Problems in Combinatorial and Analytic Number Theory

by

John Charles Saunders

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Pure Mathematics

Waterloo, Ontario, Canada, 2018

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Karl Dilcher Professor, Dept. of Mathematics & Statistics, Dalhousie University
Supervisor(s):	Kevin Hare Professor, Dept. of Pure Mathematics, University of Waterloo Yu-Ru Liu Professor, Dept. of Pure Mathematics, University of Waterloo
Internal Members:	Cam Stewart Professor, Dept. of Pure Mathematics, University of Waterloo David McKinnon Professor, Dept. of Pure Mathematics, University of Waterloo
Internal-External Member:	Jeffrey Shallit Professor, School of Computer Science, University of Waterloo

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

We focus on three problems in number theory.

The first problem studies the random Fibonacci tree, which is an infinite binary tree with non-negative integers at each node. The root consists of the number 1 with a single child, also the number 1. We define the tree recursively in the following way: if x is the parent of y, then y has two children, namely $|x - y|$ and $x + y$. This tree was studied by Benoit Rittaud [15] who proved that any pair of integers a, b that are coprime occur as a parent-child pair infinitely often. We extend his results by determining the probability that a random infinite walk in this tree contains exactly one pair $(1, 1)$, that being at the root of the tree. Also, we give tight upper and lower bounds on the number of occurrences of any specific coprime pair (a, b) at any given fixed depth in the tree.

The second problem studies sieve methods in combinatorics. We apply the Turán sieve and the simple sieve developed by Ram Murty and Yu-Ru Liu [12] to study problems in random graph theory. More specifically, we obtain bounds on the probability of a graph having diameter 2 (or diameter 3 in the case of bipartite graphs). An interesting feature revealed in these results is that the Turán sieve and the simple sieve "almost completely" complement each other.

The third problem studies the Mahler measure of a polynomial with integer coefficients. We give a lower bound of the Mahler measure on a set of polynomials that are "almost" reciprocal. Here "almost" reciprocal means that the outermost coefficients of each polynomial mirror each other in proportion, while this pattern breaks down for the innermost coefficients.

Acknowledgements

There are many people who I would like to thank for having a part in my producing this thesis. First, I would like to thank my two supervisors Dr. Kevin Hare and Dr. Yu-Ru Liu for their help, guidance, and encouragement in producing the results shown here and through my PhD degree as a whole. I would also like to thank Dr. David McKinnon in the Department of Pure Mathematics at the University of Waterloo for agreeing to be an examiner in both of my PhD oral comprehensive exams. I would also like to thank Dr. Cameron Stewart from the Department of Pure Mathematics at the University of Waterloo, Dr. Jeffrey Shallit from the School of Computer Science at the University of Waterloo and for Dr. Karl Dilcher who came all the way from Dalhousie University to be on my Graduate Committee. I would also like to thank the Department of Pure Mathematics at the University of Waterloo as a whole for accepting me into the PhD program and seeing me through to the end. Finally, I would also like to thank all of my friends, colleagues, and family in both the academic and non-academic realms for their continued support and encouragement as well. It's been a real journey!

Table of Contents

List of Figures

Chapter 1

Introduction

In this chapter, we introduce the areas of number theory in which the problems under consideration reside. We first introduce the random Fibonacci sequence, how it is portrayed in the random Fibonacci tree, and a combinatorial problem involving pairs of integers in the tree. We then introduce the Turán and simple sieves to study problems in combinatorics, especially in the area of random graph theory. Finally, we introduce the Mahler measure of a polynomial with integer coefficients and the problem of finding bounds for the Mahler measure of classes of polynomials.

1.1 Random Fibonacci Sequences

We start with the concept of the Fibonacci sequence.

Definition 1.1.1. The *Fibonacci sequence* $\{F_n\}$ is the sequence of integers defined as $F_1 = F_2 = 1$ with $F_n = F_{n-1} + F_{n-2}$ for $n \geq 3$.

The first few terms of the Fibonacci sequence are thus $1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots$ The Fibonacci sequence has many interesting properties associated with it. For instance, by induction, one can find a closed form for the nth term of the sequence. Such a closed form is known as *Binet's formula* and is

$$F_n = \frac{\phi^n - (-\phi)^{-n}}{\sqrt{5}}$$

where $\phi = \frac{1+\sqrt{5}}{2}$, which is known as the *golden ratio*. A corollary to Binet's formula is that

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \lim_{n\to\infty} F_n^{1/n} = \phi.$$

In 1960, Hillel Furstenberg and Harry Kersten [5] combined probability theory with the Fibonacci sequence and studied *random Fibonacci sequences*. These are sequences with the recursion $x_n = |x_{n-1} \pm x_{n-2}|$, where $\pm$ is chosen to be either $+$ or $-$ independently and randomly at each step. For example, we could have initial terms $1, 1$ and a $\pm$ pattern being $+, -, -, +, -, +, +, +$, which gives rise to the start of a random Fibonacci sequence:

$$1, 1, 2, 1, 1, 2, 1, 3, 4, 7.$$

Furstenberg and Kersten studied how such a sequence behaved. Intuitively, one would expect that such a random Fibonacci sequence would be bounded, since we would roughly have an equal number of subtractions and additions, and the subtractions would cancel the additions. Furstenberg and Kersten discovered that not only is this incorrect, but that a random Fibonacci sequence would almost surely grow exponentially at a constant rate. Divakar Viswanath [22] in 1999 computed this rate to be approximately $1.13198824\ldots$. In other words, if f_n is a random Fibonacci sequence, then with probability 1 we have

$$\lim_{n\to\infty} f_n^{1/n} = 1.13198824\ldots$$

Viswanath and others after him, however, have not been able to discover a closed or analytical form for this constant, which has become known as Viswanath's constant. It is not even known if it is irrational, let alone transcendental, although it is conjectured to be at least irrational.

In 2006, Jeffrey McGowan and Eran Makover [13] used the formalism of trees to give a simpler proof of Viswanath's result to evaluate the growth of the average value of the nth term. More precisely, they proved that

$$1.12095 \leq E(|t_n|)^{1/n} \leq 1.23375,$$

where $E(|t_n|)$ is the expected value of the nth term of the sequence. In 2007, Rittaud used McGowan and Makover's idea of trees to construct full binary Fibonacci trees in the following way [15]. The root, which is at the top, consists of a number g_0 with a single child g_1, with at least one of these two values not being 0. Rittaud then defined the tree recursively as follows: if x is the parent of y, then y has two children, $|x - y|$ on the left branch and $x + y$ on the right branch. Rittaud denoted this tree as $\mathbf{T}_{(g_0,g_1)}$ [15], which leads to the following definition.

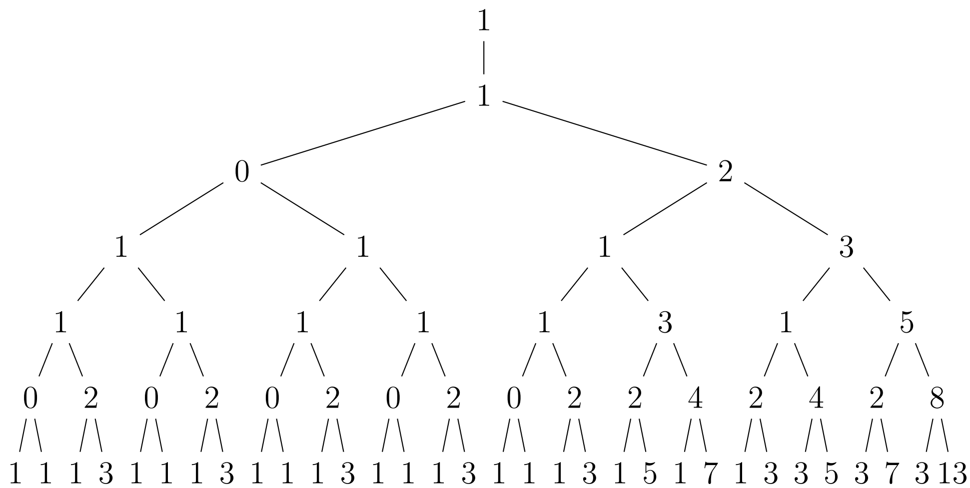

Figure 1.1: The Top of the Tree $\mathbf{T}_{(1,1)}$

Definition 1.1.2. We say a parent child pair (a, b) is at *depth* n if there exists a walk from the root of the tree $g_0, g_1, \ldots, g_{n+1}$, where $g_n = a$ and $g_{n+1} = b$.

Example 1.1.3. Figure 1.1 gives the top of the tree $\mathbf{T}_{(1,1)}$. There are five pairs $(1, 1)$ at depth 3.

Rittaud [15] showed that when $g_0 = g_1 = 1$, the resulting tree has the property that an ordered pair of natural numbers (a, b) occurs on a single branch of this tree with a being the parent of b if and only if $\gcd(a, b) = 1$.

Rittaud went on to study and characterise the *shortest walks* in the tree to an (a, b) pair. In other words, given an arbitrary coprime (a, b) pair, he characterised the walk with the smallest number of edges from the $(1, 1)$ pair at the root to an (a, b) pair in the tree and proved this walk is unique. Given a walk from the root $(1, 1)$ to a pair (a, b), this walk is the shortest walk to the pair (a, b) if and only if for any pair (c, d) appearing in the walk, the parent of c is $|c - d|$.

We extend Rittaud's result on (a, b) pairs occurring infinitely often by giving tight bounds on the number of such pairs at any specific depth in the tree. First, we present a convention.

Convention 1.1.4. In this thesis, we let $\mathbb{N} := \{1, 2, 3, ..., \}$.

Let $n \in \mathbb{N} \cup \{0\}$. We observe that if (a,b) is a pair at depth $3n$, then a and b are odd. Similarly if (a,b) is a pair at depth $3n+1$, then a is odd and b is even. Lastly if (a,b) occurs at depth $3n+2$, then a is even and b is odd. This leads to the following definition.

Definition 1.1.5. We let $A_{(a,b)}(n)$ denote the number of (a,b) pairs that are found at depth $3n+m$ in the Fibonacci tree. Here we have $m=0$ if a and b are odd, $m=1$ if a is odd and b is even, and $m=2$ if a is even and b is odd.

Example 1.1.6. In Figure 1.1, we can see that $A_{1,1}(1)=5$, $A_{(1,1)}(2)=27$, and $A_{1,2}(1)=6$.

We first consider how often can we avoid the pair $(1,1)$. In Section 2.2 we prove the following.

Theorem 1.1.7. Consider a random walk in the tree, starting at the root $(1,1)$, with probability p of choosing a right branch be p and probability $1-p$ of choosing a left branch. Then the probability the walk does not contain any $(1,1)$ pair except at the root is 0 if $p \leq 1/3$ and is

$$\frac{3p-2+\sqrt{4p-3p^2}}{2}$$

if $p > 1/3$.

In the other direction, precise asymptotics for $A_{(1,1)}(n)$ are developed in Section 2.6. In fact, we prove the following result.

Theorem 1.1.8. Letting $A_{(1,1)}(n)$ be defined as above, we have

$$\left(\frac{243\cdot(27/4)^n}{4\sqrt{3\pi}n^{3/2}}\right)\left(1-\frac{1387}{72n}\right) < A_{(1,1)}(n) < \left(\frac{243\cdot(27/4)^n}{4\sqrt{3\pi}n^{3/2}}\right)\left(1-\frac{1387}{72n}+\frac{5548}{9n^2}\right)$$

for $n \geq 100$.

In Section 2.7, we develop precise asymptotics for $A_{(a,b)}$ for all coprime pairs (a,b). That is, we prove the following.

Theorem 1.1.9. For all coprime pairs (a,b), there exists an explicitly computable positive constant $C_{(a,b)}$ and a rational constant $D_{(a,b)}$ such that

$$A_{(a,b)}(n) = \frac{C_{(a,b)}\cdot(27/4)^n}{n^{3/2}} + \frac{C_{(a,b)}D_{(a,b)}(27/4)^n}{n^{5/2}} + O\left(\frac{(27/4)^n}{n^{7/2}}\right).$$

Here the implied constant in the error term depends upon the number of branches in the shortest walk from the root $(1,1)$ to the pair (a,b).

The content of Chapter 2 has been accepted for publication in the Journal of Number Theory [8].

1.2 The Turán and Simple Sieves in Combinatorics

Another number-theoretic area of interest is sieve theory. Let $\omega^*(n)$ denote the number of distinct prime factors of n. Pál Turán in 1934 proved that [21]

$$\sum_{n\leq x}(\omega^*(n) - \log\log x)^2 = O(x\log\log x).$$

In 2004, Liu and Murty [12] generalised Turán's proof to a combinatorial setting and used it to study various problems in combinatorics. First, we present a definition.

Definition 1.2.1. A simple bipartite graph is an undirected graph whose vertices can be divided into two sets, such that there are no edges between two vertices in the same set.

Let X be a bipartite graph with finite partite sets A and B. For $a \in A$ and $b \in B$, we write $a \sim b$ if there is an edge that joins a and b. Define

$$\deg b = \#\{a \in A : a \sim b\} \quad \text{and} \quad \omega(a) = \#\{b \in B : a \sim b\}.$$

In other words, $\deg b$ is the degree of b in X and $\omega(a)$ is the degree of a in X. For $b_1, b_2 \in B$, we define

$$n(b_1, b_2) = \#\{a \in A : a \sim b_1, a \sim b_2\}.$$

Liu and Murty proved that

$$\sum_{a\in A}\left(\omega(a) - \frac{1}{|A|}\sum_{b\in B}\deg b\right)^2 = \sum_{b_1,b_2\in B} n(b_1,b_2) - \frac{1}{|A|}\left(\sum_{b\in B}\deg b\right)^2,$$

from which it follows that

$$\#\{a \in A : \omega(a) = 0\} \leq |A|^2 \cdot \frac{\sum_{b_1,b_2\in B} n(b_1,b_2)}{(\sum_{b\in B}\deg b)^2} - |A|.$$

Liu and Mury named the above result the *Turán sieve*, which gives an upper bound on the quantity $\#\{a \in A : \omega(a) = 0\}$. Liu and Murty [12] also derived an elementary sieve method, called *the simple sieve*, to give a lower bound on the quantity $\#\{a \in A : \omega(a) = 0\}$.

$$\#\{a \in A : \omega(a) = 0\} \geq |A| - \sum_{b\in B}\deg b.$$

In [12], Liu and Murty applied the sieves to study problems on characters over abelian groups, graph colouring, and Latin squares. In Chapter 3, we apply both the simple sieve and the Turán sieve to study problems in random graph theory. More precisely, we obtain upper and lower bounds on the probability of a random graph, a random directed graph, and a random k-partite graph having diameter 2 with $k \geq 3$, or diameter 3 in the case of bipartite graphs.

Definition 1.2.2. A *random graph* is a set of vertices with every edge between any pair of vertices being assigned a probability p.

We will also need the following.

Definition 1.2.3. The *diameter* d of a graph is the minimum number of edges needed in the given graph to traverse between any two vertices.

We will prove the following theorem.

Theorem 1.2.4. For $n \geq 2$, let $G^{(n)}$ denote the set of all graphs on n vertices with edge probability $p(n)$, and let $P(G^{(n)}, p(n))$ be the probability of a graph from $G^{(n)}$ having diameter 2. Then

$$\begin{aligned} & 1 - \frac{n^2(1-p(n)^2)^{n-2}(1-p(n))}{2} \\ \leq\ & P(G^{(n)}, p(n)) \\ \leq\ & \frac{2}{(n-1)^2(1-p(n)^2)^n(1-p(n))} + \frac{8}{n}\left(1 + \frac{p(n)^3}{(1-p(n))^2}\right)^n. \end{aligned}$$

Corollary 1.2.5. Let $P(G^{(n)}, p(n))$ be defined as in Theorem 1.2.4. If $p(n) = \frac{1}{2}$, then we have

$$P(G^{(n)}, 1/2) \geq 1 - \frac{4n^2(3/4)^n}{9}.$$

In the case $p(n) = \frac{1}{2}$, Gilbert [6] showed that 'almost all' graphs are connected, and since a graph with diameter 2 is connected, the above result can be viewed as an improvement of Gilbert's result.

In the situation where the edge probability $p(n) \to 0$ as $n \to \infty$, we will show the following.

Proposition 1.2.6. Let $P(G^{(n)}, p(n))$ be defined as in Theorem 1.2.4. Let $\lim_{n\to\infty} p(n) = 0$. We have

$$1 - \frac{n^2}{2} e^{-np(n)^2}(1 + o(1)) \le P(G^{(n)}, p(n)) \le (1 + o(1)) \left(\frac{2}{n^2} e^{np(n)^2}\right) \left(1 + 4ne^{np(n)^2\left(p(n)^2 - 1\right)}\right).$$

Suppose further that

$$\lim_{n\to\infty} (2\log n - np(n)^2 - \log 2) = c$$

for some $c \in \mathbb{R}\backslash\{0\}$.

1) If $c < 0$, we have

$$P(G^{(n)}, p(n)) \ge 1 - (1 + o(1))e^c.$$

2) If $c > 0$, we have

$$P(G^{(n)}, p(n)) \le (1 + o(1))e^{-c}.$$

We will also study analogous problems for a random directed graph, and a random k-partite graph having diameter 2 with $k \ge 3$, or diameter 3 in the case of bipartite graphs. As we noted in Proposition 1.2.6, the bound we obtained through the Turán sieve works effectively for $c > 0$, while the lower bound we obtained through the simple sieve gives a non-trivial result for $c < 0$. It is interesting to see that the Turán sieve and the simple sieve "almost completely" complement each other in this way.

The content of Chapter 3 is currently in preparation for submission for publication.

1.3 The Mahler Measure of a Polynomial

Our third number-theoretic area studies the Mahler measure of a polynomial. First, we present a definition.

Definition 1.3.1. The *Mahler measure* of a polynomial f with integer coefficients, denoted by $M(f)$, is defined to be the absolute value of the product of its leading coefficient and all its roots with absolute value at least 1. If no such roots exist, the Mahler measure is defined to be the absolute value of the leading coefficient. In other words, if

$$f(x) = a_n(x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n),$$

then

$$M(f) := |a_n| \prod_{i=1}^{n} \max\{1, |\alpha_i|\}.$$

It can also be defined as the geometric mean of the values of $|f|$ on the unit circle:

$$M(f) := \exp\left(\int_0^1 \log|f(e^{2\pi it})|dt\right).$$

We also define the Mahler measure of an algebraic number α as the Mahler measure of the minimal polynomial of α.

A major open problem dealing with the Mahler measure is whether it can get arbitrarily close to 1 without actually being 1. More specifically, for any $\epsilon > 0$, does there exist a polynomial f with integer coefficients such that $1 < M(f) < 1 + \epsilon$? This problem was first posed by Lehmer [11] in 1933 and has since sparked various problems in finding Mahler measures of polynomials. Lehmer was able to show that the polynomial

$$f(x) = x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1$$

has Mahler measure $M(f) = 1.1762808\ldots$. This is the smallest Mahler measure greater than 1 that is currently known.

Lehmer's conjecture has sparked the pursuit of finding lower bounds for the Mahler measures of polynomials in $\mathbb{Z}[x]$. For example, in 1971 Blanksby and Montgomery [2] showed that, given $\alpha \in \mathbb{Z}[x]$ with α not a root of unity and of degreed d, we have

$$M(\alpha) > 1 + \frac{1}{52d\log(6d)}.$$

In 1979, Dobrowolski [4] improved this lower bound to

$$M(\alpha) > 1 + \frac{1}{1200}\left(\frac{\log\log d}{\log d}\right)^3.$$

Later results showed that an important property of polynomials with regard to bounding their Mahler measures is whether they are reciprocal or not.

Definition 1.3.2. Let $f(x)$ be a polynomial of degree n. We define the reciprocal of $f(x)$ as $f^*(x) := x^n f(1/x)$. We say that $f(x)$ is a *reciprocal polynomial* if $f(x) = \pm f^*(x)$.

In 1971, Smyth [20] showed that if f is an irreducible polynomial with integer coefficients that doesn't have 0 or 1 as a root and is not reciprocal, then $M(f) \geq M(x^3 - x - 1) = 1.324717\ldots$. In 2004, Borwein, Hare, and Mossinghoff generalised the concept of being reciprocal and bounded the Mahler measure of a larger class of polynomials [3].

Theorem 1.3.3 (Borwein, Hare, Mossinghoff)**.** Let $f(x)$ be a polynomial where $f(x) \neq \pm f^*(x)$ and $f(x) \equiv \pm f^*(x) \pmod{m}$ with $m \geq 2$. Then

$$M(f) \geq \frac{m + \sqrt{m^2 + 16}}{4}$$

with this bound being sharp when m is even.

The larger m is, the more impressive this bound becomes.

In Chapter 4, we modify Borwein, Hare, and Mossinghoff's proof techniques for achieving the above bound to study a new class of polynomials that we call *k-nonreciprocal* for some integer $k \geq 1$. First, we present a definition.

Definition 1.3.4. Take a polynomial in $\mathbb{Z}[x]$, say $f(x) = \sum_{i=0}^{n} a_i x^i$. For an integer $k \geq 1$, we say that $f(x)$ is *k-nonreciprocal* if $a_n a_i = a_0 a_{n-i}$ for all $1 \leq i \leq k-1$ with $a_n a_k \neq a_0 a_{n-k}$.

As with Borwein, Hare, and Mossinghoff's result, we also prove that our bound is sharp and can get arbitrarily large, depending on the set of polynomials in question. More specifically, we prove the following.

Theorem 1.3.5. Take a polynomial in $\mathbb{Z}[x]$, say $f(x) = \sum_{i=0}^{n} a_i x^i$. Suppose for some $k \in \mathbb{N}$, $2k \leq n$ we have $a_n a_i = a_0 a_{n-i}$ for all $1 \leq i \leq k-1$. Let $M(f)$ denote the Mahler measure of f and $\alpha = |a_k a_n - a_0 a_{n-k}|$. Then

$$M(f) \geq \frac{\alpha + \sqrt{\alpha^2 + 4(|a_0| + |a_n|)^2 |a_0 a_n|}}{2(|a_0| + |a_n|)}.$$

Remark 1.3.6. Borwein, Hare, and Mossinghoff noted that a Corollary to their result is that if f is a nonreciprocal polynomial with all odd coefficients, then

$$M(f) \geq \frac{1 + \sqrt{5}}{2} = 1.618\ldots$$

By Theorem 1.3.5, however, we may replace the condition that f has all odd coefficients with the condition that for the smallest k we have $a_k a_n \neq a_0 a_{n-k}$, then $|a_k a_n - a_0 a_{n-k}| \geq 2$. Assuming that $|a_n| = |a_0| = 1$ (for otherwise $M(f) \geq \min\{|a_0|, |a_n|\} \geq 2$), this condition is substantially weaker than the condition that f is nonreciprocal and has all odd coefficients.

The content of Chapter 4 has been accepted for publication in the *Bulletin of the Australian Mathematical Society* [17].

Chapter 2

On (a, b) Pairs in Random Fibonacci Sequences

2.1 Notation

In this chapter, we prove Theorems 1.1.7, 1.1.8, and 1.1.9.

The chapter is divided up as follows. The proof of Theorem 1.1.7 is given in Section 2.2.

In Sections 2.3 and 2.4 we develop asymptotic formulas for the number of $(1, 1)$ pairs at depth n satisfying specific properties, and a weak upper bound for $A_{(1,1)}(n)$, which will be used in the proof of Theorem 1.1.8. In Section 2.5 we develop some preliminary results for other coprime pairs, which are used in both the proofs of Theorems 1.1.8 and 1.1.9. Sections 2.6 and 2.7 provide the proofs of Theorems 1.1.8 and 1.1.9 respectively.

Notation 2.1.1. Suppose we have two functions $f, g : \mathbb{N} \to \mathbb{R}$. In the rest of the chapter, we use the notation

$$f(n) \sim g(n)$$

to mean that

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1.$$

Note 2.1.2. Since Sections 2.2, 2.3, 2.4, and 2.6 deal exclusively with $(1, 1)$ pairs, we will use $A(n)$ in place of $A_{(1,1)}(n)$ for simplicity of notation in those sections.

We also need to introduce certain subsets of $(1,1)$ pairs at depth $3n$. In [15], Rittaud introduced the concept of a single 0-walk. This is a walk that consists of $3n+2$ branches, starts at the root, and ends at a node with number 0 and does not attain a 0 elsewhere in this walk. He showed that the number of these walks is

$$\frac{1}{2n+1}\binom{3n}{n}.$$

In each such walk, however, since the last node is 0, it can be seen that the third last and second last nodes form a $(1,1)$ pair at depth $3n$ in the tree. This leads to the following definition.

Definition 2.1.3. We let $B(n)$ be the number of $(1,1)$ pairs at depth $3n$ in the tree such that the walks to these pairs do not attain a 0. We define $S(n)$ as the number of $(1,1)$ pairs at depth $3n$ in the tree such that the walks to these pairs do not attain the pair $(1,1)$ in the interior of the walk. We call these $S(n)$ pairs *primitive*.

Example 2.1.4. As can be verified in Figure 1.1, $B(1)=1$, $S(1)=5$, $B(2)=3$, and $S(2)=2$.

For $n \geq 2$, we have

$$S(n) \leq B(n) = \frac{1}{2n+1}\binom{3n}{n} \leq A_{(1,1)}(n), \tag{2.1.1}$$

where $B(n) = \frac{1}{2n+1}\binom{3n}{n}$ is an equivalent definition of $B(n)$. Given a walk to any $(1,1)$ pair that isn't primitive, we know that this walk must go through an intermediate primitive $(1,1)$ pair. Thus, we have the formula

$$A_{(1,1)}(n) = \sum_{i=0}^{n-1} A_{(1,1)}(i)S(n-i), \tag{2.1.2}$$

which is an equivalent inductive definition for $A_{(1,1)}(n)$. It is also worth noting that the function $B(n)$ counts the $(1,1)$ pairs whose walks begin with a right branch and which take a right branch after every intermediate $(1,1)$ pair attained. Any walk that doesn't have this property would have to attain an intermediate 0, contradicting our definition of $B(n)$. Conversely, every walk that has this property cannot attain any node that has a 0, since the only way to attain a 0 is to take an immediate left branch after a pair $(1,1)$.

Using this fact, we also have the formula

$$B(n) = B(n-1) + \sum_{i=0}^{n-2} B(i)S(n-i) \tag{2.1.3}$$

for $n \geq 2$. Also, $B(0) = 1$, since a walk consisting of 0 branches starting at the root $(1,1)$ stays at the root $(1,1)$ without traversing 0. Also, $B(1) = 1$, since there is only one walk consisting of exactly three branches starting at the root $(1,1)$ and ending at the pair $(1,1)$, with the first branch being a right branch. We'll be using Equations (2.1.2) and (2.1.3) in the rest of the chapter.

2.2 Random Walks in the Tree

In this section we will consider the problem of how often we expect to find a random infinite walk that never attains the pair $(1,1)$.

Theorem 1.1.7. Consider a random walk in the tree, starting at the root $(1,1)$, with probability p of choosing a right branch and probability $1-p$ of choosing a left branch. Then the probability the walk does not contain any $(1,1)$ pair except at the root is 0 if $p \leq 1/3$ and is

$$\frac{3p - 2 + \sqrt{4p - 3p^2}}{2}$$

if $p > 1/3$.

Example 2.2.1. If we substitute in $p = 1/2$ into Theorem 1.1.7, it is interesting to note that for the probability the walk does not contain any $(1,1)$ pair except at the root is $\phi/2$ where

$$\phi = \frac{1+\sqrt{5}}{2},$$

is the golden ratio.

We first need some preliminary lemmas.

Lemma 2.2.2. Let $a, b, a_1, \ldots, a_n, c, d$ be a walk from (a,b) to (c,d). Then $d, c, a_n, \ldots, a_2, a_1, b, a$ is a walk from (d,c) to (b,a).

Proof. We only have to show that if a_1, a_2, a_3 occur in the given walk, then a_3, a_2, a_1 can consecutively occur in a walk in that order. We have either $a_3 = a_1 + a_2$ or $a_3 = |a_2 - a_1|$. In the first case, we have $a_1 = |a_2 - a_3|$ giving us our result. In the second case, we either have $a_3 = a_2 - a_1$, giving us $a_1 = a_2 - a_3$, or $a_3 = a_1 - a_2$, giving us $a_1 = a_2 + a_3$. $\square$

Lemma 2.2.3. The shortest walk from a non-$(1,1)$ pair to a $(1,1)$ pair is characterised as a series of left branches with no right branches.

Proof. By Lemma 2.2.2, we obtain the shortest walk by traversing backwards along the shortest walk from $(1,1)$ to the given non-$(1,1)$ pair (a,b). In [15, Corollary 5.1], Rittaud observes that the latter walk has the property that for any pair (c,d) occurring in the walk, the parent of c is $|c-d|$. Thus, the shortest walk from (a,b) to $(1,1)$ must have the property that for any pair (c,d) occurring in the walk, the child of d is $|c-d|$, a choice of a left branch. Thus, the shortest walk must contain no right branches. $\square$

Note 2.2.4. We say the walk from a pair (a,b) to another pair (c,d) consists of n branches if there are exactly n branches between the node at b of the first pair to the node at d of the second pair.

Notation 2.2.5. For a coprime pair (a,b), let $SW_{(a,b)}(c,d)$ denote the number of branches in the shortest walk from an (a,b) pair to a (c,d) pair.

