $u \in L$, select a set of B edges in $E \setminus E_1$ from u to $N(u) \setminus L$ uniformly at random to be $E_2(u)$ and let $E_2 = \cup_{u \in L} E_2(u)$. By construction, condition (a) is satisfied for any such set E_2.

For each $v \in V_\alpha \setminus L$ and $\{u_1, u_2, \ldots u_B\} \subseteq N(v) \cap L$, let $A_{v,\{u_1,u_2,\ldots,u_B\}}$ be the event that all of the edges $vu_1, vu_2, \ldots, vu_B$ belong to E_2. For each $u, v \in L$ with $uv \in E$, let $B_{u,v}$ be the event that $C_{\varphi_2}(u) = C_{\varphi_2}(v)$.

Again using the Local Lemma, Theorem 1.2.6, it is shown that the probability that none of the events $A_{v,\{u_1,u_2,\ldots,u_B\}}$ or $B_{u,v}$ occur is strictly positive and hence there is a choice of E_2 that satisfies the conditions (b) and (c).

Fix $v \in V_\alpha \setminus L$ and fix $\{u_1, \ldots, u_B\} \subseteq N(v) \cap L$. A bound on the probability of the event $A_{v,\{u_1,u_2,\ldots,u_B\}}$ is found as follows. For each $i = 1, 2, \ldots, B$,

$$\begin{aligned}
\mathbb{P}(vu_i \in E_2) &\leq \frac{\binom{\deg(u_i)-|N(u)\cap L|-M-1}{B-1}}{\binom{\deg(u_i)-|N(u)\cap L|-M}{B}} \\
&= \frac{B}{\deg(u_i) - |N(u) \cap L| - M} \\
&\leq \frac{B}{\alpha\Delta - \beta\Delta - M} \\
&= \frac{B}{(\alpha - \beta)\Delta - M}.
\end{aligned}$$

For each $i \neq j$, the events that $vu_i \in E_2$ and that $vu_j \in E_2$ are independent and since $A_{v,\{u_1,u_2,\ldots,u_B\}} = \cap_{i=1}^{B}\{vu_i \in E_2\}$,

$$\mathbb{P}(A_{v,\{u_1,u_2,\ldots,u_B\}}) \leq \left(\frac{B}{(\alpha - \beta)\Delta - M}\right)^B.$$

Now consider an event B_{uv}. Given $u, v \in L$ with $uv \in E$, fix $C_B \subseteq \varphi[E(u, V \setminus (L \cup E_1))]$ with $|C_B| = B$. Conditioning on the event $C_B = \varphi[E_2(u)]$, either

$$\mathbb{P}(C_{\varphi_2}(u) = C_{\varphi_2}(v) \mid C_B = \varphi[E_2(u)]) = 0$$

or there is exactly one set of B colours $C_{v,B}$ with $C_\varphi(u) \setminus C_B = C_\varphi(v) \setminus C_{v,B}$.

Thus,

$$\begin{aligned}\mathbb{P}(C_{\varphi_2}(u) = C_{\varphi_2}(v) \mid C_B = \varphi[E_2(u)]) &\leq \frac{1}{\binom{\deg(v)-|N(v)\cap L|-M}{B}} \\ &\leq \frac{1}{\binom{(\alpha-\beta)\Delta-M}{B}} \\ &\leq \left(\frac{B}{(\alpha-\beta)\Delta-M}\right)^B\end{aligned}$$

uniformly for all choices of C_B and hence

$$\mathbb{P}(B_{u,v}) \leq \left(\frac{B}{(\alpha-\beta)\Delta-M}\right)^B.$$

Two events of the form $A_{v_1,\{u_1,\ldots,u_B\}}$ and $A_{v_2,\{w_1,\ldots,w_B\}}$ are independent whenever $\{u_1,\ldots,u_B\} \cap \{w_1,\ldots,w_B\} = \emptyset$ and events $A_{v_1,\{u_1,\ldots,u_B\}}$ and $B_{u,w}$ are independent if $\{u_1,\ldots,u_B\} \cap \{u,w\} = \emptyset$. Similarly, two events $B_{u,w}$ and $B_{u',w'}$ are independent if $\{u,w\} \cap \{u',w'\} = \emptyset$.

Fix v_1 and $\{u_1,\ldots,u_B\}$ with $v_1 \in \cap_{i=1}^{B} N(u_i)$ and consider the number of choices for vertices v_2 and $\{w_1,\ldots,w_b\}$ so that $v_2 \in \cap_{i=1}^{B} N(w_i)$ and $\{w_1,\ldots,w_B\} \cap \{u_1,\ldots,u_B\} \neq \emptyset$. For each $i = 1,\ldots,B$, if $u_i \in \{w_1,\ldots,w_B\}$, then since $\deg(u_i) \leq \Delta$, there are at most Δ choices for the vertex $v_2 \in N(u_i)$. Given such a vertex v_2, since $\deg_L(v_2) \leq \beta\Delta$, there are at most $\binom{\beta\Delta}{B-1}$ choices for the vertices $\{w_1,\ldots,w_B\} \setminus \{u_i\}$. Thus, $A_{v_1,\{u_1,\ldots,u_B\}}$ is independent of all but at most $B\Delta\binom{\beta\Delta}{B-1}$ events of the type $A_{v_2,\{w_1,\ldots,w_B\}}$.

Consider now the number of choices for u, $w \in L$, with $uw \in E$ and such that $\{u,w\} \cap \{u_1,\ldots,u_B\} \neq \emptyset$. There are B choices for u in the set $\{u_1,\ldots,u_B\}$. Given such a vertex u, there are at most $\beta\Delta$ choices for w. Thus, $A_{v_1,\{u_1,\ldots,u_B\}}$ is independent of all but at most $B\beta\Delta$ events of the type $B_{u,w}$. Similarly, an event $B_{u,w}$ is independent of all but at most $2\Delta\binom{\beta\Delta}{B-1}$ events of the type $A_{v_2,\{w_1,\ldots,w_B\}}$ and all but $2\beta\Delta$ events of the type $B_{u',w'}$.

Therefore, by Theorem 1.2.6, since, for Δ sufficiently large in terms of α, β, and B,

$$\left(\frac{B}{(\alpha-\beta)\Delta - M}\right)^B \left(B\beta\Delta + B\left(\frac{\beta}{B-1}\right)^{B-1}\Delta^{B-1}+1\right)e \leq 1,$$

there is a choice of E_2 that satisfies conditions (b) and (c) in the statement of the lemma. $\square$

Proof of Proposition 2.3.1. Set $\varepsilon = 1/3$, $m = 10$, $d = 4$ and let $\Delta_2 > 0$ and $M > 0$ be given by Lemma 2.3.2. Set $\alpha = 1/2$, $\beta = 1/3$, $B = 2$ and let Δ_3 be given by Lemma 2.3.3. Let $G = (V, E)$ be a graph with $\Delta(G) = \Delta \geq \max\{\Delta_2, \Delta_3\}$ and let φ be a total k-colouring of G.

Let $E_1 \subseteq E$ be given by Lemma 2.3.2 and for $L = \{v \in V_h : \deg_{E_1}(v) < 8\}$ let $E_2 \subseteq E \setminus E_1$ be given by Lemma 2.3.3. As before, let $\varphi_2 = \varphi|_{V\cup E\setminus(E_1\cup E_2)}$. By the choice of E_1 and E_2, $\Delta(G[E_1 \cup E_2]) \leq M + 2$ and so by Vizing's theorem, there is a proper edge colouring, ψ, of $G[E_1 \cup E_2]$ with $M + 3$ colours. Let these $M + 3$ colours be disjoint from the set of colours used by φ. Define a total colouring φ' of G as follows

$$\varphi'(x) = \begin{cases} \varphi(x), & \text{for } x \in V \cup E \setminus (E_1 \cup E_2); \\ \psi(x), & \text{for } x \in E_1 \cup E_2. \end{cases}$$

The map φ' is a proper total $(k + M + 3)$-colouring of G. For each $u, v \in V_h$ with $uv \in E$, if $u \notin L$, then $|C_{\varphi_2}(u)\Delta C_{\varphi_2}(v)| \geq d - (B + B - 1) = 4 - (2 + 1) > 0$ and so $C_{\varphi'}(u) \neq C_{\varphi'}(v)$. If $u, v \in L$ and $uv \in E$, then $C_{\varphi_2}(u) \neq C_{\varphi_2}(v)$ by the choice of E_2 and so $C_{\varphi'}(u) \neq C_{\varphi'}(v)$. $\square$

2.4 Proof of Theorem 2.1.5

Proof of Theorem 2.1.5. Let $G = (V, E)$ be a graph with $\Delta(G) \geq \Delta_1$ and let φ be a proper total colouring of G with $\chi''(G)$ colours. By Proposition 2.3.1, there is a proper total $(\chi''(G) + C_1)$-colouring of G such that for each $u, v \in V_h$, if $uv \in E$, then $C_{\varphi'}(u) \neq C_{\varphi'}(v)$.

By Proposition 2.2.1, there is a proper total colouring φ'' with $\varphi''|_{E\cup V_h} = \varphi'|_{E\cup V_h}$ that distinguishes every vertex in V_ℓ from each of its neighbours. By the choice of φ'', if $v \in V_h$, then $C_{\varphi''}(v) = C_{\varphi'}(v)$ and hence φ'' distinguishes each vertex in V from every one of its neighbours. $\square$

Following through the calculations in the proofs carefully, it can be shown that for $\varepsilon = 1/3$, $m = 10$, $d = 4$ and $B = 2$, one can take $\lambda = 39$ and $M = 81$. While this estimate is likely not optimal, and does not seem apply to many real-world examples, it shows that for a graph G with $\Delta(G) \geq \exp(10^{58})$, then $\chi_{at}(G) \leq \chi''(G) + 84$. Given the extremely strong condition on maximum degree, it would be desirable to improve the lower bound on the maximum degree or to extend this type of result to other related problems on vertex distinguishing colourings.

Chapter 3

Modified bootstrap percolation

3.1 Introduction

In the study of 'bootstrap percolation', vertices of a graph are called sites, and these sites can be in one of two possible states: 'infected' or 'uninfected'. For any graph G and $r \in \mathbb{Z}^+$, the *bootstrap process* on G with parameter r is an update rule for the state of sites defined as follows: infected sites remain infected forever and every uninfected site with at least r infected neighbours becomes itself infected. This process is applied repeatedly and an initial configuration of infected vertices is said to *percolate* if all vertices eventually become infected. Bootstrap percolation was introduced by Chalupa, Leath, and Reich [13] who examined the behaviour of a bootstrap process on infinite regular trees.

In the questions examined here, the initial configuration of infected sites is chosen at random with each site infected, independently with some probability p. One of the central questions is for which values of p is percolation either likely or unlikely. Recall that, for any set A and $p \in (0, 1)$ a random subset $X \subseteq A$ where each element of A is included in X independently with probability p is denoted by $X \sim \mathrm{Bin}(A, p)$.

Much progress has been made on these problems when the graph G is a square grid. In this case, for some $n \in \mathbb{N}$, the vertices of G are the elements of $[1, n] \times [1, n]$ with two vertices $\mathbf{x}, \mathbf{y}$ joined by an edge iff $\|\mathbf{x} - \mathbf{y}\|_1 = 1$. The sites are often thought of, not as points, but as 1×1 squares and two sites are adjacent in the graph exactly when the two squares share an edge.

The bootstrap update rule with parameter $r = 2$ is defined as follows: Given $X \subseteq [n]^2$,

$$\mathcal{B}(X) = X \cup \{\mathbf{x} \in [n]^2 : \ |(N(\mathbf{x}) \cup \{\mathbf{x}\}) \cap X| \geq 2\}.$$

A set $X_0 \subseteq [n]^2$ of initially infected sites is said to *percolate* with respect to $\mathcal{B}$ iff

$$\bigcup_{t \geq 0} \mathcal{B}^{(t)}(X_0) = [n]^2.$$

In general, for any set of sites X_0, define $\langle X_0 \rangle_{\mathcal{B}} = \bigcup_{t \geq 0} \mathcal{B}^{(t)}(X_0)$. The set $\langle X_0 \rangle_{\mathcal{B}}$ is called the *span of* X_0 in $\mathcal{B}$.

If $X_0 \sim \text{Bin}([n]^2, p)$, write $\mathbb{P}_p(X_0 \text{ percolates in } \mathcal{B})$ for the probability that the set X_0 percolates in $\mathcal{B}$. The critical probability function for bootstrap percolation is defined as

$$p_c([n]^2, 2) = \inf\{p : \mathbb{P}_p(X_0 \text{ percolates in } \mathcal{B}) \geq 1/2\}.$$

Bounds on the critical probability function for bootstrap percolation on the grid were first given by Aizenman and Lebowitz [2]. A sharp bound for the critical probability was given by Holroyd [24] who proved that for $n \in \mathbb{N}$,

$$p_c([n]^2, 2) = \frac{\pi^2/18 + o(1)}{\log n}.$$

Even sharper results were given by Gravner and Holroyd [20] who improved the upper bound, and by Gravner, Holroyd and Morris [21] who improved the lower bound for the critical probability. They showed that there are constants $C > 0$ and $c > 0$ so that for each $n \in \mathbb{N}$,

$$\frac{\pi^2}{18 \log n} - \frac{C(\log \log n)^3}{(\log n)^{3/2}} \leq p_c([n]^2, 2) \leq \frac{\pi^2}{18 \log n} - \frac{c}{(\log n)^{3/2}}.$$

A useful feature in the analysis of bootstrap percolation is that once a site becomes infected, it remains infected. Here, a modification of the bootstrap process is considered, where sites can both become infected and also return to being uninfected. The new update rule, defined below (3.1), again depends on the infection of nearby sites.

For $n \in \mathbb{N}$, define the *modified update rule* for $[n]^2$ as follows. For any $X \subseteq [n]^2$,

$$\mathcal{M}(X) = \mathcal{B}(X) \setminus \{\mathbf{x} \in X : \ |(N(\mathbf{x}) \cup \{\mathbf{x}\}) \cap X| = 1\}. \tag{3.1}$$

From a set of initially infected sites, those with at least 2 infected neighbours become infected, but in contrast to usual bootstrap percolation, infected sites with no infected neighbours become uninfected, or 'recover'. This occurs simultaneously for all sites and the process is repeated.

Given any set $X_0 \subseteq [n]^2$, for each $t \geq 0$, define

$$X_{t+1} = \mathcal{M}(X_t).$$

The set X_0 is said to *percolate* with respect to $\mathcal{M}$ if there is a $t_{\mathcal{M}}$ such that $X_{t_{\mathcal{M}}} = [n]^2$. Unless otherwise specified, here, a set will be said to percolate if it percolates in the process $\mathcal{M}$. For any $n \in \mathbb{N}$, and $0 < p < 1$, consider $X_0 \sim \text{Bin}([n]^2, p)$ and set

$$I(n, p) = \mathbb{P}_p(X_0 \text{ percolates}) = \mathbb{P}_p(\exists\ t_{\mathcal{M}} \text{ with } X_{t_{\mathcal{M}}} = [n]^2). \tag{3.2}$$

Define the critical probability function for $\mathcal{M}$ by

$$p_c([n]^2, \mathcal{M}) = \inf\{p : I(n,p) > 1/2\}.$$

Unlike the usual bootstrap update rule $\mathcal{B}$, the sequence $X_0, \mathcal{M}(X_0), \mathcal{M}^{(2)}(X_0), \ldots$ is not, in general, monotone. For example, if the grid is initially infected with a checkerboard pattern, the the sets $(X_t)_{t\geq 0}$ alternate between X_0 and $[n]^2 \setminus X_0$, two checkerboard patterns. For this reason, it does not make sense to talk about the span of a set of infected sites in the modified process $\mathcal{M}$. However, it is sometimes helpful to compare the effect of the process $\mathcal{M}$ to that of $\mathcal{B}$ and even in the context of the modified bootstrap process, the span in $\mathcal{B}$ of a set of sites will occasionally be considered.

In joint work with T. Coker [16], bounds on the critical probability for the modified bootstrap update rule $\mathcal{M}$ were determined, together with estimates on the probability of percolation.

Theorem 3.1.1. There exists a constant $\lambda_{\mathcal{M}} > 0$ such that for every $\varepsilon > 0$ and $\{p(n)\}_{n\in\mathbb{N}} \subseteq (0,1)$,

$$I(n, p(n)) = \begin{cases} 1 - o(1) & \text{if } p(n) > \sqrt{\frac{\lambda_{\mathcal{M}}+\varepsilon}{\log n}} \\ o(1) & \text{if } p(n) < \sqrt{\frac{\lambda_{\mathcal{M}}-\varepsilon}{\log n}} \end{cases}$$

The values for $I(n, p(n))$ for each of the two ranges of values for $p(n)$ in Theorem 3.1.1 give the following immediate formulation for the critical probability for the modified bootstrap update rule.

Corollary 3.1.2. For all $n \in \mathbb{N}$,

$$p_c([n]^2, \mathcal{M}) = \left(\frac{\lambda_{\mathcal{M}} + o(1)}{\log n}\right)^{1/2}.$$

Each of the two cases of the proof of Theorem 3.1.1 are proved separately and more detailed information is given on the probability of percolation above or below the critical probability. In Chapter 4, a lower bound is given on the value $I(n, p)$ for p sufficiently small. This is used to give an upper bound for $p_c([n]^2, \mathcal{M})$.

Similarly, in Chapter 5 an upper bound on $I(n, p)$ is given for certain values of n and p that can be used to give a lower bound on $p_c([n]^2, \mathcal{M})$.

In each case, in order to analyze the process given by the update rule $\mathcal{M}$, the set of initially infected sites is altered to create a set on which the rule $\mathcal{M}$ is nearly monotone and the probability of percolation has not changed too much. This alteration is done in different ways for each case.

If an infected site $\mathbf{x} \in X_0$ has a neighbour in X_0, then when the update rule $\mathcal{M}$ is applied, both $\mathbf{x}$ and its neighbour remain infected in all subsequent sets X_t. However, if $\mathbf{x}$ shares a corner with another infected site, then for every $t \geq 0$, $\mathbf{x} \in X_{2t}$, but as in Figure 3.1, it might be the case that $\mathbf{x} \notin X_{2t+1}$.

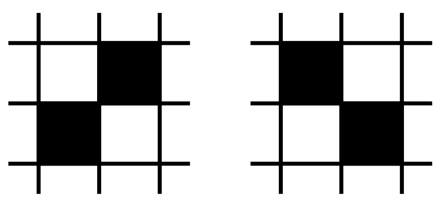

Figure 3.1: Sites whose infection status alternates

In general, the configurations of infected sites that percolate with respect to the usual bootstrap process $\mathcal{B}$ need not percolate with respect to the modified update rule $\mathcal{M}$. There are configurations which percolate with respect to $\mathcal{B}$ but for which all vertices become uninfected in the modified process $\mathcal{M}$.

With this in mind, an alternate initial infection scheme is considered in which sites are infected in pairs so that every infected site either shares an edge or a corner with another infected site. This process is detailed in Section 3.2.

Because of these two different types of pairs, it is often useful to consider both pairs of sites sharing an edge and pairs of sites sharing a corner as neighbours of different types. For $r \geq 0$ and a site $\mathbf{x} \in [n]^2$, define two different balls of radius r in the grid, centered at $\mathbf{x}$,

$$B_r(\mathbf{x}) = \{\mathbf{y} \in [n]^2 : \|\mathbf{x} - \mathbf{y}\|_1 \leq r\}, \tag{3.3}$$

$$B_r^*(\mathbf{x}) = \{\mathbf{y} \in [n]^2 : \|\mathbf{x} - \mathbf{y}\|_\infty \leq r\}. \tag{3.4}$$

For any $\mathbf{x} \in [n]^2$, the set $B_1(\mathbf{x})$ is precisely the set $\{\mathbf{x}\} \cup N(\mathbf{x})$ while the set $B_1^*(\mathbf{x})$ is the set containing $\mathbf{x}$ together with the sites either sharing an edge or corner with $\mathbf{x}$.