Lemma 2.2.6. Starting from a non-$(1,1)$ pair (a,b) in the tree, suppose $SW_{(a,b)}(1,1) = n$. Then $SW_{(b,|b-a|)}(1,1) = n-1$ and $SW_{(b,a+b)}(1,1) = n+2$.

Proof. Follows from Lemma 2.2.3. $\square$

Proof of Theorem 1.1.7. Each walk not containing a pair $(1,1)$ except at the root must begin with a right branch. From there each desirable walk can correspond to an infinite positive integer sequence, each number denoting $SW_{(a,b)}(1,1)$ for a specific pair (a,b) in the given walk. Thus, we can consider the problem of having random integer sequences beginning with 2 and either adding 2 or subtracting 1 to get the next number. We want to know the probability of such a sequence having all of its terms be positive.

For each $n \in \mathbb{N} \cup \{0\}$, denote by $P(n)$ the probability of starting a sequence with n and applying the above rules and eventually traversing 0. Thus, we have the recurrence

$$P(n) = (1-p)P(n-1) + pP(n+2), n \neq 0 \tag{2.2.1}$$

with $P(0) = 1$. Here the successor n in the sequence will either be $n-1$ with probability $1-p$ or $n+2$ with probability p.

We can prove that there exist constants A, B, and C such that for all $n \in \mathbb{N} \cup \{0\}$, we have

$$P(n) = A + Br_1^n + Cr_2^n \tag{2.2.2}$$

where

$$r_1 = \frac{-1 + \sqrt{4/p - 3}}{2} \text{ and } r_2 = \frac{-1 - \sqrt{4/p - 3}}{2}.$$

From (2.2.1), we obtain for $n \geq 3$ that

$$P(n) = \frac{P(n-2)}{p} - \frac{P(n-3)(1-p)}{p}.$$

We now split into three cases.

Case 1. $p < 1/3$.

We can work out that $r_1 > 1$ and $r_2 < -2$. Therefore, if $B \neq 0$ or $C \neq 0$, then by (2.2.2) we have

$$\limsup_{n \to \infty} P(n) = \infty,$$

a contradiction since $0 \leq P(n) \leq 1$ for all $n \in \mathbb{N}$. Therefore, $B = C = 0$ and since $P(0) = 1$, we have

$$P(n) = 1$$

for all $n \in \mathbb{N} \cup \{0\}$. Therefore, the probability of a random walk not traversing a pair $(1, 1)$ except at the root is

$$p(1 - P(2)) = p(1 - 1) = 0.$$

Case 2. $p = 1/3$.

We can work out that $r_1 = 1$ and $r_2 = -2$ so that from (2.2.1) we get

$$P(n) = A + B + C(-2)^n.$$

If $C \neq 0$, then we have

$$\limsup_{n \to \infty} P(n) = \infty,$$

a contradiction since $0 \leq P(n) \leq 1$. Therefore, $C = 0$ and since $P(0) = 1$, we have

$$P(n) = 1$$

for all $n \in \mathbb{N} \cup \{0\}$ and we proceed as in Case 1.

Case 3. $1/3 < p < 1$.

We have
$$r_2 = \frac{-1-\sqrt{4/p-3}}{2} < \frac{-1-\sqrt{4-3}}{2} = -1.$$
Therefore, if $C \neq 0$, then by (2.2.2), we have
$$\limsup_{n\to\infty} P(n) = \infty,$$
a contradiction since $0 \leq P(n) \leq 1$. Therefore, $C = 0$ and
$$P(n) = A + Br_1^n.$$
We will show that $A = 0$ by showing $\lim_{n\to\infty} P(n) = 0$. Suppose we start with $n \in \mathbb{N}$ and eventually attain 0. Then the number of times we added 2 is r and the number of times we subtracted 1 is $2r+n$ for some $r \in \mathbb{N}$. Thus, we have
$$P(n) \leq \sum_{r=0}^{\infty} \binom{3r+n}{r} p^r (1-p)^{2r+n}.$$
In [7], we have the combinatorial identity
$$\sum_{k=0}^{n} \binom{tk+r}{k}\binom{tn-tk+s}{n-k}\frac{r}{tk+r} = \binom{tn+r+s}{n},$$
which is valid for all $n \in \mathbb{N}\cup\{0\}$ and all real r, s, and t. Substituting in $t=3, r=1$ gives
$$\sum_{k=0}^{n} \binom{3k+1}{k}\binom{3n-3k+s}{n-k}\frac{1}{3k+1} = \binom{3n+s+1}{n}.$$
Using this identity, we can prove by induction on $n \in \mathbb{N}$ that
$$\sum_{r=0}^{\infty} \binom{3r+n}{r} p^r (1-p)^{2r+n} = (1-p)^n \left(\sum_{r=0}^{\infty} \binom{3r}{r} p^r (1-p)^{2r}\right) \left(\sum_{r=0}^{\infty} \binom{3r+1}{r} \frac{p^r(1-p)^{2r}}{3r+1}\right)^n.$$
Since $1/3 < p \leq 1$, we can further deduce that
$$\lim_{r\to\infty} \frac{\binom{3r}{r}}{\binom{3(r+1)}{r+1}} = \frac{4}{27} > p(1-p)^2$$

so that

$$\sum_{r=0}^{\infty} \binom{3r}{r} p^r (1-p)^{2r} < \infty.$$

Moreover, entering the command "with(SumTools)" and then the command "DefiniteSummation$\left(\binom{3r+1}{r}\frac{p^r(1-p)^{2r}}{3r+1}, r = 0\ldots\infty\right)$ assuming $0 < p \leq 1$" into Maple (the ∞ symbol is found under "Common Symbols") gives

$$\sum_{r=0}^{\infty} \binom{3r+1}{r} \frac{p^r(1-p)^{2r}}{3r+1} = \frac{2\sqrt{3}\sin\left(\frac{1}{3}\cdot\arcsin\left(\frac{3\sqrt{3}\cdot(1-p)\sqrt{p}}{2}\right)\right)}{3(1-p)\sqrt{p}}.$$

Thus

$$\sum_{r=0}^{\infty} \binom{3r+1}{r} \frac{p^r(1-p)^{2r+1}}{3r+1} = \frac{2\sin\left(\frac{1}{3}\cdot\arcsin\left(\frac{3\sqrt{3}\cdot(1-p)\sqrt{p}}{2}\right)\right)}{\sqrt{3p}}.$$

For $1/3 < p \leq 1$, we have

$$\begin{aligned}\frac{2\sin\left(\frac{1}{3}\cdot\arcsin\left(\frac{3\sqrt{3}\cdot(1-p)\sqrt{p}}{2}\right)\right)}{\sqrt{3p}} &\leq \frac{2\sin\left(\frac{\pi}{6}\right)}{\sqrt{3p}} \\ &= \frac{1}{\sqrt{3p}} \\ &< 1.\end{aligned}$$

Thus, we deduce

$$\lim_{n\to\infty} P(n) = 0. \tag{2.2.3}$$

Thus, using (2.2.3), we have

$$0 = \lim_{n\to\infty} P(n) = \lim_{n\to\infty} A + Br_1^n = \lim_{n\to\infty} A + Br_1^n = A,$$

giving us $A = 0$. Thus

$$P(n) = Br_1^n.$$

Since $P(0) = 1$, we thus have $B = 1$ so

$$P(n) = r_1^n.$$

We recall that $P(2)$ is the probability of starting at 2 and randomly adding 2 or subtracting 1 and eventually traversing 0. Thus, $1 - P(2)$ is the probability of never traversing 0. This is equal to the probability of never traversing a pair $(1,1)$ in the Fibonacci tree after taking the first branch to be a right branch. Since it could be the first branch is a left branch (from which it is unavoidable to attain another $(1,1)$ pair), we therefore have that the probability of a random walk not traversing a pair $(1,1)$ except at the root is

$$\begin{aligned} p(1-P(2)) &= p(1-r_1^2) \\ &= p - p\left(\frac{-1+\sqrt{4/p-3}}{2}\right)^2 \\ &= \frac{4p - p(\sqrt{4/p-3}-1)^2}{4} \\ &= \frac{4p - p(4/p-2-2\sqrt{4/p-3})}{4} \\ &= \frac{3p-2+p\sqrt{4/p-3}}{2}. \end{aligned}$$

□

2.3 Asymptotics for $S(n)$ and $B(n)$

Recall that $B(n)$ counts the number of $(1,1)$ pairs at depth $3n$ in the tree such that the walk does not attain a 0, whereas $S(n)$ is defined similarly except the walk does not attain an intermediate pair $(1,1)$.

Here we prove that

$$S(n) = \frac{(27/4)^n}{3\sqrt{3\pi}n^{3/2}}\left(1+\frac{17}{72n}+O\left(\frac{1}{n^2}\right)\right)$$

and

$$B(n) = \frac{\sqrt{3}\cdot(27/4)^n}{4\sqrt{\pi}n^{3/2}}\left(1-\frac{43}{72n}+O\left(\frac{1}{n^2}\right)\right).$$

Proposition 2.3.1. An equivalent definition of $S(n)$ is obtained by setting $S(1)=5$ and

$$S(n) = \frac{2}{3n-1}\binom{3n-1}{n-1}$$

for $n \geq 2$.

Proof. We see that $S(1) = 5$, and $S(2) = \frac{2}{5}\binom{5}{1} = 2$. At depth $3n$ in the tree, where $n \geq 2$, we know that if we attain a $(1,1)$ pair, then we must have taken twice as many left branches as right branches. Also, if our first branch is a left branch we will attain a $(1,1)$ pair at depth 3 in the tree. Therefore, all primitive $(1,1)$ pairs at depth $3n$ in the tree, $n \geq 2$, must occur on walks where the initial branch is a right branch. After this initial right branch, the rest of the walk must consist of $n-1$ right branches and $2n$ left branches to reach a primitive $(1,1)$ pair at depth $3n$ in the tree for $n \geq 2$. Therefore, for $n \geq 2$, we have $S(n) \leq \binom{3n-1}{n-1}$. This upper bound, however, will over-count the number of primitive $(1,1)$s since it also counts walks where the walk to an intermediate pair might have twice as many left branches as right branches. There are $\binom{3n-3k}{n-k}S(k)$ such walks where the first intermediate pair with this property occurs at depth $3k$ in the tree if $k \geq 2$. If $k = 1$, there are $\binom{3n-3}{n-1}$ such walks. For any intermediate pair we want the number of left branches to be strictly less than twice the number of right branches, and so we subtract these terms to get the recurrence

$$S(n) = \binom{3n-1}{n-1} - \binom{3n-3}{n-1} - \sum_{k=2}^{n-1} \binom{3n-3k}{n-k} S(k). \tag{2.3.1}$$

for $n \geq 2$. Assuming by induction that $S(1) = 5$ and

$$S(k) = \frac{2}{3k-1}\binom{3k-1}{k-1}$$

for $2 \leq k < n$, one can check via Maple that equation (2.3.1) is satisfied when $S(n) = \frac{2}{3n-1}\binom{3n-1}{n-1}$ in the following way. Enter the command "with(SumTools):" and then the command "DefiniteSummation$\left(\binom{3n-3k}{n-k} \cdot \frac{2}{3k-1}\binom{3k-1}{k-1}\right)$, $k = 1 \ldots n$", which gives the identity

$$\sum_{k=1}^{n} \binom{3n-3k}{n-k} \frac{2}{3k-1}\binom{3k-1}{k-1} = \frac{1}{3}\binom{3n}{n} = \binom{3n-1}{n-1}.$$

We can deduce (2.3.1) from this induction step. $\square$

Proposition 2.3.2. For all $n \in \mathbb{N}$, we have

$$\frac{\sqrt{3} \cdot (27/4)^n}{2\sqrt{\pi n}}\left(1 - \frac{7}{72n}\right) < \binom{3n}{n} < \frac{\sqrt{3} \cdot (27/4)^n}{2\sqrt{\pi n}}\left(1 - \frac{7}{72n} + \frac{1}{50n^2}\right).$$

Proof. Robbins shows [16] that, for all $n \in \mathbb{N}$, we have

$$\sqrt{2\pi} n^{n+1/2} e^{-n} \cdot e^{1/(12n+1)} < n! < \sqrt{2\pi} n^{n+1/2} e^{-n} \cdot e^{1/(12n)}. \tag{2.3.2}$$

Note the following:

$$\frac{-7}{72n} < \frac{1}{36n} - \frac{1}{24n+1} - \frac{1}{12n+1} < \frac{-7}{72n} + \frac{5}{576n^2}.$$

Thus, we have

$$\binom{3n}{n} = \frac{(3n)!}{(2n)! \cdot n!} < \frac{\sqrt{3} \cdot (27/4)^n \cdot e^{\frac{-7}{72n} + \frac{5}{576n^2}}}{2\sqrt{\pi n}}$$

For $-1 < x < 1$, we have

$$e^x = 1 + x + \frac{x^2}{2} + \cdots + \frac{x^n}{n!} + \cdots$$

Letting $x = \frac{-7}{72n} + \frac{5}{576n^2}$, we have for $n \geq 1$ that $-1 < x < 0$ and hence

$$e^x < 1 + x + \frac{x^2}{2},$$

as it is an alternating series. We have

$$\begin{aligned} \binom{3n}{n} &< \frac{\sqrt{3} \cdot (27/4)^n}{2\sqrt{\pi n}} \left(1 - \frac{7}{72n} + \frac{5}{576n^2} + \frac{\left(-\frac{7}{72n} + \frac{5}{576n^2}\right)^2}{2}\right) \\ &< \frac{\sqrt{3} \cdot (27/4)^n}{2\sqrt{\pi n}} \left(1 - \frac{7}{72n} + \frac{1}{50n^2}\right). \end{aligned}$$

The second inequality follows from

$$\frac{5}{576n^2} + \frac{\left(-\frac{7}{72n} + \frac{5}{576n^2}\right)^2}{2} = \frac{139}{10368n^2} - \frac{35}{41472n^3} + \frac{25}{663552n^4} < \frac{139}{10368n^2} < \frac{1}{50n^2}$$

with the equality following from Maple.

A similar argument can be used for the opposite inequality. □

Corollary 2.3.3. For all $n \in \mathbb{N}$, $n \geq 100$, we have

$$\frac{(27/4)^n}{3\sqrt{3\pi}n^{3/2}} \left(1 + \frac{17}{72n} + \frac{3}{40n^2}\right) < S(n) < \frac{(27/4)^n}{3\sqrt{3\pi}n^{3/2}} \left(1 + \frac{17}{72n} + \frac{1}{10n^2}\right).$$

and

$$\frac{\sqrt{3} \cdot (27/4)^n}{4\sqrt{\pi}n^{3/2}} \left(1 - \frac{43}{72n} + \frac{1}{4n^2}\right) < B(n) < \frac{\sqrt{3} \cdot (27/4)^n}{4\sqrt{\pi}n^{3/2}} \left(1 - \frac{43}{72n} + \frac{1}{3n^2}\right).$$

Proof. We can deduce our bounds from Equation (2.1.1) and Propositions 2.3.1 and 2.3.2. □

2.4 The Limit of the Ratio

Recall that $A(n)$ is the number of $(1,1)$ pairs at depth $3n$ in the tree where there are no restrictions on the walk. Here we prove a weak bound for $A(n)$, which we use to derive that

$$\lim_{n\to\infty} \frac{A(n+1)}{A(n)} = \frac{27}{4}.$$

Proposition 2.4.1. For all $n \in \mathbb{N} \cup \{0\}$, we have

$$A(n) < 2 \cdot \binom{3n}{n}.$$

Proof. We will prove the inequality by induction on n. One can check that it holds for $n = 0, 1, \ldots, 4$. Suppose for some $n \geq 5$, we have for all $0 \leq i \leq n-1$ that

$$A(i) \leq 2\binom{3i}{i}.$$

Then by (2.1.2) we have

$$A(n) < 2\sum_{i=1}^{n} \binom{3n-3i}{n-i} S(i).$$

Noticing that for $S(n) = \frac{2}{3n-1}\binom{3n-1}{n-1}$ for $n \geq 2$ and $S(1) = 5 = \frac{2}{3\cdot 1-1}\binom{3\cdot 1-1}{1-1} + 4$, we observe that

$$\frac{2\sum_{i=1}^{n}\binom{3n-3i}{n-i}S(i)}{2\binom{3n}{n}} = \frac{8\binom{3n-3}{n-1} + 2\sum_{i=1}^{n}\binom{3n-3i}{n-i}\frac{2}{3i-1}\binom{3i-1}{i-1}}{2\binom{3n}{n}}.$$

We can use Maple to evaluate the sum as in the proof of Proposition 2.3.1 and obtain that

$$\frac{8\binom{3n-3}{n-1} + 2\sum_{i=1}^{n}\binom{3n-3i}{n-i}\frac{2}{3i-1}\binom{3i-1}{i-1}}{2\binom{3n}{n}} = \frac{25n^2 - 17n + 2}{27n^2 - 27n + 6}.$$

We observe that this is less than 1 for all $n \geq 5$, proving

$$A(n) < 2\sum_{i=1}^{n}\binom{3n-3i}{n-i}S(i) < 2\binom{3n}{n},$$

as desired. $\square$

Corollary 2.4.2. We have

$$A(n) \leq (1 + o(1)) \frac{\sqrt{3} \cdot (27/4)^n}{\sqrt{\pi n}}.$$

Proof. This follows from Propositions 2.3.2 and 2.4.1. □

Note 2.4.3. In Corollary 2.4.2 no indication is giving on the speed of decay of the $o(1)$. For the purpose of how this result is used in this chapter, however, it is not necessary here to give any information on this speed of decay.

Lemma 2.4.4. For all $n \in \mathbb{N}$, $n \geq 2$ we have

$$\frac{S(n+1)}{S(n)} < \frac{S(n+2)}{S(n+1)}.$$

Proof. By Proposition 2.3.1, we have for each $n \geq 2$

$$\frac{S(n+2)S(n)}{(S(n+1))^2} = \frac{36n^4 + 126n^3 + 158n^2 + 84n + 16}{36n^4 + 126n^3 + 104n^2 - 14n - 12} > 1,$$

from which the result follows. □

Lemma 2.4.5. For all $n \in \mathbb{N}$, we have

$$\frac{A(n+1)}{A(n)} < \frac{A(n+2)}{A(n+1)}.$$

Proof. We prove this by induction on n. First, for $n = 1$, we have

$$\frac{A(2)}{A(1)} = \frac{27}{5} < \frac{152}{27} = \frac{A(3)}{A(2)}$$

where we use (2.1.2) and Proposition 2.3.1 to calculate $A(3) = 152$. Suppose by strong induction, we have

$$\frac{A(i+1)}{A(i)} < \frac{A(i+2)}{A(i+1)}$$

for all $1 \leq i \leq n-1$. Then we can deduce that

$$\frac{A(i)}{A(n)} > \frac{A(i+1)}{A(n+1)}$$

for all $1 \le i \le n-1$. Also, from (2.1.2) we have

$$\begin{aligned}
\frac{A(n+1)}{A(n)} &= 5 + \frac{S(n+1)}{S(n)} - \frac{S(n+1)(A(n)-S(n))}{A(n)S(n)} + \frac{A(n+1) - 5 \cdot A(n) - S(n+1)}{A(n)} \\
&= 5 + \frac{S(n+1)}{S(n)} - \frac{S(n+1)}{A(n)S(n)} \left(\sum_{i=1}^{n-1} A(i)S(n-i) \right) + \frac{1}{A(n)} \left(\sum_{i=1}^{n-1} A(i)S(n+1-i) \right)
\end{aligned}$$

By Lemma 2.4.4, we can derive that

$$\frac{S(n+1)}{S(n)} > \frac{S(n+1-i)}{S(n-i)},$$

or

$$\frac{S(n+1)}{S(n)} \cdot S(n-i) - S(n+1-i) > 0$$

for all $1 \le i < n$. Thus, we have the following:

$$\begin{aligned}
\frac{A(n+1)}{A(n)} &= 5 + \frac{S(n+1)}{S(n)} - \sum_{i=1}^{n-1} \frac{A(i)}{A(n)} \left(\frac{S(n+1)}{S(n)} \cdot S(n-i) - S(n+1-i) \right) \\
\frac{A(n+1)}{A(n)} &< 5 + \frac{S(n+1)}{S(n)} - \sum_{i=1}^{n-1} \frac{A(i+1)}{A(n+1)} \left(\frac{S(n+1)}{S(n)} \cdot S(n-i) - S(n+1-i) \right) \\
&= 5 + \frac{S(n+1)}{S(n)} - \frac{S(n+1)}{A(n+1)S(n)} \left(\sum_{i=1}^{n-1} A(i+1)S(n-i) \right) \\
&\quad + \frac{1}{A(n+1)} \left(\sum_{i=1}^{n-1} A(i+1)S(n+1-i) \right) \\
&= 5 + \frac{S(n+1)}{S(n)} - \frac{S(n+1)(A(n+1) - 5 \cdot S(n) - S(n+1))}{A(n+1)S(n)} \\
&\quad + \frac{A(n+2) - 5 \cdot A(n+1) - 5 \cdot S(n+1) - S(n+2)}{A(n+1)} \\
&= \frac{S(n+1)^2}{A(n+1)S(n)} + \frac{A(n+2)}{A(n+1)} - \frac{S(n+2)}{A(n+1)} \\
&= \frac{A(n+2)}{A(n+1)} - \frac{S(n+1)}{A(n+1)} \left(\frac{S(n+2)}{S(n+1)} - \frac{S(n+1)}{S(n)} \right) \\
&< \frac{A(n+2)}{A(n+1)}.
\end{aligned}$$

The last inequality follows from Lemma 2.4.4. Thus, by strong induction, we have our result. □

Proposition 2.4.6. We have

$$\lim_{n\to\infty} \frac{A(n+1)}{A(n)} = \frac{27}{4}.$$

Proof. By Lemma 2.4.5, we have $\frac{A(n+1)}{A(n)}$ is an increasing sequence in $n \in \mathbb{N} \cup \{0\}$, so the desired limit exists. First, suppose that

$$r := \lim_{n\to\infty} \frac{A(n+1)}{A(n)} < \frac{27}{4}.$$

Then we have

$$\limsup_{n\to\infty} \frac{A(n)}{r^n} < \infty.$$

Since for all $n \in \mathbb{N}$ $S(n) \leq A(n)$, we thus have

$$\limsup_{n\to\infty} \frac{S(n)}{r^n} < \infty.$$

But, by Corollary 2.3.3, we then must have that

$$\limsup_{n\to\infty} \frac{(27/4)^n}{n^{3/2} r^n} = \limsup_{n\to\infty} \frac{\left(\frac{27}{4r}\right)^n}{n^{3/2}} < \infty,$$

a contradiction since we must have

$$\lim_{n\to\infty} \frac{t^n}{n^{3/2}} = \infty$$

for all real $t > 1$. Suppose that $r > \frac{27}{4}$. Then there exists $s > \frac{27}{4}$ such that for all $n \in \mathbb{N}$ sufficiently large, we have

$$s < \frac{A(n+1)}{A(n)}.$$

Thus

$$\liminf_{n\to\infty} \frac{A(n)}{s^n} > 0.$$

But, by Corollary 2.4.2, we must then have

$$\liminf_{n\to\infty} \left(\frac{27}{4s}\right)^n > \liminf_{n\to\infty} \frac{(27/4)^n}{n^{1/2} s^n} > 0,$$

again a contradiction since

$$\lim_{n\to\infty} t^n = 0$$

for all real $0 \leq t < 1$. Thus, we have our result. □

2.5 Preliminary Results Concerning Other Coprime Pairs

In this section, we turn our attention to the behaviour of coprime pairs other than $(1, 1)$ and establish a number of useful preliminary results concerning them. But first we derive a few other useful results concerning $(1, 1)$ pairs, one of them due to Rittaud.

Rittaud constructed a subtree $\mathbf{R}$ from the Fibonacci tree consisting of all the shortest walks from the root $(1, 1)$ down to each coprime pair (a, b), calling it the *restricted tree*. The top part of this subtree is shown in Figure 2.1 (here a vertical line constitutes a right branch).

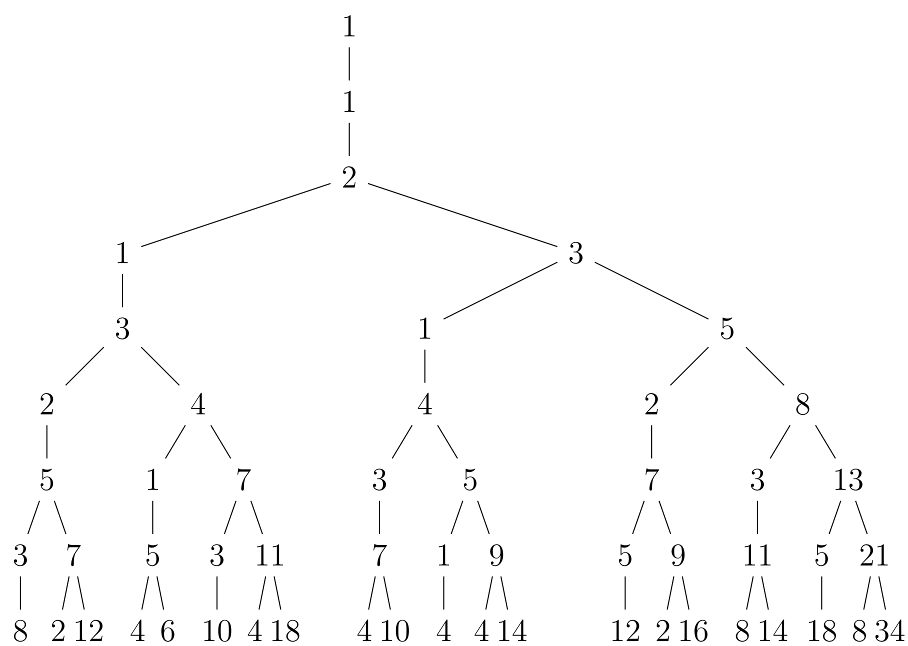

Figure 2.1: The Restricted Tree $\mathbf{R} = \mathbf{R}_{(1,1)}$

He proves the following in [15], giving another characterisation for shortest walks.

Lemma 2.5.1 (Rittaud)**.** The restricted tree $\mathbf{R}$ consists of all walks that do not have two left branches occurring with no right branch between them. Therefore, for all coprime pairs (a, b), the shortest walk from the root $(1, 1)$ to (a, b) does not have two left branches occurring with no right branch between them.

Lemma 2.5.2. Let the first occurrence of the coprime pair (a, b) be at depth k. For any integer $n \geq 0$, there exists a walk with $k+3n$ branches that ends at a pair (a, b). Moreover, if a walk of length ℓ ends at a (a, b) pair, then $\ell - k \in 3\mathbb{N}$.

Proof. If $1, 1, a_3, \ldots, a_{k-2}, a, b$ is a walk of length k, then $1, 1, 0, 1, 1, a_3, \ldots, a_{k-2}, a, b$ is a walk of length $k+3$. Hence, the first part follows by induction, whereas the second follows by the parity of a and b and the minimality of k. □

Proposition 2.5.3. Take a coprime pair (a, b) that is not $(1, 1)$ and let $1, 1, a_1, a_2, \ldots, a_m, a, b$ be a walk from $(1, 1)$ to (a, b). Then $|a - b|, a, b$ occurs as a subwalk within this walk.

Note that the terminal pair (a, b) may not be the only occurrence of the pair (a, b) that the walk traverses.

Proof. Suppose $SW_{(1,1)}(a, b) = k$. By Lemma 2.5.2, the length of all the possible walks are $k + 3n$ where $n \in \mathbb{N} \cup \{0\}$. We will prove this by induction on n.

For $n = 0$, we obtain the shortest walk from the root $(1, 1)$ to (a, b). By [15, Corollary 5.1] we have that the parent of a of the ending pair (a, b) is $|a - b|$.

Suppose now the proposition holds for all $0 \leq n < N$ for some $N \in \mathbb{N}$. Take a walk from $(1, 1)$ to (a, b) consisting of $k + 3N$ branches. If the first branch is a left branch, then we attain another $(1, 1)$ pair at depth 3 in the tree, and so we can remove these first three branches to obtain a walk of length $k + 3(N - 1)$ from which by induction the walk must consist of a pair (a, b) such that the parent of this specific a is $|a - b|$. Since we only removed the first three branches of the original walk, the original walk must have this property too. Suppose that the walk in question starts with a right branch. We know that this walk is not the shortest walk since $N \geq 1$. Therefore, by Lemma 2.5.1, we must have that the walk contains somewhere two left branches with no right branches between them. Since the first branch is a right branch, it therefore follows that somewhere in the tree we have a consecutive sequence of 3 branches consisting of a right branch followed by two left branches. Suppose the branch immediately before this right branch (in case this specific right branch is the first branch in the walk consider the root $(1, 1)$ here) consists of the pair (c, d). Then taking the right branch and then the two left branches gives us the sequence (starting with the (c, d) pair)

$$c, d, c + d, c, d$$

Therefore, the second left branch also consists of the pair (c, d). Removing the right branch and the two left branches therefore gives us a shorter walk to the pair (a, b). Since by induction this shorter walk must have a pair (a, b) with this specific a having a parent

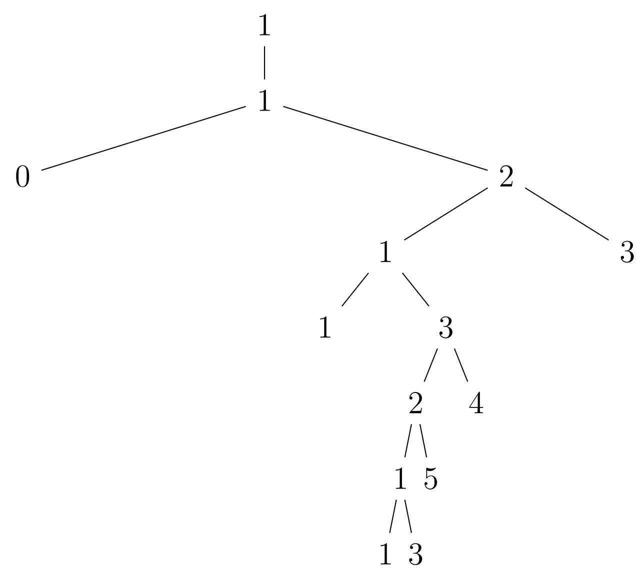

Figure 2.2: A walk from the root $(1,1)$ to a primitive $(1,1)$ pair at depth 6 in the tree

of $|a-b|$ in the walk, we therefore obtain that the original walk has this property too. By induction, we obtain our result. $\square$

Proposition 2.5.4. Let $w_1, w_2, \ldots, w_n$ be a sequence of left and right branches corresponding to a walk to a primitive $(1,1)$. Then for all $1 \le i < n$ the number of left branches in $w_1, \ldots, w_i$ is strictly less than twice the number of right branches. Further, $w_1, \ldots, w_n$ will contain exactly twice as many left branches as right branches. Moreover, all walks of this form are walks to primitive $(1,1)$s.

Proof. The fact that the first branch has to be a right branch follows from the observation that a left branch will just lead to all $(1,1)$ pairs at depth 3 in the tree. The first right branch consists of the pair $(1,2)$. From here the shortest walk to a $(1,1)$ pair consists of two left branches. Suppose we have a walk from this $(1,2)$ to a primitive $(1,1)$. Let the nodes in this walk be $1, 2, a_1, a_2, \ldots, a_k, 1, 1$. Consider the sequence $SW_{(1,2)}(1,1), SW_{(2,a_1)}(1,1), SW_{(a_1,a_2)}(1,1), \ldots, SW_{(a_k,1)}(1,1), SW_{(1,1)}(1,1)$. This is a sequence of integers starting with 2 (since $SW_{(1,2)}(1,1) = 2$). Each successive element in the sequence is obtained by adding 2 to the previous element (representing going down a right branch) or subtracting 1 from the previous element (representing going down a left branch) by Lemma 2.2.6. Finally, all integers in the sequence will be positive, except for the last being 0 since $SW_{(1,1)}(1,1) = 0$.

One property of such a sequence is that if r is the number of times you add 2, then $2r + 2$ must be the number of times you subtract 1. Moreover anywhere in the sequence except at the last element if s is the number of times you added 2 up to that point, then you cannot have subtracted 1 more than $2s + 1$ times. Moreover, it is seen that if we have a finite integer sequence starting with 2 with the above rules in play, then all the elements in the sequence will be positive except for the last one, which will be a 0.

Thus, the walks to all the primitive $(1,1)$s in the tree that have a length of more than 3 branches can be characterised as in the proposition. □

Lemma 2.5.5. Take the tree $\mathbf{T}_{(a,b)}$ for some coprime pair (a,b) where $a, b \geq 1$. Suppose we take a finite walk in the tree, starting at the root and consisting of exactly twice as many left branches as right branches, but such that at any given intermediate point the number of left branches taken is less than or equal to twice the number of right branches taken. Then the pair on the last branch will be (a,b).

Proof. Given a path as in the lemma, if at all intermediate points the number of left branches taken is strictly less than twice the number of right branches taken, then we can repeat the argument given in the proof of Proposition 2.5.4 to deduce that the ending pair will be (a,b). For the broader collection of paths given in the lemma, we may then apply induction on the number of places in the given path, where the number of left branches taken is exactly twice the number of right branches taken to obtain the result. □

Lemma 2.5.6. Take a walk in $\mathbf{T}_{(a,b)}$ that starts at the root (a,b), where $a, b \geq 1$ and $\gcd(a,b) = 1$. Suppose that the number of left branches is strictly less than twice the number of right branches in this walk. Also suppose that at any given intermediate point in the walk the number of left branches taken is less than or equal to twice the number of right branches taken. Then the pair on the final branch will not be (a,b).

Proof. Take such a path as described in the lemma. It is possible to extend this path by a series of left branches to obtain a path as described in Lemma 2.5.5, and hence the final pair on this extended path has to be (a,b). Since the values of the nodes are decreasing along this series of left branches, we must have that the ending pair of the original path cannot be (a,b). □

Lemma 2.5.7. Take the tree $\mathbf{T}_{(a,b)}$ for some coprime pair (a,b) and let $n \in \mathbb{N}$. Suppose we take all (a,b) pairs at depth $3n$ in the tree such that the walks to these (a,b) pairs satisfies the following. Let the first branch be a right branch and the branch after any intermediate pair (a, b) in the walk be a right branch. The number of such (a, b) pairs is $B(n)$.

Proof. We will first show that the walks in question are characterised as follows. There are twice as many left branches as right branches, and at any given intermediate point the number of left branches encountered is less than or equal to twice the number of right branches encountered. A walk characterised as such will begin with a right branch. Moreover, at the first point, whether it be some intermediate point or at the final branch, the number of left branches stops being less than twice the number of right branches and instead is equal to it. By Lemma 2.5.6, the pairing we encounter at this branch is (a, b). If this is an intermediate point, then we must take a right branch to preserve the inequality. This continues on until we come to the last branch that also has the pair (a, b). Thus, such a walk will satisfy the criteria in this lemma.

Conversely, a walk described as in this lemma begins with a right branch, and when it attains a (a, b) pair again, we must have twice as many left branches as right branches, by Lemmas 2.5.5 and 2.5.6. Then we take another right branch, and so on. This fits the characterisation we have given. Thus, it has become a question of counting the number of walks that are characterised as in the start of the proof. By using the definitions of $S(n)$ and $B(n)$ and Proposition 2.5.4, we can see that this is $B(n)$. □

Lemma 2.5.8. Let $a, b \geq 1$ with $\gcd(a, b) = 1$. Consider a walk in $\mathbf{T}_{(a,b)}$ that starts with a left branch and ends at a (a, b) pair with no intermediate (a, b) pair. The parent of a in the last pair is $|a - b|$.

Proof. We prove our result by induction on n where $3n$ is the length of the walk in question. For $n = 1$, we have the sequence

$$a, b, |a - b|, a, b.$$

Suppose Lemma 2.5.8 holds for some $n \in \mathbb{N}$. We want to show Lemma 2.5.8 holds for $n+1$. So consider a walk of length $3n + 3$ that starts at the root (a, b) where the first branch is a left branch and ends at a pair (a, b). Without loss of generality, we may assume that it does not attain any intermediate pair (a, b) before the final (a, b). Thus, we wish to show that the third last term in the sequence is $|a - b|$. Suppose for a contradiction that the third last term in the sequence is not $|a - b|$. Then the third last term must $a + b$. Since $b = |(a + b) - b|$, the final branch must be a left branch. Also since $a < a + b$, the second last branch must also be a left branch. Thus, somewhere in the walk there must be a right branch immediately followed by two left branches. As in the proofs of Proposition 2.5.3 and Lemma 2.5.5 such a configuration can be dropped out without affecting the pairing on the last branch (a, b). But then this smaller walk would not have any intermediate (a, b) pairs and the third last term would still be $a + b$, which isn't possible by our inductive assumption. Therefore, the third last term of the original walk had to have been $|a - b|$ as well. Thus, we have our result. $\square$

Corollary 2.5.9. Take the tree $\mathbf{T}_{(|a-b|,a)}$ for some coprime pair (a, b). Then for all walks to an (a, b), there must exist a pair (a, b) in the walk such that the parent of that specific a is $|a - b|$ in the walk.

Proof. Take such a walk to a pair (a, b) and suppose there exists no pair (a, b) in that walk such that the parent of that specific a in the walk is $|a-b|$. Suppose we lengthen the walk in front by adding a node b to be the parent of $|a - b|$ and then another node a to be the parent of b. This will give a walk that starts with two left branches if $a < b$ or a walk that starts with a left branch and then a right branch if $a \geq b$. In either case, we have a walk that contradicts Lemma 2.5.8. Therefore, the result follows. $\square$

Proposition 2.5.10. Let a and b be coprime integers and $n \in \mathbb{N}$. In $\mathbf{T}_{(a,b)}$, the number of (a, b) pairs at depth $3n$ in the tree that can be attained by a walk not containing an $(|a - b|, a)$ pair is equal to $B(n)$.

Proof. Combine Lemmas 2.5.7 and 2.5.8 and Corollary 2.5.9. $\square$

Proposition 2.5.11. Take a coprime pair (a, b) that is not the $(1, 1)$ pair and suppose that $SW_{(1,1)}(a, b) = k$. Then for all $n \geq \lfloor \frac{k}{3} \rfloor$ we have

$$A_{(a,b)}(n) = \sum_{i=\lfloor \frac{k-1}{3} \rfloor}^{n} A_{(|a-b|,a)}(i)B(n - i)$$

if either a is even or b is even, and

$$A_{(a,b)}(n) = \sum_{i=\lfloor \frac{k-1}{3} \rfloor}^{n-1} A_{(|a-b|,a)}(i)B(n - 1 - i)$$

if a and b are both odd.

Proof. By Proposition 2.5.3 any walk in the Fibonacci tree that starts at the root $(1, 1)$ and ends at the pair (a, b) must contain the pair $(|a - b|, a)$. Consider the last place in a given walk where this pair occurs and say it is at depth $3i+m$ in the tree where $0 \leq m \leq 2$

(where m depends on the parity of (a, b)). Then by Corollary 2.5.9 the next element in the walk is b and, by Proposition 2.5.10, this gives rise to $B(n-i)$ pairs of (a, b) at depth $3n+m$ or $3(n+1)+m$ in the tree (depending on the parity of a and b). Conversely, every pair $(|a-b|, a)$ that occurs at an intermediate point at depth $3i+m$ in the tree gives rise to $B(n-i)$ walks to pairs of (a, b) at depth $3n+m$ or $3(n+1)+m$ in the tree. The summation starts at $i = \lfloor \frac{k-1}{3} \rfloor$ since $SW_{(1,1)}(|a-b|, a) = k-1$ and so the pair $(|a-b|, a)$ occurs at depth $3\lfloor \frac{k-1}{3} \rfloor + m$ in the tree. Thus, the formula follows. □

Corollary 2.5.12. For all $n \in \mathbb{N} \cup \{0\}$, we have

$$A_{(1,2)}(n) = \frac{A_{(1,1)}(n+1) - B(n+1)}{4}.$$

Proof. By Proposition 2.5.11, we have for all $n \in \mathbb{N} \cup \{0\}$

$$A_{(1,2)}(n) = \sum_{i=0}^{n} A_{(1,1)}(i) B(n-i). \tag{2.5.1}$$

All walks down to a $(1, 1)$ pair at depth $3n$ in the tree must satisfy exactly one of the following two conditions. Either for all other $(1, 1)$ pairs it takes a right branch immediately afterwards, or there exists a first $(1, 1)$ pair where the walk takes a left branch immediately afterwards, consequently ending up immediately at a choice of four $(1, 1)$ pairs. Thus, for all $n \in \mathbb{N} \cup \{0\}$ we have

$$A_{(1,1)}(n) = B(n) + \sum_{i=0}^{n-1} B(i) \cdot 4 \cdot A_{(1,1)}(n-1-i).$$

Relabelling the index in the summation gives

$$A_{(1,1)}(n) = B(n) + 4 \sum_{i=0}^{n-1} A_{(1,1)}(i) B(n-1-i) \tag{2.5.2}$$

for all $n \in \mathbb{N} \cup \{0\}$. Substituting in (2.5.1), we have for all $n \in \mathbb{N}$ that

$$A_{(1,1)}(n) = B(n) + 4 \cdot A_{(1,2)}(n-1).$$

Thus, we have our result. □

Corollary 2.5.13. Take two pairs of coprime positive integers (a, b) and (c, d) and suppose that $SW_{(1,1)}(a, b) = SW_{(1,1)}(c, d)$. Then we have

$$A_{(a,b)}(n) = A_{(c,d)}(n)$$

for all $n \in \mathbb{N} \cup \{0\}$.

Proof. We can prove this by induction on the number of branches in the shortest walks, using the result of Proposition 2.5.11. $\square$

Corollary 2.5.14. For all $n \in \mathbb{N} \cup \{0\}$, we have

$$A_{(2,1)}(n) = A_{(2,3)}(n) = A_{(1,1)}(n + 1) - 4A_{(1,1)}(n).$$

Proof. By Corollary 2.5.13 it suffices to prove that

$$A_{(2,1)}(n) = A_{(1,1)}(n + 1) - 4 \cdot A_{(1,1)}(n)$$

since the pairs $SW_{(1,1)}(2, 1) = SW_{(1,1)}(2, 3) = 2$. First, consider all $(2, 1)$ pairs at depth $3n + 2$ in the tree. If we take an immediate left branch we encounter $(1, 1)$ pairs at depth $3n + 3$ in the tree. Now consider all $(1, 1)$ pairs at depth $3n$ in the tree. The walks to these $(1, 1)$ pairs must either have the element 0 or the element 2 immediately before the $(1, 1)$ pair. There are $4A(n)$ pairs $(1, 1)$ of the former type since following backwards along the walk will give us 4 $(1, 1)$ pairs at depth $3n$ in the tree. Therefore, the number of $(1, 1)$ pairs with a walk that has the element 2 immediately before the $(1, 1)$ pair is $A(n+1)-4A(n)$. Since the second and third last elements of these walks form $(2, 1)$ pairs we have, by our observation that all $(2, 1)$ pairs have a $(1, 1)$ immediately beneath them, which is our result. $\square$

Lemma 2.5.15. Take a coprime pair (a, b) that is not the $(1, 1)$ pair and suppose that $SW_{(1,1)}(a, b) = k \geq 3$. Suppose the last five numbers in the corresponding sequence of the shortest walk, including the last two numbers a and b, are

$$a_0, a_1, a_2, a, b.$$

Case 1. If a is odd, then for all $n \geq \lfloor \frac{k}{3} \rfloor$, we have

$$A_{(a,b)}(n) = A_{(a_1,a_2)}(n) - A_{(a_0,a_1)}(n)$$

Case 2. If a is even (and hence b is odd), then for all $n \geq \lfloor \frac{k}{3} \rfloor$, we have

$$A_{(a,b)}(n) = A_{(a_1,a_2)}(n + 1) - A_{(a_0,a_1)}(n).$$

Proof. We prove this by induction on $SW_{(1,1)}(a,b)$. First, suppose $SW_{(1,1)}(a,b) = 3$. Then both a and b are odd. By Proposition 2.5.11, we have

$$A_{(a,b)}(n) = \sum_{i=0}^{n-1} A_{(|a-b|,a)}(i)B(n-1-i)$$

for all $n \in \mathbb{N}$, where $(|a-b|, a)$ is a pair satisfying $SW_{(1,1)}(|a-b|, a) = 2$. There are only two pairs that $(|a-b|, a)$ can be: $(2,1)$ or $(2,3)$. Thus, by Corollary 2.5.14, we have

$$\begin{aligned} A_{(a,b)}(n) &= \sum_{i=0}^{n-1} (A_{(1,1)}(i+1) - 4 \cdot A_{(1,1)}(i))B(n-1-i) \\ &= \sum_{i=0}^{n-1} A_{(1,1)}(i+1)B(n-1-i) - 4\sum_{i=0}^{n-1} A_{(1,1)}(i)B(n-1-i). \end{aligned}$$

By Corollary 2.5.13 and (2.5.2), we have

$$\begin{aligned} A_{(a,b)}(n) &= A_{(1,2)}(n) - B(n) - 4\sum_{i=0}^{n-1} A_{(1,1)}(i)B(n-1-i) \\ &= A_{(1,2)}(n) - A_{(1,1)}(n). \end{aligned}$$

Thus, Lemma 2.5.15 holds for all $n \in \mathbb{N}$ for the pair (a,b) since $SW_{(1,1)}(1,2) = 1$ and $SW_{(1,1)}(1,1) = 0$. Suppose the proposition holds for pairs that have a shortest walk of length $k-1$ for some $k \geq 4$ and suppose we want to show it holds for pairs with the shortest walks of lengths k. Let (a,b) be a pair with $SW_{(1,1)}(a,b) = k$. Let the last six elements of the shortest walk to (a,b) be

$$a_0, a_1, a_2, |a-b|, a, b.$$

First, suppose that both a and b are odd. Then, by Proposition 2.5.11, we have

$$A_{(a,b)}(n) = \sum_{i=\lfloor \frac{k-1}{3} \rfloor}^{n-1} A_{(|a-b|,a)}(i)B(n-1-i)$$

for all $n \geq \frac{k}{3}$. By our inductive hypothesis, we have

$$A_{(|a-b|,a)}(i) = A_{(a_1,a_2)}(i+1) - A_{(a_0,a_1)}(i)$$

for all $i \geq \lfloor \frac{k-1}{3} \rfloor$ since $|a-b|$ is even and a is odd. Then we have

$$\begin{aligned}
A_{(a,b)}(n) &= \sum_{i=\lfloor \frac{k-1}{3} \rfloor}^{n-1} (A_{(a_1,a_2)}(i+1) - A_{(a_0,a_1)}(i))B(n-1-i) \\
&= \sum_{i=\lfloor \frac{k-1}{3} \rfloor}^{n-1} A_{(a_1,a_2)}(i+1)B(n-1-i) - \sum_{i=\lfloor \frac{k-1}{3} \rfloor}^{n-1} A_{(a_0,a_1)}(i)B(n-1-i) \\
&= \sum_{i=\lfloor \frac{k-1}{3} \rfloor+1}^{n} A_{(a_1,a_2)}(i)B(n-i) - \sum_{i=\lfloor \frac{k-1}{3} \rfloor}^{n-1} A_{(a_0,a_1)}(i)B(n-1-i) \\
&= \sum_{i=\lfloor \frac{k-3}{3} \rfloor+1}^{n} A_{(a_1,a_2)}(i)B(n-i) - \sum_{i=\lfloor \frac{k-4}{3} \rfloor+1}^{n-1} A_{(a_0,a_1)}(i)B(n-1-i)
\end{aligned}$$

since $\lfloor \frac{k-3}{3} \rfloor = 1 + \lfloor \frac{k-4}{3} \rfloor$ since $3|k$. Thus, by Proposition 2.5.11, we have

$$A_{(a,b)}(n) = A_{(a_2,|a-b|)}(n) - B\left(n - \left\lfloor \frac{k-3}{3} \right\rfloor\right) - \left(A_{(a_1,a_2)}(n) - B\left(n-1-\left\lfloor \frac{k-4}{3} \right\rfloor\right)\right)$$

since $SW_{(1,1)}(a_2, |a-b|) = k-2$ and $SW_{(1,1)}(a_1, a_2) = k-3$. Thus, we get our result

$$A_{(a,b)}(n) = A_{(a_2,|a-b|)}(n) - A_{(a_1,a_2)}(n)$$

for all $n \geq \frac{k}{3}$.

By a similar argument, if a is odd and b is even, then

$$A_{(a,b)}(n) = A_{(a_2,|a-b|)}(n) - A_{(a_1,a_2)}(n).$$

Also, by a similar argument, if a is even and b is odd, then

$$A_{(a,b)}(n) = A_{(a_2,|a-b|)}(n+1) - A_{(a_1,a_2)}(n).$$

$\square$

2.6 Proof of Asymptotics for $A_{(1,1)}(n)$

We establish our asymptotic results concerning $A(n) = A_{1,1}(n)$ here.

Theorem 1.1.8. Letting $A_{(1,1)}(n)$ be defined as above, we have

$$\left(\frac{3^5\cdot(27/4)^n}{4\sqrt{3\pi}n^{3/2}}\right)\left(1-\frac{1387}{72n}\right) < A_{(1,1)}(n) < \left(\frac{3^5\cdot(27/4)^n}{4\sqrt{3\pi}n^{3/2}}\right)\left(1-\frac{1387}{72n}+\frac{5548}{9n^2}\right)$$

for $n \geq 100$.

First, we prove two lemmas.

Lemma 2.6.1. For all $n \in \mathbb{N}$, we have

$$\sum_{k=0}^{n} B(k)B(n-k) = S(n+1).$$

Proof. In [7] we have the combinatorial identity

$$\sum_{k=0}^{n}\binom{tk+r}{k}\binom{tn-tk+s}{n-k}\frac{r}{tk+r}\cdot\frac{s}{tn-tk+s} = \binom{tn+r+s}{n}\frac{r+s}{tn+r+s}$$

valid for all $n \in \mathbb{N}$ and all $r, s, t \in \mathbb{R}$. Substituting in $t = 3$, $r = 1$, and $s = 1$ gives us

$$\sum_{k=0}^{n}\binom{3k+1}{k}\binom{3n-3k+1}{n-k}\frac{1}{3k+1}\cdot\frac{1}{3n-3k+1} = \binom{3n+2}{n}\frac{2}{3n+2}. \tag{2.6.1}$$

Also for all $n \in \mathbb{N}\cup\{0\}$, we have

$$B(n) = \frac{1}{2n+1}\binom{3n}{n} = \frac{1}{2n+1}\frac{(3n)!}{n!(2n)!} = \frac{(3n)!}{n!(2n+1)!} = \frac{1}{3n+1}\binom{3n+1}{n}.$$

Thus, by (2.6.1) and Proposition 2.3.1 we have for all $n \in \mathbb{N}$

$$\sum_{k=0}^{n} B(k)B(n-k) = S(n+1).$$

□

Lemma 2.6.2. For all $n \in \mathbb{N}\cup\{0\}$, we have

$$A(n+2) - 16\cdot A(n+1) + 64\cdot A(n) = B(n+2) + 4\cdot S(n+2).$$

Proof. By Proposition 2.5.11, we have, for all $n \in \mathbb{N} \cup \{0\}$,

$$A_{(2,1)}(n) = \sum_{i=0}^{n} A_{(1,2)}(i) B(n-i).$$

By Corollaries 2.5.12 and 2.5.14, we therefore have

$$A(n+1) - 4 \cdot A(n) = \sum_{i=0}^{n} \frac{A(i+1) - B(i+1)}{4} B(n-i).$$

Thus, we obtain the following:

$$A(n+1) - 4 \cdot A(n) = \frac{1}{4} \sum_{i=1}^{n+1} A(i) B(n+1-i) - \frac{1}{4} \sum_{i=0}^{n} B(i+1) B(n-i).$$

By (2.5.2), we thus have

$$\begin{aligned} A(n+1) - 4 \cdot A(n) &= \frac{1}{4} \left(\frac{A(n+2) - B(n+2)}{4} - B(n+1) \right) - \frac{1}{4} \sum_{i=0}^{n} B(i+1) B(n-i) \\ &= \frac{A(n+2) - B(n+2)}{16} - \frac{B(n+1)}{4} - \frac{1}{4} \sum_{i=0}^{n} B(i+1) B(n-i). \end{aligned} \tag{2.6.2}$$

Lemma 2.6.1 then gives

$$\sum_{i=0}^{n} B(i+1) B(n-i) = \sum_{i=1}^{n+1} B(i) B(n+1-i) = S(n+2) - B(n+1).$$

Thus, by (2.6.2) we have for all $n \in \mathbb{N}$

$$\begin{aligned} A(n+1) - 4 \cdot A(n) &= \frac{A(n+2) - B(n+2)}{16} - \frac{B(n+1)}{4} - \frac{1}{4}(S(n+2) - B(n+1)) \\ &= \frac{A(n+2) - B(n+2)}{16} - \frac{S(n+2)}{4}. \end{aligned}$$

Thus, for all $n \in \mathbb{N}$, this proves our result. □

Proposition 2.6.3. For all $n \in \mathbb{N}$, we have

$$A(n) = \frac{243 \cdot (27/4)^n}{4\sqrt{3\pi} n^{3/2}} (1 + o(1)).$$

Proof. For all $n \in \mathbb{N}$, by Proposition 2.3.1, we have

$$\begin{aligned}
\frac{B(n)}{S(n)} &= \frac{3n-1}{2(2n+1)} \cdot \frac{\binom{3n}{n}}{\binom{3n-1}{n-1}} \\
&= \frac{3n-1}{4n+2} \cdot \frac{(3n)!}{n!(2n)!} \cdot \frac{(n-1)!(2n)!}{(3n-1)!} \\
&= \frac{3n-1}{4n+2} \cdot \frac{3n}{n} \\
&= \frac{9n-3}{4n+2}.
\end{aligned}$$

Thus

$$\lim_{n\to\infty} \frac{B(n)}{S(n)} = \frac{9}{4}. \tag{2.6.3}$$

Thus, by Corollary 2.3.3, we have

$$B(n) \sim \frac{9 \cdot (27/4)^n}{4\sqrt{27\pi}n^{3/2}}. \tag{2.6.4}$$

Thus, by Corollary 2.3.3, we have

$$\begin{aligned}
\frac{B(n+2)}{S(n)} + \frac{4 \cdot S(n+2)}{S(n)} &\sim \frac{(27/4)^2 \cdot B(n)}{S(n)} + 4 \cdot (27/4)^2 \\
&\sim \frac{9}{4} \cdot (27/4)^2 + 4 \cdot (27/4)^2 \\
&= (5 \cdot 3^3/2^3)^2.
\end{aligned}$$

Thus, by Lemma 2.6.2, we have

$$\lim_{n\to\infty} \frac{A(n+2) - 16 \cdot A(n+1) + 64 \cdot A(n)}{S(n)} = (5 \cdot 3^3/2^3)^2.$$

By Proposition 2.4.6, we must therefore have

$$\lim_{n\to\infty} \frac{(27/4)^2 \cdot A(n) - 16 \cdot (27/4) \cdot A(n) + 64 \cdot A(n)}{S(n)} = (5 \cdot 3^3/2^3)^2.$$

or

$$\lim_{n\to\infty} \frac{5^2 \cdot A(n)}{2^4 \cdot S(n)} = (5 \cdot 3^3/2^3)^2.$$

Thus, we have

$$\lim_{n\to\infty}\frac{A(n)}{S(n)}=\frac{(5\cdot 3^3/2^3)^2}{5^2/2^4}=3^6/2^2. \tag{2.6.5}$$

By Corollary 2.3.3, we thus have

$$A(n)\sim\frac{3^6\cdot(27/4)^n}{4\sqrt{27\pi}n^{3/2}}.$$

$\square$

Note 2.6.4. For the rest of this section let

$$f(n):=\frac{25\cdot(27/4)^n}{12\sqrt{3\pi}n^{3/2}}.$$

Proposition 2.6.5. For all $n\in\mathbb{N}$, $n\geq 100$, we have

$$f(n+2)\left(1-\frac{23}{360(n+2)}+\frac{69}{500(n+2)^2}\right)<B(n+2)+4\cdot S(n+2)$$

and

$$B(n+2)+4\cdot S(n+2)<f(n+2)\left(1-\frac{23}{360(n+2)}+\frac{23}{125(n+2)^2}\right).$$

Proof. For all $n\in\mathbb{N}$, $n\geq 100$, we have, by Corollary 2.3.3:

$$\begin{aligned}
B(n+2)+4\cdot S(n+2) &< \frac{\sqrt{3}\cdot(27/4)^{n+2}}{4\sqrt{\pi}(n+2)^{3/2}}\left(1-\frac{43}{72(n+2)}+\frac{21}{20(n+2)^2}\right)\\
&\quad+\frac{4\cdot(27/4)^{n+2}}{3\sqrt{3\pi}(n+2)^{3/2}}\left(1+\frac{17}{72(n+2)}+\frac{5}{6(n+2)^2}\right)\\
&=\frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+2)^{3/2}}\left(1-\frac{23}{360(n+2)}+\frac{1367}{1500(n+2)^2}\right)\\
&\leq\frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+2)^{3/2}}\left(1-\frac{23}{360(n+2)}+\frac{23}{25(n+2)^2}\right),
\end{aligned}$$

and

$$\begin{aligned}
B(n+2)+4\cdot S(n+2) &> \frac{\sqrt{3}\cdot(27/4)^{n+2}}{4\sqrt{\pi}(n+2)^{3/2}}\left(1-\frac{43}{72(n+2)}+\frac{1}{4(n+2)^2}\right)\\
&\quad+\frac{4\cdot(27/4)^{n+2}}{3\sqrt{3\pi}(n+2)^{3/2}}\left(1+\frac{17}{72(n+2)}+\frac{3}{40(n+2)^2}\right)\\
&=\frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+2)^{3/2}}\left(1-\frac{23}{360(n+2)}+\frac{69}{500(n+2)^2}\right).
\end{aligned}$$