Often, it is not just square grids that are of interest, but also 'rectangles' contained in the grid. A set $R \subseteq [1, n]^2$ is called a *rectangle* if there are $a_1 \leq a_2$ and $b_1 \leq b_2$ with $R = [a_1, a_2] \times [b_1, b_2]$. A rectangle R is said to be *internally spanned* by the initially infected sites X_0 if there is a t_R so that $\mathcal{M}^{(t_R)}(X_0 \cap R) = R$. In other words, based only on the initially infected sites inside the rectangle R, every site in R eventually becomes infected. For $p \in (0, 1)$, and $X_0 \sim \text{Bin}(R, p)$, let $I(R, p)$ denote the probability that the rectangle R is internally spanned.

A rectangle $R = [a_1, a_2] \times [b_1, b_2]$ is said to be *horizontally traversable from left to right* by X_0 if $R \setminus (\{a_2\} \times [b_1, b_2]) \cup (\{a_1 - 1\} \times [b_1, b_2])$ is internally spanned by $X_0 \cup \{a_1 - 1\} \times [b_1, b_2]$. That is, if all sites in the column $\{a_1 - 1\} \times [b_1, b_2]$ are infected then the sites in X_0 will cause the infection to spread to all of R except possibly the final column, depending only on the sites that are infected inside the rectangle R. The events that the rectangle R is horizontally traversable from right to left, vertically traversable from bottom to top, or vertically traversable from top to bottom are defined similarly.

The following notation for rectangles is used throughout. For a rectangle $R = [a_1, a_2] \times [b_1, b_2]$, the *dimensions of* R, denoted by $\dim(R)$ is the pair of side-lengths of R: $\dim(R) = (a_2 - a_1 + 1, b_2 - b_1 + 1)$. The length of the shorter side of R is denoted $\text{short}(R) = \min\{a_2 - a_1 + 1, b_2 - b_1 + 1\}$, the length of the longer side of R is denoted $\text{long}(R) = \max\{a_2 - a_1 + 1, b_2 - b_1 + 1\}$ and the semi-perimeter of R is $\phi(R) = (a_2 - a_1 + 1) + (b_2 - b_1 + 1) = \text{short}(R) + \text{long}(R)$.

3.2 Infection with pairs of sites

As described in Section 3.1, the effect of the modified bootstrap process $\mathcal{M}$ on infected sites that have an infected neighbour can be more easily understood than the effect of $\mathcal{M}$ on sites with no infected neighbours. There is still considerable difficulty in dealing with sites that have infected neighbours since the events that two particular sites both have infected neighbours are not, in general, independent.

With this in mind, a new infection scheme is defined where pairs of neighbouring sites are infected simultaneously.

For each $\mathbf{x} \in [n]^2$, consider the four pairs of sites

$$\begin{aligned} T_{(1,1)}(\mathbf{x}) &= \{\mathbf{x}, \mathbf{x} + (1,1)\}, & T_{(1,-1)}(\mathbf{x}) &= \{\mathbf{x}, \mathbf{x} + (1,-1)\}, \\ T_{(1,0)}(\mathbf{x}) &= \{\mathbf{x}, \mathbf{x} + (1,0)\}, \text{ and} & T_{(0,1)}(\mathbf{x}) &= \{\mathbf{x}, \mathbf{x} + (0,1)\}. \end{aligned} \tag{3.5}$$

Call each of these pairs of sites a 2-*tile*. In order to be precise about the position of such pairs, for each 2-tile in (3.5), call $\mathbf{x}$ the *anchor* of the 2-tile. The anchor of a 2-tile is the left-most, bottom-most site. In Figure 3.2, these are the black squares while the non-anchor sites are grey squares.

The 2-tiles of the first three types, $T_{(1,1)}(\mathbf{x})$, $T_{(1,0)}(\mathbf{x})$, and $T_{(1,-1)}(\mathbf{x})$ are said to be of *width 2* while the 2-tiles of the last type, $T_{(0,1)}(\mathbf{x})$ are said to be of *width 1*.

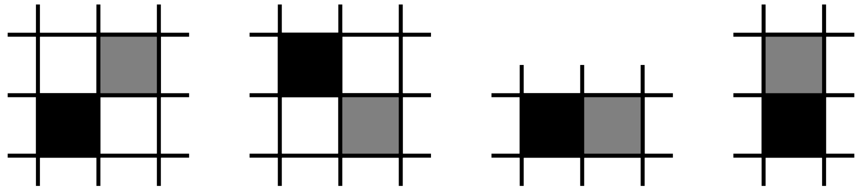

Figure 3.2: Pairs of sites forming 2-tiles

Given $p > 0$, let X_{tiles} be a random configuration of 2-tiles on the grid $[n]^2$ with each of the 2-tiles with anchor in $[n]^2$:

$$\bigcup_{\mathbf{x}\in[n]^2} \{T_{(1,1)}(\mathbf{x}), T_{(1,-1)}(\mathbf{x}), T_{(1,0)}(\mathbf{x}), T_{(0,1)}(\mathbf{x})\}$$

included independently with probability p^2. Note that, in general, X_{tiles} might contain many overlapping 2-tiles. To avoid confusion, the measure on configurations of 2-tiles on the grid is denoted $\mathbb{P}_2$.

Any configuration of 2-tiles is naturally associated with the set of sites in the grid that are contained in some 2-tile. A configuration of 2-tiles, X_{tiles} is said to *percolate* if the set of sites in some 2-tile of X_{tiles} percolates. Similarly a rectangle R will be said to be traversable in any one of the four directions with respect to X_{tiles} if R is traversable in that direction by the set of sites in some 2-tile.

It is shown, in Chapters 4 and 5, that the probability that a random configuration of infected sites percolates (in $\mathcal{M}$) can be approximated by the probability that a random configuration of 2-tiles percolates.

One advantage to working with 2-tiles is that since every infected site has a neighbour, either along an edge or at a corner, if X_{tiles} is a configuration of 2-tiles then any rectangle R is traversable by X_{tiles} under $\mathcal{M}$ exactly when R is traversable by X_{tiles} with respect to the usual bootstrap process, $\mathcal{B}$. Thus, as in usual bootstrap percolation, the only obstacle to crossing R is a pair of adjacent columns containing no infected sites.

Given a configuration of 2-tiles, a column is called 2-*occupied* if it contains the anchor of a 2-tile of width 2, a column is 1-*occupied* if it contains the anchor of a 2-tile of width 1, and a column is *unoccupied* if it does not contain the anchor of any 2-tile. A column is said to be *occupied* if it is either 1-occupied or 2-occupied. Note that a column might be unoccupied and yet contain the non-anchor of some 2-tile. Call a column *empty* if it does not contain any sites from any 2-tiles. A pair of empty adjacent columns is called a *double gap*.

As in the study of usual bootstrap percolation (see for example, Holroyd [24, Lemma 7]), the probability that a rectangle R contains no double gaps is defined recursively in terms of the number of columns in R. The following function appears as the characteristic function of the recurrence relation that arises in the analysis of infection by 2-tiles and a few helpful facts about it are first proved.

Definition 3.2.1. For each $u \in (0, 1)$, set

$$\begin{aligned} F(u, x) = F_u(x) &= x^3 - (1 - u^4)x^2 - u^4(1 - u^4)x - u^8(1 - u^3) \\ &= (x - 1)(x^2 + u^4x + u^8) + u^{11} \end{aligned}$$

and let $\beta(u)$ be the largest real root of $F_u(x)$.

In fact, for any u, the polynomial $F_u(x)$ has exactly one root in $(0, 1)$, and this root is the largest. Since $F_u(0) = -u^8(1 - u^3) < 0$ and $F_u(1) = u^{11} > 0$, there is at least one root in $(0, 1)$. Consider the derivative

$$\frac{d}{dx}F(u, x) = F'_u(x) = 3x^2 - 2(1 - u^4)x - u^4(1 - u^4).$$

As $F'_u(0) = -u^4(1 - u^4) < 0$ and $F'_u(1) = 1 + u^4 + u^8 > 0$, the function F_u has a relative maximum less than zero and a relative minimum between 0 and 1. Since $F_u(x)$ is a polynomial of degree 3 in x, $F_u(x)$ has exactly one root in $(0, 1)$. In order to obtain bounds on the value of $\beta(u)$ in terms of u, note that

$$F_u(1-u^{11}) = -u^{15}(1+u^4-2u^7-u^{11}+u^{18}) < 0,$$

and hence $1-u^{11} \leq \beta(u) \leq 1$. With a little more work, it can be shown that for any $u \in (0,1)$, $\beta(u) \in (1-u^{11}, (6(1-u))^{1/3})$ and that if $0 < u \leq 1/2$, then $\beta(u) \in (1-u^{11}, 1-u^{12})$.

Lemma 3.2.2. Fix $p \in (0,1)$ and let R be a rectangle of dimension (m,h). Set $u = (1-p^2)^h$ and let X_{tiles} be a random configuration of 2-tiles, each included independently with probability p^2. Then,

$$(1-u^8)\beta(u)^m \leq \mathbb{P}_2(R \text{ is horizontally traversable by } X_{\text{tiles}}) \leq \beta(u)^{m-1}.$$

Proof. Fix $h \geq 1$ and set $u = u(p,h) = (1-p^2)^h$. Let C be any column of sites of height h, a rectangle of dimension $(1,h)$, then

$$\begin{aligned}
\mathbb{P}_2(C \text{ is 1-occupied}) &= 1-(1-p^2)^h = 1-u \\
\mathbb{P}_2(C \text{ is 2-occupied}) &= 1-(1-p^2)^{3h} = 1-u^3 \\
\mathbb{P}_2(C \text{ is either 1 or 2-occupied}) &= 1-(1-p^2)^{4h} = 1-u^4 \\
\mathbb{P}_2(C \text{ is unoccupied}) &= (1-p^2)^{4h} = u^4
\end{aligned}$$

Considering only the squares inside the relevant rectangle, for each $m \geq 0$, let $R_m = [m] \times [h]$ and set

$$A_m = \mathbb{P}_2(R_m \text{ horiz. trav.}) = \mathbb{P}_2(X_{\text{tiles}} \text{ has no double gaps in } R_m).$$

In order to obtain bounds on the value of A_m, a recursion for the sequence $\{A_m\}_{m\geq 0}$ is defined. Let the columns of R be denoted $C_1, C_2, \ldots, C_m$. When

$m \geq 3$, there are three distinct ways to traverse a rectangle of width m:

(a) either C_m is occupied and $R \setminus C_m$ is traversable, or

(b) C_m is unoccupied, C_{m-1} is occupied and $R \setminus (C_{m-1} \cup C_m)$ is traversable, or finally

(c) C_{m-1} and C_m are both unoccupied, the column C_{m-2} is 2-occupied and the first $m-3$ columns of R are traversable.

The first few values of A_m can be calculated exactly: $A_0 = 1$, $A_1 = 1$, and $A_2 = 1 - \mathbb{P}_2(C_1 \text{ is unoccupied})^2 = 1 - u^8$.

Considering the three cases above, for each $m \geq 3$, a recurrence relation for the sequence $\{A_m\}_{m\geq 0}$ is given by

$$A_m = (1-u^4)A_{m-1} + u^4(1-u^4)A_{m-2} + u^8(1-u^3)A_{m-3}.$$

Recall that $\beta(u)$ is a real root in $(0,1)$ of the polynomial $F_u(x) = x^3 - (1-u^4)x^2 - u^4(1-u^4)x - u^8(1-u^3)$. Instead of solving the recursion exactly, the goal is to show that for all m, the value of A_m is close to $\beta(u)^m$. The proof proceeds by induction on m.

The base cases can be checked directly,

$$(1-u^8)\beta(u)^0 = (1-u^8) < 1 = A_0 < \beta(u)^{-1} \qquad (\text{since } \beta(u)^{-1} > 1)$$
$$(1-u^8)\beta(u) < 1 = A_1 = \beta(u)^0$$

Since $\beta(u) < 1$, then $(1-u^8)\beta(u)^2 < (1-u^8) = A_2$ and since $u \in (0,1)$, $A_2 = 1 - u^8 < 1 - u^{11} < \beta(u)$.

The rest follows by induction, using the fact that $\beta(u)$ is a zero of the characteristic equation for the recurrence for the sequence $\{A_m\}$. □

In what follows, a few basic properties of the function $\beta(u)$ are used: the rough bounds already given and the properties stated in the following lemma.

Lemma 3.2.3. In the interval $(0,1)$, the function $\beta(u)$ is continuous, decreasing, and concave.

Proof. By the implicit function theorem, since $F(u,x)$ is a continuously differentiable function, so is $\beta(u)$ on any open interval for which $\frac{\partial F}{\partial x}(u,\beta(u)) \neq 0$. Now

$$\begin{aligned}
\frac{\partial F}{\partial x}(u,x) &= 3x^2 - 2(1-u^4)x - u^4(1-u^4) \\
&= \frac{3}{x}\left(x^3 - \frac{2}{3}(1-u^4)x^2 - \frac{1}{3}u^4(1-u^4)x\right) \\
&= \frac{3}{x}\left(F(u,x) + \frac{1}{3}(1-u^4)x^2 + \frac{2}{3}u^4(1-u^4)x + u^8(1-u^3)\right) \\
&> \frac{3}{x}F(u,x) \qquad\qquad \text{(for } u,x \in (0,1))
\end{aligned}$$

Thus, for any $u \in (0,1)$, $\frac{\partial F}{\partial x}(u,\beta(u)) > \frac{3}{\beta(u)}F(u,\beta(u)) = 0$. Further,

$$\begin{aligned}
\frac{\partial F}{\partial u}(u,x) &= 4u^3x^2 - 4u^3x + 8u^7x - 8u^7 + 11u^{10} \\
&= 4u^3(x^2 - (1-2u^4)x - 2u^4(1-11/8u^3)).
\end{aligned}$$

Since for all $u \in (0,1)$, $\beta(u) > 1 - u^{11} > 1 - 2u^4$,

$$\begin{aligned}
\frac{\partial F}{\partial u}(u,\beta(u)) &= 4u^3\big(\beta(u)(\beta(u) - (1-2u^4)) - 2u^4(1-11/8u^3)\big) \\
&> 4u^3\big((1-u^{11})(1-u^{11}-1+2u^4) - 2u^4 + 11/4u^7\big) \\
&= 4u^{10}(11/4 - u^4 - 2u^8 + u^{15}) \\
&= 4u^{10}(3/4 + (1-u^7)(1-u^8) + (1-u^4) + u^7(1-u)) \\
&> 0.
\end{aligned}$$

Thus, for all $u \in (0,1)$, the function $\beta(u)$ is differentiable and hence continuous with

$$\beta'(u) = -\frac{\frac{\partial F}{\partial u}(u,\beta(u))}{\frac{\partial F}{\partial x}(u,\beta(u))} < 0$$

and hence $\beta(u)$ is decreasing.

To see that β is concave note that by differentiating implicitly,

$$\beta''(u) = -\left(\frac{\frac{\partial^2 F}{\partial u^2}\left(\frac{\partial F}{\partial x}\right)^2 + \frac{\partial^2 F}{\partial x^2}\left(\frac{\partial F}{\partial u}\right)^2 - 2\frac{\partial^2 F}{\partial u \partial x}\frac{\partial F}{\partial x}\frac{\partial F}{\partial u}}{\left(\frac{\partial F}{\partial x}\right)^3}\right)(u,\beta(u)).$$

As above, $\left(\frac{\partial F}{\partial x}\right)^3(u,\beta(u)) > 0$ and to see that

$$\left(\frac{\partial^2 F}{\partial u^2}\left(\frac{\partial F}{\partial x}\right)^2 + \frac{\partial^2 F}{\partial x^2}\left(\frac{\partial F}{\partial u}\right)^2 - 2\frac{\partial^2 F}{\partial u \partial x}\frac{\partial F}{\partial x}\frac{\partial F}{\partial u}\right)(u,\beta(u)) > 0, \tag{3.6}$$

expanding the above expression yields

$$\begin{aligned}
&\left(\frac{\partial^2 F}{\partial u^2}\left(\frac{\partial F}{\partial x}\right)^2 + \frac{\partial^2 F}{\partial x^2}\left(\frac{\partial F}{\partial u}\right)^2 - 2\frac{\partial^2 F}{\partial u \partial x}\frac{\partial F}{\partial x}\frac{\partial F}{\partial u}\right)(u,x)\\
&= 2u^6(-1+u^4+3x)(11u^7+8u^4(-1+x)+4(-1+x)x)^2\\
&\quad - 8u^6(-1+2u^4+2x)(11u^7+8u^4(-1+x)+4(-1+x)x)(u^8\\
&\quad + u^4(-1+2x)+x(-2+3x)) + 2u^2(55u^7+28u^4(-1+x)\\
&\quad + 6(-1+x)x)(u^8+u^4(-1+2x)+x(-2+3x))^2\\
&= -120u^{14}+374u^{17}+48u^{18}-242u^{20}-308u^{21}+72u^{22}+242u^{24}\\
&\quad - 66u^{25}-204u^{10}x+440u^{13}x+624u^{14}x-1760u^{17}x+36u^{18}x+726u^{20}x\\
&\quad + 264u^{21}x-72u^{22}x-240u^6x^2+440u^9x^2+924u^{10}x^2-1628u^{13}x^2\\
&\quad - 792u^{14}x^2+1452u^{17}x^2-84u^{18}x^2-48u^2x^3+984u^6x^3-1320u^9x^3\\
&\quad - 1224u^{10}x^3+1320u^{13}x^3+288u^{14}x^3+192u^2x^4-1296u^6x^4+990u^9x^4\\
&\quad + 504u^{10}x^4-252u^2x^5+552u^6x^5+108u^2x^6
\end{aligned}$$

$$= F(u,x)(108u^2x^3 + (-144u^2 + 444u^6)x^2 + (48u^2 - 600u^6 + 990u^9$$
$$- 48u^{10})x + 192u^6 - 330u^9 + 24u^{10} + 222u^{13} - 108u^{14})+$$
$$+ 2u^9x^2((55 - 18u + 29u^4 + 36u^5 - 102u^8 + 36u^9)+$$
$$+ ux(18 + 31u^3 - 72u^4 + 191u^7 - 72u^8 - 132u^{10} + 45u^{11} + 18u^{12})+$$
$$+ u^5(1-u)(36 + 36u + 36u^2 - 38u^3 - 2u^4 - 2u^5 + 42u^6 - 13u^7+$$
$$- 31u^8 - 31u^9 - 21u^{10}))).$$

Then inequality (3.6) follows since $F(u,\beta(u)) = 0$ and

$$55 - 18u + 29u^4 + 36u^5 - 102u^8 + 36u^9$$
$$= 36 + (1-u)(19 + u(1 + u + u^2 + 30u^3 + 66u^4 + 66u^5 + 30u^6$$
$$+ 36u^6(1-u))) > 0,$$
$$18 + 31u^3 - 72u^4 + 191u^7 - 72u^8 - 132u^{10} + 45u^{11} + 18u^{12}$$
$$= 18 + u^3(31 - 72u + 191u^4 - 72u^5 - 132u^7 + 45u^8 + 18u^9)$$
$$= 18 + u^3\left(\frac{283}{28} + 63\left(u - \frac{3}{14}\right)^2 + 50u^4 + 72u^4(1-u) + 69u^4(1-u^3)\right.$$
$$\left. + 45u(1-u^6)(1-u) + 18(1-u^7)(1-u^2)\right) > 0$$
$$36 + 36u + 36u^2 - 38u^3 - 2u^4 - 2u^5 + 42u^6 - 13u^7 - 31u^8 - 31u^9 - 21u^{10}$$
$$= 33 + (1-u)(3 + 39u + 75u^2 + 37u^3 + 35u^4 + 33u^5 + 75u^6$$
$$+ 62u^7 + 31u^8) > 0.$$

□

Following similar notation to that used in the study of the usual bootstrap percolation, it is often convenient to use the following functions in place of $\beta(u)$. Set

$$g(x) = -\log(\beta(e^{-x})).$$

For $p \in (0,1)$, define

$$q = q(p) = (-\log(1-p^2))^{1/2}.$$

When p is small, $q^2 \sim p^2$, with the advantage that for any $p > 0$ and $h \in \mathbb{Z}^+$,

$$\beta((1-p^2)^h) = e^{-g(hq^2)}. \tag{3.7}$$

Since $\beta(u)$ is defined for $u \in [0,1]$, then $g(x)$ is defined for $x \in (0,\infty)$ and has the following useful properties.

Fact 3.2.4. The function $g(x) = -\log(\beta(e^{-x}))$ is decreasing, convex and integrable on $(0,\infty)$.

Proof of Fact. Since $\beta(u)$ is a decreasing function of u, then $\frac{\partial}{\partial x}(\beta(e^{-x})) = -e^{-x}\beta(e^{-x}) > 0$. Thus, since $-\log x$ is decreasing in x, the function g is decreasing.

The function $\beta(e^{-x})$ is concave since

$$\frac{\partial^2}{\partial x^2}(\beta(e^{-x})) = e^{-x}(\beta''(e^{-x})e^{-x} + \beta'(e^{-x})) < 0.$$

This is used to show that g is convex as follows. Let $a, b > 0$ and fix $t \in [0,1]$. Since $\beta(e^{-x})$ is concave, $\beta(e^{-(ta+(1-t)b)}) \geq t\beta(e^{-a}) + (1-t)\beta(e^{-b})$ and so since the function $-\log x$ is both decreasing and convex in x then

$$\begin{aligned} g(ta+(1-t)b) &= -\log\left(\beta(e^{-(ta+(1-t)b)})\right) \\ &\leq -\log\left(t\beta(e^{-a}) + (1-t)\beta(e^{-b})\right) \\ &\leq t(-\log\beta(e^{-a})) + (1-t)(-\log\beta(e^{-b})) \\ &= tg(a) + (1-t)g(b). \end{aligned}$$

To see that g is integrable, note that since $\beta(e^{-x}) \geq 1 - e^{-11x}$, then

$g(x) \le -\log\left(1 - e^{-11x}\right)$ and so

$$\begin{aligned}
\int_0^\infty g(x)\ dx &\le \int_0^\infty -\log(1 - e^{-11x})\ dx \\
&= \int_0^\infty \left(\sum_{k \ge 1} \frac{e^{-11kx}}{k} \right)\ dx \\
&= \sum_{k \ge 1} \left(\int_0^\infty \frac{e^{-11kx}}{k}\ dx \right) \\
&= \sum_{k \ge 1} \frac{1}{11k^2} = \frac{\pi^2}{66} < \infty.
\end{aligned}$$

Thus, g is convex and $\int_0^\infty g(x)\ dx < \infty$.

Set $\lambda = \lambda_{\mathcal{M}} = \int_0^\infty g(x)\ dx$ and for $n > 0$, set $\lambda_n = \int_{1/n}^n g(x)\ dx$. The exact value of λ is not used in any of the proofs that follow, but it can be shown that

$$\lambda \approx 0.0779.$$

The results of Holroyd [24] on the critical probability for usual bootstrap percolation can be directly applied to the model of infection by 2-tiles with the function g as given in equation (3.7) and λ as above. For $\{p(n)\}_{n \ge 1} \subseteq (0, 1)$, let $X_{\text{tiles}}(n)$ a random configuration of 2-tiles in $[n]^2$, with each 2-tile included in X_{tiles} independently with probability $p(n)^2$. Then, for all $\varepsilon > 0$, if for all $n \ge 1$, $p(n)^2 < \frac{\lambda - \varepsilon}{\log n}$ then

$$\mathbb{P}_2(X_{\text{tiles}}(n) \text{ percolates}) = o(1).$$

Similarly, for all $\varepsilon > 0$, if for all $n \ge 1$, $p(n)^2 > \frac{\lambda + \varepsilon}{\log n}$, then

$$\mathbb{P}_2(X_{\text{tiles}}(n) \text{ percolates}) = 1 - o(1).$$

It remains to show that, indeed, the model of infection by 2-tiles is a good approximation for the probability of percolation in the modified process $\mathcal{M}$ when single sites are initially infected. In Chapters 4 and 5, two different alterations of an initially infected set of sites are given to obtain lower and upper bounds, respectively, on the probability of percolation.

Chapter 4

Lower bound for probability of percolation

4.1 Traversing rectangles

In this chapter, it is shown that for certain values of p and n, it is very likely that the grid, $[n]^2$, percolates in the modified bootstrap process when sites are initially infected independently with probability p.

Given a configuration of infected sites X, a new configuration X^- is defined so that $X^- \subseteq X$ and with the property that every site in X^- has a neighbour in X^- either sharing an edge or a corner. The configuration X^- can then be compared to configurations of 2-tiles. This is accomplished most simply in the cases where there is no ambiguity with regards to assigning 2-tiles to pairs of sites in X^-.

Throughout, let $X_1 \subseteq R$ be the set of initially infected sites; each site infected independently with probability p. As before, let X_{tiles} be a configuration of 2-tiles on the sites in R with each 2-tile occurring independently with probability p^2. Given a configuration of 2-tiles, X_{tiles}, define $|X_{\text{tiles}}|$ to be the number of squares in the grid that are contained in at least one 2-tile. If there is a site is contained in more than one 2-tile, of the configuration X_{tiles}, the site is only counted once for $|X_{\text{tiles}}|$.

A configuration of sites, X_0, where every site is contained in some 2-tile can be most easily compared to a configuration of 2-tiles if X_0 determines exactly one configuration of 2-tiles. With this in mind, it will be useful to keep track of pairs of 2-tiles that could cause ambiguity. Recall that for any site (x_1, x_2) in the grid, $B_1^*((x_1, x_2)) = \{(y_1, y_2) : \max\{|x_1 - y_1|, |x_2 - y_2|\} \leq 1\}$.

Definition 4.1.1. A pair of 2-tiles $\{\mathbf{x}, \mathbf{x_1}\}$ and $\{\mathbf{y}, \mathbf{y_1}\}$ forms a *triple* if

$$B_1^*(\{\mathbf{x}, \mathbf{x_1}\}) \cap \{\mathbf{y}, \mathbf{y_1}\} \neq \emptyset.$$

Thus, two tiles that overlap form a triple and also two 2-tiles that touch, either along an edge or at a corner, form a triple. These are called triples since they involve at least 3 sites and so occur in the set X_0 with probability at most p^3.

Definition 4.1.2. For any $n \in \mathbb{N}$ and $X \subseteq [n]^2$, define $X^- \subseteq X$ as follows:

$$X^- = \{\mathbf{x} \in X : B_1^*(\mathbf{x}) \cap X \neq \{\mathbf{x}\}\}.$$

If $\mathbf{x} \in X$ and $B_1^*(\mathbf{x}) \cap X = \{\mathbf{x}\}$, call $\mathbf{x}$ an *isolated site.*

Since $X^- \subseteq X$, if X^- percolates, then so does X. However, since every site in X^- has a neighbour in X^- sharing an edge or a corner, the set X^- can be compared to a configuration of 2-tiles and the estimates from Lemma 3.2.2 on the probability of traversing a rectangle can be used. In the following lemma, a lower bound is given for the probability that a rectangle of a particular scale is horizontally traversable. In further proofs, this lower bound is used for rectangles of height either slightly smaller or slightly larger than p^{-2} and so rectangles are considered whose height is in the interval $[p^{-15/8}, p^{-17/8}]$. In order to better control the errors that occur, only rectangles of width at most $p^{-1/4}$ are considered.

Lemma 4.1.3. There is a $p_0 > 0$ so that for all $p < p_0$, $h = h(p)$ with $p^{-15/8} \leq h \leq p^{-17/8}$, $m = m(p)$ with $1 \leq m \leq p^{-1/4}$ and rectangle R of dimension (m, h), if $X_1 \sim \text{Bin}(R, p)$ then,

$$\mathbb{P}(R \text{ is horiz. trav. by } X_1) \geq e^{-463hmp^{5/2}}(1 - (1 - p^2)^{8h})e^{-mg(q^2h)}.$$

Proof. Fix p, h, and m with $p^{-15/8} \leq h \leq p^{-17/8}$ and $1 \leq m \leq p^{-1/4}$. Let R be a rectangle of dimension (m, h) and define a set of configurations of 2-tiles

$$\mathcal{Q} = \{A \ \mid A \text{ is a configuration of 2-tiles in } R\text{, containing no triples with } |A| \leq hmp^{3/2}\}.$$

The configurations of 2-tiles in $\mathcal{Q}$ are, essentially, those that can be unambiguously compared to configurations of infected sites. In later estimates, it is useful to assume that $|X_{\text{tiles}}|$ is not too large and so the condition $|A| \leq hmp^{3/2}$ is included also.

Given a configuration A of 2-tiles, let A_1 be the set of sites that are contained in some 2-tile from A and let $\mathcal{Q}_1 = \{A_1 \ \mid A \in \mathcal{Q}\}$ be the configurations of infected sites corresponding to the configurations of 2-tiles in $\mathcal{Q}$.

First, it is shown that the probability of the event $|X_{\text{tiles}}| > hmp^{3/2}$ is relatively small. If at least $hmp^{3/2}$ sites are covered by 2-tiles in X_{tiles}, then at least $\frac{1}{2}hmp^{3/2}$ different 2-tiles were included in X_{tiles}. Note that, for p sufficiently small, $4hmp^2 < \frac{hmp^{3/2}}{2}$. Thus, by tail estimates for binomial random variables given in Lemma 1.2.1,

$$\begin{aligned}\mathbb{P}_2(|X_{\text{tiles}}| > hmp^{3/2}) &\le \mathbb{P}\left(\text{Bin}(4hm, p^2) > \frac{1}{2}hmp^{3/2}\right)\\ &\le \exp\left(\frac{hmp^{3/2}}{2}\right)(8p^{1/2})^{hmp^{3/2}/2}\\ &\le \exp\left(\frac{hmp^{3/2}}{2}\right)(e^{-5})^{hmp^{3/2}/2} \qquad \text{(for } p \le e^{-10}/64)\\ &= e^{-2hmp^{3/2}}.\end{aligned}$$

In order to compare this term with those involving $\beta(u)$, note that since $u = (1-p^2)^h \le e^{-11p^2h}$ and $\beta(u) \ge 1 - u^{11}$,

$$\beta(u) \ge 1 - e^{-11p^2h} \ge 1 - e^{-11p^{1/8}} \ge e^{-p^{-3/8}} \ge e^{-hp^{3/2}}.$$

Thus, $\mathbb{P}_2(|X_{\text{tiles}}| > hmp^{3/2}) \le e^{-hmp^{3/2}}\beta(u)^m$.

Fix $A \in \mathcal{Q}$. Since A contains no triples, the configuration A consists of exactly $|A|/2$ tiles. Thus,

$$\begin{aligned}\mathbb{P}_2(X_{\text{tiles}} = A) &= (p^2)^{|A|/2}(1-p^2)^{4|R|-|A|/2}\\ &= p^{|A|}(1-p^2)^{4|R|-|A|/2}.\end{aligned}$$

In order to bound the probability that $X_1^- = A_1$, note that $X_1^- = A_1$ if the following three events occur:

- E_1: the event that $A_1 \subseteq X_1$,
- E_2: the event $(B_1^*(A_1) \setminus A_1) \cap X_1 = \emptyset$, and
- E_3: the event that every site $\mathbf{x} \in X_1 \setminus A_1$ is isolated.

Since E_1 is independent of $E_2 \cap E_3$,

$$\mathbb{P}(X_1^- = A_1) = \mathbb{P}(E_1 \cap E_2 \cap E_3) = \mathbb{P}(E_1)\mathbb{P}(E_2 \cap E_3) = p^{|A|}\mathbb{P}(E_2 \cap E_3).$$

To obtain an upper bound on $|B_1^*(A_1) \setminus A_1|$, note that if $\mathbf{x}_1$ and $\mathbf{x}_2$ are two sites sharing an edge, then $|B_1^*(\{\mathbf{x}_1, \mathbf{x}_2\}) \setminus \{\mathbf{x}_1, \mathbf{x}_2\}| = 10$ whereas if $\mathbf{x}_1$ and $\mathbf{x}_2$ are two sites sharing a corner, then $|B_1^*(\{\mathbf{x}_1, \mathbf{x}_2\}) \setminus \{\mathbf{x}_1, \mathbf{x}_2\}| = 12$. In both cases, $|B_1^*(\{\mathbf{x}_1, \mathbf{x}_2\}) \setminus \{\mathbf{x}_1, \mathbf{x}_2\}| \leq 6|\{\mathbf{x}_1, \mathbf{x}_2\}|$ and in general, there are at most $6|A|$ sites in $B_1^*(A_1) \setminus A_1$,

$$\mathbb{P}(E_2) = (1-p)^{|B_1^*(A_1)\setminus A_1|} \geq (1-p)^{6|A|}.$$

The event E_3 is the intersection of a collection of decreasing events: that for each site outside of A_1, none of the 4 possible sets of sites forming tiles is included in X_1. Thus, by Lemma 1.2.2, $\mathbb{P}(E_3) \geq (1-p^2)^{4(|R|-|A|)}$. Since E_2 and E_3 are both decreasing events, applying Lemma 1.2.2 again yields

$$\mathbb{P}(E_2 \cap E_3) \geq \mathbb{P}(E_2)\mathbb{P}(E_3) \geq (1-p)^{6|A|}(1-p^2)^{4(|R|-|A|)}.$$

Thus,

$$\begin{aligned}
\mathbb{P}(X_1^- = A_1) &\geq p^{|A|}(1-p)^{6|A|}(1-p^2)^{4(|R|-|A|)} \\
&= p^{|A|}(1-p^2)^{4|R|-|A|/2}(1-p)^{6|A|} \\
&\geq \mathbb{P}_2(X_{\text{tiles}} = A)(1-p)^{6|A|} \\
&\geq \mathbb{P}_2(X_{\text{tiles}} = A)e^{-7p|A|} \\
&\geq \mathbb{P}_2(X_{\text{tiles}} = A)e^{-7hmp^{5/2}} \qquad \text{(since } |A| \leq hmp^{3/2}\text{)} \qquad (4.1)
\end{aligned}$$

Let $\mathcal{C} = \{A : A \subseteq R \text{ and } R \text{ is traversable by } A\}$. Then,

$$\begin{aligned}
\mathbb{P}(X_1 \in \mathcal{C}) &\geq \mathbb{P}(X_1^- \in \mathcal{C} \cap \mathcal{Q}_1) \\
&= \sum_{A_1 \in \mathcal{Q}_1 \cap \mathcal{C}} \mathbb{P}(X_1^- = A_1) \\
&\geq \sum_{A \in \mathcal{Q} \cap \mathcal{C}} \mathbb{P}_2(X_{\text{tiles}} = A) e^{-7hmp^{5/2}} && \text{(by inequality (4.1))} \\
&= e^{-7hmp^{5/2}} \mathbb{P}_2(X_{\text{tiles}} \in \mathcal{C} \cap \mathcal{Q}) \\
&= e^{-7hmp^{5/2}} \left[\mathbb{P}_2(X_{\text{tiles}} \in \mathcal{C}) - \mathbb{P}(X_{\text{tiles}} \in \mathcal{C} \setminus \mathcal{Q})\right].
\end{aligned}$$

By Lemma 3.2.2, with $u = (1-p^2)^h$, the probability of traversing satisfies $\mathbb{P}_2(X_{\text{tiles}} \in \mathcal{C}) \geq \beta(u)^m (1-u^8)$ and so it remains to find an appropriate upper bound for $\mathbb{P}(X_{\text{tiles}} \in \mathcal{C} \setminus \mathcal{Q})$. First,

$$\begin{aligned}
&\mathbb{P}_2(X_{\text{tiles}} \in \mathcal{C} \setminus \mathcal{Q}) \\
&\qquad \leq \mathbb{P}_2(X_{\text{tiles}} \in \mathcal{C} \text{ and } X_{\text{tiles}} \text{ contains a triple}) + \mathbb{P}_2(|X_{\text{tiles}}| > hmp^{3/2}) \\
&\qquad \leq \sum_{T \text{ a triple}} \mathbb{P}_2(X_{\text{tiles}} \in \mathcal{C} \text{ and } T \subseteq X_{\text{tiles}}) + e^{-hmp^{3/2}} \beta(u)^m.
\end{aligned}$$

Fix a triple T and consider $\mathbb{P}_2(X_{\text{tiles}} \in \mathcal{C} \text{ and } T \subseteq X_{\text{tiles}})$. Note that the sites in the triple T are contained in at most 4 different columns of R. Removing the columns containing sites from T produces two smaller rectangles R_1 and R_2 both of height h. If X_{tiles} crosses R, there are rectangles R_1' and R_2' obtained from R_1 and R_2 by removing at most one column from each so that R_1' and R_2' are of height h and of width m_1 and m_2 (respectively) with $m_1 + m_2 \geq m - 6$ and with the property that R_1' is traversable by $X_{\text{tiles}} \cap R_1'$ and R_2' is traversable by $X_{\text{tiles}} \cap R_2'$. Thus,

$$\mathbb{P}_2(R \text{ is traversable by } X_{\text{tiles}} \text{ and } T \subseteq X_{\text{tiles}})$$

$$\le \mathbb{P}_2(R_1' \text{ is trav. by } X_{\text{tiles}} \cap R_1')\mathbb{P}_2(R_2' \text{ is trav. by } X_{\text{tiles}} \cap R_2')\mathbb{P}_2(T \subseteq X_{\text{tiles}})$$
$$\le \beta(u)^{m_1+m_2-2}(p^2)^2$$
$$\le \beta(u)^{m-8}p^4.$$

For each site $\mathbf{x} \in R$, consider the possible number of triples that contain $\mathbf{x}$ as one of the anchor sites. There are 4 different 2-tiles that contain $\mathbf{x}$ as the anchor site. If $\{\mathbf{x}, \mathbf{x}_2\}$ is a 2-tile, then $|B_1^*(\{\mathbf{x}, \mathbf{x}_2\})| \le 14$ and for each of the sites $\mathbf{y} \in B_1^*(\{\mathbf{x}, \mathbf{x}_2\})$, there are at 8 different 2-tiles that contain $\mathbf{y}$. Since $\mathbf{x}$ could be the anchor site of one of two 2-tiles, the number of triples that contain $\mathbf{x}$ as one of the anchor sites is at most $4 \cdot 14 \cdot 8/2 = 224$. Thus

$$\sum_{T \text{ a triple}} \mathbb{P}_2(R \text{ is traversable by } X_{\text{tiles}} \text{ and } T \subseteq X_{\text{tiles}}) \le 224hmp^4\beta(u)^{m-8}$$

and so

$$\begin{aligned}\mathbb{P}_2(X_{\text{tiles}} \in \mathcal{C} \setminus \mathcal{Q}) &\le e^{-hmp^{3/2}}\beta(u)^m + 224hmp^4\beta(u)^{m-8} \\ &= \beta(u)^m(e^{-hmp^{3/2}} + 224hmp^4\beta(u)^{-8})\end{aligned}$$

Since

$$\beta(u) \ge 1 - e^{-11p^2h} \ge \begin{cases} (1-e^{-11})p^2h & p^{1/8} \le p^2h \le 1 \\ 1 - e^{-11} & 1 \le p^2h \end{cases}$$

It follows that $\beta(u)^{-8} \le (1-e^{-11})^{-8}p^{-1}$.