$\square$

Definition 2.6.6. Define $D : \mathbb{N} \cup \{0\} \to \mathbb{Z}$ as $D(0) = -3$ and

$$D(n+1) = 8 \cdot D(n) + B(n+2) + 4 \cdot S(n+2)$$

for all $n \in \mathbb{N} \cup \{0\}$. It can be verified with the help of Lemma 2.6.2 that, for all $n \in \mathbb{N} \cup \{0\}$, we have

$$D(n) = A(n+1) - 8 \cdot A(n). \tag{2.6.6}$$

Lemma 2.6.7. We have

$$D(n) \sim \frac{-405\sqrt{3} \cdot (27/4)^n}{16\sqrt{\pi} n^{3/2}}.$$

Proof. The lemma can be verified with (2.6.6) and Proposition 2.6.3. $\square$

Proposition 2.6.8. For all $n \geq 100$, we have

$$\left(-\frac{405\sqrt{3} \cdot (27/4)^n}{16\sqrt{\pi} n^{3/2}}\right)\left(1 - \frac{4019}{360n} + \frac{207}{n^2}\right) < D(n) < \left(-\frac{405\sqrt{3} \cdot (27/4)^n}{16\sqrt{\pi} n^{3/2}}\right)\left(1 - \frac{4019}{360n}\right).$$

Proof of the upper bound for $D(n)$. Note that

$$-\frac{405\sqrt{3} \cdot (27/4)^n}{16\sqrt{\pi} n^{3/2}} = -\frac{729 f(n)}{20}.$$

Suppose, to obtain a contradiction, that for some $n \geq 100$, we have

$$D(n) \geq -\frac{729 f(n)}{20}\left(1 - \frac{4019}{360n}\right).$$

The right-hand side of the above inequality is a transcendental number for all $n \in \mathbb{N}$, and since $D(n) \in \mathbb{Z}$ for all $n \in \mathbb{N}$, we must therefore have that

$$D(n) > -\frac{729 f(n)}{20}\left(1 - \frac{4019}{360n}\right).$$

Fix such an $n \in \mathbb{N}$. Thus, we have

$$D(n) = -\frac{729 f(n)}{20}\left(1 - \frac{C}{n}\right).$$

where $C > \frac{4019}{360}$ depends on n. Note that since we have fixed n, we have fixed C and so can be treated as a constant. Then, by Proposition 2.6.5, we have

$$\begin{aligned}
D(n+1) &= 8 \cdot D(n) + B(n+2) + 4 \cdot S(n+2) \\
&> \left(-\frac{405\sqrt{3} \cdot (27/4)^n}{2\sqrt{\pi}(n+1)^{3/2}}\right)\left(\frac{n+1}{n}\right)^{3/2}\left(1 - \frac{C}{n}\right) \\
&\quad + \frac{25 \cdot (27/4)^{n+2}}{12\sqrt{3\pi}(n+1)^{3/2}}\left(\frac{n+1}{n+2}\right)^{3/2}\left(1 - \frac{23}{360(n+2)}\right).
\end{aligned}$$

We have

$$\begin{aligned}
\left(\frac{n+1}{n}\right)^{3/2} &= \left(\frac{n}{n+1}\right)^{-3/2} \\
&= \left(1 - \frac{1}{n+1}\right)^{-3/2}. \qquad (2.6.7)
\end{aligned}$$

For all $0 < x \leq \frac{1}{101}$, we have, by the binomial theorem,

$$\begin{aligned}
(1-x)^{-3/2} &= 1 + \frac{3x}{2} + \frac{15x^2}{8} + \cdots + \frac{\frac{3}{2} \cdot \frac{5}{2} \cdots \frac{2k+1}{2} x^k}{k!} + \ldots \qquad (2.6.8) \\
&< 1 + \frac{3x}{2} + \frac{19x^2}{10}. \qquad (2.6.9)
\end{aligned}$$

Let $f(x) = (x+1)^{-3/2}$ and $g(x) = 1 - \frac{3x}{2}$. We have $f(0) = g(0) = 1$, and for all $x > 0$, we have

$$f'(x) = \frac{-3}{2}(x+1)^{-5/2} > \frac{-3}{2} = g'(x),$$

where $f'(x)$ and $g'(x)$ are the derivatives of $f(x)$ and $g(x)$ respectively. Thus, for all $x > 0$, we must have $f(x) > g(x)$ or

$$(x+1)^{-3/2} > 1 - \frac{3x}{2}. \qquad (2.6.10)$$

Thus, we have

$$
\begin{aligned}
D(n+1) &> \left(-\frac{405\sqrt{3}\cdot(27/4)^n}{2\sqrt{\pi}(n+1)^{3/2}}\right)\left(1+\frac{3}{2(n+1)}+\frac{19}{10(n+1)^2}\right)\left(1-\frac{C}{(n+1)}\right) \\
&\quad + \frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+1)^{3/2}}\left(1-\frac{3}{2(n+1)}\right)\left(1-\frac{23}{360(n+1)}\right) \\
&= \left(-\frac{405\sqrt{3}\cdot(27/4)^n}{2\sqrt{\pi}(n+1)^{3/2}}\right)\left(1+\frac{3-2C}{2(n+1)}+\frac{19-15C}{10(n+1)^2}-\frac{19C}{10(n+1)^3}\right) \\
&\quad + \frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+1)^{3/2}}\left(1-\frac{563}{360(n+1)}+\frac{11}{48(n+1)^2}-\frac{1}{5(n+1)^3}\right) \\
&> \left(-\frac{405\sqrt{3}\cdot(27/4)^n}{2\sqrt{\pi}(n+1)^{3/2}}\right)\left(1+\frac{3-2C}{2(n+1)}\right) \\
&\quad + \frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+1)^{3/2}}\left(1-\frac{563}{360(n+1)}\right).
\end{aligned}
$$

Thus, we have

$$D(n+1) > -\frac{729f(n+1)}{20}\left(1-\frac{2304C-4019}{1944(n+1)}\right).$$

We deduce that

$$r := \frac{2304C-4019}{1944C} > 1.$$

Let $C_1 := rC$. Repeating the argument with C_1 in place of C and $n+1$ in place of n gives us

$$D(n+2) = \left(-\frac{405\sqrt{3}\cdot(27/4)^{n+2}}{16\sqrt{\pi}(n+2)^{3/2}}\right)\left(1-\frac{C_2}{(n+2)}\right),$$

where

$$C_2 := \frac{2304C_1-4019}{1944}.$$

Thus, $C_2 > rC_1 = r^2C$. Repeating the argument as many times as necessary, we thus have, for all $k \in \mathbb{N}$,

$$D(n+k) = \left(-\frac{405\sqrt{3}\cdot(27/4)^{n+k}}{16\sqrt{\pi}(n+k)^{3/2}}\right)\left(1-\frac{C_k}{(n+k)}\right)$$

where $C_k > r^k C$. This leads to

$$\lim_{k\to\infty} \frac{r^k}{k} = 0,$$

which does not hold since $r > 1$, a contradiction. Thus, we have our first desired inequality for all $n \geq 100$. $\square$

Proof of the lower bound for $D(n)$. Suppose for a contradiction that there exists $n \geq 100$ such that

$$D(n) = -\frac{729f(n)}{20}\left(1 - \frac{4019}{360n} + \frac{C}{n}\right).$$

where $\frac{207}{n} \leq C$. We can derive that

$$\frac{4019}{360n} < \frac{4019}{360(n+1)} + \frac{C}{32(n+1)} \tag{2.6.11}$$

and

$$\frac{2228257}{97200(n+1)} < \frac{C}{9}. \tag{2.6.12}$$

By Proposition 2.6.5, we have

$$D(n+1) < \left(-\frac{405\sqrt{3}\cdot(27/4)^n}{2\sqrt{\pi}(n+1)^{3/2}}\right)\left(\frac{n+1}{n}\right)^{3/2}\left(1 - \frac{4019}{360n} + \frac{C}{n}\right)$$
$$+ \frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+1)^{3/2}}\left(\frac{n+1}{n+2}\right)^{3/2}\left(1 - \frac{23}{360(n+2)} + \frac{23}{125(n+2)^2}\right).$$

By (2.6.7), (2.6.8), and (2.6.11), we have

$$D(n+1) < \left(-\frac{405\sqrt{3}\cdot(27/4)^n}{2\sqrt{\pi}(n+1)^{3/2}}\right)\left(1 + \frac{3}{2(n+1)}\right)\left(1 - \frac{4019}{360(n+1)} + \frac{C}{(n+1)} - \frac{C}{32(n+1)}\right)$$
$$+ \frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+1)^{3/2}}\left(1 + \frac{1}{(n+1)}\right)^{-3/2}\left(1 - \frac{23}{360(n+1)} + \frac{23}{360(n+1)^2} + \frac{23}{125(n+1)^2}\right)$$

For all $0 < x < \frac{8}{9}$, we have, by the binomial theorem,

$$(1+x)^{-3/2} = 1 - \frac{3x}{2} + \frac{15x^2}{8} - \cdots + \frac{(-1)^k \frac{3}{2}\cdot\frac{5}{2}\cdots\frac{2k+1}{k}x^k}{k!}.$$

For all $k \geq 3$, k odd, we have

$$\begin{aligned}
\frac{(-1)^k \frac{3}{2}\cdot\frac{5}{2}\cdots\frac{2k+1}{k}x^k}{k!} + \frac{(-1)^{k+1}\frac{3}{2}\cdot\frac{5}{2}\cdots\frac{2k+1}{k}x^{k+1}}{(k+1)!} &= -\frac{\frac{3}{2}\cdot\frac{5}{2}\cdots\frac{2k+1}{2}x^k}{k!} + \frac{\frac{3}{2}\cdot\frac{5}{2}\cdot\ldots\cdot\frac{2k+3}{2}x^{k+1}}{(k+1)!} \\
&= \frac{\frac{3}{2}\cdot\frac{5}{2}\cdots\frac{2k+1}{k}x^k}{k!}\left(-1+\frac{(2k+3)x}{2k+2}\right) \\
&\leq \frac{\frac{3}{2}\cdot\frac{5}{2}\cdots\frac{2k+1}{k}x^k}{k!}\left(-1+\frac{9x}{8}\right) \\
&< 0.
\end{aligned}$$

Thus,

$$(1+x)^{-3/2} < 1 - \frac{3x}{2} + \frac{15x^2}{8}. \tag{2.6.13}$$

Thus

$$\begin{aligned}
D(n+1) < &\left(-\frac{405\sqrt{3}\cdot(27/4)^n}{2\sqrt{\pi}(n+1)^{3/2}}\right)\left(1+\frac{3}{2(n+1)}\right)\left(1-\frac{4019}{360(n+1)}+\frac{C}{(n+1)}-\frac{C}{32(n+1)}\right) \\
&+\frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+1)^{3/2}}\left(1-\frac{3}{2(n+1)}+\frac{15}{8(n+1)^2}\right) \\
&\left(1-\frac{23}{360(n+1)}+\frac{23}{360(n+1)^2}+\frac{23}{125(n+1)^2}\right) \\
&< \left(-\frac{405\sqrt{3}\cdot(27/4)^n}{2\sqrt{\pi}(n+1)^{3/2}}\right)\left(1-\frac{4019}{360(n+1)}+\frac{31C+48}{32(n+1)}+\frac{1395C-18236}{960(n+1)^2}\right) \\
&+\frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+1)^{3/2}}\left(1-\frac{563}{360(n+1)}+\frac{39937}{18000(n+1)^2}-\frac{3933}{8000(n+1)^3}+\frac{2231}{4800(n+1)^4}\right) \\
&< \left(-\frac{405\sqrt{3}\cdot(27/4)^n}{2\sqrt{\pi}(n+1)^{3/2}}\right)\left(1-\frac{4019}{360(n+1)}+\frac{31C+48}{32(n+1)}+\frac{1395C-18236}{960(n+1)^2}\right) \\
&+\frac{25\cdot(27/4)^{n+2}}{12\sqrt{3\pi}(n+1)^{3/2}}\left(1-\frac{563}{360(n+1)}+\frac{39937}{18000(n+1)^2}\right) \\
&= \left(-\frac{405\sqrt{3}\cdot(27/4)^{n+1}}{16\sqrt{\pi}(n+1)^{3/2}}\right)\left(1-\frac{4019}{360(n+1)}+\frac{31C}{27(n+1)}+\frac{167400C-2228257}{97200(n+1)^2}\right).
\end{aligned}$$

By (2.6.12), we have

$$
\begin{aligned}
D(n+1) &< \left(-\frac{405\sqrt{3}\cdot(27/4)^{n+1}}{16\sqrt{\pi}(n+1)^{3/2}}\right)\left(1-\frac{4019}{360(n+1)}+\frac{31C}{27(n+1)}-\frac{2228257}{97200(n+1)^2}\right)\\
&< \left(-\frac{405\sqrt{3}\cdot(27/4)^{n+1}}{16\sqrt{\pi}(n+1)^{3/2}}\right)\left(1-\frac{4019}{360(n+1)}+\frac{31C}{27(n+1)}-\frac{C}{9(n+1)}\right)\\
&= \left(-\frac{405\sqrt{3}\cdot(27/4)^{n+1}}{16\sqrt{\pi}(n+1)^{3/2}}\right)\left(1-\frac{4019}{360(n+1)}+\frac{28C}{27(n+1)}\right).
\end{aligned}
$$

From the fact that $\frac{207}{n} \leq C$ we can deduce that $\frac{207}{(n+1)} \leq \frac{28C}{27}$. Thus, we can repeat the above argument with $\frac{28C}{27}$ in place of C and $n+1$ in place of n to derive that

$$
D(n+2) < \left(-\frac{405\sqrt{3}\cdot(27/4)^{n+2}}{16\sqrt{\pi}(n+2)^{3/2}}\right)\left(1-\frac{4019}{360(n+2)}+\frac{\left(\frac{28}{27}\right)^2 C}{(n+2)}\right).
$$

Repeating the argument as many times as necessary, we get that, for all $k \in \mathbb{N}$,

$$
D(n+k) < \left(-\frac{405\sqrt{3}\cdot(27/4)^{n+k}}{16\sqrt{\pi}(n+k)^{3/2}}\right)\left(1-\frac{4019}{360(n+k)}+\frac{\left(\frac{10}{9}\right)^k C}{(n+k)}\right).
$$

We know that

$$
\lim_{k\to\infty} \frac{-D(n+k)16\sqrt{\pi}(n+k)^{3/2}}{405\sqrt{3}\cdot(27/4)^{n+k}} = 1
$$

so that

$$
\lim_{k\to\infty} \frac{-4019}{360(n+k)} - \frac{\left(\frac{28}{27}\right)^k C}{n+k} = 0.
$$

Thus

$$
\lim_{k\to\infty} \frac{\left(\frac{28}{27}\right)^k}{n+k} = 0
$$

so that

$$
\lim_{k\to\infty} \frac{\left(\frac{28}{27}\right)^k}{k} = 0,
$$

which does not hold, a contradiction. Thus, we have our second desired inequality for all $n \geq 100$. $\square$

Proof of Theorem 1.1.8. The proof of these inequalities follows the same procedure as in the proof of Proposition 2.6.8. We prove by contradiction in the following way. We first assume that the desired upper bound does not hold for some value of $n \geq 100$. Using (2.6.6) and the lower bound for $D(n)$ in Proposition 2.6.8, we derive a lower bound for $A(n+1)$. Again, we see that $A(n+1)$ does not satisfy the desired upper bound given in the Theorem so that we can repeat the argument to get a lower bound for $A(n+2)$ and so on. As $k \to \infty$, we see that the error term for $A(n+k)$ grows too big, overwhelming the main term, contradicting Proposition 2.6.3. The proof for the lower bound works the same way, using (2.6.6) and the upper bound for $D(n)$. □

2.7 Proof of Asymptotics for $A_{(a,b)}(n)$

Finally, we establish our asymptotic results for other coprime pairs $A_{(a,b)}(n)$ for all coprime ordered pairs (a, b).

Theorem 1.1.9. For all coprime pairs (a, b), there exists an explicitly computable positive constant $C_{(a,b)}$ and a rational constant $D_{(a,b)}$ such that

$$A_{(a,b)}(n) = \frac{C_{(a,b)} \cdot (27/4)^n}{n^{3/2}} + \frac{C_{(a,b)} D_{(a,b)} (27/4)^n}{n^{5/2}} + O\left(\frac{(27/4)^n}{n^{7/2}}\right).$$

Here the implied constant in the error term depends upon the number of branches in the shortest walk from the root $(1, 1)$ to the pair (a, b).

First, from Theorem 1.1.8 and using results from Section 2.4, we can derive the asymptotic formulas for the pairs $(1, 2)$, $(2, 1)$, and $(2, 3)$.

Proposition 2.7.1. For all $n \geq 100$, we have

$$\frac{405 \cdot (27/4)^n}{4\sqrt{3\pi} n^{3/2}} \left(1 - \frac{7559}{360n} + \frac{29}{n^2}\right) < A_{(1,2)}(n) < \frac{405 \cdot (27/4)^n}{4\sqrt{3\pi} n^{3/2}} \left(1 - \frac{7559}{360n} + \frac{657}{n^2}\right)$$

and

$$\frac{2673 \cdot (27/4)^n}{16\sqrt{3\pi} n^{3/2}} \left(1 - \frac{18173}{792n} - \frac{16072}{99n^2}\right) < A_{(2,1)}(n) = A_{(2,3)}(n) < \frac{2673 \cdot (27/4)^n}{16\sqrt{3\pi} n^{3/2}} \left(1 - \frac{18173}{792n} + \frac{88768}{99n^2}\right).$$

Proof. From Corollaries 2.3.3 and 2.5.12 and Theorem 1.1.8, we can deduce the bounds for $A_{(1,2)}(n)$. By Corollary 2.5.14, we have

$$A_{(2,1)} = A_{(2,3)} = D(n) + 4 \cdot A_{(1,1)}(n).$$

Applying Proposition 2.6.8 and Theorem 1.1.8 gives us the desired bounds for $A_{(2,1)}(n) = A_{(2,3)}(n)$. □

We are now ready to prove our main result concerning the asymptotic formulas for all coprime pairs (a, b).

Proof of Theorem 1.1.9. First, we claim that the constants $C_{(a,b)}$ in the Theorem have the form

$$C_{(a,b)} = \frac{243 t_k}{4\sqrt{3\pi}}$$

if $SW_{(1,1)}(a, b) = k$ where for all $k \in \mathbb{N} \cup 0$ we have

$$t_{3k} = \left(\frac{1}{2}\right)^k \left(1 + \frac{k}{3}\right),$$

$$t_{3k+1} = \left(\frac{1}{2}\right)^k \left(\frac{5}{3} + \frac{k}{2}\right),$$

and

$$t_{3k+2} = \left(\frac{1}{2}\right)^k \left(\frac{11}{4} + \frac{3k}{4}\right).$$

Note that $t_0 = 1$, $t_1 = \frac{5}{3}$, and $t_2 = \frac{11}{4}$. One can verify that, for all $k \geq 3$, t_k satisfies the recurrence

$$t_k = t_{k-2} - t_{k-3}$$

if $3 \nmid k + 1$ and

$$t_k = \frac{27}{4} \cdot t_{k-2} - t_{k-3}$$

if $3|k + 1$. Also, for the constants $D_{(a,b)}$, we define the sequence $(s_k)_k \in \mathbb{N} \cup \{0\}$ and $s_k := D_{(a,b)}$ if $SW_{(1,1)}(a, b) = k$. By Corollary 2.5.13, this sequence is well-defined. We further claim that for all $k \geq 3$, we have

$$s_k = \frac{t_{k-2}s_{k-2}}{t_k} - \frac{s_{k-3}t_{k-3}}{t_k}$$

if $3 \nmid k + 1$ and

$$s_k = \frac{t_{k-2}\left(s_{k-2} - \frac{3}{2}\right)}{t_k} - \frac{s_{k-3}t_{k-3}}{t_k}$$

if $3|k+1$. We prove the first claim by induction. In the induction step in our proof we also prove the claimed relations between the sequences $(t_k)_k \in \mathbb{N} \cup \{0\}$ and $(s_k)_k \in \mathbb{N} \cup \{0\}$. Theorem 1.1.8 and Proposition 2.7.1 provide the cases for $k = 0$, $k = 1$, and $k = 2$. Suppose the result holds for pairs with the shortest walks consisting of $k-2$ branches and $k-3$ branches for some $k \geq 3$, and we want to show it also holds for k. Let (a, b) be a pair with $SW_{(1,1)}(a, b) = n$ and let the fourth last and third last branches in this walk have the pairs (a_0, a_1) and $(a_1, |a-b|)$ respectively. By our inductive hypothesis, we have

$$A_{(a_0,a_1)} = \frac{C_{(a_0,a_1)} \cdot (27/4)^n}{n^{3/2}} \left(1 + \frac{s_{k-3}}{n} + O\left(\frac{1}{n^2}\right)\right)$$

and

$$A_{(a_1,|a-b|)} = \frac{C_{(a,|a-b|)} \cdot (27/4)^n}{n^{3/2}} \left(1 + \frac{s_{k-2}}{n} + O\left(\frac{1}{n^2}\right)\right)$$

where

$$C_{(a_0,a_1)} = \frac{243 \cdot t_{k-3}}{4\sqrt{3\pi}}$$

and

$$C_{(a_1,|a-b|)} = \frac{243 \cdot t_{k-2}}{4\sqrt{3\pi}}.$$

Suppose first that $3 \nmid n+1$. Then we have either a and b are both odd or a is odd and b is even. By Lemma 2.5.15, we have for all $n \geq \lfloor \frac{k}{3} \rfloor$,

$$A_{(a,b)}(n) = A_{(a_1,|a-b|)}(n) - A_{(a_0,a_1)}(n).$$

Then we have

$$\begin{aligned}
A_{(a,b)}(n) &= \frac{C_{(a_1,|a-b|)} \cdot (27/4)^n}{n^{3/2}} \left(1 + \frac{s_{k-2}}{n} + O\left(\frac{1}{n^2}\right)\right) - \frac{C_{(a_0,a_1)} \cdot (27/4)^n}{n^{3/2}} \left(1 + \frac{s_{k-3}}{n} + O\left(\frac{1}{n^2}\right)\right) \\
&= \frac{(C_{(a_1,|a-b|)} - C_{(a_0,a_1)}) \cdot (27/4)^n}{n^{3/2}} \left(1 + \left(\frac{t_{k-2}s_{k-2}}{t_k} - \frac{s_{k-3}t_{k-3}}{t_k}\right) \cdot \frac{1}{n} + O\left(\frac{1}{n^2}\right)\right) \\
&= \frac{(t_{k-2} - t_{k-3}) \cdot (27/4)^n}{4\sqrt{3\pi} n^{3/2}} \left(1 + \left(\frac{t_{k-2}s_{k-2}}{t_k} - \frac{s_{k-3}t_{k-3}}{t_k}\right) \cdot \frac{1}{n} + O\left(\frac{1}{n^2}\right)\right).
\end{aligned}$$

Thus, $t_k - t_{k-2} = t_{k-3}$ and

$$s_k = \frac{t_{k-2}s_{k-2}}{t_k} - \frac{s_{k-3}t_{k-3}}{t_k}.$$

The case when $3|n+1$ is similar. This proves our claims. Let $a_k = s_k t_k$. By our recursive formulas for s_k, we have

$$a_k = a_{k-2} - a_{k-3}.$$

Solving this recurrence relation in much the same way we solved the recurrence relation in Theorem 1.1.7 gives the asymptotic estimate

$$a_k \approx C \cdot (-1.3247\ldots)^k$$

for some constant C where $-1.3247\ldots$ is the only real root of

$$x^3 - x + 1.$$

Thus, we obtain

$$s_k \approx \frac{C' \cdot (-2.6494\ldots)^k}{k}$$

where C' depends on k mod 3. $\square$

Chapter 3

Sieve Methods in Random Graph Theory

3.1 The Sieves

We recall the setup for the Turán and simple sieves. First, recall the definition of a simple bipartite graph.

Definition 1.2.1. A simple bipartite graph is an undirected graph whose vertices can be divided into two sets, such that there are no edges between two vertices in the same set.

Let X be a simple bipartite graph with finite partite sets A and B. For $a \in A$ and $b \in B$, we denote by $a \sim b$ if there is an edge that joins a and b. Define

$$\deg b = \#\{a \in A : a \sim b\} \quad \text{and} \quad \omega(a) = \#\{b \in B : a \sim b\}.$$

In other words, $\deg b$ is the degree of b in X and $\omega(a)$ is the degree of a in X. For $b_1, b_2 \in B$, we define

$$n(b_1, b_2) = \#\{a \in A : a \sim b_1, a \sim b_2\}.$$

Then the Turán sieve states

$$\#\{a \in A : \omega(a) = 0\} \leq |A|^2 \cdot \frac{\sum_{b_1, b_2 \in B} n(b_1, b_2)}{(\sum_{b \in B} \deg b)^2} - |A|,$$

giving an upper bound of the quantity $\#\{a \in A : \omega(a) = 0\}$. Also, the simple sieve states

$$\#\{a \in A : \omega(a) = 0\} \geq |A| - \sum_{b \in B} \deg b,$$

giving a lower bound of the quantity $\#\{a \in A : \omega(a) = 0\}$.

3.2 The Set of all Graphs on n Vertices

In this section, we first prove Theorem 1.2.4, which is as follows.

Theorem 1.2.4. For $n \geq 2$, let $G^{(n)}$ denote the set of all graphs on n vertices with edge probability $p(n)$, and let $P(G^{(n)}, p(n))$ be the probability of a graph from $G^{(n)}$ having diameter 2. Then

$$\begin{aligned}
&1 - \frac{n^2(1-p(n)^2)^{n-2}(1-p(n))}{2} \\
&\leq P(G^{(n)}, p(n)) \\
&\leq \frac{2}{(n-1)^2(1-p(n)^2)^n(1-p(n))} + \frac{8}{n}\left(1 + \frac{p(n)^3}{(1-p(n))^2}\right)^n.
\end{aligned}$$

Proof. For a fixed $n \in \mathbb{N}$, let $G^{(n)}$ denote the set of all graphs on n vertices with edge probability $p(n)$, and let $P(G^{(n)}, p(n))$ be the probability of a graph from $G^{(n)}$ having diameter 2. Consider the function $g_n : [0,1] \to [0,1]$ defined as $g_n(x) := P(G^{(n)}, x)$. There are $2^{\frac{n(n-1)}{2}}$ graphs in total in $G^{(n)}$. Let us say M of these have diameter 2 and label these as G_1, G_2,…,G_M. For $1 \leq i \leq M$, let k_i denote the number of edges in G_i. Then the probability of selecting the graph G_i from $G^{(n)}$ according to the edge probability x is $x^{k_i}(1-x)^{\frac{n(n-1)}{2}-k_i}$. Therefore,

$$g_n(x) = x^{k_1}(1-x)^{\frac{n(n-1)}{2}-k_1} + x^{k_2}(1-x)^{\frac{n(n-1)}{2}-k_2} + \cdots + x^{k_M}(1-x)^{\frac{n(n-1)}{2}-k_M}.$$

Thus, for each $n \in \mathbb{N}$ the function g_n is continuous. Therefore, we may assume that $p(n) \in \mathbb{Q} \cap (0,1)$ since $\mathbb{Q} \cap (0,1)$ is dense in $[0,1]$.

Let $p(n) = \frac{r}{s}$ where $r = r(n), s = s(n) \in \mathbb{N}$. We let A be the set of all graphs in $G^{(n)}$, allowing for a number of duplicates of each possible graph to accommodate the edge probability $p(n)$. We accomplish this by letting there be $r^{\binom{n}{2}}$ copies of the complete graph, $r^{\binom{n}{2}}\left(\frac{s}{r}-1\right)$ copies of each graph with $\binom{n}{2}-1$ edges, $r^{\binom{n}{2}}\left(\frac{s}{r}-1\right)^2$ copies of each graph with $\binom{n}{2}-2$ edges, and so on. By the binomial theorem we have

$$|A| = \sum_{k=0}^{\binom{n}{2}} \binom{\binom{n}{2}}{k} r^k (s-r)^{\binom{n}{2}-k} = s^{\binom{n}{2}}.$$

We let B be all pairs of vertices so $|B| = \binom{n}{2}$. For a graph $a \in A$ and a pair of vertices $b \in B$, we say $a \sim b$ if the pair of vertices b in a do not share a common neighbouring vertex and are not neighbours themselves. Thus, we will have $\omega(a) = 0$ (or a doesn't match up with any pair of vertices) if and only if a is connected with diameter at most 2.

Pick a pair of vertices $b \in B$ and call them v_1 and v_2. To calculate $\deg b$, we need to calculate the number of graphs in A such that the pair of vertices do not have a common neighbouring vertex and are not neighbours themselves. For each of the potential $(n-2)$ neighbouring vertices, we need to consider two edges, making sure at least one of them is not in the graph. Since each potential edge contributes a factor of r or $(s-r)$ depending on whether it is in a specified graph, we have

$$\begin{aligned} D(r,s,n) := \deg b &= ((s-r)^2 + 2r(s-r))^{n-2}(s-r)(s^{\binom{n}{2}-2(n-2)-1}) \\ &= (s^2-r^2)^{n-2}(s-r)s^{\binom{n}{2}-2(n-2)-1}. \end{aligned}$$

It follows that

$$\sum_{b \in B} \deg b = \frac{s^{\binom{n}{2}} n(n-1)(1-p(n)^2)^{n-2}(1-p(n))}{2}.$$

By the simple sieve, we obtain

$$P(G^{(n)}, p(n)) \geq 1 - \frac{n(n-1)(1-p(n)^2)^{n-2}(1-p(n))}{2} > 1 - \frac{n^2(1-p(n)^2)^{n-2}(1-p(n))}{2}. \tag{3.2.1}$$

We now try to get an upper bound for $P(G^{(n)}, p(n))$, in which we need to estimate $\sum_{b_1,b_2 \in B} n(b_1,b_2)$. In the following, we calculate $n(b_1,b_2)$, depending on how many vertices b_1 and b_2 have in common.

Case 1. Suppose that b_1 and b_2 are two pairs of vertices that have no vertices in common, i.e., b_1 and b_2 consist of 4 distinct vertices. For each of b_1 and b_2, the probability that the pair of vertices in question are not connected by an edge nor have any common neighbouring vertices is

$$\frac{D(r,s,n)}{s^{\binom{n}{2}}}.$$

As is the case for calculating $\deg b$, for each of the pair of vertices b_1 and b_2, we need to consider pairs of edges for each potential neighbouring vertex. If the potential neighbouring vertex is among the remaining $n-4$ vertices, then the pair of edges to consider with respect to b_1 will be disjoint from the pair of edges to consider with respect to b_2. The only real problem to consider is when the potential neighbouring vertex is among the pair of vertices b_1 and b_2 where we have four possible edges to consider. These observations give rise to

$$n(b_1, b_2) = \frac{D(r,s,n)^2}{s^{\binom{n}{2}}} \cdot \frac{s^4((s-r)^4 + 4r(s-r)^3 + 2r^2(s-r)^2)}{(s^2-r^2)^4},$$

and thus

$$\begin{aligned}\sum_{b_1, b_2 \in B,\ 4 \text{ vertices}} n(b_1, b_2) &< \binom{n}{2}^2 \frac{D(r,s,n)^2}{s^{\binom{n}{2}}} \\ &\quad \cdot \frac{p(n)^{-4}((p(n)^{-1}-1)^4 + 4(p(n)^{-1}-1)^3 + 2(p(n)^{-1}-1)^2)}{(p(n)^{-2}-1)^4} \\ &< \binom{n}{2}^2 \frac{D(r,s,n)^2}{s^{\binom{n}{2}}} \cdot \left(1 + \frac{4p(n)^3}{(1-p(n))^2}\right).\end{aligned}$$

Case 2. Take two pairs of vertices b_1 and b_2 that have exactly one vertex in common, i.e., b_1 and b_2 consist of 3 distinct vertices. We can do a similar kind of analysis of edge selection as in Case 1 to calculate

$$\begin{aligned}\sum_{b_1, b_2 \in B,\ 3 \text{ vertices}} n(b_1, b_2) &= \frac{D(r,s,n)^2 n(n-1)(n-2)}{s^{\binom{n}{2}}} \left(1 + \frac{1}{p(n)^{-3} + p(n)^{-2} - p(n)^{-1} - 1}\right)^{n-3} \\ &\leq \frac{D(r,s,n)^2 n(n-1)(n-2)}{s^{\binom{n}{2}}} \left(1 + \frac{p(n)^3}{(1-p(n))}\right)^{n-3}.\end{aligned}$$

Case 3. Suppose b_1 and b_2 have two vertices in common. Then the two pairs are identical, and we have

$$n(b_1, b_2) = \deg b.$$

It follows that

$$\sum_{b_1, b_2 \in B,\ 2 \text{ vertices}} n(b_1, b_2) = \sum_{b \in B} \deg b = \frac{s^{\binom{n}{2}} n(n-1)(1-p(n)^2)^{n-2}(1-p(n))}{2}.$$

Combining Cases $1-3$, we get

$$\sum_{b_1,b_2\in B} n(b_1,b_2) < \binom{n}{2}^2 \frac{D(r,s,n)^2}{s^{\binom{n}{2}}}\cdot\left(1+\frac{4p(n)^3}{(1-p(n))^2}\right)$$
$$+\frac{D(r,s,n)^2 n(n-1)(n-2)}{s^{\binom{n}{2}}}\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-3}$$
$$+\frac{s^{n_1n_2}n(n-1)(1-p(n)^2)^{n-2}(1-p(n))}{2}.$$

By the Turán sieve, we deduce

$$P(G^{(n)},p(n)) \le \frac{2}{n(n-1)(1-p(n)^2)^{n-2}(1-p(n))}+\frac{4}{n}\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-3}+\frac{4p(n)^3}{(1-p(n))^2}.$$

Notice that

$$\frac{p(n)^3}{(1-p(n))^2} < \frac{1}{n}\left(1+n\frac{p(n)^3}{(1-p(n))^2}\right) < \frac{1}{n}\left(1+\frac{p(n)^3}{(1-p(n))^2}\right)^n.$$

It follows that

$$P(G^{(n)},p(n)) \le \frac{2}{(n-1)^2(1-p(n)^2)^n(1-p(n))}+\frac{8}{n}\left(1+\frac{p(n)^3}{(1-p(n))^2}\right)^n. \tag{3.2.2}$$

By (3.2.1) and (3.2.2) Theorem 1.2.4 follows. □

We now prove Proposition 1.2.6, which states as follows.

Proposition 1.2.6. Let $P(G^{(n)},p(n))$ be defined as in Theorem 1.2.4. Let $\lim_{n\to\infty} p(n) = 0$. We have

$$1-\frac{n^2}{2}e^{-np(n)^2}(1+o(1)) \le P(G^{(n)},p(n)) \le (1+o(1))\left(\frac{2}{n^2}e^{np(n)^2}\right)\left(1+4ne^{np(n)^2\left(p(n)^2-1\right)}\right). \tag{3.2.3}$$

Suppose further that

$$\lim_{n\to\infty}(2\log n - np(n)^2-\log 2) = c$$

for some $c\in\mathbb{R}\backslash\{0\}$.

1) If $c<0$, we have

$$P(G^{(n)},p(n)) \ge 1-(1+o(1))e^{c}.$$

2) If $c>0$, we have

$$P(G^{(n)},p(n)) \le (1+o(1))e^{-c}.$$

Proof. By Theorem 1.2.4 we have

$$P(G^{(n)}, p(n)) > 1 - \frac{n^2 \left(1-p(n)^2\right)^{p(n)^{-2}\cdot np(n)^2} \left(1-p(n)^2\right)^{-2} (1-p(n))}{2}.$$

Since $p(n)^{-2} \geq 1$, we have

$$e^{-np(n)^2} \left(1-p(n)^2\right)^{np(n)^2} < \left(1-p(n)^2\right)^{p(n)^{-2}np(n)^2} < e^{-np(n)^2}. \tag{3.2.4}$$

Since $0 = \lim_{n\to\infty} p(n)$, we have that

$$\lim_{n\to\infty} \left(1-p(n)^2\right)^{-2} = 1$$

and

$$\lim_{n\to\infty} (1-p(n)) = 1,$$

from which we get

$$P(G^{(n)}, p(n)) \geq 1 - \frac{n^2}{2} e^{-np(n)^2}(1+o(1)). \tag{3.2.5}$$

For the upper bound, first note that

$$\frac{8}{n}\left(1 + \frac{p(n)^3}{(1-p(n))^2}\right)^n < \frac{2e^{np(n)^2}}{n^2} \cdot 4ne^{\left(\frac{np(n)^3}{(1-p(n))^2} - np(n)^2\right)}.$$

Combining this with Equations (3.2.2) and (3.2.4), we get

$$P(G^{(n)}, p(n)) < \frac{2e^{np(n)^2}}{(n-1)^2 \left(1-p(n)^2\right)^{np(n)^2} (1-p(n))} + \frac{2e^{np(n)^2}}{n^2} \cdot 4ne^{\left(\frac{np(n)^3}{(1-p(n))^2} - np(n)^2\right)}.$$

Note that for $n \in \mathbb{N}$ with $\frac{2}{n^2}e^{np(n)^2} \geq 1$, we have

$$\left(\frac{2}{n^2}e^{np(n)^2}\right)\left(1 + 4ne^{np(n)^2(p(n)-1)}\right) > 1. \tag{3.2.6}$$

In particular for those n, the bound in Theorem 1.2.4 is trivial. Thus, it suffices to consider $n \in \mathbb{N}$ such that

$$\frac{2}{n^2}e^{np(n)^2} < 1.$$

Label all such $n \in \mathbb{N}$ as $n_1, n_2, \ldots, n_j, \ldots$ such that $n_1 < n_2 < \cdots$ If there are only finitely many, then for sufficiently large n, we will have (3.2.6) and so the bound in Theorem 1.2.4

is trivial. Thus, we may assume that $n_1, n_2, ..., n_j, ...$ is an infinite list. Then for all $j \in \mathbb{N}$, we have

$$n_j p(n_j)^2 < 2\log n_j - \log 2,$$

and so

$$\lim_{j\to\infty} n_j p(n_j)^4 = 0, \lim_{j\to\infty} p(n_j)^2 = 0, \text{ and } \lim_{j\to\infty} n_j p(n_j)^3 = 0. \tag{3.2.7}$$

Note that if $0 \le x \le 1$ and $y \ge 1$, then $(1-x)^y \ge 1 - xy$. Thus, if $n_j p(n_j)^2 \ge 1$, then

$$\left(1 - p(n_j)^2\right)^{n_j p(n_j)^2} \ge 1 - n_j p(n_j)^4.$$

Suppose that $n_j p(n_j)^2 < 1$. Then we have

$$\left(1 - p(n_j)^2\right)^{n_j p(n_j)^2} \ge 1 - p(n_j)^2.$$

Thus, by Equation (3.2.7), we have

$$\lim_{j\to\infty} \left(1 - p(n_j)^2\right)^{n_j p(n_j)^2} = 1$$

and

$$\lim_{j\to\infty} n_j p(n_j)^3 \left(1 - \frac{1}{\left(1 - p(n_j)\right)^2}\right) = 0.$$

Also, notice that

$$n_j p(n_j)^3 \left(1 - \frac{1}{\left(1 - p(n_j)\right)^2}\right) = n_j p(n_j)^2 \left(p(n_j) - 1\right) - \left(\frac{n_j p(n_j)^3}{(1 - p(n_j))^2} - n_j p(n_j)^2\right).$$

We thus obtain

$$P(G^{(n)}, p(n)) \le (1 + o(1)) \left(\frac{2}{n^2} e^{np(n)^2}\right) \left(1 + 4ne^{np(n)^2\left(p(n)^2 - 1\right)}\right). \tag{3.2.8}$$

Now we suppose further that

$$\lim_{n\to\infty} (2\log n - np(n)^2 - \log 2) = c \tag{3.2.9}$$

for some $c \in \mathbb{R}\backslash\{0\}$. Then we have

$$\lim_{n\to\infty} \left(\log n - \frac{np(n)^2}{2}\right) = \tilde{c}$$

for some $\tilde{c} \in \mathbb{R}$. Since $\lim_{n\to\infty} p(n) = 0$, it follows that

$$\begin{aligned}
&\lim_{n\to\infty} \left(\log n + np(n)^3 - np(n)^2\right) \\
&= \lim_{n\to\infty} \left(\left(\log n - \frac{np(n)^2}{2}\right) + \left(np(n)^2 - \frac{np(n)^2}{2}\right) \right) \\
&= -\infty. \qquad (3.2.10)
\end{aligned}$$

By Equation (3.2.10), we have

$$ne^{np(n)^2\left(p(n)^2-1\right)} = o(1), \qquad (3.2.11)$$

and by Equation (3.2.9), we have

$$\frac{2}{n^2} e^{np(n)^2} = e^{-c}(1 + o(1)) \qquad (3.2.12)$$

and

$$\frac{n^2}{2} e^{-np(n)^2} = e^{c}(1 + o(1)). \qquad (3.2.13)$$

By Equations (3.2.5) and (3.2.13), we obtain

$$P(G^{(n)}, p(n)) \geq 1 - (1 + o(1))e^{c},$$

and by Equations (3.2.8), (3.2.12), and (3.2.11), we obtain

$$P(G^{(n)}, p(n)) \leq (1 + o(1))e^{-c}.$$

□

Remark 3.2.1. Assume that $n \geq 200$ and $p(n) \leq 1/2$. The $o(1)$ in the lower bound in (3.2.3) can be made explicit as $4p(n)^2$ and the $o(1))$ in the upper bound in (3.2.3) can be made explicit as $\frac{4(\log n)^2+2}{n} + p(n) + \frac{3e^8(2\log n)^{3/2}}{n^{1/2}}$. With the additional assumption that

$$c - 1 < 2\log n - np(n)^2 - \log 2 < c + 1$$

In Proposition 1.2.6 1), the $o(1))$ can be made explicit as

$$(1 + |c + \log 2 - 2\log n + np(n)^2|e)(1 + 4p(n)^2) - 1$$

and in Proposition 1.2.6 2), the $o(1))$ can be made explicit as

$$\begin{aligned}
&(1 + |c + \log 2 - 2\log n + np(n)^2|e) \left(1 + \frac{8e^{c+1+\frac{(2\log n - \log 2 - c+1)^{3/2}}{n^{1/2}}}}{n}\right) \\
&\cdot \left(1 + \frac{4(\log n)^2 + 2}{n} + p(n) + \frac{3e^8(2\log n)^{3/2}}{n^{1/2}}\right) - 1.
\end{aligned}$$

Using the above methods, we can obtain similar results about the probability of a random directed graph having diameter 2.

Theorem 3.2.2. Let $G^{(n),\to}$ denote the set of all directed graphs on $n \geq 2$ vertices with edge probability $p(n)$, and let $P(G^{(n),\to}, p(n))$ be the probability of a graph from $G^{(n),\to}$ having diameter 2. Then

$$\begin{aligned}
&1 - n^2(1-p(n)^2)^{n-2}(1-p(n)) \\
&\leq P(G^{(n),\to}, p(n)) \\
&\leq \frac{1}{(n-1)^2(1-p(n)^2)^n(1-p(n))} + \frac{4}{n}\left(1 + \frac{p(n)^3}{(1-p(n))^2}\right)^n.
\end{aligned}$$

Corollary 3.2.3. Let $P(G^{(n),\to}, p(n))$ be defined as in Theorem 3.2.2. If $p(n) = \frac{1}{2}$, then we have

$$P(G^{(n),\to}, 1/2) \geq 1 - \frac{2n^2(3/4)^n}{9}.$$

In the case when $p(n) \to 0$ as $n \to \infty$, we can prove the following.

Proposition 3.2.4. Let $P(G^{(n),\to}, p(n))$ be defined as in Theorem 3.2.2. Also, let $\lim_{n\to\infty} p(n) = 0$. We have

$$1-n^2e^{-np(n)^2}(1+o(1)) \leq P(G^{(n),\to}, p(n)) \leq (1+o(1))\left(\frac{1}{n^2}e^{np(n)^2}\right)\left(1 + 4ne^{np(n)^2\left(p(n)^2-1\right)}\right). \tag{3.2.14}$$

Suppose further that

$$\lim_{n\to\infty}(2\log n - np(n)^2) = c$$

for some $c \in \mathbb{R}\backslash\{0\}$.

1) If $c < 0$, we have

$$P(G^{(n),\to}, p(n)) \geq 1 - (1+o(1))e^c.$$

2) If $c > 0$, we have

$$P(G^{(n),\to}, p(n)) \leq (1+o(1))e^{-c}.$$

Remark 3.2.5. Assume that $n \geq 200$ and $p(n) \leq 1/2$. The $o(1)$ in the lower bound in (3.2.14) can be made explicit as $4p(n)^2$ and the $o(1)$ in the upper bound in (3.2.14) can be made explicit as $\frac{4(\log n)^2+2}{n} + p(n) + \frac{3e^8(2\log n)^{3/2}}{n^{1/2}}$. With the additional assumption that

$$c - 1 < 2\log n - np(n)^2 < c + 1$$

in Proposition 3.2.4 1), the $o(1)$ can be made explicit as

$$(1+|c-2\log n+np(n)^2|e)(1+4p(n)^2)-1$$

and in Proposition 3.2.4 2), the $o(1)$ can be made explicit as

$$(1+|c-2\log n+np(n)^2|e)\left(1+\frac{8e^{c+1+\frac{(2\log n-c+1)^{3/2}}{n^{1/2}}}}{n}\right)$$
$$\cdot\left(1+\frac{4(\log n)^2+2}{n}+p(n)+\frac{3e^8(2\log n)^{3/2}}{n^{1/2}}\right)-1.$$

3.3 Analysis of k-partite Graphs

Here we apply our analysis to k-partite graph sets for $k \geq 3$. First, we present a definition.

Definition 3.3.1. Let $k \geq 2$. A simple k-partite graph is an undirected graph whose vertices can be divided into k sets, such that there are no edges between two vertices in the same set.

We exclude the bipartite case ($k = 2$) because the only bipartite graph that has diameter 2 is the complete bipartite graph; we analyze that case by itself in the next section.

Convention 3.3.2. For each k-partite graph, we label the k partite sets of the graph in a non-decreasing order in terms of the number of vertices each set contains. Thus, the ith set is a set containing n_i vertices.

Theorem 3.3.3. Fix $k \geq 3$ and for each $n \in \mathbb{N}$, $n \geq k+2$, pick $n_1, n_2, ..., n_k \in \mathbb{N}$ such that $n_1 \leq n_2 \leq \ldots \leq n_k$, $n_{k-1} \geq 2$, and $n_1 + n_2 + \cdots + n_k = n$. Let $\mathbf{k} = (n_1, n_2, \ldots, n_k)$ and let $G^{(n),\mathbf{k}}$ denote the set of all k-partite graphs with the partite sets having $n_1, n_2, \ldots, n_k$ vertices respectively, with edge probability $p(n)$, and let $P(G^{(n),\mathbf{k}}, p(n))$ be the probability of a graph from $G^{(n),\mathbf{k}}$ having diameter 2. Then

$$1-\frac{n_k^2(1-p(n)^2)^{n-n_k}}{2}\left(1+\frac{2n_{k-1}(1-p(n)^2)^{-n_{k-1}}}{n_k}+\frac{7k^2n_{k-1}^2(1-p(n)^2)^{n_k-n_{k-1}-n_{k-2}}}{3n_k^2}\right)$$
$$\leq P(G^{(n),\mathbf{k}}, p(n))$$
$$\leq \frac{2}{n_k(n_k-1)(1-p(n)^2)^{n-n_k}}\left(1+\frac{2n_{k-1}(1-p(n)^2)^{-n_{k-1}}(1-p(n))}{(n_k-1)}\right)^{-1}$$
$$+\frac{3k^3\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k}(1-p(n)^2)^{-2}}{(n_{k-1}-1)}.$$

Proof. As in the proof of Theorem 1.2.4, we may assume that $p(n) \in \mathbb{Q} \cap (0,1)$ for all $n \in \mathbb{N}$.

Let $p(n) = \frac{r}{s}$ where $r, s \in \mathbb{N}$. As in the proof of Theorem 1.2.4, we let A be the set of all graphs in $G^{(n),\mathbf{k}}$, allowing for a number of duplicates of each possible graph to accommodate the edge probability $p(n)$. Since the complete k-partite graph has $t := \sum_{1 \leq i < j \leq k} n_i n_j$ edges, we have r^t copies of the complete bipartite graph and $|A| = s^t$.

We let B be all pairs of vertices. Thus, $|B| = \frac{n(n-1)}{2}$. For a graph $a \in A$ and a pair of vertices $b \in B$, we say $a \sim b$ if the pair of vertices b in a do not share a common neighbouring vertex and are not connected by a single edge. Thus, we will have $\omega(a) = 0$ if and only if a is connected with diameter at most 2. For each pair of vertices $b \in B$ that are in the ith partite set for some $1 \leq i \leq k$, we will have

$$\begin{aligned} D(r, s, n, n_i) := \deg b &= ((s-r)^2 + 2r(s-r))^{n-n_i}((s-r)+r)^{t-2n+2n_i} \\ &= (1 - p(n)^2)^{n-n_i} s^t. \end{aligned}$$

For each pair of vertices $b \in B$ with one vertex being in the ith partite set and the other in the jth partite set where $i < j$, we have

$$\begin{aligned} D(r, s, n, n_i, n_j) := \deg b &= ((s-r)^2 + 2r(s-r))^{n-n_i-n_j}((s-r)+r)^{t-2n+2n_i+2n_j}(1-p(n)) \\ &= (1 - p(n)^2)^{n-n_i-n_j}(1-p(n)) s^t. \end{aligned}$$

It follows that

$$\sum_{b \in B} \deg b = s^t \sum_{i=1}^{k} \binom{n_i}{2} (1 - p(n)^2)^{n-n_i} + s^t \sum_{1 \leq i < j \leq k} n_i n_j (1 - p(n)^2)^{n-n_i-n_j}(1 - p(n)).$$

By the simple sieve, we obtain

$$
\begin{aligned}
P(G^{(n),\mathbf{k}}, p(n)) &> 1 - \sum_{i=1}^{k} \binom{n_i}{2} (1-p(n)^2)^{n-n_i} - \sum_{1 \le i < j \le k} n_i n_j (1-p(n)^2)^{n-n_i-n_j} (1-p(n)) \\
&> 1 - \frac{n_k^2 (1-p(n)^2)^{n-n_k}}{2} - n_k n_{k-1} (1-p(n)^2)^{n-n_k-n_{k-1}} - \frac{k n_{k-1}^2 (1-p(n)^2)^{n-n_{k-1}}}{2} \\
&- k^2 n_{k-1} n_{k-2} (1-p(n)^2)^{n-n_{k-1}-n_{k-2}} \\
&> 1 - \frac{n_k^2 (1-p(n)^2)^{n-n_k}}{2} - n_k n_{k-1} (1-p(n)^2)^{n-n_k-n_{k-1}} \\
&\quad - \frac{7k^2 n_{k-1}^2 (1-p(n)^2)^{n-n_{k-1}-n_{k-2}}}{6} \\
&= 1 - \frac{n_k^2 (1-p(n)^2)^{n-n_k}}{2} \\
&\quad \cdot \left(1 + \frac{2n_{k-1}(1-p(n)^2)^{-n_{k-1}}}{n_k} + \frac{7k^2 n_{k-1}^2 (1-p(n)^2)^{n_k-n_{k-1}-n_{k-2}}}{3n_k^2}\right).
\end{aligned}
$$

We now try to get an upper bound for $P(G^{(n),\mathbf{k}}, p(n))$, in which we need to estimate $\sum_{b_1, b_2 \in B} n(b_1, b_2)$. In the following, we calculate $n(b_1, b_2)$, depending on how many vertices b_1 and b_2 have in common.

Case 1. Suppose that b_1 and b_2 are two pairs of vertices with no overlapping vertices. As in the proof of Theorem 1.2.4, we can calculate

$$
\begin{aligned}
n(b_1, b_2) &\le \frac{\deg b_1 \deg b_2}{s^{n_1 n_2}} \cdot \frac{p(n)^{-4}((p(n)^{-1}-1)^4 + 4(p(n)^{-1}-1)^3 + 2(p(n)^{-1}-1)^2)}{(p(n)^{-2}-1)^4} \\
&< \frac{\deg b_1 \deg b_2}{s^{n_1 n_2}} \cdot \left(1 + \frac{4p(n)^3}{(1-p(n))^2}\right).
\end{aligned}
$$

The next four cases consider when the two pairs of vertices overlap by exactly one vertex.

Case 2. Suppose the overlapping vertex occurs in the ith set, and the other two vertices occur in the jth set and the lth set with all three sets being different from one another.

Again, as in the proof of Theorem 1.2.4, we can calculate

$$\begin{aligned} n(b_1, b_2) &= \frac{D(r,s,n,n_i,n_j)D(r,s,n,n_i,n_l)}{s^t}\left(\frac{p(n)^{-4} - 2p(n)^{-2} + p(n)^{-1}}{(p(n)^{-2}-1)^2}\right)^{n-n_i-n_j-n_l} (1-p(n)^2)^{-2} \\ &< \frac{D(r,s,n,n_k,n_{k-1})D(r,s,n,n_k,n_{k-2})}{s^t}\left(1 + \frac{p(n)^3}{(1-p(n))}\right)^{n-n_k-n_{k-1}-n_{k-2}} (1-p(n)^2)^{-2}. \end{aligned} \tag{3.3.1}$$

Case 3. If two pairs have an overlapping vertex in the ith set and the other two vertices are in the jth set, then again we have

$$\begin{aligned} n(b_1, b_2) &= \frac{D(r,s,n,n_i,n_j)^2}{s^t}\left(\frac{p(n)^{-4} - 2p(n)^{-2} + p(n)^{-1}}{(p(n)^{-2}-1)^2}\right)^{n-n_i-n_l} \\ &< \frac{D(r,s,n,n_k,n_{k-1})^2}{s^t}\left(1 + \frac{p(n)^3}{(1-p(n))}\right)^{n-n_k-n_{k-1}}. \end{aligned} \tag{3.3.2}$$

Case 4. If one pair has both vertices in the ith set and the other pair has the other vertex in the jth set then we have

$$\begin{aligned} n(b_1, b_2) &= \frac{D(r,s,n,n_i)D(r,s,n,n_i,n_j)}{s^t}\left(\frac{p(n)^{-4} - 2p(n)^{-2} + p(n)^{-1}}{(p(n)^{-2}-1)^2}\right)^{n-n_i-n_j} (1-p(n)^2)^{-1} \\ &< \frac{D(r,s,n,n_k)D(r,s,n,n_k,n_{k-1})}{s^t}\left(1 + \frac{p(n)^3}{(1-p(n))}\right)^{n-n_k-n_{k-1}} (1-p(n)^2)^{-1}. \end{aligned} \tag{3.3.3}$$

with the last factor accommodating the undercount of the edge between the overlapping vertex and the vertex in the ith set.

Case 5. If all 3 vertices are in the same set, say the ith set, then we have

$$\begin{aligned} n(b_1, b_2) &= \frac{D(r,s,n,n_i)^2}{s^t}\left(\frac{p(n)^{-4} - 2p(n)^{-2} + p(n)^{-1}}{(p(n)^{-2}-1)^2}\right)^{n-n_i} \\ &< \frac{D(r,s,n,n_k)^2}{s^t}\left(1 + \frac{p(n)^3}{(1-p(n))}\right)^{n-n_k}. \end{aligned} \tag{3.3.4}$$

Case 6. Finally, if the two pairs of vertices are identical, then we have

$$n(b_1, b_2) = \deg b_1 = \deg b_2.$$

Note that

$$\left(\sum_{b\in B}\deg b\right)^2 > \frac{D(r,s,n,n_k)^2 n_k^2(n_k-1)^2}{4s^{2t}}, \tag{3.3.5}$$

$$\left(\sum_{b\in B}\deg b\right)^2 > \frac{n_k^2 n_{k-1}^2 D(r,s,n,n_k,n_{k-1})^2}{s^{2t}}, \tag{3.3.6}$$

and

$$\left(\sum_{b\in B}\deg b\right)^2 > \frac{D(r,s,n,n_k)D(r,s,n,n_k,n_{k-1})n_k^2(n_k-1)n_{k-1}}{2s^{2t}}. \tag{3.3.7}$$

Dividing each of (3.3.1) and (3.3.2) by the right-hand side of (3.3.6), dividing (3.3.4) by the right-hand side of (3.3.5), and dividing (3.3.3) by the right-hand side of (3.3.7), we deduce by the Turán sieve that

$$\begin{aligned}
P(G^{(n),\mathbf{k}},p(n)) &< \left(\frac{n_k(n_k-1)(1-p(n)^2)^{n-n_k}}{2} + n_k n_{k-1}(1-p(n)^2)^{n-n_k-n_{k-1}}(1-p(n))\right)^{-1} \\
&\quad + \frac{4p(n)^3}{(1-p(n))^2} + \frac{k^3 n_{k-2}\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k-n_{k-1}-n_{k-2}}(1-p(n)^2)^{n_{k-1}-n_{k-2}-2}}{n_k n_{k-1}} \\
&\quad + \frac{k^2\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k-n_{k-1}}}{n_{k-1}} + \frac{2k^2\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k-n_{k-1}}(1-p(n)^2)^{-1}}{(n_k-1)} \\
&\quad + \frac{4kn_k\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k}}{(n_k-1)^2} \\
&< \left(\frac{n_k(n_k-1)(1-p(n)^2)^{n-n_k}}{2} + n_k n_{k-1}(1-p(n)^2)^{n-n_k-n_{k-1}}(1-p(n))\right)^{-1} \\
&\quad + \frac{4p(n)^3}{(1-p(n))^2} + \frac{k^3\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k-n_{k-1}-n_{k-2}}(1-p(n)^2)^{-2}}{n_{k-1}} \\
&\quad + \frac{k^2\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k-n_{k-1}}}{n_{k-1}} + \frac{2k^2\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k-n_{k-1}}(1-p(n)^2)^{-1}}{(n_k-1)} \\
&\quad + \frac{6k\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k}}{(n_k-1)}.
\end{aligned}$$

Notice that

$$\frac{4p(n)^3}{(1-p(n))^2} < \frac{4\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n_{k-1}}}{n_{k-1}(1-p(n))} < \frac{4\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k}}{(n_{k-1}-1)(1-p(n))}$$

so that

$$\begin{aligned}P(G^{(n),\mathbf{k}}, p(n)) < {} & \frac{2}{n_k(n_k-1)(1-p(n)^2)^{n-n_k}}\left(1+\frac{2n_{k-1}(1-p(n)^2)^{-n_{k-1}}(1-p(n))}{(n_k-1)}\right)^{-1} \\ & + \frac{3k^3\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n-n_k}(1-p(n)^2)^{-2}}{(n_{k-1}-1)}.\end{aligned}$$

$\square$

By substituting $p(n) = \frac{1}{2}$, we deduce from Theorem 3.3.3 the following.