Thus,

$$\begin{aligned} e^{-hmp^{3/2}} + 224hmp^4\beta(u)^{-8} &\le e^{-hmp^{3/2}} + 224hmp^4(1-e^{-11})^{-8}p^{-1} \\ &\le e^{-hmp^{3/2}} + 225hmp^3 \\ &\le e^{-p^{-3/8}} + 225hmp^3 \qquad \text{(since } hm \ge p^{-15/8}\text{)} \end{aligned}$$

$$
\begin{aligned}
&\le p^{9/8} + 225hmp^3 \\
&\le 226hmp^3. && \text{(since } hm \ge p^{-15/8}\text{)}
\end{aligned}
$$

Combining these bounds yields

$$
\begin{aligned}
\mathbb{P}(X_1 \in \mathcal{C}) &\ge e^{-7hmp^{5/2}}(1 - u^8 - 226hmp^3)\beta(u)^m \\
&\ge e^{-7hmp^{5/2}}(1 - u^8)\left(1 - \frac{226hmp^3}{1 - u^8}\right)\beta(u)^m.
\end{aligned}
$$

Now, by calculations similar to those in previous estimates,

$$
\begin{aligned}
\frac{226hmp^3}{1 - u^8} &\le \frac{226hmp^3}{1 - e^{-8p^2h}} \\
&\le \begin{cases} 227p^{3/4} & p^{1/8} \le p^2h \le 1 \\ 227p^{5/8} & 1 \le p^2h \le p^{-1/8} \end{cases} \\
&\le 227p^{5/8} && \text{(since } m \le p^{-1/4} \text{ and } p^2h \le p^{-1/8}\text{)}
\end{aligned}
$$

and so

$$
\begin{aligned}
\mathbb{P}(X_1 \in \mathcal{C}) &\ge e^{-7hmp^{5/2}}(1 - u^8)\left(1 - \frac{226hmp^3}{1 - u^8}\right)\beta(u)^m \\
&\ge e^{-7hmp^{5/2}}(1 - u^8)(1 - 227p^{5/8})\beta(u)^m \\
&\ge e^{-7hmp^{5/2}}(1 - u^8)e^{-454p^{5/8}}\beta(u)^m && \text{(for } p \text{ small)} \\
&\ge e^{-7hmp^{5/2} - 454hmp^{15/8}p^{5/8}}(1 - u^8)\beta(u)^m \\
&\ge e^{-463hmp^{5/2}}(1 - u^8)\beta(u)^m.
\end{aligned}
$$

Therefore, for p sufficiently small,

$$
\mathbb{P}(R \text{ is horiz. trav. by } X_1) \ge e^{-463hmp^{5/2}}(1 - u^8)\beta((1 - p^2)^h)^m
$$

$$= e^{-463hmp^{5/2}}(1-u^8)e^{-mg(q^2h)}$$

yielding the desired lower bound. □

4.2 Lower bound on $I(n,p)$

In the previous section, a bound on the crossing probability of a rectangle is given in terms of the function β. This is used here to establish a bound on the probability that a large, but not arbitrarily large, rectangle is internally spanned when sites are initially infected at random.

Lemma 4.2.1. There exists a $p_1 > 0$ such that if $p < p_1$, then

$$I(\lfloor p^{-17/8} \rfloor, p) \geq \exp\left(-\frac{2\lambda + 2p^{1/9}}{p^2}\right).$$

Proof. Fix $p \in (0,1)$ and set $m = \lfloor p^{-1/4} \rfloor$ and let h_0 be the smallest integer in $[p^{-15/8}, 2p^{-15/8}]$ such that $\lfloor p^{-17/8} \rfloor - h_0$ is divisible by m. Set

$$n = (\lfloor p^{-17/8} \rfloor - h_0)/m$$

and for $j = 1, 2, \ldots, n$, set $h_j = j \cdot m + h_0$. In particular $h_n = \lfloor p^{-17/8} \rfloor$ and $p^{-15/8}(1 - 3p^{1/4}) \leq n \leq p^{-15/8}$.

Setting $N = \lfloor p^{-17/8} \rfloor$, the square $[N]^2$ is internally spanned if the following three events all occur:

- The sites $(1,1), (2,2), \ldots, (h_0, h_0)$, and $(1,2)$ are initially infected,
- for $j = 1, 2, \ldots, n-1$, the rectangles $[h_j + 1, h_{j+1}] \times [1, h_j]$ are horizontally traversable from left to right and the rectangles $[1, h_j] \times [h_j + 1, h_{j+1}]$ are vertically traversable from bottom to top, and

- for each $j = 1, 2, \ldots, n$, the rectangle $\{h_j\} \times [1, h_{j-1}]$ and the rectangle $[1, h_{j-1}] \times \{h_j\}$ each contain two adjacent infected sites.

Let S denote the intersection of these three events. Note that

$$\mathbb{P}(\{h_{j+1}\} \times [1, h_j] \text{ contains two adjacent infected sites}) \geq 1 - (1-p^2)^{(h_j-1)/2}.$$

Since S is the intersection of increasing events, by Lemma 4.1.3 and Harris's Lemma (Lemma 1.2.2),

$$\begin{aligned}
\mathbb{P}(S) &\geq p^{h_0+1}\left(\prod_{j=0}^{n-1} \mathbb{P}(X_1 \text{ crosses } [h_j+1, h_{j+1}] \times [1, h_j])(1-(1-p^2)^{(h_j-1)/2})\right)^2 \\
&\geq p^{h_0+1}\left(\prod_{j=0}^{n-1} e^{-463h_j m p^{5/2}}(1-(1-p^2)^{8h_j})e^{-mg(q^2h_j)}(1-(1-p^2)^{(h_j-1)/2})\right)^2 \\
&= p^{h_0+1}\left(\prod_{j=0}^{n-1} e^{-463h_j m p^{5/2}}(1-e^{-8q^2h_j})e^{-mg(q^2h_j)}(1-e^{-q^2(h_j-1)/2})\right)^2.
\end{aligned}$$

Each of the terms in the above expression is simplified separately. First, since $m = h_j - h_{j-1}$, and $q \geq p$,

$$\begin{aligned}
\sum_{j=0}^{n-1} mg(q^2h_j) &= \frac{1}{q^2}\sum_{j=0}^{n-1} mq^2 g(q^2h_j) \\
&\leq \frac{1}{p^2}\int_0^\infty g(x)\ dx = \frac{\lambda}{p^2}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\sum_{j=0}^{n-1} h_j m p^{5/2} &\leq \sum_{j=0}^{n-1} h_j p^{9/4} && (\text{since } m \leq p^{-1/4}) \\
&\leq n h_n p^{9/4} \\
&\leq p^{-15/8} p^{-17/8} p^{9/4} \\
&= \frac{p^{1/4}}{p^2},
\end{aligned}$$

$$\sum_{j=0}^{n-1} -\log(1-e^{-8q^2 h_j}) \leq \frac{1}{8mq^2}\int_0^\infty \big(-\log(1-e^{-x})\big)\ dx$$
$$\leq \frac{p^{1/4}(\pi^2/24)}{p^2}$$
$$\sum_{j=0}^{n-1} -\log(1-e^{-q^2(h_j-1)/2}) \leq \frac{2}{q^2 m}\int_0^\infty \big(-\log(1-e^{-x})\big)\ dx$$
$$\leq \frac{p^{1/4}(2\pi^2/3)}{p^2}.$$

Finally,

$$\begin{aligned} p^{h_0+1} &= \exp((h_0+1)\log p) \\ &\geq \exp((p^{-15/8}+1)\log p) \\ &= \exp\left(-\frac{(-p^{1/8}\log p)(1+p^{15/8})}{p^2}\right) \\ &\geq \exp\left(-\frac{p^{1/9}}{p^2}\right). && \text{(for } p \text{ small enough)} \end{aligned}$$

Combining these yields

$$\begin{aligned} I(\lfloor p^{-17/8}\rfloor, p) &\geq \mathbb{P}(S) \\ &\geq \exp\left(-\frac{(p^{1/9}+2(463p^{1/4}+p^{1/4}\pi^2/24+p^{1/4}2\pi^2/3+\lambda)}{p^2}\right) \\ &\geq \exp\left(-\frac{2p^{1/9}+2\lambda}{p^2}\right), \end{aligned}$$

completing the proof of the lemma.

□

The bound from Lemma 4.2.1 can be further extended to estimating the probability that an arbitrarily large rectangle is internally spanned.

Lemma 4.2.2. There is a $p_2 > 0$ such that if $p < p_2$ and $n > p^{-17/8}$,

$$I(n,p) \geq \exp\left(-\frac{(2\lambda + 3p^{1/9})}{p^2}\right).$$

Proof. The idea of the proof that the grid $[n]^2$ is internally spanned if the sub-square $[1, \lfloor p^{-17/8} \rfloor]^2$ is internally spanned and the rest of the grid contains many rows and columns with pairs of adjacent, initially infected sites that allow the infection to spread one row and column at a time from this sub-square.

In particular, $[n]^2$ is internally spanned if the following events occur

- the square $[\lfloor p^{-17/8} \rfloor]^2$ is internally spanned, and
- for each $j = \lfloor p^{-17/8} \rfloor + 1, \ldots, n$, the rectangles $\{j\} \times [1, j-1]$ and $[1, j-1] \times \{j\}$ both contains pairs of adjacent sites that are initially infected.

Let S' be the above event and note that S' is the intersection of many independent events.

For each $j = \lfloor p^{-17/8} \rfloor + 1, \ldots, n$, let S_j be the event that $\{j\} \times [1, j-1]$ contains a pair of adjacent sites that are initially infected. Note that $\mathbb{P}(S_j) = \mathbb{P}([1, j-1] \times \{j\}$ contains a pair of adj. initially inf. sites) also. Then, as in the proof of the previous lemma, $\mathbb{P}(S_j) \geq 1 - (1-p^2)^{\lfloor j/2 \rfloor}$ and hence

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{j=\lfloor p^{-17/8} \rfloor + 1}^{n} S_j\right)^2 &= \prod_{j=\lfloor p^{-17/8} \rfloor + 1}^{n} \mathbb{P}(S_j)^2 \\
&\geq \prod_{j=\lfloor p^{-17/8} \rfloor + 1}^{n} (1 - (1-p^2)^{\lfloor j/2 \rfloor})^2 \\
&= \exp\left(2 \sum_{j=\lfloor p^{-17/8} \rfloor + 1}^{n} \log(1 - e^{-q^2(j-1)/2})\right) \\
&\geq \exp\left(-\frac{2}{p^2} \int_{q^2(\lfloor p^{-17/8} \rfloor + 1)/2}^{\infty} -\log(1 - e^{-x})\, dx\right).
\end{aligned}$$

It is straightforward to check that if $k \geq 1$, then $\int_k^\infty \left(-\log(1-e^{-x})\right)\ dx \leq \frac{5}{4}e^{-k}$. Thus,

$$\begin{aligned}
\mathbb{P}(\cap_{j=\lfloor p^{-17/8}\rfloor+1}^{n} S_j)^2 &\geq \exp\left(-\frac{5/2e^{-q^2(\lfloor p^{-17/8}\rfloor+1)/2}}{p^2}\right) \\
&\geq \exp\left(-\frac{5/2e^{-p^{-1/8}/4}}{p^2}\right) \\
&\geq \exp\left(-\frac{p^{1/9}}{p^2}\right).
\end{aligned}$$

and so $I(n,p) \geq \exp\left(-\frac{p^{1/9}}{p^2}\right) I(\lfloor p^{-17/8}\rfloor, p)$ and by the previous lemma,

$$I(n,p) \geq \exp\left(-\frac{p^{1/9}}{p^2}\right)\exp\left(-\frac{2p^{1/9}+2\lambda}{p^2}\right) = \exp\left(-\frac{3p^{1/9}+2\lambda}{p^2}\right)$$

as claimed. □

Following an argument similar to that used by Holroyd [24] for the analysis of the usual bootstrap process, Lemma 4.2.2 is used to show that if $p^2 \log n > \lambda$, then $I(n,p)$ is close to 1.

Theorem 4.2.3. For every $\varepsilon > 0$, there exists $n_0 \in \mathbb{Z}^+$ such that if $n \geq n_0$ and $p \in (0,1)$ with $p \geq \sqrt{\frac{\lambda+\varepsilon}{\log n}}$ then

$$I(n,p) \geq 1 - 3\exp(-n^{\varepsilon/6}).$$

Proof. Fix $\varepsilon > 0$ and $n_0 \geq 0$ large enough so that Lemma 4.2.2 applies for any p with $p \leq \sqrt{\frac{\lambda+\varepsilon/2}{\log n_0}}$.

Fix $n \geq n_0$ and $p \in (0,1)$ with $p \geq \sqrt{\frac{\lambda+\varepsilon}{\log n}}$. Note that if $p' < p$ then $I(n,p') \leq I(n,p)$ and so it suffices to prove the theorem for $p = \sqrt{\frac{\lambda+\varepsilon}{\log n}}$.

Instead of randomly infecting all sites at once, sites are infected in two 'rounds'. Two random configurations of infected sites are independently coupled so that a large sub-rectangle of $[n]^2$ is likely to be internally spanned by sites from the first configuration and that, using only sites from the second configuration, the infection is able to spread row by row and column by column from this rectangle to the entire grid.

Set $p_1 = \sqrt{\frac{\lambda+\varepsilon/2}{\log n}}$ and $p_2 = \frac{\varepsilon/2}{\log n}$. Let $X_0 \sim \text{Bin}([n]^2, p)$, $X_1 \sim \text{Bin}([n]^2, p_2)$, and $X_2 \sim \text{Bin}([n]^2, p_2)$. Let the two probability spaces associated with X_1 and X_2 be coupled independently so that $X_1 \cup X_2 \subseteq X_0$ since $p_1 + (1-p_1)p_2 \leq p$ for $n \geq 75$.

Set $\ell = \left\lfloor \exp\left(\frac{\varepsilon}{8p_1^2}\right)\right\rfloor$. Note that since $\lambda < 1/8$,

$$\ell = \left\lfloor \exp\left(\frac{\varepsilon}{8p_1^2}\right)\right\rfloor = \lfloor n^{\frac{\varepsilon}{8(\lambda+\varepsilon/2)}} \rfloor \leq n^{\varepsilon} < n.$$

Divide the grid $[n]^2$ into $\lfloor n/\ell \rfloor^2$ disjoint $\ell \times \ell$ sub-grids, with potentially some remainder: $\{[k\ell+1, (k+1)\ell] \times [j\ell+1, (j+1)\ell] : \ k, j \in [0, \lfloor n/\ell \rfloor - 1]\}$. For each of these $\ell \times \ell$ sub-grids, the probability that the sub-grid is internally spanned by X_1 is $I(\ell, p_1)$. The probability that none of these $\ell \times \ell$ sub-grids are internally spanned is

$$\begin{aligned}(1 - I(\ell, p_1))^{\lfloor n/\ell \rfloor^2} &\leq (1 - I(\ell, p_1))^{\frac{n^2}{2\ell^2}} \\ &\leq \exp\left(-\frac{n^2}{2\ell^2} I(\ell, p_1)\right).\end{aligned}$$

Now,

$$\begin{aligned}\frac{n^2}{2\ell^2} I(\ell, p_1) &\geq \frac{n^2}{2n^{\varepsilon/8\lambda}} \exp\left(-\frac{2\lambda + 3p_1^{1/4}}{p_1^2}\right) && \text{(by Lemma 4.2.2)} \\ &\geq \frac{1}{2} n^{2-\frac{\varepsilon}{8\lambda}} \exp\left(-\frac{2\lambda + 3p_1^{1/4}}{\lambda + \varepsilon/2} \log n\right) \\ &\geq \frac{1}{2} n^{2-\frac{\varepsilon}{8\lambda}} \exp\left(-\left(2 - \frac{\varepsilon}{\lambda}\right) \log n\right) \\ &= \frac{1}{2} n^{2-\frac{\varepsilon}{8\lambda}} n^{-2+\frac{\varepsilon}{\lambda}}\end{aligned}$$

$$= \frac{1}{2} n^{\frac{7\varepsilon}{8\lambda}}$$

$$\geq n^{\frac{3\varepsilon}{4\lambda}}. \qquad \left(\text{for } n \geq \exp\left(\frac{1}{2\varepsilon}\right)\right)$$

Let S be the event that at least one $\ell \times \ell$ sub-grid is internally spanned by X_1. Then, since $\lambda \leq 1/12$,

$$\mathbb{P}(S) \geq 1 - \exp(-n^{9\varepsilon}). \tag{4.2}$$

Next, consider the probability that an internally spanned $\ell \times \ell$ sub-grid, together with sites in X_2 will percolate in $[n]^2$. As in Lemma 4.2.2 the probability of this occurring is bounded below by the probability that, in many rows and columns, there are pairs of adjacent infected sites.

Let A_r be the event that for every k and j with $0 \leq k \leq \lfloor n/\ell \rfloor - 1$ and $1 \leq j \leq n$, the row $[k\ell + 1, (k+1)\ell] \times \{j\}$ contains at least two adjacent infected sites in X_2. Then

$$\begin{aligned}\mathbb{P}(A_r) \geq (1 - (1 - p_2^2)^{(\ell-1)/2})^{n\lfloor n/\ell \rfloor} &\geq (1 - \exp(-p_2^2(\ell-1)/2))^{n^2/\ell} \\ &\geq \exp\left(-\frac{2n^2}{\ell} e^{-p_2^2 \ell/3}\right)\end{aligned}$$

Now, for n large enough, $(\log n)^2 \leq \frac{\varepsilon^2}{12} n^{\frac{\varepsilon}{72(\lambda+\varepsilon/2)}}$ and so

$$\frac{p_2^2 \ell}{3} = \frac{\varepsilon^2 \lfloor n^{\varepsilon/8(\lambda+\varepsilon/2)} \rfloor}{12(\log n)^2} \geq n^{\frac{\varepsilon}{9(\lambda+\varepsilon/2)}}.$$

Similarly, for n sufficiently large, depending on ε,

$$2n^{2-\frac{\varepsilon}{8(\lambda+\varepsilon/2)}} \exp(-n^{\frac{\varepsilon}{9(\lambda+\varepsilon/2)}}) \leq \exp(-n^{\frac{\varepsilon}{10(\lambda+\varepsilon/2)}}) \leq \exp(-n^{\varepsilon/6}),$$

and hence

$$\mathbb{P}(A_r) \geq \exp(-e^{-n^{\varepsilon/6}}). \tag{4.3}$$

Define, similarly, A_c to be the event that for every k and j with $0 \le k \le \lfloor n/\ell \rfloor - 1$ and $1 \le j \le n$, the column $\{j\} \times [k\ell + 1, (k+1)\ell]$ contains at least two adjacent infected sites in X_2. Then $\mathbb{P}(A_c) = \mathbb{P}(A_r)$ and since the events A_c and A_r are both increasing events, by Harris's Lemma (Lemma 1.2.2), $\mathbb{P}(A_c \cap A_r) \ge \mathbb{P}(A_c)\mathbb{P}(A_r)$. Now, if both events S and $A_c \cap A_r$ occur, then $[n]^2$ is internally spanned by the set of initially infected sites $X_1 \cup X_2$. Thus,

$$\begin{aligned}
I(n,p) &\ge \mathbb{P}(S)\mathbb{P}(A_c \cap A_r) \\
&\ge \mathbb{P}(S)\mathbb{P}(A_c)\mathbb{P}(A_r) \\
&\ge (1 - \exp(-n^{9\varepsilon}))\exp(-2e^{-n^{\varepsilon/6}}) \qquad \text{(by eqns. (4.2) and (4.3))} \\
&\ge 1 - 2\exp(-n^{\varepsilon/6}) - \exp(-n^{9\varepsilon}) \\
&\ge 1 - 3\exp(-n^{\varepsilon/6}).
\end{aligned}$$

For n sufficiently large, depending on ε and if $p \ge \sqrt{\frac{\lambda+\varepsilon}{\log n}}$, then $I(n,p) \ge 1 - 3\exp(-n^{\varepsilon/6})$. □

In particular, for every $\varepsilon > 0$ and any sequence $\{p(n)\}_{n\in\mathbb{N}} \subseteq (0,1)$ with the property that for all $n \in \mathbb{N}$, $p(n) \ge \sqrt{\frac{\lambda+\varepsilon}{\log n}}$, then

$$I(n, p(n)) \ge 1 - 3\exp(-n^{\varepsilon/6}) = 1 - o(1)$$

and so with high probability, a random set of initially infected sites $X_0 \sim \mathrm{Bin}([n]^2, p(n))$ percolates in the modified bootstrap process. Thus the critical probability satisfies

$$p_c([n]^2, \mathcal{M}) \le \sqrt{\frac{\lambda + o(1)}{\log n}}.$$

Chapter 5

Upper bound for probability of percolation

5.1 Traversing rectangles and growing rectangles

To obtain an upper bound for the probability of percolation in the modified bootstrap process, an alteration of the initial configuration is defined that is different from the one given in Chapter 4. An initial configuration of infected sites X is altered to produce a new configuration X^+ that can be more easily compared to the process of infecting sites with 2-tiles, but in such a way that if X percolates in $\mathcal{M}$, then so does X^+. The idea is to uninfect isolated sites that do not affect the final infection status of any of their neighbours, while including some new infected sites next to isolated sites that have a chance of affecting whether or not their neighbours become infected.