Corollary 3.3.4. Let $P(G^{(n),\mathbf{k}}, p(n))$ be defined as in Theorem 3.3.3. If $p(n) = \frac{1}{2}$, then we have

$$P(G^{(n),\mathbf{k}}, 1/2) \geq 1 - \frac{n_k^2(3/4)^{n-n_k}}{2}\left(1+\frac{2n_{k-1}(3/4)^{-n_{k-1}}}{n_k} + \frac{7k^2n_{k-1}^2(3/4)^{n_k-n_{k-1}-n_{k-2}}}{3n_k^2}\right).$$

In the case when $p(n) \to 0$ as $n \to \infty$, we have the following.

Proposition 3.3.5. Let $P(G^{(n),\mathbf{k}}, p(n))$ be defined as in Theorem 3.3.3. Let $\lim_{n\to\infty} p(n)^4(n-n_k) = 0$. We have

$$\begin{aligned}& 1 - \frac{n_k^2 e^{-p(n)^2(n-n_k)}}{2}\left(1+\frac{2n_{k-1}}{n_k}e^{p(n)^2 n_{k-1}}\left(1+\frac{7k^2n_{k-1}e^{-p(n)^2(n_k-n_{k-2})}}{6n_k}\right)\right) \\ & \leq P(G^{(n),\mathbf{k}}, p(n)) \\ & \leq \frac{2e^{p(n)^2(n-n_k)}}{n_k^2}\left(1+\frac{2n_{k-1}}{n_k}e^{p(n)^2n_{k-1}}\right)^{-1} \\ & \left(1+\frac{3k^3n_k^2e^{\left(p(n)^3-p(n)^2\right)(n-n_k)}}{2(n_{k-1}-1)} + \frac{3k^3n_kn_{k-1}e^{\left(p(n)^3-p(n)^2\right)(n-n_k)+p(n)^2n_{k-1}}}{(n_{k-1}-1)}\right)(1+o(1)).\end{aligned} \tag{3.3.8}$$

Suppose further that

$$\lim_{n\to\infty} \log n_{k-1} - \log n - p(n)^2 n_{k-1} = -\infty,$$
$$\lim_{n\to\infty} 2\log n + (p(n)^3 - p(n)^2)(n-n_k) - \log n_{k-1} = -\infty,$$
$$\lim_{n\to\infty} \left(p(n)^3 - p(n)^2\right)(n-n_k) + p(n)^2 n_{k-1} + \log n = -\infty,$$

and that

$$\lim_{n\to\infty} 2\log n_k - p(n)^2(n-n_k) - \log 2 + \log\left(1 + \frac{2n_{k-1}}{n_k} e^{p(n)^2 n_{k-1}}\right) = c$$

for some $c \in \mathbb{R}$.

1) If $c < 0$, we have

$$P(G^{(n)}, p(n)) \geq 1 - (1+o(1))e^c.$$

2) If $c > 0$, we have

$$P(G^{(n)}, p(n)) \leq (1+o(1))e^{-c}.$$

Proof. By Theorem 3.3.3, we have

$$P(G^{(n),\mathbf{k}}, p(n)) > 1 - \frac{n_k^2 e^{-p(n)^2(n-n_k)}}{2} - n_k n_{k-1} e^{-p(n)^2(n-n_k-n_{k-1})}\left(1 + \frac{k n_{k-1} e^{-p(n)^2 n_k}}{2n_k} + \frac{k^2 n_{k-2} e^{-p(n)^2(n_k - n_{k-2})}}{n_k}\right).$$

Also, we can deduce that

$$P(G^{(n),\mathbf{k}}, p(n)) < \frac{2}{n_k(n_k-1)\left(1-p(n)^2\right)^{p(n)^{-2}\cdot p(n)^2(n-n_k)}}\left(1 + \frac{2n_{k-1}\left(1-p(n)^2\right)^{\frac{-n}{np(n)^2}\cdot p(n)^2 n_{k-1}}(1-p(n))}{(n_k-1)}\right)^{-1} + \frac{3k^3 e^{\frac{p(n)^3(n-n_k)}{(1-p(n))}}\left(1-p(n)^2\right)^{-2}}{(n_{k-1}-1)}.$$

Then by similar reasoning as in the proof of Theorem 1.2.4 and the facts that $\lim_{n\to\infty} p(n)^4(n-n_k) = 0$ and $p(n)^{-2} \geq 1$, for $n \to \infty$, we have

$$e^{-p(n)^2(n-n_k)} > \left(1-p(n)^2\right)^{p(n)^{-2}\cdot p(n)^2(n-n_k)} > e^{-p(n)^2(n-n_k)}\left(1-p(n)^2\right)^{p(n)^2(n-n_k)} = e^{-p(n)^2(n-n_k)}(1-o(1))$$

and

$$e^{p(n)^2 n_{k-1}} < \left(1-p(n)^2\right)^{-p(n)^{-2}\cdot p(n)^2 n_{k-1}} < e^{p(n)^2 n_{k-1}} \left(1-p(n)^2\right)^{-p(n)^2 n_{k-1}}$$
$$= e^{p(n)^2 n_{k-1}}(1-o(1)).$$

Thus

$$\frac{2}{n_k(n_k-1)\left(1-p(n)^2\right)^{p(n)^{-2}\cdot p(n)^2(n-n_k)}} \left(1+\frac{2n_{k-1}\left(1-p(n)^2\right)^{\frac{-n}{np(n)^2}\cdot p(n)^2 n_{k-1}}(1-p(n))}{(n_k-1)}\right)^{-1}$$
$$= \frac{2e^{p(n)^2(n-n_k)}}{n_k^2}\left(1+\frac{2n_{k-1}}{n_k}e^{p(n)^2 n_{k-1}}\right)^{-1}(1+o(1)).$$

Since $\lim_{n\to\infty} p(n)^4(n-n_k)=0$, we have

$$\begin{aligned}
&\lim_{n\to\infty}(n-n_k)p(n)^2\left(\frac{p(n)}{(1-p(n))}-1\right)-(n-n_k)p(n)^2\left(p(n)-1\right)\\
&=\lim_{n\to\infty}(n-n_k)p(n)^3\left(\frac{1}{(1-p(n))}-1\right)\\
&=\lim_{n\to\infty}\frac{(n-n_k)p(n)^4}{(1-p(n))}\\
&=0.
\end{aligned}$$

Thus

$$P(G^{(n),\mathbf{k}},p(n)) < \frac{2e^{p(n)^2(n-n_k)}}{n_k^2}\left(1+\frac{2n_{k-1}}{n_k}e^{p(n)^2 n_{k-1}}\right)^{-1}$$
$$\left(1+\frac{3k^3n_k^2e^{\left(p(n)^3-p(n)^2\right)(n-n_k)}}{2(n_{k-1}-1)}+\frac{3k^3n_kn_{k-1}e^{\left(p(n)^3-p(n)^2\right)(n-n_k)+p(n)^2n_{k-1}}}{(n_{k-1}-1)}\right)(1+o(1)).$$

Statements (1) and (2) follow as in the proof of Proposition 1.2.6. □

Remark 3.3.6. Assume that $p(n)\le 1/2$. The $o(1)$ in (3.3.8) can be made explicit as

$$\left(1-p(n)^4(n-n_k)-p(n)^2\right)^{-1}\left(1-\frac{k}{n}\right)^{-1}(1+2p(n))(1+4p(n)^2)e^{2(n-n_k)p(n)^4}-1.$$

Also, in Proposition 3.3.5 1), the $o(1)$ can be made explicit as

$$\left(\frac{7k^3e^{\log n_{k-1}-\log n-p(n)^2n_{k-1}}}{6}-1\right)e^{2\log n_k-p(n)^2(n-n_k)-\log 2+\log\left(1+\frac{2n_{k-1}}{n_k}e^{p(n)^2n_{k-1}}\right)-c}-1$$

and under the additional assumption that

$$2\log n_k - p(n)^2(n-n_k) - \log 2 + \log\left(1+\frac{2n_{k-1}}{n_k}e^{p(n)^2 n_{k-1}}\right) < c+1$$

in Propositon 3.3.5 2), the $o(1)$ can be made explicit as

$$\begin{aligned}&\left(1-p(n)^4(n-n_k)-p(n)^2\right)^{-1}\left(1-\frac{k}{n}\right)^{-1}(1+2p(n))(1+4p(n)^2)e^{2(n-n_k)p(n)^4}\\ &\cdot\left(1+3k^3e^{2\log n+(p(n)^3-p(n)^2)(n-n_k)-\log n_{k-1}}+6k^3e^{(p(n)^3-p(n)^2)(n-n_k)+p(n)^2n_{k-1}+\log n}\right)\\ &\cdot e^{c-2\log n_k+p(n)^2(n-n_k)+\log 2-\log\left(1+\frac{2n_{k-1}}{n_k}e^{p(n)^2n_{k-1}}\right)}-1\end{aligned}$$

We consider one more application of the sieves to random k-partite graphs.

Definition 3.3.7. The *k-partite Turán graph* (named after the same Pál Turán) on n vertices is defined as the k-paritite graph on n vertices such that the partitioned sets are as equal as possible. In other words, for each $1 \leq i \leq k$, we have $n_i = \lfloor \frac{n}{k} \rfloor$ or $n_i = \lceil \frac{n}{k} \rceil$.

In the case of k-partite Turán graphs, we can calculate $\sum_{b\in B}\deg b$ a lot more precisely, using the above methods. Then we can prove the following.

Theorem 3.3.8. Let $G^{(n),\mathbf{k},t}$ denote the set of all Turán k-partite graphs with edge probability $p(n)$, and let $P(G^{(n),\mathbf{k},t}, p(n))$ be the probability of a graph from $G^{(n),\mathbf{k},t}$ having diameter 2. For $n > 2k$, we have

$$\begin{aligned}&1-\frac{n^2(1-p(n)^2)^{n(1-1/k)-1}}{2k}(1+(k-1)(1-p(n)^2)^{-n/k-1})\left(1+\frac{k}{n}\right)\\ &\leq P(G^{(n),\mathbf{k},t},p(n))\\ &\leq \frac{2k}{n^2(1-p(n)^2)^{n(1-1/k)+1}}\left(1+(k-1)(1-p(n)^2)^{1-n/k}(1-p(n))\right)^{-1}\left(1-\frac{2k}{n}\right)^{-1}\\ &\quad+\frac{4k^3\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n(1-1/k)+1}(1-p(n)^2)^{-2}}{n(k-1)}\left(1-\frac{2k}{n}\right)^{-4}.\end{aligned}$$

Corollary 3.3.9. Let $G^{(n),\mathbf{k},t}$ be defined as in Theorem 3.3.8. If $p(n)=\frac{1}{2}$, we have

$$P(G^{(n),\mathbf{k},t},1/2) \geq 1-\frac{4n^2(3/4)^{n(1-1/k)}}{6k}\left(1+(k-1)(4/3)^{n/k+1}\right)\left(1+\frac{k}{n}\right).$$

In the case when $p(n) \to 0$ as $n \to \infty$, we can prove the following.

Proposition 3.3.10. Let $G^{(n),\mathbf{k},t}$ be as in Theorem 3.3.8. Let $\lim_{n\to\infty} p(n)^4 n = 0$. As $n \to \infty$, we have

$$\begin{aligned} &1 - \frac{n^2 e^{-np(n)^2\left(1-\frac{1}{k}\right)}}{2k}\left(1+(k-1)e^{\frac{np(n)^2}{k}}\right)(1+o(1)) \\ &\leq P(G^{(n),\mathbf{k},t}, p(n)) \\ &\leq \frac{2ke^{np(n)^2\left(1-\frac{1}{k}\right)}}{n^2}\left(1+(k-1)e^{\frac{np(n)^2}{k}}\right)^{-1} \\ &\cdot\left(1+\frac{2k^2ne^{\left(np(n)^3-np(n)^2\right)(1-1/k)}}{(k-1)}+2k^2ne^{\left(np(n)^3-np(n)^2\right)(1-1/k)+\frac{np(n)^2}{k}}\right)(1+o(1)) \end{aligned} \quad (3.3.9)$$

as $n \to \infty$. Suppose further that

$$\lim_{n\to\infty} 2\log n - \log k - np(n)^2\left(1-\frac{1}{k}\right) - \log 2 + \log\left(1+(k-1)e^{\frac{np(n)^2}{k}}\right) = c$$

for some $c \in \mathbb{R}\backslash\{0\}$.

1) If $c < 0$, we have

$$P(G^{(n),\mathbf{k},t}, p(n)) \geq 1-(1+o(1))e^{c}.$$

2) If $c > 0$, we have

$$P(G^{(n),\mathbf{k},t}, p(n)) \leq (1+o(1))e^{-c}.$$

Remark 3.3.11. Assume that $p(n) \leq 1/4$. The $o(1)$ in the lower bound of (3.3.9) can be made explicit as

$$(1+4p(n)^2)\left(1+\frac{k}{n}\right)-1,$$

while the $o(1)$ in the upper bound of (3.3.9) can be made explicit as

$$\left(1-p(n)^4n(1+1/k)-p(n)^2\right)^{-1}\left(1+\frac{4p(n)^2}{3}\right)^2(1+2p)\left(1-\frac{2k}{n}\right)^{-5}e^{2p(n)^4n(1-1/k)+p(n)^3+2p(n)^4}-1.$$

Also, in Proposition 3.3.10 1), the $o(1)$ can be made explicit as

$$(1+4p(n)^2)\left(1+\frac{k}{n}\right)e^{c-2\log n+\log k+np(n)^2\left(1-\frac{1}{k}\right)+\log 2-\log\left(1+(k-1)e^{\frac{np(n)^2}{k}}\right)}-1,$$

and under the additional assumption that

$$c - 1 < 2\log n - \log k - np(n)^2\left(1 - \frac{1}{k}\right) - \log 2 + \log\left(1 + (k-1)e^{\frac{np(n)^2}{2}}\right) < c + 1$$

in Proposition 3.3.10 2), the $o(1)$ can be made explicit as

$$\left(1 - p(n)^4 n(1+1/k) - p(n)^2\right)^{-1}\left(1 + \frac{4p(n)^2}{3}\right)^2 (1+2p)\left(1 - \frac{2k}{n}\right)^{-5} e^{2p(n)^4 n(1-1/k) + p(n)^3 + 2p(n)^4}$$
$$\cdot\left(1 + \frac{8k^3 e^{\frac{(k-1)c(1-p)}{(k-2)} + 2 - \frac{(k-1)\log n}{2(k-2)}}}{(k-1)^2}\right)\cdot e^{c - 2\log n + \log k + np(n)^2\left(1-\frac{1}{k}\right) + \log 2 - \log\left(1 + (k-1)e^{\frac{np(n)^2}{k}}\right)} - 1.$$

We can similarly derive all of the above results for directed k-partite graphs.

Theorem 3.3.12. For each $n \geq k$, choose $n_1, \ldots, n_k \in \mathbb{N}$ such that $n_1 \leq n_2 \leq \ldots \leq n_k$, $n_{k-1} \geq 2$, and $n_1 + n_2 + \cdots + n_k = n$. Let $G^{(n),\mathbf{k},\to}$ denote the set of all directed k-partite graphs with the partite sets having $n_1, n_2, \ldots, n_k$ vertices respectively with edge probability $p(n)$, and let $P(G^{(n),\mathbf{k},\to}, p(n))$ be the probability of a graph from $G^{(n),k,,\to}$ having diameter 2. Then

$$\begin{aligned}
&1 - n_k^2(1-p(n)^2)^{n-n_k}\left(1 + \frac{2n_{k-1}(1-p(n)^2)^{-n_{k-1}}}{n_k} + \frac{7k^2 n_{k-1}^2(1-p(n)^2)^{n_k - n_{k-1} - n_{k-2}}}{3n_k^2}\right)\\
&\leq P(G^{(n),\mathbf{k},\to}, p(n))\\
&\leq \frac{1}{n_k(n_k-1)(1-p(n)^2)^{n-n_k}}\left(1 + \frac{n_{k-1}(1-p(n)^2)^{-n_{k-1}}(1-p(n))}{(n_k-1)}\right)^{-1}\\
&\quad + \frac{3k^3\left(1 + \frac{p(n)^3}{(1-p(n))}\right)^{n-n_k}(1-p(n)^2)^{-2}}{2(n_{k-1}-1)}.
\end{aligned}$$

Corollary 3.3.13. Let $P(G^{(n),\mathbf{k},\to}, p(n))$ be defined as in Theorem 3.3.12. If $p(n) = \frac{1}{2}$, then

$$P(G^{(n),\mathbf{k}}, 1/2) \geq 1 - n_k^2(3/4)^{n-n_k}\left(1 + \frac{2n_{k-1}(3/4)^{-n_{k-1}}}{n_k} + \frac{7k^2 n_{k-1}^2(3/4)^{n_k - n_{k-1} - n_{k-2}}}{3n_k^2}\right).$$

Proposition 3.3.14. Let $P(G^{(n),\mathbf{k},\to}, p(n))$ be as in Theorem 3.3.12. Let $\lim_{n\to\infty} p(n)^4(n-$

$n_k) = 0$. Then

$$\begin{aligned}
&1 - n_k^2 e^{-p(n)^2(n-n_k)}\left(1 + \frac{2n_{k-1}}{n_k}e^{p(n)^2 n_{k-1}}\left(1 + \frac{7k^2 n_{k-1}e^{-p(n)^2(n_k-n_{k-2})}}{6n_k}\right)\right)\\
&\leq P(G^{(n),\mathbf{k}}, p(n))\\
&\leq \frac{e^{p(n)^2(n-n_k)}}{n_k^2}\left(1 + \frac{2n_{k-1}}{n_k}e^{p(n)^2 n_{k-1}}\right)^{-1}\\
&\left(1 + \frac{3k^3 n_k^2 e^{\left(p(n)^3-p(n)^2\right)(n-n_k)}}{2(n_{k-1}-1)} + \frac{3k^3 n_k n_{k-1} e^{\left(p(n)^3-p(n)^2\right)(n-n_k)+p(n)^2 n_{k-1}}}{(n_{k-1}-1)}\right)(1+o(1)).
\end{aligned} \tag{3.3.10}$$

Suppose further that

$$\lim_{n\to\infty} \log n_{k-1} - \log n - p(n)^2 n_{k-1} = -\infty,$$

$$\lim_{n\to\infty} 2\log n + (p(n)^3 - p(n)^2)(n-n_k) - \log n_{k-1} = -\infty,$$

$$\lim_{n\to\infty} \left(p(n)^3 - p(n)^2\right)(n-n_k) + p(n)^2 n_{k-1} + \log n = -\infty,$$

and that

$$\lim_{n\to\infty} 2\log n_k - p(n)^2(n-n_k) + \log\left(1 + \frac{2n_{k-1}}{n_k}e^{p(n)^2 n_{k-1}}\right) = c$$

for some $c \in \mathbb{R}\backslash\{0\}$.

1) If $c < 0$, we have

$$P(G^{(n)}, p(n)) \geq 1 - (1+o(1))e^c.$$

2) If $c > 0$, we have

$$P(G^{(n)}, p(n)) \leq (1+o(1))e^{-c}.$$

Remark 3.3.15. Assume that $p(n) \leq 1/2$. The $o(1)$ in (3.3.10) can be made explicit as

$$\left(1 - p(n)^4(n-n_k) - p(n)^2\right)^{-1}\left(1 - \frac{k}{n}\right)^{-1}(1+2p(n))(1+4p(n)^2)e^{2(n-n_k)p(n)^4} - 1.$$

Also, in Proposition 3.3.14 1), the $o(1)$ can be made explicit as

$$\left(\frac{7k^3 e^{\log n_{k-1} - \log n - p(n)^2 n_{k-1}}}{6} - 1\right)e^{2\log n_k - p(n)^2(n-n_k) + \log\left(1+\frac{2n_{k-1}}{n_k}e^{p(n)^2 n_{k-1}}\right) - c} - 1$$

and under the additional assumption that

$$2\log n_k - p(n)^2(n-n_k) + \log\left(1+\frac{2n_{k-1}}{n_k}e^{p(n)^2 n_{k-1}}\right) < c+1$$

in Proposition 3.3.14 2), the $o(1)$ can be made explicit as

$$\begin{aligned}
&\left(1-p(n)^4(n-n_k)-p(n)^2\right)^{-1}\left(1-\frac{k}{n}\right)^{-1}(1+2p(n))(1+4p(n)^2)e^{2(n-n_k)p(n)^4} \\
&\cdot\left(1+3k^3e^{2\log n+(p(n)^3-p(n)^2)(n-n_k)-\log n_{k-1}}+6k^3e^{(p(n)^3-p(n)^2)(n-n_k)+p(n)^2 n_{k-1}+\log n}\right) \\
&\cdot e^{c-2\log n_k+p(n)^2(n-n_k)-\log\left(1+\frac{2n_{k-1}}{n_k}e^{p(n)^2 n_{k-1}}\right)}-1.
\end{aligned}$$

Theorem 3.3.16. Let $G^{(n),\mathbf{k},t,\rightarrow}$ denote the set of all Turán directed k-partite graphs with edge probability $p(n)$, and let $P(G^{(n),\mathbf{k},t,\rightarrow},p(n))$ be the probability of a graph from $G^{(n),\mathbf{k},t}$ having diameter 2. For $n>2k$,

$$\begin{aligned}
&1-\frac{n^2(1-p(n)^2)^{n(1-1/k)-1}}{k}\left(1+(k-1)(1-p(n)^2)^{-n/k-1}\right)\left(1+\frac{k}{n}\right) \\
&\leq P(G^{(n),\mathbf{k},t,\rightarrow},p(n)) \\
&\leq \frac{k}{n^2(1-p(n)^2)^{n(1-1/k)+1}}\left(1+(k-1)(1-p(n)^2)^{1-n/k}(1-p(n))\right)^{-1}\left(1-\frac{2k}{n}\right)^{-1} \\
&\quad+\frac{2k^3\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n(1-1/k)+1}(1-p(n)^2)^{-2}}{n(k-1)}\left(1-\frac{2k}{n}\right)^{-4}.
\end{aligned}$$

Corollary 3.3.17. Let $G^{(n),\mathbf{k},t,\rightarrow}$ be defined as in Theorem 3.3.16. If $p(n)=\frac{1}{2}$, then

$$P(G^{(n),\mathbf{k},t,\rightarrow},1/2)\geq 1-\frac{4n^2(3/4)^{n(1-1/k)}}{3k}\left(1+\frac{4(k-1)(4/3)^{n/k+1}}{3}\right)\left(1+\frac{k}{n}\right).$$

Proposition 3.3.18. Let $G^{(n),\mathbf{k},t,\rightarrow}$ be as in Theorem 3.3.16. Let $\lim_{n\to\infty}p(n)^4 n=0$.

Then

$$
\begin{aligned}
&1-\frac{n^2e^{-np(n)^2\left(1-\frac{1}{k}\right)}}{k}\left(1+(k-1)e^{\frac{np(n)^2}{k}}\right)(1+o(1))\\
&\le P(G^{(n),\mathbf{k},t},p(n))\\
&\le \frac{ke^{np(n)^2\left(1-\frac{1}{k}\right)}}{n^2}\left(1+(k-1)e^{\frac{np(n)^2}{k}}\right)^{-1}\\
&\quad\cdot\left(1+\frac{2k^2ne^{\left(np(n)^3-np(n)^2\right)(1-1/k)}}{(k-1)}+2k^2ne^{\left(np(n)^3-np(n)^2\right)(1-1/k)+\frac{np(n)^2}{k}}\right)(1+o(1))
\end{aligned}
\tag{3.3.11}
$$

as $n\to\infty$. Suppose further that

$$
\lim_{n\to\infty} 2\log n-\log k-np(n)^2\left(1-\frac{1}{k}\right)+\log\left(1+(k-1)e^{\frac{np(n)^2}{k}}\right)=c
$$

for some $c\in\mathbb{R}\backslash\{0\}$.

1) If $c<0$, we have

$$
P(G^{(n),\mathbf{k},t},p(n))\ge 1-(1+o(1))e^{c}.
$$

2) If $c>0$, we have

$$
P(G^{(n),\mathbf{k},t},p(n))\le (1+o(1))e^{-c}.
$$

Remark 3.3.19. Assume that $p(n)\le 1/4$. The $o(1)$ in the lower bound of (3.3.11) can be made explicit as

$$
(1+4p(n)^2)\left(1+\frac{k}{n}\right)-1,
$$

while the $o(1)$ in the upper bound of (3.3.11) can be made explicit as

$$
\left(1-p(n)^4n(1+1/k)-p(n)^2\right)^{-1}\left(1+\frac{4p(n)^2}{3}\right)^2(1+2p)\left(1-\frac{2k}{n}\right)^{-5}e^{2p(n)^4n(1-1/k)+p(n)^3+2p(n)^4}-1.
$$

Also, in Proposition 3.3.18 1), the $o(1)$ can be made explicit as

$$
(1+4p(n)^2)\left(1+\frac{k}{n}\right)e^{c-2\log n+\log k+np(n)^2\left(1-\frac{1}{k}\right)-\log\left(1+(k-1)e^{\frac{np(n)^2}{k}}\right)}-1,
$$

and under the additional assumption that

$$c-1<2\log n-\log k-np(n)^2\left(1-\frac{1}{k}\right)+\log\left(1+(k-1)e^{\frac{np(n)^2}{2}}\right)<c+1$$

in Proposition 3.3.18 2), the $o(1)$ can be made explicit as

$$\left(1-p(n)^4n(1+1/k)-p(n)^2\right)^{-1}\left(1+\frac{4p(n)^2}{3}\right)^2(1+2p)\left(1-\frac{2k}{n}\right)^{-5}e^{2p(n)^4n(1-1/k)+p(n)^3+2p(n)^4}$$
$$\cdot\left(1+\frac{4k^3e^{\frac{(k-1)c(1-p(n)}{(k-2)}+2-\frac{(k-1)\log n}{2(k-2)}}}{(k-1)^2}\right)\cdot e^{c-2\log n+\log k+np(n)^2\left(1-\frac{1}{k}\right)+\log 2-\log\left(1+(k-1)e^{\frac{np(n)^2}{k}}\right)}-1.$$

3.4 Bipartite Graphs with Diameter 3

Here we analyze bipartite graphs in a similar way to k-partite graphs, but instead of considering diameter 2, we consider diameter 3 since, except for the complete bipartite graph, all bipartite graphs have diameter at least 3.

Theorem 3.4.1. For each $n\in\mathbb{N}$, $n\geq 4$, pick $n_1,n_2\in\mathbb{N}$ such that $2\leq n_1\leq n_2$ and $n_1+n_2=n$. Let $\mathbf{b}=(n_1,n_2)$ and let $G^{(n),\mathbf{b}}$ denote the set of all bipartite graphs with the partite sets having n_1 and n_2 vertices respectively, with edge probability $p(n)$, and let $P(G^{(n),\mathbf{b}},p(n))$ be the probability of a graph from $G^{(n),\mathbf{b}}$ having diameter 3. Then

$$\begin{aligned}
&1-\frac{n_2^2(1-p(n)^2)^{n_1}}{2}\left(1+\frac{n_1^2(1-p(n)^2)^{n_2-n_1}}{n_2^2}\right)\\
&\leq P(G^{(n),\mathbf{b}},p(n))\\
&\leq\left(\frac{2}{n_2(n_2-1)(1-p(n)^2)^{n_1}}+\frac{\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n_1}}{n_2}\left(8+\frac{8}{(1-p(n))}\right)\right)\\
&\quad\cdot\left(1+\frac{n_1(n_1-1)(1-p(n)^2)^{n_2-n_1}}{n_2(n_2-1)}\right)^{-1}.
\end{aligned}$$

Proof. As in the proof of Theorem of 1.2.4, we may assume that $p(n)\in\mathbb{Q}\cap(0,1)$ for all $n\in\mathbb{N}$.

Let $p(n)=\frac{r}{s}$ where $r,s\in\mathbb{N}$. As in the proof of Theorem 1.2.4, we let A be the set of all graphs in $G^{(n),\mathbf{b}}$, allowing for a number of duplicates of each possible graph to accommodate the edge probability $p(n)$. Since the complete bipartite graph has n_1n_2 edges, we have $r^{n_1n_2}$ copies of the complete bipartite graph and $|A| = s^{n_1n_2}$.

We let B be the set of all pairs of vertices such that both vertices of a pair occur in the same partite set. Thus, $|B| = \binom{n_1}{2}\binom{n_2}{2}$. For $a \in A$ and $b \in B$, we write $a \sim b$ if the pair of vertices b in the graph a do not share a common neighbouring vertex. Thus, we will have $\omega(a) = 0$ if and only if a is connected with diameter at most 3. For each pair of vertices $b \in B$ in the set containing n_1 vertices, we have

$$D(r, s, n, n_1) := \deg b = ((s-r)^2 + 2r(s-r))^{n_2}((s-r)+r)^{n_1n_2-2n_2}.$$

For each pair of vertices $b \in B$ in the set containing n_2 vertices, the n_1 and n_2 are switched in the above equality. It follows that

$$\sum_{b\in B} \deg b = \frac{s^{n_1n_2}n_1(n_1-1)(1-p(n)^2)^{n_2}}{2} + \frac{s^{n_1n_2}n_2(n_2-1)(1-p(n)^2)^{n_1}}{2}.$$

By the simple sieve, we obtain

$$\begin{aligned} P(G^{(n),\mathbf{b}}, p(n)) &> 1 - \frac{n_1^2(1-p(n)^2)^{n_2}}{2} - \frac{n_2^2(1-p(n)^2)^{n_1}}{2} \\ &= 1 - \frac{n_2^2(1-p(n)^2)^{n_1}}{2}\left(1 + \frac{n_1^2(1-p(n)^2)^{n_2-n_1}}{n_2^2}\right) \end{aligned}$$

We now try to get an upper bound for $P(G^{(n),\mathbf{b}}, p(n))$, in which we need to estimate $\sum_{b_1,b_2\in B} n(b_1, b_2)$. In the following, we calculate $n(b_1, b_2)$, depending on how many vertices b_1 and b_2 have in common.

Case 1. Suppose that b_1 and b_2 are two pairs of vertices lying in the same partite set that have no vertices in common. Let there be n_i vertices in total in this partite set for $i = 1$ or 2. Since there are no edges between any of these four vertices, we then have that

$$n(b_1, b_2) = \frac{D(r, s, n, n_i)^2}{s^{n_1n_2}},$$

and thus

$$\sum_{\substack{b_1,b_2\in B \\ \text{4 vertices, all in 1 set}}} n(b_1, b_2) = \binom{n_1}{2}\binom{n_1-2}{2}\frac{D(r, s, n, n_2)^2}{s^{n_1n_2}} + \binom{n_2}{2}\binom{n_2-2}{2}\frac{D(r, s, n, n_1)^2}{s^{n_1n_2}}. \tag{3.4.1}$$

Case 2. Suppose that b_1 and b_2 are two pairs of vertices lying in the two different partite sets. By independent and dependent selections of edges, we can calculate

$$n(b_1, b_2) = \frac{D(r,s,n,n_1)D(r,s,n,n_2)}{s^{n_1 n_2}} \frac{s^4((s-r)^4 + 4r(s-r)^3 + 2r^2(s-r)^2)}{(s^2-r^2)^4}$$

and thus

$$\sum_{\substack{b_1,b_2 \in B \\ \text{4 vertices, in both sets}}} n(b_1, b_2)$$
$$= 2\binom{n_1}{2}\binom{n_2}{2}\frac{D(r,s,n,n_1)D(r,s,n,n_2)}{s^{n_1 n_2}} \cdot \frac{p(n)^{-4}((p(n)^{-1}-1)^4 + 4(p(n)^{-1}-1)^3 + 2(p(n)^{-1}-1)^2)}{(p(n)^{-2}-1)^4}$$
$$< 2\binom{n_1}{2}\binom{n_2}{2}\frac{D(r,s,n,n_1)D(r,s,n,n_2)}{s^{n_1 n_2}} \cdot \left(1 + \frac{4p(n)^3}{(1-p(n))^2}\right). \tag{3.4.2}$$

Case 3. Suppose that b_1 and b_2 are two pairs of vertices lying in the same partite set and have exactly one vertex in common. Again, by independent and dependent selections of edges, we can calculate

$$\sum_{\substack{b_1,b_2 \in B \\ \text{3 vertices}}} < \frac{D(r,s,n,n_1)^2 n_1(n_1-1)(n_1-2)}{s^{n_1 n_2}}\left(1 + \frac{1}{p(n)^{-3}+p(n)^{-2}-p(n)^{-1}-1}\right)^{n_2}$$
$$+ \frac{D(r,s,n,n_2)^2 n_2(n_2-1)(n_2-2)}{s^{n_1 n_2}}\left(1 + \frac{1}{p(n)^{-3}+p(n)^{-2}-p(n)^{-1}-1}\right)^{n_1}$$
$$< \frac{D(r,s,n,n_1)^2 n_1(n_1-1)(n_1-2)}{s^{n_1 n_2}}\left(1 + \frac{p(n)^3}{(1-p(n))}\right)^{n_2}$$
$$+ \frac{D(r,s,n,n_2)^2 n_2(n_2-1)(n_2-2)}{s^{n_1 n_2}}\left(1 + \frac{p(n)^3}{(1-p(n))}\right)^{n_1}. \tag{3.4.3}$$

Case 4. Finally, if b_1 and b_2 are identical, then

$$n(b_1, b_2) = \deg b_1 = \deg b_2,$$

so that

$$\sum_{\substack{b_1,b_2 \in B \\ \text{2 vertices}}} n(b_1, b_2) = \sum_{b \in B} \deg b = \frac{s^{n_1 n_2} n_1(n_1-1)(1-p(n)^2)^{n_2}}{2} + \frac{s^{n_1 n_2} n_2(n_2-1)(1-p(n)^2)^{n_1}}{2}$$
$$= \frac{s^{n_1 n_2} n_2(n_2-1)(1-p(n)^2)^{n_1}}{2}\left(1 + \frac{n_1(n_1-1)(1-p(n)^2)^{n_2-n_1}}{n_2(n_2-1)}\right). \tag{3.4.4}$$

Using (3.4.4), we have that

$$\left(\sum_{b\in B}\deg b\right)^2 > \frac{s^{2n_1n_2}n_2^2(n_2-1)^2(1-p(n)^2)^{2n_1}}{4}\left(1+\frac{n_1(n_1-1)(1-p(n)^2)^{n_2-n_1}}{n_2(n_2-1)}\right)^2. \tag{3.4.5}$$

Note that adding (3.4.1) and (3.4.2) together, multiplying the result by $s^{n_1n_2}$, and then dividing the result by (3.4.5) will give an expression that is bounded above by

$$1+\frac{8n_1^2\left(\frac{p(n)^3}{(1-p(n))^2}\right)(1-p(n)^2)^{n_2-n_1}}{n_2^2}\left(1+\frac{n_1(n_1-1)(1-p(n)^2)^{n_2-n_1}}{n_2(n_2-1)}\right)^{-2}.$$

Dividing (3.4.3) and (3.4.4) by (3.4.5), we therefore deduce, by the Turán sieve, that

$$\begin{aligned}
&P(G^{(n),\mathbf{b}},p(n))\\
&< \frac{2}{n_2(n_2-1)(1-p(n)^2)^{n_1}}\left(1+\frac{n_1(n_1-1)(1-p(n)^2)^{n_2-n_1}}{n_2(n_2-1)}\right)^{-1}\\
&\quad+\left(\frac{4n_1^3\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n_2}(1-p(n)^2)^{2n_2-2n_1}}{n_2^4}+\frac{4\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n_1}}{n_2}+\frac{8n_1^2\left(\frac{p(n)^3}{(1-p(n))^2}\right)(1-p(n)^2)^{n_2-n_1}}{n_2^2}\right)\\
&\quad\cdot\left(1+\frac{n_1(n_1-1)(1-p(n)^2)^{n_2-n_1}}{n_2(n_2-1)}\right)^{-2}.
\end{aligned}$$

Notice that

$$\frac{p(n)^3}{(1-p(n))}<\frac{1}{n_1}\left(1+n_1\frac{p(n)^3}{(1-p(n))}\right)<\frac{1}{n_1}\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n_1}.$$

and

$$\begin{aligned}
\left(1+\frac{p(n)^3}{(1-p(n))}\right)(1-p(n)^2)^2 &= 1-2p^2+p^3+2p^4-p^5-p^6\\
&= 1-p^2(2-p-2p^2+p^3+p^4)\\
&< 1-p^2(2-p-2p^2+p^3)\\
&= 1-p^2(2-p)(1-p^2)\\
&< 1.
\end{aligned}$$

It follows that

$$
\begin{aligned}
& P(G^{(n),\mathbf{b}}, p(n)) \left(1 + \frac{n_1(n_1-1)(1-p(n)^2)^{n_2-n_1}}{n_2(n_2-1)}\right) \\
< & \frac{2}{n_2(n_2-1)(1-p(n)^2)^{n_1}} + \frac{4n_1^3 \left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n_2} (1-p(n)^2)^{2n_2-2n_1}}{n_2^4} \\
& + \frac{\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n_1}}{n_2} \left(4 + \frac{8n_1(1-p(n)^2)^{n_2-n_1}}{n_2(1-p(n))}\right) \\
< & \frac{2}{n_2(n_2-1)(1-p(n)^2)^{n_1}} + \frac{\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n_1}}{n_2} \left(4 + \frac{8n_1(1-p(n)^2)^{n_2-n_1}}{n_2(1-p(n))} + \frac{4n_1^3}{n_2^3}\right) \\
\leq & \frac{2}{n_2(n_2-1)(1-p(n)^2)^{n_1}} + \frac{\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n_1}}{n_2} \left(8 + \frac{8}{(1-p(n))}\right)
\end{aligned}
$$

from which we obtain our upper bound. $\square$

By substituting in $p(n) = \frac{1}{2}$, we deduce from Theorem 3.4.1 the following.

Corollary 3.4.2. Let $P(G^{(n),\mathbf{b}}, p(n))$ be defined as in Theorem 3.4.1. If $p(n) = \frac{1}{2}$, then we have

$$
\begin{aligned}
& 1 - \frac{n_2^2(3/4)^{n_1}}{2}\left(1 + \frac{n_1^2(3/4)^{n_2-n_1}}{n_2^2}\right) \\
& \leq P(G^{(n),\mathbf{b}}, 1/2) \\
& \leq \left(\frac{2(4/3)^{n_1}}{n_2(n_2-1)} + \frac{24(5/4)^{n_1}}{n_2}\right)\left(1 + \frac{n_1(n_1-1)(3/4)^{n_2-n_1}}{n_2(n_2-1)}\right)^{-1}.
\end{aligned}
$$

Remark 3.4.3. The upper bound given for $P(G^{(n),\mathbf{b}}, 1/2)$ in Corollary 3.4.2 will in general only be non-trivial, i.e., less than 1, when n_2 much larger than n_1. For instance, if $n_1 < \frac{2\log n_2 - \log 8}{\log(4/3)}$ and $n_1 < \frac{\log n_2 - \log 48}{\log(5/4)}$, then the upper bound will be less than 1.

In the situation where the edge probability $p(n) \to 0$ as $n \to \infty$, we will show the following.

Proposition 3.4.4. Let $P(G^{(n),\mathbf{b}}, p(n))$ be as in Theorem 3.4.1. Let $\lim_{n\to\infty} np(n)^4 = 0$. We have

$$\begin{aligned}
&1 - \frac{n_2^2 e^{-n_1 p(n)^2}}{2}\left(1 + e^{2\log n_1 - 2\log n_2 - (n_2-n_1)p(n)^2}\right) \\
&\leq P(G^{(n)}, p(n)) \\
&\leq (1+o(1))\left(\frac{2}{n_2^2}e^{n_1 p(n)^2}\right)\left(1 + e^{2\log n_1 - 2\log n_2 - (n_2-n_1)p(n)^2}\right)^{-1}\left(1 + 8n_2 e^{n_1 p(n)^2(p(n)-1)}\right).
\end{aligned} \tag{3.4.6}$$

Suppose further that

$$\lim_{n\to\infty} 2\log n_1 - 2\log n_2 - (n_2 - n_1)p(n)^2 = -\infty,$$

and

$$\lim_{n\to\infty} 2\log n_2 - n_1 p(n)^2 - \log 2 = c$$

for some $c \in \mathbb{R}$.

1) If $c < 0$, we have

$$P(G^{(n)}, p(n)) \geq 1 - (1+o(1))e^{c}.$$

2) If $c > 0$, we have

$$P(G^{(n)}, p(n)) \leq (1+o(1))e^{-c}.$$

Proof. We can get

$$\begin{aligned}
&P(G^{(n),\mathbf{b}}, p(n))\left(1 + \frac{n_1(n_1-1)\left(1-p(n)^2\right)^{n_2-n_1}}{n_2(n_2-1)}\right) \\
&< \frac{2}{n_2(n_2-1)\left(1-p(n)^2\right)^{n_1}} + \left(8 + \frac{8}{1-p(n)^2}\right)\frac{\left(1+\frac{p(n)^3}{1-p(n)^2}\right)^{n_1}}{n_2}.
\end{aligned}$$

Furthermore, since $\lim_{n\to\infty} np(n)^4 = 0$, we have $\lim_{n\to\infty} n_1 p(n)^4 = 0$. Thus, since $p(n)^{-2} \geq 1$, we have

$$\left(1-p(n)^2\right)^{n_1} > e^{-n_1 p(n)^2}\left(1-p(n)^2\right)^{n_1 p(n)^2} = e^{-n_1 p(n)^2}(1-o(1))$$

and

$$\frac{\left(1+\frac{p(n)^3}{1-p(n)^2}\right)^{n_1}}{n_2} < \frac{e^{\frac{n_1 p(n)^3}{1-p(n)^2}}}{n_2} = \frac{e^{n_1 p(n)^2}}{n_2^2} \cdot n_2 e^{n_1 p(n)^2\left(\frac{p(n)}{1-p(n)}-1\right)}.$$

Since $\lim_{n\to\infty} np(n)^4 = 0$, we have

$$\begin{aligned}
&\lim_{n\to\infty} n_1 p(n)^2 \left(\frac{p(n)}{1-p(n)}-1\right) - n_1 p(n)^2 \left(p(n)-1\right) \\
&= \lim_{n\to\infty} n_1 p(n)^3 \left(\frac{1}{1-p(n)}-1\right) \\
&= \lim_{n\to\infty} n_1 p(n)^4 \left(\frac{1}{1-p(n)}\right) \\
&= 0.
\end{aligned}$$

Also,

$$\begin{aligned}
\frac{n_1^2\left(1-p(n)^2\right)^{n_2-n_1}}{n_2^2\left(1-p(n)^2\right)} &= e^{2\log n_1 - 2\log n_2}\left(1-p(n)^2\right)^{p(n)^{-2}\cdot(n_2-n_1)p(n)^2}\left(1-p(n)^2\right)^{-1} \\
&> e^{2\log n_1 - 2\log n_2 - (n_2-n_1)p(n)^2}\left(1-p(n)^2\right)^{(n_2-n_1)p(n)^2}\left(1-p(n)^2\right)^{-1}.
\end{aligned}$$

Since $\lim_{n\to\infty} np(n)^4 = 0$, we have

$$\lim_{n\to\infty}\left(1-p(n)^2\right)^{(n_2-n_1)p(n)^2} = 1.$$

We thus obtain our bounds. Statements (1) and (2) follow as in the proof of Proposition 1.2.6. $\square$

Remark 3.4.5. Assume that $p(n) \le \frac{1}{4}$. The $o(1)$ in (3.4.6) can be made explicit as

$$\left(1+\frac{2}{(n-2)}\right)(1-p(n)^4 n_1 - p(n)^2)^{-1}(1-p(n)^4(n_2-n_1)-p(n)^2)^{-1} - 1.$$

Also, in Proposition 3.4.4 1), the o(1) can be made explicit as

$$\left(1+e^{2\log n_1 - 2\log n_2 - (n_2-n_1)p(n)^2}\right)e^{2\log n_2 - n_1 p(n)^2 - \log 2 - c} - 1$$

and under the additional assumption that

$$c-1 \le 2\log n_2 - n_1 p(n)^2 - \log 2 \le c+1$$

in Proposition 3.4.4 2), the $o(1)$ can be made explicit as

$$\left(1+\frac{2}{(n-2)}\right)(1-p(n)^4n_1-p(n)^2)^{-1}(1-p(n)^4(n_2-n_1)-p(n)^2)^{-1}$$
$$\cdot e^{c+\log 2-2\log n_2+n_1p(n)^3}\left(1+e^{2\log n_1-2\log n_2-(n_2-n_1)p(n)^2}\right)^{-1}\left(1+128e^{c(1-p)+1+p(1-\log 2)-\frac{\log n}{2}}\right)-1.$$

Substituting in $n_1=n_2=\frac{n}{2}$ or $n_1=\frac{n-1}{2}$ and $n_2=\frac{n+1}{2}$ can lead to similar asymptotics for Turán bipartite graphs.

Theorem 3.4.6. Let $G^{(n),\mathbf{b},t}$ denote the set of all Turán bipartite graphs with edge probability $p(n)$, and let $P(G^{(n),\mathbf{b},t},p(n))$ be the probability of a graph from $G^{(n),\mathbf{b},t}$ having diameter 3. For $n\geq 4$, we have

$$\begin{aligned}
&1-\frac{(n+1)^2(1-p(n)^2)^{(n-1)/2}}{8}\\
&\leq P(G^{(n),\mathbf{b},t},p(n))\\
&\leq\left(\frac{8}{n(n-2)(1-p(n)^2)^{n/2}}+\frac{2\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n/2}}{n}\left(8+\frac{8}{(1-p(n))}\right)\right)\\
&\quad\cdot\left(1+\frac{(n-3)(1-p(n)^2)}{(n+1)}\right)^{-1}.
\end{aligned}$$

Substituting $p(n)=\frac{1}{2}$ gives the following.

Corollary 3.4.7. Let $G^{(n),\mathbf{b},t}$ be defined as in Corollary 3.4.6. If $p(n)=\frac{1}{2}$, then we have

$$P(G^{(n),\mathbf{b}},1/2)\geq 1-\frac{(n+1)^2(3/4)^{(n-1)/2}}{4}.$$

In the situation where the edge probability $p(n)\to 0$ as $n\to\infty$, we have the following.

Proposition 3.4.8. Let $G^{(n),\mathbf{b},t}$ be defined as in Corollary 3.4.6. Let $\lim_{n\to\infty} np(n)^4=0$. We have

$$1-\frac{n^2e^{-\frac{np(n)^2}{2}}}{4}(1+o(1))\leq P(G^{(n),\mathbf{b},t},p(n))\leq(1+o(1))\left(\frac{4}{n^2}e^{\frac{np(n)^2}{2}}\right)\left(1+8ne^{\frac{np(n)^2}{2}\left(p(n)^2-1\right)}\right).\tag{3.4.7}$$

Suppose further that

$$\lim_{n\to\infty}\left(2\log n - \log 4 - \frac{np(n)^2}{2}\right) = c$$

for some $c \in \mathbb{R}$.

1) If $c < 0$, we have

$$P(G^{(n)}, p(n)) \geq 1 - (1 + o(1))e^c.$$

2) If $c > 0$, we have

$$P(G^{(n)}, p(n)) \leq (1 + o(1))e^{-c}.$$

Remark 3.4.9. The $o(1)$ in the lower bound of (3.4.7) can be made explicit as

$$\left(1 + \frac{1}{n}\right)^2 (1 - p(n)^2)^{-1/2} - 1,$$

while the $o(1)$ in the upper bound of (3.4.7) can be made explicit as

$$\frac{n}{(n-2)}\left(1 - \frac{np(n)^4}{2} - p(n)^2\right)^{-1}(1 - p(n)^{-1}\left(1 - \frac{p(n)^2}{2} - \frac{(2 - 2p(n)^2)}{(n+1)}\right)^{-1} e^{np(n)^4} - 1.$$

Also, in Proposition 3.4.8 1), the $o(1)$ can be made explicit as

$$\left(1 + \frac{1}{n}\right)^2 (1 - p(n)^2)^{-1/2} e^{2\log n - \log 4 - \frac{np(n)^2}{2} - c}$$

and under the additional assumption that $p(n) \leq \frac{1}{4}$ and

$$c - 1 < 2\log n - \log 4 - \frac{np(n)^2}{2} < c + 1$$

in Proposition 3.4.8 2), the $o(1)$ can be made explicit as

$$\frac{n}{(n-2)}\left(1 - \frac{np(n)^4}{2} - p(n)^2\right)^{-1}(1 - p(n)^{-1}\left(1 - \frac{p(n)^2}{2} - \frac{(2 - 2p(n)^2)}{(n+1)}\right)^{-1} e^{np(n)^4}$$
$$\cdot\left(1 + e^{1 + p + (c + \log 4)(1 - p(n)) - \frac{\log n}{2}}\right) e^{c - 2\log n + \log 4 + \frac{np(n)^2}{2}} - 1.$$

Again, we can give analogous results for directed bipartite graphs.

Theorem 3.4.10. For each $n \in \mathbb{N}$, $n \geq 4$, pick $n_1, n_2 \in \mathbb{N}$ such that $2 \leq n_1 \leq n_2$ and $n_1 + n_2 = n$. Let $G^{(n),\mathbf{b},\rightarrow}$ denote the set of all bipartite graphs with the partite sets having n_1 and n_2 vertices respectively with edge probability $p(n)$, and let $P(G^{(n),\mathbf{b},\rightarrow}, p(n))$ be the probability of a graph from $G^{(n),\mathbf{b},\rightarrow}$ having diameter 3. We have

$$\begin{aligned}
&1 - n_2^2(1-p(n)^2)^{n_1}\left(1 + \frac{n_1^2(1-p(n)^2)^{n_2-n_1}}{n_2^2}\right)\\
&\leq P(G^{(n),\mathbf{b},\rightarrow}, p(n))\\
&\leq \left(\frac{1}{n_2(n_2-1)(1-p(n)^2)^{n_1}} + \frac{\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n_1}}{n_2}\left(4 + \frac{4}{(1-p(n))}\right)\right)\\
&\quad \cdot \left(1 + \frac{n_1(n_1-1)(1-p(n)^2)^{n_2-n_1}}{n_2(n_2-1)}\right)^{-1}.
\end{aligned}$$

Corollary 3.4.11. Let $P(G^{(n),\mathbf{b},\rightarrow}, p(n))$ be defined as in Theorem 3.4.10. If $p(n) = \frac{1}{2}$, then we have

$$\begin{aligned}
&1 - n_2^2(3/4)^{n_1}\left(1 + \frac{n_1^2(3/4)^{n_2-n_1}}{n_2^2}\right)\\
&\leq P(G^{(n),\mathbf{b},\rightarrow}, 1/2)\\
&\leq \left(\frac{(4/3)^{n_1}}{n_2(n_2-1)} + \frac{12(5/4)^{n_1}}{n_2}\right)\left(1 + \frac{n_1(n_1-1)(3/4)^{n_2-n_1}}{n_2(n_2-1)}\right)^{-1}.
\end{aligned}$$

Remark 3.4.12. The upper bound given for $P(G^{(n),\mathbf{b}}, 1/2)$ in Corollary 3.4.2 will in general only be non-trivial, i.e., less than 1, when n_2 much larger than n_1. For instance, if $n_1 < \frac{2\log n_2 - \log 4}{\log(4/3)}$ and $n_1 < \frac{\log n_2 - \log 24}{\log(5/4)}$, then the upper bound will be less than 1.

In the situation where the edge probability $p(n) \to 0$ as $n \to \infty$, we have the following.

Proposition 3.4.13. Let $P(G^{(n),\mathbf{b},\rightarrow}, p(n))$ be as in Theorem 3.4.10. Let $\lim_{n\to\infty} np(n)^4 = 0$. Then

$$\begin{aligned}
&1 - n_2^2 e^{-n_1 p(n)^2}\left(1 + e^{2\log n_1 - 2\log n_2 - (n_2-n_1)p(n)^2}\right)(1+o(1))\\
&\leq P(G^{(n),\mathbf{b},\rightarrow}, p(n))\\
&\leq (1+o(1))\left(\frac{1}{n_2^2}e^{n_1 p(n)^2}\right)\left(1 + e^{2\log n_1 - 2\log n_2 - (n_2-n_1)p(n)^2}\right)^{-1}\left(1 + 8n_2 e^{n_1 p(n)^2\left(p(n)^2-1\right)}\right).
\end{aligned}$$

Suppose further that

$$\lim_{n\to\infty} 2\log n_1 - 2\log n_2 - (n_2 - n_1)p(n)^2 = -\infty$$

and

$$\lim_{n\to\infty} (2\log n_2 - n_1 p(n)^2) = c$$

for some $c \in \mathbb{R}$.

1) If $c < 0$, we have

$$P(G^{(n),\mathbf{b},\to}, p(n)) \geq 1 - (1 + o(1))e^{c}.$$

2) If $c > 0$, we have

$$P(G^{(n),\mathbf{b},\to}, p(n)) \leq (1 + o(1))e^{-c}.$$

Remark 3.4.14. Assume that $p(n) \leq \frac{1}{4}$. The $o(1)$ in (3.4.6) can be made explicit as

$$\left(1 + \frac{2}{(n-2)}\right)(1 - p(n)^4 n_1 - p(n)^2)^{-1}(1 - p(n)^4(n_2 - n_1) - p(n)^2)^{-1} - 1.$$

Also, in Proposition 3.4.13 1), the o(1) can be made explicit as

$$\left(1 + e^{2\log n_1 - 2\log n_2 - (n_2 - n_1)p(n)^2}\right) e^{2\log n_2 - n_1 p(n)^2 - c} - 1$$

and under the additional assumption that

$$c - 1 \leq 2\log n_2 - n_1 p(n)^2 \leq c + 1$$

in Proposition 3.4.13 2), the $o(1)$ can be made explicit as

$$\left(1 + \frac{2}{(n-2)}\right)(1 - p(n)^4 n_1 - p(n)^2)^{-1}(1 - p(n)^4(n_2 - n_1) - p(n)^2)^{-1}$$
$$\cdot\, e^{c - 2\log n_2 + n_1 p(n)^3}\left(1 + e^{2\log n_1 - 2\log n_2 - (n_2 - n_1)p(n)^2}\right)^{-1}\left(1 + 32e^{c(1-p)+1+p-\frac{\log n}{2}}\right) - 1.$$

Substituting in $n_1 = n_2 = \frac{n}{2}$ or $n_1 = \frac{n-1}{2}$ and $n_2 = \frac{n+1}{2}$ can obtain similar asymptotics for directed Turán bipartite graphs.

Theorem 3.4.15. Let $G^{(n),\mathbf{b},t,\to}$ denote the set of all Turán directed bipartite graphs with edge probability $p(n)$, and let $P(G^{(n),\mathbf{b},t,\to}, p(n))$ be the probability of a graph from $G^{(n),\mathbf{b},\to,t}$ having diameter 3. If $n \geq 4$, then

$$\begin{aligned}
&1-\frac{(n+1)^2(1-p(n)^2)^{(n-1)/2}}{2}\\
&\leq P(G^{(n),\mathbf{b},t,\to}, p(n))\\
&\leq \left(\frac{4}{n(n-2)(1-p(n)^2)^{n/2}}+\frac{2\left(1+\frac{p(n)^3}{(1-p(n))}\right)^{n/2}}{n}\left(4+\frac{4}{(1-p(n))}\right)\right)\\
&\quad\cdot\left(1+\frac{(n-3)(1-p(n)^2)}{(n+1)}\right)^{-1}.
\end{aligned}$$

Corollary 3.4.16. Let $P(G^{(n),\mathbf{b},t,\to}, p(n))$ be defined as in Corollary 3.4.15. If $p(n) = \frac{1}{2}$, then we have

$$P(G^{(n),\mathbf{b},t,\to}, 1/2) \geq 1-\frac{(n+1)^2(3/4)^{(n-1)/2}}{2}.$$

In the situation where the edge probability $p(n) \to 0$ as $n \to \infty$, we have the following.

Proposition 3.4.17. Let $P(G^{(n),\mathbf{b},t,\to}, p(n))$ be defined as in Theorem 3.4.15. Let $\lim_{n\to\infty} np(n)^3 = 0$. We have

$$1-\frac{n^2e^{-\frac{np(n)^2}{2}}}{2}(1+o(1)) \leq P(G^{(n),\mathbf{b},t,\to}, p(n)) \leq (1+o(1))\left(\frac{2}{n^2}e^{\frac{np(n)^2}{2}}\right)\left(1+8ne^{\frac{np(n)^2}{2}(p(n)^2-1)}\right). \tag{3.4.8}$$

Suppose further that

$$\lim_{n\to\infty}(2\log n-\log 2-\frac{np(n)^2}{2}) = c$$

for some $c \in \mathbb{R}\backslash\{0\}$.

1) If $c < 0$, we have

$$P(G^{(n),\mathbf{b},t,\to}, p(n)) \geq 1-(1+o(1))e^{c}.$$

2) If $c > 0$, we have

$$P(G^{(n),\mathbf{b},t,\to}, p(n)) \leq (1+o(1))e^{-c}.$$

Remark 3.4.18. The $o(1)$ in the lower bound of (3.4.8) can be made explicit as

$$\left(1+\frac{1}{n}\right)^2 (1-p(n)^2)^{-1/2} - 1,$$

while the $o(1)$ in the upper bound of (3.4.8) can be made explicit as

$$\frac{n}{(n-2)}\left(1-\frac{np(n)^4}{2}-p(n)^2\right)^{-1}(1-p(n)^{-1}\left(1-\frac{p(n)^2}{2}-\frac{(2-2p(n)^2)}{(n+1)}\right)^{-1}e^{np(n)^4}-1.$$

Also, in Proposition 3.4.17 1), the $o(1)$ can be made explicit as

$$\left(1+\frac{1}{n}\right)^2 (1-p(n)^2)^{-1/2} e^{2\log n-\log 2-\frac{np(n)^2}{2}-c}$$

and under the additional assumption that $p(n) \leq \frac{1}{4}$ and

$$c-1 < 2\log n - \log 2 - \frac{np(n)^2}{2} < c+1$$

in Proposition 3.4.17 2), the $o(1)$ can be made explicit as

$$\frac{n}{(n-2)}\left(1-\frac{np(n)^4}{2}-p(n)^2\right)^{-1}(1-p(n)^{-1}\left(1-\frac{p(n)^2}{2}-\frac{(2-2p(n)^2)}{(n+1)}\right)^{-1}e^{np(n)^4}$$
$$\cdot\left(1+e^{1+p+(c+\log 2)(1-p(n))-\frac{\log n}{2}}\right)e^{c-2\log n+\log 2+\frac{np(n)^2}{2}}-1.$$

3.5 Generalisation to Closed Graph Sets

In this remaining section, we generalise our use of the simple sieve on random graphs and random k-partite graphs to more general collections of graphs called *closed graph sets*. Let $n \in \mathbb{N}$ and consider the set of graphs on n labeled vertices denoted by $G^{(n)}$ and the set of directed graphs on n labeled vertices denoted by $G^{(n),\to}$. First, a definition.