For convenience, set

$$e_1 = (1, 0), e_2 = (0, 1), e_3 = (-1, 0) \text{ and } e_4 = (0, -1).$$

Recall the definitions of two different types of distances on the grid: balls in the ℓ_∞ metric are written $B_r^*(\mathbf{x})$ while balls in the ℓ_1 metric are written $B(\mathbf{x})$ (see 3.3 and 3.4).

Definition 5.1.1. For any $X \subseteq \mathbb{Z}^2$, define $X^+ \subseteq \mathbb{Z}^2$ as follows:

- if $\mathbf{x} \in X$ with $B_1^*(\mathbf{x}) \cap X \neq \{\mathbf{x}\}$, then $\mathbf{x} \in X^+$.
- if $\mathbf{x} \in X$ with $B_2(\mathbf{x}) \cap X = \{\mathbf{x}\}$ then $\mathbf{x} \notin X^+$.
- if $\mathbf{x} \in X$ is isolated and for some $i \in \{1, 2, 3, 4\}$, $\mathbf{x} + 2e_i \in X$, then
  - if $B_2(\{\mathbf{x}, \mathbf{x} + e_i, \mathbf{x} + 2e_i\}) \cap X \setminus \{\mathbf{x}, \mathbf{x} + 2e_i\} = \emptyset$ then $\mathbf{x} \notin X^+$, and
  - if $B_2(\{\mathbf{x}, \mathbf{x} + e_i, \mathbf{x} + 2e_i\}) \cap X \setminus \{\mathbf{x}, \mathbf{x} + 2e_i\} \neq \emptyset$ then $\mathbf{x}, \mathbf{x} + e_i \in X^+$.

Figure 5.1 shows the configurations of infected sites in X that are uninfected in X^+. The shaded sites represent infected sites and uninfected sites are represented by sites containing empty circles.

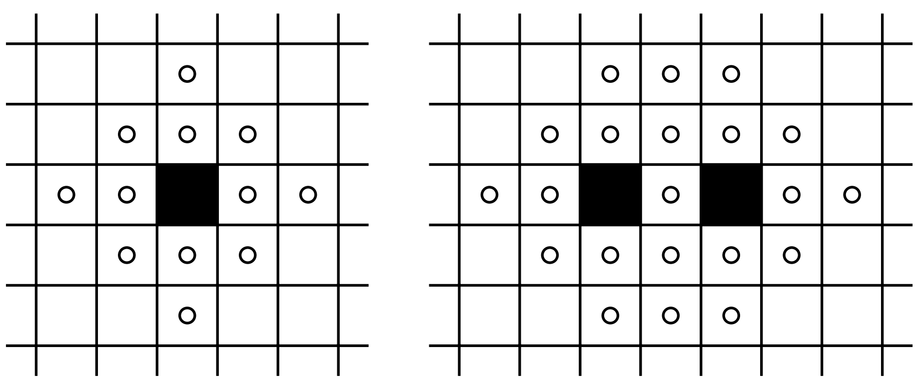

Figure 5.1: Sites from X that are uninfected

Figure 5.2 shows an isolated site $\mathbf{x}$ with $\mathbf{x}, \mathbf{x} + 2e_1 \in X$. If any other site inside the outlined region is infected (in X), then $\mathbf{x}$ and $\mathbf{x} + e_1$ (the site containing a shaded circle) are included in X^+.

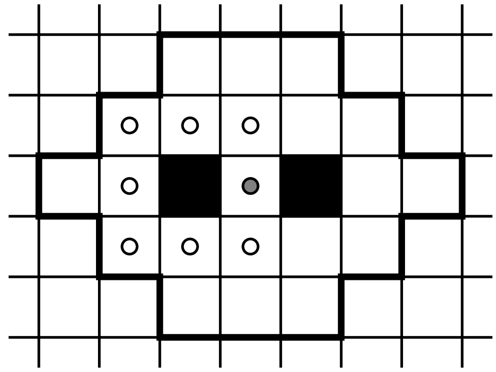

Figure 5.2: Sites included in X^+

Call any such configuration of three infected sites in X a *triplet*. In Figure 5.3, the different types of triplets are shown with the associated sites marked with an empty circle.

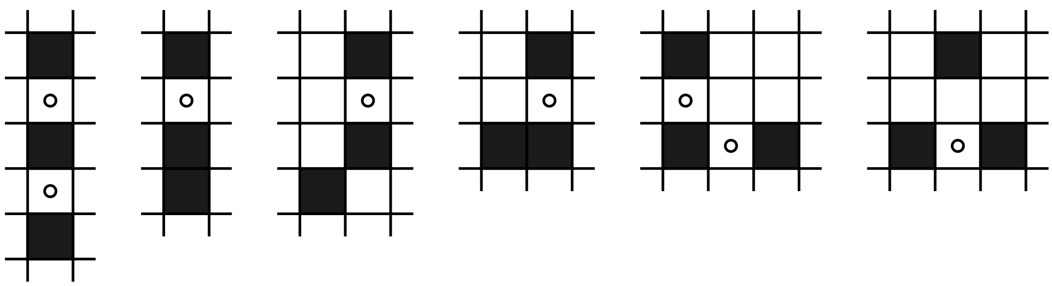

Figure 5.3: Six types of triplets

Considering rotations and reflections, there are 2 triplets of each of the first and second type, 8 triplets of each of the third and fourth type, and 4 triplets of each of the fifth and sixth types. Thus, in total, there are 28 different triplets.

As before, pairs of infected points that form one of the four 2-tiles are called a *double*. By definition, every site contained in a configuration X^+ is either contained in a double in X, or associated with a set of three points in X that form a triplet.

In the next lemma, it is shown how the probability that a rectangle R is traversable by the set X^+ can be compared to the probability that R is traversable by a configuration on 2-tiles. As in Chapter 4, first, rectangles with height close to p^{-2} are considered.

Lemma 5.1.2. Let $B > 1$, $Z > 0$, $m \in \mathbb{N}$ and set

$$Q_1(B, Z, m) = 1500Bm + \frac{30Bm + \frac{3\cdot 25^3}{Z}}{(1 - e^{-11Z})^{m-1}}.$$

There exists $p_0 = p_0(B, Z, m) > 0$ such that for all $h \in \mathbb{N}$ with $Z \leq hp^2 \leq B$ and every rectangle R of dimension (m, h), if $p < p_0$ and $X \sim \text{Bin}(R, p)$ then

$$\mathbb{P}(R \text{ is horizontally traversable}) \leq (1 + pQ_1(B, Z, m))e^{-g(hq^2)(m-1)}.$$

Proof. Fix $p > 0$, $B > 1$, $Z > 0$, $m \in \mathbb{N}$ and let $h \in \mathbb{N}$ be such that $Z \leq hp^2 \leq B$. Let R be a rectangle of dimension (m, h). Similarly to the proof of Lemma 4.1.3, let

$$\mathcal{Q} = \{A \subseteq R : \text{ every site in } A \text{ has a neighbour, } A \text{ contains no triplets and } |A| \leq |R|p\}$$

and let $\mathcal{C}$ be the collection of configurations of infected sites for which R is horizontally traversable from left to right in the process $\mathcal{M}$. Fix $A \in \mathcal{Q}$ and let $X \sim \text{Bin}(R, p)$. Since R will be horizontally traversable by X^+ if R is horizontally traversable by X,

$$\mathbb{P}(R \text{ is horiz. trav. by } X) \leq \mathbb{P}(R \text{ is horiz. trav. by } X^+).$$

If $X^+ = A$, then since A contains no triplets and any site in $X^+ \setminus X$ is contained in a triple, $A \subseteq X$. Further, any site in $X \setminus A$ is isolated and not contained in a triplet. In order to deal with independent events, consider the following two events:

- E_1: every site in A is in X, and
- E_2: there are no doubles or triplets in $X \cap (R \setminus B_3(A))$.

Since E_1 and E_2 are independent, $\mathbb{P}(X^+ = A) \leq \mathbb{P}(E_1)\mathbb{P}(E_2) = p^{|A|}\mathbb{P}(E_2)$. In order to bound $\mathbb{P}(E_2)$, a version of Janson's inequality (Theorem 1.2.5) is used.

Let $(B_i)_{i \in I}$ be the sequence of events that a particular double or triplet occurs in $X \cap (R \setminus B_3(A))$. For each site $\mathbf{x}$, there are 4 different doubles containing $\mathbf{x}$ as the left-most and bottom-most site and there are 28 different triplets containing $\mathbf{x}$ as the left-most and bottom-most site. Thus, there are at most $4|R \setminus B_3(A)|$ such doubles and at most $28|R \setminus B_3(A)|$ such triplets. Consider the number of sites in $|B_3(A)|$. For any double $\{\mathbf{x}_1, \mathbf{x}_2\}$, one can verify that

$$|B_3(\{\mathbf{x}_1, \mathbf{x}_2\})| = 32 = 16|\{\mathbf{x}_1, \mathbf{x}_2\}|.$$

Thus, since every site in the configuration A is contained in a double, $|B_3(A)| \leq 16|A|$.

Then $E_2 = \cap_{i \in I} \overline{B}_i$ and this event depends only on the $|R \setminus B_3(A)| \geq |R| - 16|A|$ independent events that a particular site is initially infected or not. In order to apply Theorem 1.2.5 a bound is required for the sum of probabilities of events $B_i \cap B_j$ for which B_i and B_j are not independent.

Consider the number of overlapping doubles and triples. For each site $\mathbf{x}$, there are 4 doubles containing $\mathbf{x}$ as the anchor and 2 sites in the double that could be overlapping with another double. For the sites in the first double, there are 8 different doubles containing that site. In this way each pair of overlapping doubles is counted twice and so there are at most $32|R \setminus B_3(A)|$ different pairs of overlapping doubles.

Similarly, since there are 28 different triplets, there are at most $4 \cdot 2 \cdot 3 \cdot 28 = 672$ different pairs of a triple and an overlapping double that $\mathbf{x}$ as its anchor at at most $28 \cdot 28 \cdot 3^2/2 = 3528$ pairs of triples that contain $\mathbf{x}$ as the lowest left-most site of one of the triplets. Therefore, in all, there are $672|R \setminus B_3(A)|$ different pairs of a double and a triplet that overlap and at most $3528|R \setminus B_3(A)|$ pairs of overlapping triplets. Since a pair of distinct doubles that overlap contain at least 3 sites, a double and a triple that overlap contain at least 3 sites and a pair of distinct triplets that overlap contain at least 4 sites,

$$\begin{aligned}\sum_{B_i, B_j \text{ not indep.}} \mathbb{P}(B_i \cap B_j) &\leq (32 + 672)|R \setminus B_3(A)|p^3 + 3528|R \setminus B_3(A)|p^4 \\ &\leq 710|R \setminus B_3(A)|p^3\end{aligned}$$

when $p \leq 1/588$. Similarly, $\sum_{i \in I} \mathbb{P}(B_i) \leq (4p^2 + 28p^3)|R \setminus B_3(A)|$ and applying Theorem 1.2.5,

$$\begin{aligned}\mathbb{P}(E_2) = \mathbb{P}(\cap_{i \in I} \overline{B}_i) &\leq \exp(-(4p^2 + 28p^3)|R \setminus B_3(A)|) \exp(710|R \setminus B_3(A)|p^3) \\ &= \exp((-4p^2 + 682p^3)|R \setminus B_3(A)|) \\ &\leq \exp((-4p^2 + 682p^3)(|R| - 16|A|)) \\ &\leq \exp(-4|R|p^2 + 64p^2|A| + 682p^3|R|) \\ &\leq \exp(-4p^2|R| + 746p^3|R|). \qquad \text{(since } |A| \leq |R|p)\end{aligned}$$

Thus, since the event that $X^+ = A$ is contained in the intersection of independent events E_1 and E_2,

$$\begin{aligned}\mathbb{P}(X^+ = A) &\leq p^{|A|} \exp(-4p^2|R| + 746|R|p^3) \\ &= p^{|A|}(1 - p^2)^{4|R| - |A|/2}(1 - p^2)^{-4|R| + |A|/2} \exp(-4p^2|R| + 746|R|p^3) \\ &= \mathbb{P}_2(X_{\text{tiles}} = A)(1 - p^2)^{-4|R| + |A|/2} \exp(-4p^2|R| + 746|R|p^3)\end{aligned}$$

$$\leq \mathbb{P}_2(X_{\text{tiles}} = A)(1-p^2)^{-4|R|} \exp(-4p^2|R| + 746|R|p^3).$$

For p sufficiently small, $1 - p^2 \geq e^{-(p^2+p^4)}$ and for x small enough, $e^x \leq 1 + 2x$. Thus,

$$\begin{aligned}
(1-p^2)^{-4|R|} \exp(-4p^2|R| + 746|R|p^3) &\leq \exp(4|R|(p^2+p^4) - 4p^2|R| + 746|R|p^3) \\
&\leq \exp(750|R|p^3) \\
&\leq \exp\left(750\frac{B}{p^2}mp^3\right) \\
&= \exp(750Bmp) \\
&\leq 1 + 1500Bmp.
\end{aligned}$$

Therefore,

$$\mathbb{P}(X^+ = A) \leq \mathbb{P}_2(X_{\text{tiles}} = A)(1 + 1500Bmp).$$

This inequality can be used to compare the probability that R is traversable by X^+ to that of R being traversable by a random configuration of 2-tiles, conditioned on either configuration being in the collection $\mathcal{Q}$.

Consider now the probability that $X^+ \notin \mathcal{Q}$. Since every site in X^+ has a neighbour, if $X^+ \notin \mathcal{Q}$ then either X^+ contains a triplet or $|X^+| > |R|p$. Let $\{T_j\}_{j\in J}$ be the collection of sets of sites in R that form triplets and consider first the probability that X^+ contains one of the triplets T_j. If $T_j \subseteq X^+$ then either $T_j \subseteq X$ or else one of the sites in T_j is associated with another triplet contained in X. In particular, if $T_j \nsubseteq X$, then every site in $T_j \setminus X$ is adjacent to at least 2 sites in X and together with sites in $T_j \cap X$, there are at least 4 sites in X. If a site $\mathbf{x} \in T_j \setminus X$ is associated with another triplet in X, then either $\{\mathbf{x} + (-1,0), \mathbf{x} + (1,0)\} \subseteq X$ or $\{\mathbf{x} + (0,-1), \mathbf{x} + (0,1)\} \subseteq X$. Very roughly then $\mathbb{P}(T_j \subseteq X^+) \leq p^3 + 3^3p^4$. Since there are at most $28|R|$ different triplets in R and

$|R| = hm \leq Bm/p^2$,

$$\mathbb{P}(\cup_{j\in J}\{T_j \subseteq X^+\}) \leq \sum_{j\in J} \mathbb{P}(T_j \subseteq X^+) \leq 28|R|(p^3 + 27p^3) \leq 30Bmp \tag{5.1}$$

as long as $p \leq 1/378$.

It is slightly more complicated to determine the probability that $|X^+| \geq |R|p$ since the events that any two sites are included in X^+ are not independent.

Since the membership in X^+ of any site is determined by at most 25 independent events, the initial infection of sites in X within a ball of radius 3, then a version of Talagrand's inequality [34] can be used to bound the probability that X^+ is large.

For every site $\mathbf{x} \in R$,

$$\mathbb{P}(\mathbf{x} \in X^+) \leq 8p^2 + 100p^3 \leq 10p^2$$

when p is sufficiently small. Thus $\mathbb{E}(|X^+|) \leq 10|R|p^2$. Changing the initial infection status of one site changes the value of $|X^+|$ by at most 25 and for any r, the event that $|X^+| \geq r$ can be certified by the initial infection status of $25r$ sites. Thus, applying Talagrand's inequality (Theorem 1.2.4),

$$\begin{aligned}
\mathbb{P}(|X^+| \geq |R|p) &\leq \exp\left(\frac{-(|R|p - 10|R|p^2)^2}{2 \cdot 25^3|R|p}\right) \\
&\leq \exp\left(-\frac{|R|p}{3 \cdot 25^3}\right) && \text{(for } p \leq 1/30\text{)} \\
&\leq \exp\left(-\frac{Zm}{3 \cdot 25^3 p}\right) && \text{(since } h \geq Z/p^2\text{)} \\
&\leq \frac{3 \cdot 25^3 p}{Zm} && \text{(using } e^{-x} \leq 1/x\text{)} \qquad (5.2)
\end{aligned}$$

Thus, the probability that X^+ is not a configuration in $\mathcal{Q}$ can be estimated as follows. Combining the two inequalities (5.1) and (5.2), yields

$$\begin{aligned}\mathbb{P}(X^+ \notin \mathcal{Q}) &\leq \mathbb{P}(X^+ \text{ contains a triplet}) + \mathbb{P}(|X^+| \geq |R|p) \\ &\leq 30Bmp + \frac{3 \cdot 25^3 p}{Zm}. \end{aligned} \tag{5.3}$$

Finally, it is possible to bound from above the probability that R is horizontally traversable by X^+ using Lemma 3.2.2,

$$\begin{aligned}\mathbb{P}(X^+ \in \mathcal{C}) &\leq \mathbb{P}(X^+ \in \mathcal{C} \cap \mathcal{Q}) + \mathbb{P}(X^+ \notin \mathcal{Q}) \\ &\leq \sum_{A \in \mathcal{C} \cap \mathcal{Q}} \mathbb{P}(X^+ = A) + 30Bmp + \frac{3 \cdot 25^3 p}{Zm} \\ &\leq \sum_{A \in \mathcal{C} \cap \mathcal{Q}} \mathbb{P}_2(X_{\text{tiles}} = A)(1 + 1500Bmp) + 30Bmp + \frac{3 \cdot 25^3 p}{Zm} \\ &\leq \mathbb{P}(X_{\text{tiles}} \in \mathcal{C})(1 + 1500Bmp) + 30Bmp + \frac{3 \cdot 25^3 p}{Zm} \\ &\leq e^{-g(hq^2)(m-1)}(1 + 1500Bmp) + 30Bmp + \frac{3 \cdot 25^3 p}{Zm} \\ &= e^{-g(hq^2)(m-1)} \left(1 + 1500Bmp + \frac{30Bmp + \frac{3 \cdot 25^3 p}{Zm}}{(1 - e^{-11hp^2})^{m-1}}\right) \\ &\leq e^{-g(hq^2)(m-1)} \left(1 + p\left(1500Bm + \frac{30Bm + \frac{3 \cdot 25^3}{Z}}{(1 - e^{-11Z})^{m-1}}\right)\right) \\ &= e^{-g(hq^2)(m-1)} \left(1 + pQ_1(B, Z, m)\right). \end{aligned}$$

Thus $\mathbb{P}(R \text{ is horiz. trav. by } X) \leq e^{-g(hq^2)(m-1)} \left(1 + pQ_1(B, Z, m)\right)$. □

In Lemma 5.1.2, the width of the rectangle being traversed is arbitrary. However, for large values of m, part of the error term, $Q_1(B, Z, m)$, might become too large for this lemma to be useful for upper bounds on the the probability of percolation. Instead of considering the probability of traversing a large rectangle all at once, it is useful to consider traversing 'strips' of a fixed width one at a time. There can potentially be dependence between the probability of crossing adjacent strips, but this can be dealt with by ignoring the infection configuration in a few columns. The following lemma gives the details.

Lemma 5.1.3. Fix $B > 1$, $Z > 0$, $p < p_0$ and $h \in [Z/p^2, B/p^2]$. For any $m, w \in \mathbb{N}$ with $w < m$ and any rectangle R of dimensions (m, h),

$$\mathbb{P}(R \text{ is horiz. trav.}) \leq (1 + pQ_1(B, Z, w))^{m/w+1} e^{-g(hq^2)m(1-11/w)}.$$

Proof. Fix $w < m$ and a rectangle R of dimension (m, h). Let $\ell \in \mathbb{N}$ and $0 \leq r < w$ be such that $m = \ell w + r$. Let R be any rectangle of dimension (h, m) and divide R into ℓ sub-rectangles, $R_1, R_2, \ldots, R_\ell$, each of height h and width w, with a remainder sub-rectangle of width r, denoted R_0.

For each $i = 0, 1, 2, \ldots, \ell$, it might not be the case that R_i is horizontally traversable by X since this event might depend on sites in adjacent sub-rectangles.

Since membership in the set X^+ depends only on the initial infection of sites within distance 3, at least it is true that the sub-rectangle of R_i obtained by deleting 3 columns from each side is horizontally traversable by $(X \cap R_i)^+$. Denote these sub-rectangles by $R'_0, R'_1, \ldots, R'_\ell$. Set

$$Q_1 = \max\{Q_1(B, Z, r-3), Q_1(B, Z, w-6)\}.$$

Applying Lemma 5.1.2 to the sub-rectangles $R'_0, R'_1, \ldots, R'_\ell$,

$$\begin{aligned}
&\mathbb{P}(R \text{ is horiz. trav. by } X^+) \\
&\quad \leq \prod_{i=0}^{\ell} \mathbb{P}(R'_i \text{ is horiz. trav. by } X^+) \\
&\quad \leq (1 + pQ_1)^\ell e^{-g(hq^2)(w-7)\ell} (1 + pQ_1) e^{-g(hq^2)(r-4)} \\
&\quad \leq (1 + pQ_1)^{\ell+1} e^{-g(hq^2)(w\ell+r-7\ell-4)}
\end{aligned}$$

$$\leq (1 + pQ_1)^{m/w+1} e^{-g(hq^2)(m-(7\ell+4))}$$

$$\leq (1 + pQ_1)^{m/w+1} e^{-g(hq^2)m(1-11/w)}$$

yielding the desired bound on the probability that R is traversable. □

In the analysis of the upper bound on the probability of percolation, rather than considering only the probability that rectangles are traversable, the probability that an infected rectangle grows into a larger infected rectangle because of the infected sites around it is also used.

Definition 5.1.4. For any two rectangles $R \subseteq R'$ and $X \sim \mathrm{Bin}(R', p)$, let $D(R, R')$ be the event that R' is internally spanned by $R \cup X$.

Essentially, this is the event that the four rectangles surrounding R in R' are traversable by the sites in $X \setminus R$. The following lemma shows that even though these events are not independent, they are nearly so.

Lemma 5.1.5. For every $B \geq 1$, $Z \geq 0$ and $c \in (0, 1/6)$, there exist $T \geq 0$ and $p_1 = p_1(Z, c)$ such that for all $p \leq p_1$ and all m, n, s and t with $Z/p^2 \leq m, n \leq B/p^2$, and $s, t \leq T/p^2$ if $R \subset R'$ are two rectangles with dimensions (m, n) and $(m + s, n + t)$, respectively, then

$$\mathbb{P}(D(R, R')) \leq 3(1 + pQ_1(B, Z, \lceil 11/c \rceil))^{\frac{12}{11}c(s+t)+4} e^{16g(Z)-(1-6c)(sg(nq^2)+tg(mq^2))}.$$

Proof. Fix $B > 1$, $Z > 0$, $c \in (0, 1/6)$, $p > 0$ and let R be a rectangle of dimension (m, n) and let R' be a rectangle of dimension $(m + s, n + t)$ with $R \subseteq R'$. Suppose, without loss of generality that $s \leq t$. Let $R' = [a_1, a_2] \times [b_1, b_2]$ and $R = [c_1, c_2] \times [d_1, d_2]$. The rectangle R' is decomposed into R together with the following 8 sub-rectangles, as in Figure 5.4,

$$R_1 = [a_1, c_1 - 1] \times [b_1, d_1 - 1] \qquad R_2 = [c_1, c_2] \times [b_1, d_1 - 1]$$

$$R_3 = [c_2 + 1, a_2] \times [b_1, d_1 - 1] \qquad R_4 = [c_2 + 1, a_2] \times [d_1, d_2]$$

$$R_5 = [c_2 + 1, a_2] \times [d_2 + 1, b_2] \qquad R_6 = [c_1, c_2] \times [d_2 + 1, b_2]$$

$$R_7 = [a_1, c_1 - 1] \times [d_2 + 1, b_2] \qquad R_8 = [a_1, c_1 - 1] \times [d_1, d_2].$$

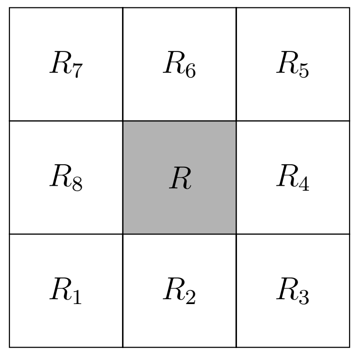

Figure 5.4: Decomposition of the rectangle R'

Let $X \sim \text{Bin}(R', p)$. If the event $D(R, R')$ occurs, then each of the rectangles $R_3 \cup R_4 \cup R_5$ and $R_7 \cup R_8 \cup R_1$ are horizontally traversable and each of the rectangles $R_5 \cup R_6 \cup R_7$ and $R_1 \cup R_2 \cup R_3$ are vertically traversable. The probability of each of these events individually can be approximated by Lemma 5.1.3, but these events are not independent. Conditioning on the infected sites in the corner rectangles, R_1, R_3, R_5, and R_7, it is possible to approximate the probability of these events by slightly different events that are independent of each other.

Set $Y = X^+ \cap (R_1 \cup R_3 \cup R_5 \cup R_7)$. Since $|R_1 \cup R_3 \cup R_5 \cup R_7| = st$, then $\mathbb{E}|Y| \leq st(8p^2 + 100p^3) \leq 9stp^2$ for $p \leq 1/100$.

The events that two particular sites are contained in X^+ are not independent, however, if $d(\mathbf{x}, \mathbf{y}) \geq 7$, then the events $\{\mathbf{x} \in X^+\}$ and $\{\mathbf{y} \in X^+\}$ are independent since they each depend on the initial infection of disjoint sets of sites.

The grid, $\mathbb{Z}^2$, can be decomposed into 25 disjoint sets $C_1, \ldots, C_{25}$ such that for each $i \in [1, 25]$ and $\mathbf{x}, \mathbf{y} \in C_i$, $d(\mathbf{x}, \mathbf{y}) \geq 7$. Indeed, set $B = B_5(\mathbf{0})$ and for each

$\mathbf{b} \in B$, define $C_{\mathbf{b}} = \{\mathbf{b} + x(4,3) + y(3,-4) : \ x, y \in \mathbb{Z}\}$. These sets $\{C_{\mathbf{b}} : \ \mathbf{b} \in B\}$ are disjoint, $|B| = 25$ and for any $\mathbf{x}, \mathbf{y} \in C_{\mathbf{b}}$, if $\mathbf{x} \neq \mathbf{y}$, then $d(\mathbf{x}, \mathbf{y}) \geq 7$ and hence the events $\{\mathbf{x} \in X^+\}$ and $\{\mathbf{y} \in X^+\}$ are independent.

Now, if $|Y| \geq cs$, then for some $\mathbf{b} \in B$, the expected number of sites in $C_{\mathbf{b}} \cap Y$ satisfies $|C_{\mathbf{b}} \cap Y| \geq cs/25$. For each $\mathbf{b} \in B$, $\mathbb{E}(|C_{\mathbf{b}} \cap Y|) \leq 9p^2 st/25$ and thus by Lemma 1.2.1, for $T \leq c/9$,

$$\begin{aligned}
\mathbb{P}\left(|C_{\mathbf{b}} \cap Y| \geq \frac{cs}{25}\right) &\leq \left(\frac{9p^2 st/25}{cs/25}\right)^{cs/25} \\
&= \left(\frac{9p^2 t}{c}\right)^{cs/25} \\
&\leq \left(\frac{9T}{c}\right)^{cs/25}.
\end{aligned}$$

Thus, $\mathbb{P}(|Y| \geq cs) \leq 25\left(\frac{9T}{c}\right)^{cs/25}$. Chose $T = T(c, Z) \leq \frac{c}{9}\left(\frac{1}{25}e^{-2(1-6c)g(Z))}\right)^{25/c}$. Since g is a decreasing function, $g(mq^2), g(nq^2) \leq g(Z)$ and hence

$$25\left(\frac{9T}{c}\right)^{cs/25} \leq e^{-2s(1-6c)g(Z)} \leq e^{-(1-6c)(sg(nq^2)+tg(mq^2))} \leq e^{-(1-6c)sg(nq^2)} \leq 1.$$

Similarly, for $s \leq T/p^2 \leq c/(9p^2)$,

$$\mathbb{P}(|Y| \geq ct) \leq 25\left(\frac{9T}{c}\right)^{ct/25} \leq e^{-2t(1-6c)g(Z)}.$$

Consider the probability of the event $D(R, R')$ conditioned on $|Y| \leq cs$. If every column of R' that contained sites of $X^+ \cap (R_1 \cup R_3 \cup R_5 \cup R_7)$ were removed, the rectangles R_4 and R_8 would be split into as most $cs + 2$ sub-rectangles of height n and total width at least $s - cs$.

If $D(R, R')$ occurs, then in particular, each of these sub-rectangles is horizontally traversable by the sites in X^+. However, the membership of sites in X^+ might depend on initially infected sites in the deleted columns or adjacent rectangles. In order to obtain a set of rectangles for which the events that each are horizontally traversable are independent, two further columns on either side of each sub-rectangle are removed. Since this might also depend on sites in Y, delete 2 further rows from the top and bottom of each sub-rectangle to ensure that the events are independent of the sites in Y. Let the sub-rectangles be of widths $s_1, s_2, \ldots, s_j$ and note that $\sum_{i=1}^{j} s_i \geq s - cs - 4(cs+2) = s(1-5c) - 8$. Set $w = \lceil 11/c \rceil$, let $Q_1 = Q_1(B, Z, w)$ and apply Lemma 5.1.3 using w for the widths of the strips. Then by the choice of w, and since g is decreasing,

$$\begin{aligned}
&\mathbb{P}(R_4 \text{ and } R_8 \text{ are horiz. trav.} \mid |Y| \leq cs) \\
&\quad \leq \prod_{i=1}^{j} (1+pQ_1)^{s_i/w+1} e^{-g((n-4)q^2)s_i(1-11/w)} \\
&\quad \leq (1+pQ_1)^{s/w+cs+2} e^{-(s(1-5c)-8)(1-11/w)g(nq^2)} \\
&\quad \leq (1+pQ_1)^{s/w+cs+2} e^{-g(nq^2)(s(1-6c)-8)} \\
&\quad \leq (1+pQ_1)^{\frac{12}{11}sc+2} e^{8g(Z)-g(nq^2)s(1-6c))}.
\end{aligned}$$

Similarly, conditioning on the event that $|Y| \leq ct$,

$$\mathbb{P}(R_2 \text{ and } R_6 \text{ are vert. trav.} \mid |Y| \leq ct) \leq (1+pQ_1)^{\frac{12}{11}ct+2} e^{8g(Z)-t(1-6c)g(mq^2)}.$$

Consider the event $D(R, R')$ conditioned on the following three possible ranges for the values of $|Y|$: $|Y| \leq cs$, $cs < |Y| \leq ct$, and $|Y| > ct$.

$$\begin{aligned}
&\mathbb{P}(D(R,R') \mid |Y| \leq cs)\mathbb{P}(|Y| \leq cs) \\
&\qquad \leq \mathbb{P}(D(R,R') \mid |Y| \leq cs) \\
&\qquad \leq (1+pQ_1)^{\frac{12}{11}c(s+t)+4} \exp\left(16g(Z) - (1-6c)(tg(mq^2) + sg(nq^2))\right), \\
&\mathbb{P}(D(R,R') \mid cs < |Y| \leq ct)\mathbb{P}(cs < |Y| \leq ct)
\end{aligned}$$

$$
\begin{aligned}
&\leq \mathbb{P}(D(R,R') \mid cs < |Y| \leq ct)\mathbb{P}(cs < |Y|) \\
&\leq (1+pQ_1)^{\frac{12}{11}ct+2} \exp\left(8g(Z) - t(1-6c)g(mq^2)\right) e^{-(1-6c)sg(nq^2)} \\
&\leq (1+pQ_1)^{\frac{12}{11}c(s+t)+4} \exp\left(16g(Z) - (1-6c)(tg(mq^2)+sg(nq^2))\right),
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{P}(D(R,R')|\ |Y| > ct)\mathbb{P}(|Y| > ct)& \\
&\leq \mathbb{P}(|Y| > ct) \\
&\leq e^{-(1-6c)(sg(nq^2)+tg(mq^2))} \\
&\leq (1+pQ_1)^{\frac{12}{11}c(s+t)+4} \exp\left(16g(Z) - (1-6c)(sg(nq^2)+tg(mq^2))\right).
\end{aligned}
$$

Combining these yields,

$$
\mathbb{P}(D(R,R')) \leq 3(1+pQ_1)^{\frac{12}{11}c(s+t)+4} \exp\left(16g(Z) - (1-6c)(sg(nq^2)+tg(mq^2))\right),
$$

the desired upper bound for the probability that the infection grows from the rectangle R to the rectangle R'. □

5.2 Hierarchies

As in the study of the usual bootstrap process, the notion of a 'hierarchy' is used to account for the different ways in which small internally spanned rectangles can either join together or grow into larger rectangles through the update process. The definitions and results in this section are similar to the notion of hierarchies in [24], though on a different scale with respect to the parameter p and with some small changes to deal with sites that could become uninfected.

Definition 5.2.1. A *hierarchy for a rectangle* R, is a pair $\mathcal{H} = (G_{\mathcal{H}}, \{R_u\}_{u \in V(G_{\mathcal{H}})})$, where $G_{\mathcal{H}}$ is a finite directed rooted tree with all edges directed away from the root and with maximum out-degree 3, together with a collection of rectangles $\{R_u\}_{u \in V(G_{\mathcal{H}})}$ such that

- if r is the root of $G_{\mathcal{H}}$, then $R_r = R$,
- if $u \to v$ in $G_{\mathcal{H}}$, then $R_u \supseteq R_v$
- if u has three children, then at least one child has as its corresponding rectangle a single site.
- if u has two or three children and at least one child v has $\text{short}(R_v) > 2$, then R_u is internally spanned by the rectangles corresponding to its children.

Vertices with out-degree 0 are called *seeds*, vertices with out-degree 1 are called *normal* and vertices with out-degree 2 or 3 are called *splitters*.

As in the analysis of usual bootstrap percolation, hierarchies are thought of as constructed 'bottom up' using initially infected sites: two rectangles are joined to create a 'parent' when their sites span a single larger rectangle. There is a slight modification to deal with the case when one of these rectangles is a single site. In this case, in order to remain consistent with the definition of X^+, a single site is only joined to another rectangle if the site is part of a triplet among the initially infected sites. In this case, the rectangles joined will be those that correspond to sites in the triplet.

Proposition 5.2.2. Let R be a rectangle, $X \subseteq R$ and set $X^* = X \cap X^+$. Suppose R is internally spanned by X. Then, there exists a hierarchy $\mathcal{H} = (G, \{R_u\}_{u \in V(G)})$ for R and $\{X_u\}_{u \in V(G)}$ with $X_u \subseteq X^* \cap R_u$ such that

- the root $r \in V(G)$ has $R_r = R$,
- the rectangles corresponding to the seeds of $\mathcal{H}$ are all the individual sites in X^*,
- every vertex that is not a seed has out degree at least 2,

- if u and w are both children of a vertex v, then $X_u \cap X_w = \emptyset$,
- if v is not a seed and has at least one child u with $\text{short}(R_u) > 2$, then R_v is internally spanned by the rectangles corresponding to its children.

Proof. Note that by the definition of X^+, every site in the set X^* is either part of a double or a triplet of sites in X and the only sites in $X \setminus X^*$ are those that do not contribute to the final infection of any other sites before they recover themselves. Thus, R is internally spanned by X iff R is internally spanned by X^*.

The hierarchy $\mathcal{H}$ can be constructed recursively. Let $R_1^0, R_2^0, \ldots, R_k^0$ be the individual sites in X^* and let these correspond to the seeds of the hierarchy $\mathcal{H}$. Given a partially constructed hierarchy $\mathcal{H}$, if there exist two vertices u and v with no parent so that $d(R_u, R_v) \leq 2$, add a new vertex to $G_{\mathcal{H}}$ by the following rules:

Case 1: If neither R_u nor R_v is a single site add a new vertex w as the parent of u and v with R_w the smallest rectangle that contains $R_u \cup R_v$ and set $X_w = X_u \cup X_v$.

Case 2: If $R_u = \mathbf{x}$ is a single site, then by the choice of X^*, the site $\mathbf{x}$ is part of either a double or a triplet. The sites that form either the double or triplet containing $\mathbf{x}$ might already be a part of another rectangle, but in either case, there is either another rectangle R_u' or two rectangles R_u' and R_u'' with no parents that contain the sites associated with the double or triplet containing $\mathbf{x}$. Add a new vertex w as in the previous case and join either the rectangle and the site or the two rectangles and the site.

This process continues until there are no more sites or rectangles that have yet to be joined. Since R is internally spanned by X^*, this process will stop only when the last remaining vertex with no parent is a root that corresponds to the rectangle R. The resulting directed graph and collection of rectangles have the desired properties, by induction. □

Definition 5.2.3. Given an initial infection X of R, the hierarchy $\mathcal{H}$ is said to *occur* (with respect to X) iff

- for every seed u, if the short side of R_u is the horizontal side then in every 4 adjacent columns in the the rectangle R_u, there are at least 2 initially infected sites within distance 2 (similarly for sets of 4 adjacent rows if the short side of R_u is vertical),
- for every normal vertex u with $u \to v \in E(G_{\mathcal{H}})$, the event $D(R_v, R_u)$ holds,

and these events occur disjointly.

For any rectangle R_u let $J(R_u)$ be the event, as above, that in every 4 adjacent columns, there are at least 2 initially infected sites within distance 2 if the short side of R_u is horizontal and similarly for set of 4 adjacent rows if the short side of R_u is vertical.

The condition that these events occur disjointly is included so that by the van den Berg-Kesten inequality (Lemma 1.2.3),

$$\mathbb{P}(\mathcal{H} \text{ occurs}) \leq \prod_{w \text{ seed}} \mathbb{P}(J(R_w)) \prod_{\substack{u \text{ normal} \\ u \to v}} \mathbb{P}(D(R_u, R_v)).$$

Note that the event that some hierarchy for a rectangle R occurs is not equivalent to the event that the rectangle R is internally spanned. Rectangles corresponding to seeds might have two initially infected sites within distance two in every set of 4 adjacent columns without being internally spanned. However, as long as the rectangles corresponding to seeds are not too large, the difference will be small. The definition is made in this way because, by Proposition 5.2.2, if R is internally spanned by X, then there is a hierarchy $\mathcal{H}$ for R that occurs. The number of these hierarchies might be too large compared to the probability that a particular hierarchy occurs to give reasonable estimates on the probability that

R is internally spanned. For this reason, it is useful to consider the following types of hierarchies where the difference in dimensions between parent and child rectangles are not arbitrarily small.

Definition 5.2.4. Given $Z > T > 0$ and $p > 0$, the hierarchy $\mathcal{H}$ is said to be *good* for Z, T, and p if the rectangles $\{R_u\}_{u \in V}$ satisfy the following additional conditions on their dimensions:

- if v is a seed, then $\text{short}(R_v) < 2Z/p^2$,
- if v is not a seed, then $\text{short}(R_v) \geq 2Z/p^2$,
- if u is normal with child v, then $\phi(R_u) - \phi(R_v) \leq T/p^2$
- if u is normal with $u \to v$ and v is also normal, then $\phi(R_u) - \phi(R_v) \geq \frac{T}{2p^2}$, and
- if u is a splitter and v is a child of u, then $\phi(R_u) - \phi(R_v) \geq \frac{T}{2p^2}$.

Next, it is shown that there exist hierarchies that are both good and occur for rectangles that are internally spanned.

Proposition 5.2.5. Let $Z > T > 0$, $p > 0$ and let R be a rectangle and let $X \subseteq R$. If R is internally spanned by X, then there exists a hierarchy $\mathcal{H}$ that is good for Z, T and p and that occurs.

Proof. The proof proceeds by induction on R. If $\text{short}(R) < 2Z/p^2$, then take $G_{\mathcal{H}}$ to be a single isolated vertex r and $R_r = R$. If R is internally spanned, then $\mathcal{H} = (G_{\mathcal{H}}, \{R_r\})$ is a good hierarchy that occurs.

Assume now that $\text{short}(R) \geq 2Z/p^2$. Then $\phi(R) \geq 4Z/p^2$. Construct a sequence $R \supseteq R_1 \supseteq \ldots$ from Proposition 5.2.2 going down the tree from the root, always talking R_i to be the largest rectangle. Let $m \geq 1$ be the smallest such that $\phi(R) - \phi(R_m) \geq \frac{T}{2p^2}$ and consider the following three cases.

Case 1: If $\frac{T}{2p^2} \leq \phi(R) - \phi(R_m) \leq \frac{T}{p^2}$, then let $\mathcal{H}' = (G', \{R_u\}_{u \in V'})$ be a good hierarchy rooted at r' that occurs for R_m. Let r be a new vertex and define a new hierarchy rooted at r with $R_r = R$ as follows. Set $G = (V' \cup \{r\}, E(G') \cup \{r \to r'\})$ and then $\mathcal{H} = (G, \{R_u\}_{u \in V(G)})$ is the desired hierarchy.

Case 2: If $\phi(R) - \phi(R_m) > T/p^2$ and $m = 1$, let R'_1 be the other rectangle from the tree in Proposition 5.2.2. Note that by construction, $\phi(R'_1) \leq \phi(R_1) \leq \phi(R) - \frac{T}{2p^2}$. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be good hierarchies that occur for R_1 and R'_1, respectively. Construct a good hierarchy for R by adding a new vertex r as the root, with edges joining it to the roots of the trees for $\mathcal{H}_1$ and $\mathcal{H}_1$.

Case 3: If $\phi(R) - \phi(R_m) > T/p^2$ and $m \geq 2$, let R'_m be the other rectangle contained in R_{m-1} from the tree in Proposition 5.2.2. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be good hierarchies that occur for R_m and R'_m respectively. For $i = 1, 2$, denote the root of $\mathcal{H}_i$ by r_i. Let r and u be two new vertices and set $R_r = R$ and $R_u = R_{m-1}$. Define a new hierarchy $\mathcal{H}$ with $G_{\mathcal{H}} = G_{\mathcal{H}_1} \cup G_{\mathcal{H}_2} \cup \{r \to u, u \to r_1, u \to r_2\}$, rooted at r. Since $\phi(R) - \phi(R_{m-1}) < \frac{T}{2p^2}$ and $\phi(R) - \phi(R_m) \geq T/p^2$ imply that

$$\phi(R_{m-1}) - \phi(R'_m) \geq \phi(R_{m-1}) - \phi(R_m) \geq \frac{T}{p^2} - \frac{T}{2p^2} = \frac{T}{2p^2}$$

then, $\mathcal{H}$ is a good hierarchy for T, Z and p. □

Good hierarchies are useful because there are not too many of them for rectangles of certain dimensions. Fix $B \geq 1$, $p > 0$ and let R be a rectangle with $\text{short}(R) \leq \text{long}(R) \leq B/p^2$. Let $Z, T > 0$ and let $\mathcal{H}$ be a hierarchy for R that is good for Z, T and p. By the definition of good hierarchies, for every directed path of length two in $G_{\mathcal{H}}$, $u \to v \to w$, the rectangles R_u and R_w satisfy

$\phi(R_u) - \phi(R_w) \geq \frac{T}{2p^2}$. Thus, the height of the tree $G_{\mathcal{H}}$ is at most

$$2\frac{2B/p^2}{T/(2p^2)} + 1 = \frac{8B}{T} + 1.$$

Since the out-degree of each vertex is at most 3, there are at most $3^{8B/T+2}$ vertices in $G_{\mathcal{H}}$. Set $M = M(B,T) = 3^{8B/T+2}$.

The number of rooted trees on M vertices is $M \cdot M^{M-2} = M^{M-1}$ and so there are at most M^{M-1} different rooted trees among all those belonging to a good hierarchy for R. Consider now the number of different collections of rectangles corresponding to hierarchies. In R, the number of different rectangles is

$$\binom{\text{long}(R)+1}{2}\binom{\text{short}(R)+1}{2} \leq \frac{(B/p^2+1)^4}{4} \leq \left(\frac{B}{p^2}\right)^4.$$

Thus, for any rooted tree G on at most M vertices, there are at most $(B/p^2)^{4M}$ different collections $\{R_u\}_{u\in V(G)}$ such that for each $u \in V(G)$, R_u is a rectangle contained in R. Therefore, in total, there are at most

$$M^{M-1}\left(\frac{B}{p^2}\right)^{4M} = M^{M-1}B^{4M}p^{-8M} \tag{5.4}$$

different good hierarchies for the rectangle R. While this number might be very large, it turns out to be small enough compared to the probability that a given hierarchy occurs to give a reasonable upper bound on the probability that the rectangle R is internally spanned.

5.3 Upper bound on $I(n,p)$

The following definitions and lemmas can be found in the paper by Holroyd [24]. Although, in that article, the function g is different, the proofs use only the properties that the function g is continuously differentiable, positive, decreasing and convex. The function g, given by equation (3.7), has these properties, by definition and by Fact 3.2.4.

Definition 5.3.1. Let $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = (b_1, b_2)$ with for $i = 1, 2$, $0 \leq a_i \leq b_i$. Define

$$W(\mathbf{a}, \mathbf{b}) = \inf \left\{ \int_\gamma g(y)\ dx + g(x)dy \mid \gamma : \mathbf{a} \to \mathbf{b} \text{ piecewise linear path} \right\}.$$

The function W and its properties are used to bound the term $\exp(tg(mq^2) + sg(nq^2))$ arising in Lemma 5.1.5. The following, Lemmas 5.3.2, 5.3.3, 5.3.4, 5.3.5, are from Holroyd [24] (Propositions 12, 13, 14, and 15).

Lemma 5.3.2. Let $\mathbf{a}, \mathbf{b}, \mathbf{c} \in (\mathbb{R}^+)^2$ with $\mathbf{a} \leq \mathbf{b} \leq \mathbf{c}$. Then

$$W(\mathbf{a}, \mathbf{b}) + W(\mathbf{b}, \mathbf{c}) \geq W(\mathbf{a}, \mathbf{c}).$$

Lemma 5.3.3. If $\mathbf{a} \leq \mathbf{b}$, then $W(\mathbf{a}, \mathbf{b}) \leq (b_1 - a_1)g(a_2) + (b_2 - a_2)g(a_1)$.

Lemma 5.3.4. If $\mathbf{a} = (a_1, a_2)$ with $a_1 + a_2 = A$ and $\mathbf{b} = (B, B)$ with $A \leq B$, then

$$W(\mathbf{a}, \mathbf{b}) \geq 2 \int_A^B g(x)\ dx.$$

Lemma 5.3.5. For every z, Z with $0 < z \leq Z$ and $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{r} \in (\mathbb{R}^+)^2$ with $\mathbf{a} \leq \mathbf{b}$, $\mathbf{c} \leq \mathbf{d}$, $\mathbf{r} \geq \mathbf{b}, \mathbf{d}, (2Z, 2Z)$ and $\mathbf{r} \leq \mathbf{b} + \mathbf{d} + (a, a)$, there exists $\mathbf{s} \in (\mathbb{R}^+)^2$ with $\mathbf{s} \leq \mathbf{r}$ and $\mathbf{s} \leq \mathbf{a} + \mathbf{c}$ such that

$$W(\mathbf{a}, \mathbf{b}) + W(\mathbf{c}, \mathbf{d}) \geq W(\mathbf{s}, \mathbf{r}) - 2zg(Z).$$

Definition 5.3.6. For rectangles $R \subseteq R'$, set

$$U(R, R') = W(q^2 \dim(R), q^2 \dim(R')).$$

This definition is useful since if R is a rectangle of dimension (m, n) and R' is a rectangle of dimension $(m+s, n+t)$ with $R \subseteq R'$, then by Lemma 5.3.3,

$$\frac{U(R, R')}{q^2} \leq sg(mq^2) + tg(nq^2).$$

The following lemma, adapted from a corresponding result in [24], shows that every hierarchy is associated with a rectangle called a 'pod' that can be used to bound the sum of the values $U(R_v, R_u)$ over all normal vertices in the hierarchy.

Lemma 5.3.7. Fix Z, T, q with $3q^2 < Z$, let $\mathcal{H}$ be a good hierarchy for the rectangle R with root r and let $N_s(\mathcal{H})$ be the number of vertices in $G_{\mathcal{H}}$ that are splitters. There exists a rectangle $S = S(\mathcal{H})$ with $S \subseteq R$ and

$$\dim(S) \leq \sum_{w \text{ seed}} \dim(R_w),$$

and

$$\sum_{\substack{u \to v \\ u \text{ normal}}} U(R_v, R_u) \geq U(S, R) - 6N_s(\mathcal{H})q^2 g(Z).$$

Proof. The proof proceeds by induction on the number of vertices in $G_{\mathcal{H}}$. If $|V(G_{\mathcal{H}})| = 1$, then take $S = R$.

If $|V(G_{\mathcal{H}})| > 1$, consider separately the cases where the root r is a normal vertex or a splitter. If r is a normal vertex with child u, let $\mathcal{H}'$ be the sub-hierarchy with root u and apply the induction hypothesis to $\mathcal{H}'$ to get a rectangle $S' \subseteq R_u$. The hierarchy $\mathcal{H}'$ has the same number of splitters as the hierarchy $\mathcal{H}$ and the same seeds. Thus $\dim(S') \leq \sum_{w \in V(G_{\mathcal{H}'}) \text{ seed}} \dim(R_w)$ and

$$\sum_{v \to w} U(R_w, R_v) \geq U(R_u, R) + U(S, R_u) - 6N_s(\mathcal{H}')q^2 g(Z)$$
$$\geq U(S, R) - 6N_s(\mathcal{H})q^2 g(Z). \qquad \text{(by Lemma 5.3.2)}$$

Suppose now that r is a splitter and let u and v be two children of r that correspond to the two largest rectangles among the children of r. If r has a third child that corresponds to a single site, it is disregarded. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be the two sub-hierarchies with roots u and v respectively. Then, since r is a splitter, $N_s(\mathcal{H}) = N_s(\mathcal{H}_1) + N_s(\mathcal{H}_2) + 1$. Also, $\dim(R_u) + \dim(R_v) \geq \dim(R) - (3, 3)$, accounting for the case when there is a third vertex that corresponds to a site in a triplet.

Let $S_1 \subseteq R_u$ and $S_2 \subseteq R_v$ be given by the induction hypothesis and for $i = 1, 2$ set $\mathbf{s}_i = q^2 \dim(S_i)$, $\mathbf{r}_1 = q^2 \dim(R_u)$, $\mathbf{r}_2 = q^2 \dim(R_v)$ and $\mathbf{r} = q^2 \dim(R)$. By Lemma 5.3.5, there exists $\mathbf{s} \leq \mathbf{r}$ with $\mathbf{s} \leq \mathbf{s}_1 + \mathbf{s}_2$ such that

$$W(\mathbf{s}_1, \mathbf{r}_1) + W(\mathbf{s}_2, \mathbf{r}_2) \geq W(\mathbf{s}, \mathbf{r}) - 2(3q^2)g(Z).$$

Let S be a rectangle in R of dimension $\frac{1}{q^2}\mathbf{s}$. Then

$$\begin{aligned} \dim(S) &\leq \dim(S_1) + \dim(S_2) \\ &\leq \sum_{w \text{ seed in } \mathcal{H}_1} \dim(R_w) + \sum_{w \text{ seed in } \mathcal{H}_2} \dim(R_w) \\ &= \sum_{w \text{ seed in } \mathcal{H}} \dim(R_w). \end{aligned}$$

Also, by the choice of $\mathbf{s}$, $U(S_1, R_1) + U(S_2, R_2) \geq U(S, R) - 6q^2 g(Z)$ and by the choice of S_1 and S_2,

$$\begin{aligned}
&\sum_{\substack{x,y\in V(\mathcal{H})\\ x\to y}} U(R_y, R_x)\\
&\qquad = \sum_{\substack{x,y\in V(\mathcal{H}_1)\\ x\to y}} U(R_y, R_x) + \sum_{\substack{x,y\in V(\mathcal{H}_1)\\ x\to y}} U(R_y, R_x)\\
&\qquad \geq U(S_1, R_u) - 6q^2 N_s(\mathcal{H}_1)g(Z) + U(S_2, R_v) - 6q^2 N_s(\mathcal{H}_1)g(Z)\\
&\qquad \geq U(S, R) - 6q^2 g(Z) - 6q^2 N_s(\mathcal{H}_1)g(Z) - 6q^2 N_s(\mathcal{H}_2)g(Z)\\
&\qquad = U(S, R) - (N_s(\mathcal{H}_1) + N_s(\mathcal{H}_2) + 1)6q^2 g(Z)\\
&\qquad = U(S, R) - 6q^2 N_s(\mathcal{H})g(Z).
\end{aligned}$$

By induction, the result holds for all good hierarchies, $\mathcal{H}$. □

Using the notion of pods, the following upper bound is given on the probability that squares of a particular size are internally spanned. Recall that in Section 3.2, for $B \geq 1$ the value λ_B was defined to be $\lambda_B = \int_{1/B}^{B} g(x)\,dx \leq \lambda$.

Theorem 5.3.8. For every $\varepsilon > 0$, there is a $B_0 = B_0(\varepsilon) > 0$ such that for $B > B_0$ there exists $p_0 = p_0(\varepsilon, B)$ such that if $0 < p < p_0$ then

$$I(\lfloor B/p^2 \rfloor, p) \leq \exp\left(\frac{\varepsilon - 2(1 - 7/B)\lambda_B}{p^2}\right).$$

Proof. Fix $\varepsilon > 0$, $B > 1$ and $p > 0$. Set $L = \lfloor B/p^2 \rfloor$, $c = 1/B$, fix $Z > 0$ and let T be given by Lemma 5.1.5. By the van den Berg-Kesten inequality (Lemma 1.2.3), for any hierarchy $\mathcal{H}$ that is good for Z, T and p with respect to $[L]^2$,

$$\mathbb{P}(\mathcal{H} \text{ occurs}) \leq \prod_{w \text{ seed}} \mathbb{P}(J(R_w)) \prod_{\substack{u \text{ normal}\\ u\to v}} \mathbb{P}(D(R_u, R_v)).$$

Consider first the terms $\mathbb{P}(D(R_v, R_u))$. Set $Q_1 = Q_1(B, Z, \lceil 11/c \rceil)$. By

Lemma 5.1.5, for a normal vertex u and $u \to v$ with $\dim(R_u) = (m+s, n+t)$ and $\dim(R_v) = (m, n)$,

$$\begin{aligned}
&\mathbb{P}(D(R_v, R_u)) \\
&\qquad \leq 3(1+pQ_1)^{\frac{12}{11}c(s+t)+4} e^{16g(Z)-(1-6c)(sg(nq^2)+tg(mq^2))} \\
&\qquad \leq 3(1+pQ_1)^{\frac{24}{11}cT/p^2+4} \exp\left(16g(Z) - (1-6c)\frac{U(R_v, R_u)}{q^2}\right) \\
&\qquad \leq \exp\left(\log 3 + \left(\frac{24cT}{11p^2} + 4\right) pQ_1 + 16g(Z) - (1-6c)\frac{U(R_v, R_u)}{q^2}\right) \\
&\qquad \leq \exp\left(\frac{(\log 3 + (24cT/11+4)Q_1 + 16g(Z))}{p} - \frac{(1-6c)U(R_v, R_u)}{q^2}\right).
\end{aligned}$$

Set $Q_2 = Q_2(B, Z) = \log 3 + (24cT/11 + 4)Q_1 + 16g(Z)$. Then, Q_2 is a constant that depends only B, Z since T and c depend only on B and Z and

$$\mathbb{P}(D(R_v, R_u)) \leq \exp(Q_2/p) \exp\left(-(1-6c)\frac{U(R_u, R_v)}{q^2}\right). \tag{5.5}$$

Let $N_1(\mathcal{H})$ be the number of normal vertices in $G_{\mathcal{H}}$ and let $N_0(\mathcal{H})$ be the number of seeds. Recall that the number of vertices in the hierarchy $\mathcal{H}$ is at most $M = 3^{8B/T+2}$ and so $N_1(\mathcal{H})$ and $N_0(\mathcal{H})$ are both at most a constant that depends only on B and Z, since T depends on B and Z. Let S be a pod rectangle for $\mathcal{H}$ given by Lemma 5.3.7. Then, $\dim(S) \leq \sum_{w \text{ seed}} \dim(R_w)$ and

$$\sum_{\substack{u \text{ normal} \\ u \to v}} U(R_v, R_u) \geq U(S, R) - 6N_s(\mathcal{H})q^2 g(Z).$$

Combining this with inequality (5.5),

$$\prod_{\substack{u \text{ normal} \\ u \to v}} \mathbb{P}(D(R_u, R_v)) \leq \prod_{\substack{u \text{ normal} \\ u \to v}} \exp\left(\frac{Q_2}{p} - \frac{(1-6c)U(R_u, R_v)}{q^2}\right)$$

$$
\begin{aligned}
&= \exp\left(\frac{N_1(\mathcal{H})Q_2}{p} - \frac{(1-6c)}{q^2}\sum_{\substack{u \text{ normal} \\ u\to v}} U(R_u, R_v)\right) \\
&\le \exp\left(\frac{N_1(\mathcal{H})Q_2}{p} - \frac{(1-6c)}{q^2}\left(U(S,R) - 6N_s(\mathcal{H})q^2 g(Z)\right)\right) \\
&\le \exp\left(\frac{N_1(\mathcal{H})Q_2 p + 6N_s(\mathcal{H})g(Z)q^2}{p^2} - (1-6c)\frac{U(S,R)}{q^2}\right).
\end{aligned}
$$

Let p be small enough so that $N_1(\mathcal{H})Q_2 p + 6N_s(\mathcal{H})g(Z)q^2 \le \varepsilon/3$, then

$$
\prod_{\substack{u \text{ normal} \\ u\to v}} \mathbb{P}(D(R_u, R_v)) \le \exp\left(\frac{\varepsilon/3}{p^2} - (1-6c)\frac{U(S,R)}{q^2}\right). \tag{5.6}
$$

To estimate the probability of the events $J(R_w)$, suppose without loss of generality that the short side of R_w is horizontal and consider $\lfloor \text{long}(R_w)/4 \rfloor$ disjoint sets of 4 adjacent columns in R_w. In one set of 4 adjacent columns, there are at most 32 short$(R_w) \le 32Z/p^2$ pairs of sites within distance 2. The probability that at least one of these pairs are both initially infected is at most $32Z$ and hence

$$
\mathbb{P}(J(R_w)) \le (32Z)^{\lfloor \text{long}(R_w)/4 \rfloor} \le (32Z)^{\phi(R_w)/8-1}.
$$

Thus, using the fact that $\dim(S) \le \sum_{w \text{ seed}} \dim(R_w)$,

$$
\begin{aligned}
\prod_{w \text{ seed}} \mathbb{P}(J(R_w)) &\le (32Z)^{\sum_{w \text{ seed}} \phi(R_w)/8 - 3/4} \\
&\le (32Z)^{\phi(S)/8 - N_0(\mathcal{H})} \\
&= \exp\left(\frac{\log(32Z)\phi(S)}{8} - \log(32Z)N_0(\mathcal{H})\right).
\end{aligned}
$$

Let p be small enough so that $-\log(32Z)N_0(\mathcal{H}) \leq \frac{\varepsilon}{3p^2}$. Then

$$\prod_{w \text{ seed}} \mathbb{P}(J(R_w)) \leq \exp\left(\frac{\log(32Z)\phi(S)}{8} + \frac{\varepsilon/3}{p^2}\right)$$

and so

$$\begin{aligned}\mathbb{P}(\mathcal{H} \text{ occurs}) &\leq \exp\left(\frac{\log(32Z)\phi(S)}{8} + \frac{\varepsilon/3}{p^2} + \frac{\varepsilon/3}{p^2} - (1-6c)\frac{U(S,R)}{q^2}\right) \\ &= \exp\left(\frac{2\varepsilon/3}{p^2} + \frac{\log(32Z)\phi(S)}{8} - \frac{(1-6c)U(S,R)}{q^2}\right).\end{aligned}$$

Consider two different cases, depending on the size of the semi-perimeter of the rectangle S.

Case 1: If $\phi(S) \leq \frac{1}{Bq^2}$, then applying Lemma 5.3.4 with $q^2\phi(S) = A \leq 1/B$,

$$U(S,R) = W(q^2 \dim(S), q^2 \dim(R)) \geq 2\int_{1/B}^{B} g(x)\ dx = 2\lambda_B$$

and so

$$\exp\left(-\frac{(1-6c)}{q^2}U(S,R)\right) \leq \exp\left(-\frac{2(1-6c)}{q^2}\lambda_B\right).$$

Then,

$$\begin{aligned}\mathbb{P}(\mathcal{H} \text{ occurs}) &\leq \exp\left(\frac{2\varepsilon/3}{p^2} - \frac{(1-6c)U(S,R)}{q^2}\right) \\ &\leq \exp\left(\frac{2\varepsilon/3}{p^2} - \frac{2(1-6c)\lambda_B}{q^2}\right).\end{aligned}$$

Case 2: If, on the other hand, $\phi(S) > \frac{1}{Bq^2}$, then

$$\frac{-\log(32Z)\phi(S)}{8} \geq \frac{-\log(32Z)}{8Bq^2}.$$

Choose $Z > 0$ to be small enough so that

$$\frac{-\log(32Z)}{8B} \geq 2\lambda \geq 2(1-6c)\lambda_B.$$

Then,

$$\begin{aligned}
\mathbb{P}(\mathcal{H} \text{ occurs}) &\leq \exp\left(\frac{2\varepsilon/3}{p^2} + \frac{\log(32Z)\phi(S)}{8}\right) \\
&\leq \exp\left(\frac{2\varepsilon/3}{p^2} - \frac{2(1-6c)\lambda_B}{q^2}\right).
\end{aligned}$$

Finally, recall that for $M = 3^{8B/T+2}$, there are at most $M^{M-1}B^{4M}p^{-8M}$ different good hierarchies $\mathcal{H}$. Let p be small enough so that $M^{M-1}B^{4M}p^{-8M} \leq \exp\left(\frac{\varepsilon/3}{p^2}\right)$. Then,

$$\begin{aligned}
I(L,p) &\leq \sum_{\mathcal{H}} \mathbb{P}(\mathcal{H} \text{ occurs}) \\
&\leq M^{M-1}B^{4M}p^{-8M} \exp\left(\frac{2\varepsilon/3}{p^2} - \frac{2(1-6c)\lambda_B}{q^2}\right) \\
&= \exp\left(\frac{\varepsilon/3}{p^2} + \frac{2\varepsilon/3}{p^2} - \frac{2(1-6c)\lambda_B}{q^2}\right) \\
&= \exp\left(\frac{\varepsilon}{p^2} - \frac{2(1-6c)\lambda_B}{q^2}\right).
\end{aligned}$$

Let p be small enough so that $\frac{q^2}{p^2} \leq \frac{1-6c}{1-7c}$. Then,

$$I(L,p) \leq \exp\left(\frac{\varepsilon - 2(1-7c)\lambda_B}{p^2}\right).$$

□

In order to extend Theorem 5.3.8 to give an upper bound on the probability that an arbitrarily large rectangle percolates, the following lemma is used. If a large rectangle R is internally spanned, it might not be possible to guarantee that

R will contain internally spanned squares of a particular scale, but the following shows that it is at least possible to guarantee the existence of internally spanned rectangles of a particular scale. Lemma 5.3.9 is an immediate analogue to a result on usual bootstrap percolation given in [2].

Lemma 5.3.9. Fix a rectangle R, $k \in \mathbb{N}$ with $\mathrm{long}(R) \geq 2k$, and $X_0 \subseteq R$. If R is internally spanned by X_0, then there exists a rectangle $T \subseteq R$ with $\mathrm{long}(T) \in [k, 2k]$ that is internally spanned by X_0.

While Theorem 5.3.8 gives an upper bound on the probability of percolation for any large enough rectangle and small enough probability of initial infection, it remains to show how this can be used to give a bound on the critical probability.

Theorem 5.3.10. For every $\varepsilon > 0$, there exists $n_0 = n(\varepsilon)$ such that for every $n \geq n_0$, if $p < 0$ is such that $p \leq \sqrt{\frac{\lambda-\varepsilon}{\log n}}$, then

$$I(n,p) \leq n^{-\frac{\varepsilon}{2(\lambda-\varepsilon)}}.$$

Proof. Fix $\varepsilon > 0$ and let $B = B(\varepsilon)$ and $p_0 = p_0(\varepsilon)$ be given by Theorem 5.3.8 and with B large enough so that $\lambda - (1 - 6/B)\lambda_B < \varepsilon/12$. Let $n_0 = n_0(B, \varepsilon)$ be large enough so that $\sqrt{\frac{\lambda-\varepsilon}{\log n_0}} < p_0$ and if $n \geq n_0$, then $n \geq \frac{B \log n}{\lambda-\varepsilon}$.

Fix $n > n_0$ and $p > 0$ with $p \leq \sqrt{\frac{\lambda-\varepsilon}{\log n}}$. Note that if $p \leq p'$, then by coupling, $I(n,p) \leq I(n,p')$ and so it suffices to prove the result assuming that $p = \sqrt{\frac{\lambda-\varepsilon}{\log n}}$. By the choice of n_0, $n > \frac{B \log n}{\lambda-\varepsilon} = B/p^2$. Set $R = [n]^2$.

Set $K = \lfloor B/p^2 \rfloor$ and $k = \lfloor B/2p^2 \rfloor$ so that $2k \leq K < n$. By Lemma 5.3.9, if R is internally spanned, then there is an internally spanned rectangle $T \subseteq R$ with $\mathrm{long}(T) \in [k, 2k]$. Thus,

$$I(n,p) \leq \sum_{\substack{T \subseteq R \\ \mathrm{long}(T) \in [k,2k]}} I(T,p).$$

In R, there are at most $n^2(2k)^2 \leq n^2K^2$ such rectangles T. By the choice of K, and for n sufficiently large,

$$K^2 \leq \frac{B^2}{p^4} = \frac{B^2(\log n)^2}{(\lambda-\varepsilon)^2} \leq n^{\frac{\varepsilon}{6(\lambda-\varepsilon)}}.$$

It remains to determine an upper bound on the probability of such a rectangle being internally spanned. Fix such a rectangle T of dimension (a, b) and suppose without loss of generality that $a \leq b$ and that $T = [1, a] \times [1, b]$. Consider one particular way in which the rectangle $[1, K]^2$ can be internally spanned. The rectangle $[K]^2$ is internally spanned if T is internally spanned and every column of the rectangle $[a+1, K] \times [K]$ contains two adjacent initially infected sites and every row of the rectangle $[a] \times [b+1, K]$ contains two adjacent initially infected sites. Since these events are all independent,

$$\begin{aligned}
I(K,p) &\geq I(T,p)(1-(1-p^2)^{\lfloor K/2 \rfloor})^{K-b}(1-(1-p^2)^{\lfloor b/2 \rfloor})^{K-a} \\
&\geq I(T,p)(1-e^{-p^2(K-1)/2})^K(1-e^{-p^2(k-1)/2})^K \\
&\geq I(T,p)(1-e^{-p^2(k-1)/2})^{2K} \\
&\geq I(T,p)(1-e^{-(B/4-1)})^{2K} \\
&\geq I(T,p)\exp(-4Ke^{-B/4+1}) \qquad \text{(for } B \geq 5) \\
&\geq I(T,p)\exp\left(\frac{-4Be^{-B/4+1}}{p^2}\right).
\end{aligned}$$

Hence for any $T \subseteq R$ with $\text{long}(T) \in [k, 2k]$, by Theorem 5.3.8 applied to $[K]^2$,

$$I(T,p) \leq \exp\left(\frac{4Be^{-B/4+1}}{p^2} + \frac{\varepsilon - 2(1-7/B)\lambda_B}{p^2}\right).$$

Let B be large enough so that $4Be^{-B/4+1} \leq \varepsilon/6$. Since $(1 - 7/B)\lambda_B > \lambda - \frac{\varepsilon}{12}$,

$$\begin{aligned} I(T, p) &\leq \exp\left(\frac{\varepsilon/6 + \varepsilon - 2(\lambda - \varepsilon/(12))}{p^2}\right) \\ &= \exp\left(\frac{4\varepsilon/3 - 2\lambda}{p^2}\right) \\ &= \exp\left(\frac{(4\varepsilon/3 - 2\lambda)\log n}{\lambda - \varepsilon}\right) \\ &= n^{\frac{-(2\lambda - 4\varepsilon/3)}{\lambda - \varepsilon}} = n^{-2 - \frac{2\varepsilon/3}{\lambda - \varepsilon}}. \end{aligned}$$

Therefore, the probability that $[n]^2$ is internally spanned can be bounded above as

$$\begin{aligned} I(n, p) &\leq n^2 K^2 n^{-(2+2\varepsilon/3)} \\ &\leq n^2 n^{\frac{\varepsilon}{6(\lambda - \varepsilon)}} n^{-2 - \frac{2\varepsilon/3}{\lambda - \varepsilon}} \\ &= n^{\frac{-\varepsilon}{2(\lambda - \varepsilon)}} \end{aligned}$$

as claimed. □

In particular, by Theorem 5.3.10 for each $\varepsilon > 0$ and sequence $\{p(n)\}_{n \in \mathbb{N}}$ with the property that for each $n \in \mathbb{N}$,

$$p(n) \leq \sqrt{\frac{\lambda - \varepsilon}{\log n}}$$

then $I(n, p(n)) = o(1)$.

This implies that for every $\varepsilon > 0$, there is an n_ε such that for all $n \geq n_\varepsilon$,

$$p_c([n]^2, \mathcal{M}) \geq \sqrt{\frac{\lambda - \varepsilon}{\log n}}.$$

Combining Theorems 5.3.10 and 4.2.3, this shows that the critical probability for the update rule $\mathcal{M}$ satisfies

$$p_c([n]^2, \mathcal{M}) = \sqrt{\frac{\lambda + o(1)}{\log n}}.$$

A remaining open problem would be to determine a more exact expression for the critical probability $p_c([n]^2, \mathcal{M})$ following the results of Gravner and Holroyd [20] and Gravner, Holroyd and Morris [21] for the critical probability for usual bootstrap percolation.

Recently, Balogh, Bollobás and Morris [5, 6] gave sharp thresholds for bootstrap processes in grids of any dimension. It would also be of interest to consider the effect of a modification of the bootstrap update rules in higher dimensions to allow for the possibility of recovery.

Bibliography

[Note:] Numbers following each bibliography item indicate where that item is cited in the text.

[1] M. Aigner, E. Triesch, and Z. Tuza, Irregular assignments and vertex-distinguishing edge-colorings of graphs, in *Combinatorics '90, Ann. Discrete Math.* **52**, North-Holland, Amsterdam, 1992, pp 1–9. 2, 12

[2] M. Aizenman and J. L. Lebowitz, Metastability effects in bootstrap percolation, *J. Phys. A* **21** (1988), 3801–3813. 4, 36, 97

[3] N. Alon and J. H. Spencer, *The probabilistic method*, Second edition, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. 6

[4] P. N. Balister, E. Győri, J. Lehel, and R. H. Schelp, Adjacent vertex distinguishing edge-colorings, *SIAM J. Discrete Math.* **21** (2007), 237–250. 12

[5] J. Balogh, B. Bollobás and R. Morris, Bootstrap percolation in high dimensions, arXiv:0907.3097v3 [math.PR] (2009). 100

[6] J. Balogh, B. Bollobás and R. Morris, Bootstrap percolation in three dimensions, *Ann. Probab.* **37** (2009), 1329–1380. 100

[7] M. Behzad, *Graphs and their chromatic numbers*, Diss., Michigan State University, 1965. 14

[8] J. van den Berg and H. Kesten, Inequalities with applications to percolation and reliability, *J. Appl. Probab.* **22** (1985), 556–569. 2, 7

[9] B. Bollobás, *Random graphs*, Second edition, Cambridge Studies in Advanced Mathematics **73**, Cambridge University Press, Cambridge, 2001. 9

[10] B. Bollobás and O. Riordan, *Percolation*, Cambridge University Press, New York, 2006. 7

[11] A. C. Burris and R. H. Schelp, Vertex-distinguishing proper edge-colorings, *J. Graph Theory* **26** (1997), 73–82. 2, 12

[12] J. Černý, M. Horňák, and R. Soták, Observability of a graph, *Math. Slovaca* **46** (1996), 21–31. 2, 12

[13] J. Chalupa, P. L. Leath and G. R. Reich, Bootstrap percolation on a Bethe lattice, *J. Phys. C: Solid State Phys.* **12** (1979), L31. 4, 35

[14] X. Chen, On the adjacent vertex distinguishing total coloring numbers of graphs with $\Delta = 3$, *Discrete Math.* **308** (2008), 4003–4007. 15

[15] T. Coker and K. Johannson, The adjacent vertex distinguishing total chromatic number, submitted (2010), available at arXiv:1009.1785v2. 15

[16] T. Coker and K. Johannson, A sharp threshold for a modified bootstrap percolation with recovery, in preparation (2011). 38

[17] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53** (1947), 292–294. 1

[18] P. Erdős, and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, *Infinite and finite sets (Colloq., Keszthely, 1973), Vol. II*, 609–627. 2, 9

[19] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* **6** (1959), 290–297. 1

[20] J. Gravner and A. E. Holroyd, Slow convergence in bootstrap percolation, *Ann. Appl. Probab.* **18** (2008), 909-928. 36, 100

[21] J. Gravner, A. E. Holroyd and R. Morris, A sharper threshold for bootstrap percolation in two dimensions, arXiv:1002.3881v1 [math.PR] (2010). 36, 100

[22] T. E. Harris, A lower bound of the critical probability in a certain percolations process, *Proc. Cambridge Philos. Soc.* **56** (1960), 13–20. 2, 7

[23] H. Hatami, $\Delta + 300$ is a bound on the adjacent vertex distinguishing edge chromatic number, *J. Combin. Theory Ser. B* **95** (2005), 246–256. 3, 12, 15

[24] A. Holroyd, Sharp metastability threshold for two-dimensional bootstrap percolation, *Probab. Theory Related Fields* **125** (2003), 195–224. 4, 5, 36, 43, 50, 64, 82, 88, 89, 90

[25] J. Hulgan, Concise proofs for adjacent vertex-distinguishing total colorings, *Discrete Math.* **309** (2009), 2548–2550. 15

[26] J. Hulgan, *Graph Colorings with Constraints*, Diss. University of Memphis, 2010. 15

[27] S. Janson, Poisson approximations for large deviations, *Random Structures Algorithms*, **1** (1990), 221–229. 2, 9

[28] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience, New York, 2000. 9

[29] X. S. Liu, M. Q. An, and Y. Gao, An upper bound for the adjacent vertex-distinguishing total chromatic number of a graph, *J. Math. Res. Exposition* **29** (2009), 343–348. 15

[30] C. McDiarmid, and B. Reed, Concentration for self-bounding functions and an inequality of Talagrand, *Random Structures Algorithms* **29** (2006), 549–557. 8

[31] M. Molloy, and B. Reed, A bound on the total chromatic number, *Combinatorica* **18** (1998), 241–280. 14, 16

[32] J. von Neumann, *Theory of Self-Reproducing Automata*, Univ. Illinois Press, Urbana, 1966. 4

[33] D. Reimer, Proof of the van den Berg-Kesten conjecture, *Combin. Probab. Comput.* **9** (2000), 27-32. 7

[34] M. Talagrand, Concentration of measure and isoperimetric inequalities in product spaces, *Inst. Hautes Études Sci. Publ. Math.* **81** (1995), 73–205. 2, 8, 75

[35] M. Talagrand, New concentration inequalities in product spaces, *Invent. Math.* **126** (1996), 505–563. 8

[36] S. Ulam, Random processes and transformations, *Proc. Internat. Congr. Math.* (1950), 264–275. 4

[37] V. Vizing, Some unsolved problems in graph theory, *Russ. Math. Surv.* **23** (1968), 125–141. 14

[38] H. Wang, On the adjacent vertex-distinguishing total chromatic numbers of the graphs with $\Delta(G) = 3$, *J. Comb. Optim.* **14** (2007), 87–109. 15

[39] Z. Zhang, X. Chen, J. Li, B. Yao, X. Lu, and J. Wang, On adjacent-vertex-distinguishing total coloring of graphs, *Sci. China Ser. A* **48** (2005), 289–299. 2, 3, 13, 14

[40] Z. Zhang, L. Liu, and J. Wang, Adjacent strong edge coloring of graphs, *Appl. Math. Lett.* **15** (2002), 623–626. 2, 12