Definition 3.5.1. We say $S \subseteq G^{(n)}$ or $S \subseteq G^{(n),\to}$ is a *closed graph set* if the following hold

1. If $G \in S$, then G contains all n labeled vertices.

2. If $G, H \in S$, then the graphs having edge sets $E(G) \cup E(H)$, $E(G) \cap E(H)$ are in S. Thus, S contains a *maximal graph* L containing every possible edge that occurs in any graph in S, which we will call the maximal graph of S.

3. If $G \in S$, then the graph having edge set $E(L) \backslash E(G)$ is in S.

Remark 3.5.2. It is important to note here that if we're dealing with directed graphs, then between any pair of vertices, say v_1 and v_2, we distinguish the edge going from v_1 to v_2 and from the directed going from v_2 to v_1. Thus, there can be two edges between any pair of vertices in the directed case.

Lemma 3.5.3. *Let $S \subseteq G^{(n)}$ or $S \subseteq G^{(n),\to}$ be a closed graph set and L be the maximal graph of S. Then $E(L)$ can be partitioned into pairwise t disjoint subsets E_1, E_2,...,E_t, such that for each $G \in S$ we have*

$$E(G) = \bigcup_{1 \leq i \leq y} E_{k_i},$$

where $\{k_1, k_2, \ldots, k_y\} \subseteq \{1, 2, \ldots, t\}$.

Proof. Let $G_1, \ldots, G_t \in S$, where each G_i is minimal in the sense that the only proper subgraph of G_i that appears in S is possibly the graph with no edges. For each $1 \leq i \leq t$, let $E_i = E(G_i)$. The E_i are pairwise disjoint. Let $G \in S$. For each $1 \leq i \leq t$, we have $G \cap G_i \subseteq G_i$ and so $E(G \cap G_i) = E_i$ or $E(G \cap G_i) = \emptyset$. Thus, $E_i \subseteq E(G)$ or $E_i \cap E(G) = \emptyset$. Let

$$\bigcup_{1 \leq i \leq y} E_{k_i} \subseteq E(G),$$

with the union including all E_i such that $E_i \subseteq E(G)$. Suppose $E(G) \backslash \bigcup_{1 \leq i \leq y} E_{k_i} \neq \emptyset$ and let $H \in S$ such that $E(H) = E(G) \backslash \bigcup_{1 \leq i \leq y} E_{k_i}$. Then $E_j \subseteq E(H)$ for some $1 \leq j \leq t$ and so $E_j \subseteq E(G)$, a contradiction. Thus, the claim follows. $\square$

To help clarify the above definition, we provide a few examples of closed graph sets.

Example 3.5.4. $G^{(n)}$**: The set of all graphs on** n **vertices**

$G^{(n)}$, the set of all graphs on n vertices, is a closed graph set. It is easily seen that the maximal graph L in this case is the complete graph on n vertices and that properties 1), 2), and 3) in Definition 3.5.1 all hold. Also, applying Lemma 1, we have t being the total number of edges so that $t = \binom{n}{2} = \frac{n(n-1)}{2}$ and each E_i represents one edge.

Example 3.5.5. $G^{(n),\to}$**: The set of all directed graphs on** n **vertices**

$G^{(n),\to}$ the set of all directed graphs on n vertices is a closed graph set. It is easily seen that the maximal graph L in this case is the complete directed graph on n vertices and that properties 1), 2), and 3) in Definition 3.5.1 all hold. Also, applying Lemma 1, we have t being the total number of edges so that $t = 2\binom{n}{2} = n(n-1)$, and each E_i represents one directed edge.

Example 3.5.6. Sets of k**-partite graphs**

The k-partite graphs are graphs for which the set of vertices can be partitioned into k different sets with each set being independent. An independent set is a set of vertices where no two vertices in the set are connected by an edge. Thus bipartite graphs are a special case of k-partite graphs with $k = 2$. If we fix $n_1 \le n_2 \le \ldots \le n_k$ and have $n_1 + n_2 + \cdots + n_k = n$, then the set of all k-partite graphs with n_i vertices in the ith independent set forms a closed graph set. Here the maximal graph L would be the complete k-partite graph on these independent sets. Moreover, if G and H are two graphs in this set, then the graph having exactly the edges $E(G) \cup E(H)$ will still maintain the independent sets as independent. The same is true for the graph having exactly the edges $E(G) \cap E(H)$ and $E(L) \backslash E(G)$ Also, applying Lemma 1, we have t being the total number of edges and each E_i represents one edge.

Example 3.5.7. Sets of Circulant graphs

A *circulant graph* on n vertices is a graph where the set of vertices is labeled from 1 to n, with the ith vertex being connected by an edge to the jth vertex if and only if

$$(|i - j| \bmod n) \in D',$$

where D' is a certain subset of $D := \{1, 2, \ldots, \lfloor \frac{n}{2} \rfloor\}$. The set of all circulant graphs on n vertices is a closed graph set. Here the maximal graph L is the complete graph. Every circulant graph on n vertices is uniquely defined by a $D' \subset D$. We can easily see that if the graphs G and H are defined by $D_1, D_2 \subset D$, then the graph having exactly the edges $E(G) \cap E(H)$ is defined by $D_1 \cap D_2$, the graph having exactly the edges is defined by $E(G) \cup E(H)$ is defined by $D_1 \cup D_2$, and the graph having edge set $E(L) \backslash E(G)$ is defined by $D \backslash D_1$ and so 1), 2), and 3) in Definition 3.5.1 all hold. Furthermore, each E_i here represents the set of edges of a particular distance, or an element of D.

We use the simple sieve to give a lower bound on the probability of a randomly selected graph from a closed graph set having diameter 2 and give applications on circulant graphs.

Theorem 3.5.8. Let $n \in \mathbb{N}$ and let $S^{(n)} \subseteq G^{(n)}$ be a closed graph set and E_1, E_2,...,E_t be the disjoint edge sets and the graph L_n be the maximal graph of S_n. Let $r \leq \frac{t}{2}$. Suppose that between any two vertices (or from any one vertex to another vertex with directed graphs), there are r paths of length 2 in L such that there is at most one edge used from each E_i for the r paths. Also suppose that for each E_i the probability that the graph will have the edges from E_i is $p(n)$. Let $P(S^{(n)}, p(n))$ be the probability that the selected graph has diameter 2. Then

$$P(S_n, p(n)) \geq 1 - \frac{n^2}{2}\left(1 - p(n)^2\right)^r - p(n)^t$$

if $S_n \subseteq G^{(n)}$.

Proof. As in the proof of Theorem 1.2.4, we may assume that $p(n) \in \mathbb{Q} \cap (0,1)$ for all $n \in \mathbb{N}$. Let $p(n) = \frac{r}{s}$ where $r = r(n), s = s(n) \in \mathbb{N}$. As in the proof of Theorem 1.2.4, we let A be the set of all graphs in $S^{(n)}$, allowing for a number of duplicates of each possible graph to accommodate the edge probability $p(n)$. Since there are t groups of edges, we have r^t copies of the maximal graph and $|A| = s^t$.

We let B be all pairs of vertices, so $|B| = \binom{n}{2}$. We will say that a graph $a \in A$ matches up with a pair of vertices $b \in B$ if the pair of vertices do not share a common neighbouring vertex in a. Thus, $\omega(a) = 0$ (or a doesn't match up with any pair of vertices) if and only if a is connected with diameter at most 2.

There are at least r paths of length 2 between the two vertices or from the first vertex to the second vertex in the case of directed graphs. For a graph $G \in A$ to not have a path shorter than 3, at least one edge from each of the r paths must not be in G. Since only one edge is used from each group of edges for the r paths, each pair of edges multiplies the number of graphs by $(\frac{m_2}{m_1})^2$, but we do not want to have graphs that include both edges in the specific pair, so instead we multiply by $(\frac{m_2}{m_1})^2 - 1$ for each pair. Thus, we have

$$\deg b \leq \left(\left(\frac{m_2}{m_1}\right)^2 - 1\right)^r m_2^t \left(\frac{m_1}{m_2}\right)^{2r}.$$

By the simple sieve, we have the following as a lower bound if $S \subseteq G^{(n)}$:

$$|A| - \sum_{b \in B} \deg b \geq m_2^t - \binom{n}{2} \left(\left(\frac{m_2}{m_1} \right)^2 - 1 \right)^r m_2^t \left(\frac{m_1}{m_2} \right)^{2r}$$
$$= m_2^t \left(1 - \binom{n}{2} \left(1 - \left(\frac{m_1}{m_2} \right)^2 \right)^r \right).$$

Since the only graph on n vertices with diameter 1 is the complete graph, we have

$$P(S^{(n)}, p(n)) \geq 1 - \frac{n^2 \left(1 - p(n)^2\right)^r}{2} - p(n)^t.$$

□

In the situation where the edge probability $p(n) \to 0$ as $n \to \infty$, we will show the following result.

Proposition 3.5.9. Let $n \in \mathbb{N}$ and let $S^{(n)} \subseteq G^{(n)}$ be a closed graph set and E_1, E_2,. . .,E_t be the disjoint edge sets and the graph L_n be the maximal graph of S_n. Let $r \leq \frac{t}{2}$. Suppose that between any two vertices (or from any one vertex to another vertex with directed graphs), there are r paths of length 2 in L such that there is at most one edge used from each E_i for the r paths. Also suppose that for each E_i the probability that the graph will have the edges from E_i is $p(n)$. Let $P(S^{(n)}, p(n))$ be the probability that the selected graph has diameter 2. Suppose that $\lim_{n\to\infty} p(n) = 0$. Then as $n \to \infty$, we have

$$P(S^{(n)}, p(n)) \geq 1 - \frac{n^2 e^{-rp(n)^2}}{2}(1 + o(1)).$$

Moreover, if

$$\lim_{n\to\infty} \left(2 \log n - rp(n)^2 - \log 2\right) = c \tag{3.5.1}$$

where $-\infty \leq c < 0$, then we have the non-trivial lower bound

$$\liminf_{n\to\infty} P(n) \geq 1 - e^c,$$

replacing e^c with 0 if $c = -\infty$.

Proof. By Theorem 3.5.8 we have

$$P(S^{(n)}, p(n)) \geq 1 - \frac{n^2 \left(1 - p(n)^2\right)^{p(n)^{-2} \cdot rp(n)^2} \left(1 - p(n)^2\right)^{-2}}{2}.$$

Since $p(n)^{-2} \geq 1$, we have

$$\left(1-p(n)^2\right)^{p(n)^{-2}rp(n)^2} < e^{-rp(n)^2}.$$

Since $\lim_{n\to\infty} p(n) = 0$, we have that

$$\lim_{n\to\infty}\left(1-p(n)^2\right)^{-2} = 1,$$

from which we get

$$P(S^{(n)}, p(n)) \geq 1 - \frac{n^2}{2} e^{-rp(n)^2}(1+o(1)).$$

□

We consider an application to circulant graphs.

Corollary 3.5.10. Let $S^{(n)}$ denote the set of all circulant graphs on n vertices with edge probability $p(n)$, and let $P(S^{(n)}, p(n))$ be the probability of a graph from $S^{(n)}$ having diameter 2. For all $n \in \mathbb{N}$, we have

$$P(S^{(n)}, p(n)) \geq 1 - \frac{n^2}{2}\left(1-p(n)^2\right)^{\frac{n-10}{12}} - p(n)^{\frac{n-2}{2}}.$$

Proof. Here we have S being the set of all circulant graphs. It is routine to show that this set satisfies the properties in Lemma 3.5.3. For every possible distance $1 \leq d \leq \lfloor \frac{n-1}{2} \rfloor$ we have the edges connecting two vertices at a d distance to constitute a partition. Thus, we have $t = \lfloor \frac{n-1}{2} \rfloor \geq \frac{n-2}{2}$.

To determine a suitable value for r, note that all integers $1 \leq x \leq d-1$ can be put into groups of two such that x_1 and x_2 occur in the same group if and only if $x_1 + x_2 = d$. The only possible value excluded from the pairing is $\frac{d}{2}$, which occurs if d is even. Thus, for this value of d, we can have $r = \lfloor \frac{d-1}{2} \rfloor \geq \frac{d-2}{2}$. Also, by a simple counting argument, we can divide up all integers $1 \leq x \leq \lfloor \frac{n-1}{2} \rfloor$ in pairs so that for any x_1 and x_2 in a pair, we have $x_1 - x_2 = d$, excluding fewer than $2d$ numbers. Thus, we can also have

$$r = \frac{\lfloor \frac{n-1}{2} \rfloor - 2d}{2} \geq \frac{n-2}{4} - d.$$

For each d, we want to take the maximum of $\frac{d-2}{2}$ and $\frac{n-2}{4} - d$. We then take the minimum of the $\lfloor \frac{n-1}{2} \rfloor$ resulting values to obtain an appropriate value for r. It turns out this minimum value is bounded below by $r = \frac{n-10}{12}$. Thus, we may take this to be our value for r and by

Theorem 3.5.8, we have the probability of a circulant graph on n vertices having diameter 2 is bounded below by

$$1 - \frac{n^2}{2}\left(1 - p(n)^2\right)^{\frac{n-10}{12}} - p(n)^{\frac{n-2}{2}}.$$

$\square$

In the situation where the edge probability $p(n) \to 0$ as $n \to \infty$, we will show the following.

Corollary 3.5.11. Let $S^{(n)}$ denote the set of all circulant graphs on n vertices with edge probability $p(n)$, and let $P(S^{(n)}, p(n))$ be the probability of a graph from $S^{(n)}$ having diameter 2. Let $\lim_{n\to\infty} p(n) = 0$. Then as $n \to \infty$, we have

$$P(S^{(n)}, p(n)) \geq 1 - \frac{n^2}{2} e^{-\frac{np(n)^2}{12}}(1 + o(1)).$$

If

$$\lim_{n\to\infty}\left(2\log n - \frac{np(n)^2}{12} - \log 2\right) = d,$$

where $-\infty \leq d < 0$, then we have the non-trivial lower bound

$$\liminf_{n\to\infty} P(n) \geq 1 - e^d,$$

replacing e^d with 0 if $d = -\infty$.

Proof. As in Corollary 3.5.10, we can take $r = \frac{n-10}{12}$, which, upon substitution into Proposition 3.5.9 and simplifying, gives the desired result. $\square$

Definition 3.5.12. A *directed circulant graph* on n vertices is a graph where the set of vertices is labeled from 1 to n with a directed edge from the ith vertex to the jth vertex if and only if

$$(i - j \bmod n) \in D',$$

where D' is a certain subset of $D := \{1, 2, \ldots, \lfloor \frac{n}{2} \rfloor\}$.

We also obtain similar results about the probability of a random directed circulant graph having diameter 2.

Corollary 3.5.13. Let $S^{(n)}$ denote the set of all directed circulant graphs on n vertices with edge probability $p(n)$, and let $P(S^{(n)}, p(n))$ be the probability of a graph from $S^{(n)}$ having diameter 2. For all $n \in \mathbb{N}$, we have

$$P(S^{(n)}, p(n)) \geq 1 - n^2(1 - p(n)^2)^{\frac{n-4}{2}} - p(n)^{n-1}.$$

Corollary 3.5.14. Let $S^{(n)}$ denote the set of all circulant graphs on n vertices with edge probability $p(n)$, and let $P(S^{(n)}, p(n))$ be the probability of a graph from $S^{(n)}$ having diameter 2. Let $\lim_{n\to\infty} p(n) = 0$. Then as $n \to \infty$, we have

$$P(S^{(n)}, p(n)) \geq 1 - n^2 e^{-\frac{np(n)^2}{2}}(1 + o(1)).$$

If

$$\lim_{n\to\infty}\left(2\log n - \frac{np(n)^2}{2}\right) = d$$

where $-\infty \leq d < 0$, then we have the non-trivial lower bound

$$\liminf_{n\to\infty} P(n) \geq 1 - e^d,$$

replacing e^d with 0 if $d = -\infty$.

Chapter 4

Mahler Measure of "Almost" Reciprocal Polynomials

4.1 k-nonreciprocal Polynomials and Result

Recall the definition of a k-nonreciprocal polynomial.

Definition 1.3.4. Take a polynomial in $\mathbb{Z}[x]$, say $f(x) = \sum_{i=0}^{n} a_i x^i$. For an integer $k \geq 1$, we say that $f(x)$ is *k-nonreciprocal* if $a_n a_i = a_0 a_{n-i}$ for all $1 \leq i \leq k-1$ with $a_n a_k \neq a_0 a_{n-k}$.

Also, recall our result on a lower bound for the Mahler measure of a k-nonreciprocal polynomial and our remark concerning our result.

Theorem 1.3.5. Take a polynomial in $\mathbb{Z}[x]$, say $f(x) = \sum_{i=0}^{n} a_i x^i$. Suppose for some $k \in \mathbb{N}$, $2k \leq n$ we have $a_n a_i = a_0 a_{n-i}$ for all $1 \leq i \leq k-1$. Let $M(f)$ denote the Mahler measure of f and $\alpha = |a_k a_n - a_0 a_{n-k}|$. Then

$$M(f) \geq \frac{\alpha + \sqrt{\alpha^2 + 4(|a_0| + |a_n|)^2 |a_0 a_n|}}{2(|a_0| + |a_n|)}. \tag{4.1.1}$$

Remark 1.3.6. Borwein, Hare, and Mossinghoff noted that a consequence of Theorem 1.3.3 is that if f is a nonreciprocal polynomial with all odd coefficients, then

$$M(f) \geq \frac{1 + \sqrt{5}}{2} = 1.618\ldots$$

By Theorem 1.3.5, however, we may replace the condition that f has all odd coefficients with the condition that for the smallest k for which $a_k a_n \neq a_0 a_{n-k}$, we have $|a_k a_n - a_0 a_{n-k}| \geq 2$. Assuming that $|a_n| = |a_0| = 1$ (for otherwise $M(f) \geq \min\{|a_0|, |a_n|\} \geq 2$), this condition is substantially weaker than the condition that f is nonreciprocal and has all odd coefficients.

4.2 Proof and Example

In this section, we prove Theorem 1.3.5.

Note 1.3.6. We can see that if $f(x) \in \mathbb{Z}[x]$ is k-nonreciprocal for some $k \geq 1$, then $\pm(x-1)f(x)$ is also k-nonreciprocal. Therefore, it is enough to consider polynomials where both the leading coefficient and the constant term are both positive.

Our proof follows that of Borwein, Hare, and Mossinghoff in [3]. Unlike their result, however, we allow the innermost coefficients to not necessarily adhere to the reciprocal strucutre. We use the following result by Wiener, found in [19, pg. 392].

Lemma 1.3.6 (Wiener)**.** Suppose that $\phi(z) = \sum_{i\geq 0} \gamma_i z^i$, with $\gamma_i \in \mathbb{C}$ is analytic in an open disk containing $|z| \leq 1$ and satisfies $|\phi(z)| \leq 1$ on $|z| = 1$. Then $|\gamma_i| \leq 1 - |\gamma_0|^2$ for $i \geq 1$.

We now prove Theorem 1.3.5.

Proof of Theorem 1.3.5. Let $f(z) = \sum_{i=0}^{n} a_i z^i = a_n(z-\alpha_1)\cdots(z-\alpha_n)$ satisfy the hypothesis in the theorem, with a_0 and a_n both being positive. Write $f^*(z) = \sum_{i=0}^{n} d_i z^i$, so that $a_0 d_i = a_n a_i$ for all $1 \leq i \leq k-1$. Let the power series of $1/f^*(z)$ be $\sum_{i\geq 0} e_i z^i$. Then we have $e_0 = 1/a_n$. Let

$$G(z) = f(z)/f^*(z) = \sum_{i\geq 0} q_i z^i.$$

It does not matter if $q_i \in \mathbb{Z}$ for all $i \in \mathbb{N} \cup \{0\}$ or not. We have $q_0 = \frac{a_0}{a_n}$. From $f^*(z)G(z) = f(z)$, we obtain $\sum_{i=0}^{j} d_i q_{j-i} = a_j$. Thus, for $j \geq 1$, we have

$$a_n q_j = (a_j - q_0 d_j) - \sum_{i=1}^{j-1} d_i q_{j-i}.$$

From $a_0 d_i = a_n a_i$, we can see by induction that $q_i = 0$ for all $1 \leq i \leq k-1$ and $q_k = \frac{a_k}{a_n} - \frac{a_o a_{n-k}}{a_n^2} \neq 0$.

Let $\epsilon = -1$ if $f(z)$ has a zero of odd mulitiplicity at $z = 1$ and $\epsilon = 1$ otherwise. Since

$$\prod_{|\alpha_i|=1} \frac{z-\alpha_i}{1-\overline{\alpha_i}z} = \prod_{|\alpha_i|=1} \frac{-\alpha_i(1-z/\alpha_i)}{1-z/\alpha_i} = \prod_{|\alpha_i|=1} (-\alpha_i) = \epsilon,$$

we define

$$g(z) := \epsilon \prod_{|\alpha_i|<1} \frac{z-\alpha_i}{1-\overline{\alpha_i}z}$$

and

$$h(z) := \prod_{|\alpha_i|>1} \frac{1-\overline{\alpha_i}z}{z-\alpha_i}.$$

so that

$$\frac{g(z)}{h(z)} = \frac{\prod_{i=1}^n (z-\alpha_i)}{\prod_{i=1}^n (1-\overline{\alpha_i}z)} = \frac{\prod_{i=1}^n (z-\alpha_i)}{\prod_{i=1}^n (1-\alpha_i z)} = \frac{f(z)}{f^*(z)} = G(z).$$

Since all poles of both $g(z)$ and $h(z)$ lie outside the unit disk, both functions are analytic in a region including $|z| \leq 1$. Also, if $|z| = 1$ and $\beta \in \mathbb{C}$, then

$$\left(\frac{z-\beta}{1-\overline{\beta}z}\right)\overline{\left(\frac{z-\beta}{1-\overline{\beta}z}\right)} = \left(\frac{z-\beta}{1-\overline{\beta}z}\right)\left(\frac{1/z-\overline{\beta}}{1-\beta/z}\right) = 1$$

so $|g(z)| = |h(z)| = 1$ on $|z| = 1$. Let

$$g(z) = \sum_{i \geq 0} b_i z^i$$

and

$$h(z) = \sum_{i \geq 0} c_i z^i.$$

Since $g(z) = h(z)G(z)$, we have $b_i = c_i q_0$ for $0 \leq i < k$ and $b_k = c_0 q_k + c_k q_0$. Thus

$$\left| c_0 \left(\frac{a_k}{a_n} - \frac{a_0 a_{n-k}}{a_n^2} \right) \right| = |c_0 q_k| = |b_k - c_k q_0| \leq |b_k| + |c_k| q_0.$$

Notice that

$$c_0 = |h(0)| = \prod_{|\alpha_i|>1} 1/|\alpha_i| = |a_n|/M(f), \tag{4.2.1}$$

so that

$$\left|\frac{1}{M(f)}\left(a_k - \frac{a_0 a_{n-k}}{a_n}\right)\right| = |c_0 q_k| \leq |b_k| + |c_k| q_0. \tag{4.2.2}$$

We now consider two cases. By Lemma 1.3.6, we have $|c_k| \leq 1 - c_o^2$ and $|b_k| \leq 1 - b_0^2$. Notice that $b_0 = c_0 q_0$. Combining (4.2.1) and (4.2.2), we have

$$\begin{aligned}
\frac{\left|a_k - \frac{a_0 a_{n-k}}{a_n}\right|}{M(f)} &\leq (1 - b_0^2) + (1 - c_0^2) q_0 \\
&= (1 - c_0^2 q_0^2) + (1 - c_0^2) q_0 \\
&= (1 + q_0)(1 - q_0 c_0^2) \\
&= (q_0 + 1)\left(1 - \frac{q_0 a_n^2}{M(f)^2}\right) \\
&= (q_0 + 1)\left(1 - \frac{a_0 a_n}{M(f)^2}\right).
\end{aligned}$$

Thus, we have

$$M(f)\left|a_k - \frac{a_0 a_{n-k}}{a_n}\right| \leq (q_0 + 1)(M(f)^2 - a_0 a_n).$$

This gives

$$M(f) \geq \frac{\left|a_k - \frac{a_0 a_{n-k}}{a_n}\right| + \sqrt{\left|a_k - \frac{a_0 a_{n-k}}{a_n}\right|^2 + 4(q_0+1)^2 a_0 a_n}}{2(q_0 + 1)}.$$

The result follows. $\square$

Note 1.3.6. If $|a_n a_k - a_0 a_{n-k}| > |a_0^2 - a_n^2|$, then the bound in (4.1.1) is non-trivial since then it is greater than

$$\begin{aligned}
\frac{|a_0^2 - a_n^2| + \sqrt{|a_0^2 - a_n^2|^2 + 4(a_0 + a_n)^2 |a_0 a_n|}}{2(a_0 + a_n)} &= \frac{|a_0 - a_n| + \sqrt{(a_0 - a_n)^2 + 4a_0 a_n}}{2} \\
&= \frac{|a_0 - a_n| + \sqrt{(a_0 + a_n)^2}}{2} \\
&= \frac{|a_0 - a_n| + a_0 + a_n}{2} \\
&= \max\{a_n, a_0\},
\end{aligned}$$

which is the trivial bound.

Note 1.3.6. If $f(x)$ is a reciprocal polynomial, the bound in (4.1.1) is trivial, for we would have $a_n = a_0$ and $a_k = a_{n-k}$ so that

$$|a_n a_k - a_0 a_{n-k}| = 0.$$

Thus, by Theorem 1.3.5, we have

$$M(f) \geq \frac{\sqrt{4(a_0 + a_n)^2 a_n a_0}}{2(a_0 + a_n)} = a_n,$$

which is trivial.

We now show some examples, indicating that our bound in Theorem 1.3.5 is sharp.

Example 1.3.6. Let $k, n \in \mathbb{N}$ where $n > 2k$ and $n \neq 3k$ and $a, b, c \in \mathbb{Z}$ such that $a > 0 > c$, and $a - |b| \leq -c \leq a + |b|$. Consider the polynomial $f(x) = (ax^{2k} + bx^k + c)(x^{n-2k} - 1)$, which satisfies $a_n a_i = a_0 a_{n-i}$ for all $1 \leq i \leq k-1$ and $a_n a_k \neq a_0 a_{n-k}$ where $f(x) = \sum_{i=0}^{n} a_i x^i$. Let $\alpha = |a_k a_n - a_0 a_{n-k}|$ We have

$$M(f) = \frac{\alpha + \sqrt{\alpha^2 + 4(a_0 + a_n)^2 a_0 a_n}}{2(a_0 + a_n)}.$$

Let $k, n \in \mathbb{N}$ where $n \geq 2k$ and $a, b, c \in \mathbb{Z}$ satisfying the given conditions. We have

$$f(x) = ax^n + bx^{n-k} + cx^{n-2k} - ax^{2k} - bx^k - c$$

if $n > 4k$,

$$f(x) = ax^{4k} + bx^{3k} + (c - a)x^{2k} - bx^k - c,$$

if $n = 4k$,

$$f(x) = ax^n + bx^{n-k} - ax^{2k} + cx^{n-2k} - bx^k - c$$

if $4k > n > 3k$, and

$$f(x) = ax^n - ax^{2k} + bx^{n-k} - bx^k + cx^{n-2k} - c$$

if $3k > n > 2k$. In all cases, we can easily see that if we write $f(x) = \sum_{i=0}^{n} a_i x^n$, then we have $a_n = a$, $a_0 = -c$, $a_k = -b$, $a_{n-k} = b$, $a_n a_i = a_0 a_{n-i}$ for all $1 \leq i \leq k-1$ and $a_n a_k \neq a_0 a_{n-k}$. We therefore have that $\alpha = |a_n a_k - a_0 a_{n-k}| = |b(a - c)|$.

Since all the roots of $x^{n-2k}-1$ have absolute value 1, we have $M(f) = M(ax^{2k} + bx^k + c)$. By the quadratic formula, the roots of $ax^{2k} + bx^k + c$ are the kth roots of the numbers

$$\frac{-b \pm \sqrt{b^2-4ac}}{2a}.$$

Since $c < 0 < a$, the absolute values of these numbers are

$$\frac{\pm|b| + \sqrt{b^2-4ac}}{2a}$$

First, consider

$$\frac{|b| + \sqrt{b^2-4ac}}{2a} \tag{4.2.3}$$

If $|b| > a$, then clearly this number is greater than 1 so we may assume that $|b| \leq a$. By our assumption that $-c \geq a - |b|$, we then have

$$\begin{aligned}\frac{|b| + \sqrt{b^2-4ac}}{2a} &\geq \frac{|b| + \sqrt{b^2+4a(a-|b|)}}{2a} \\ &= \frac{|b| + \sqrt{4a^2-4a|b|+b^2}}{2|a|} \\ &= \frac{|b| + 2a - |b|}{2a} \\ &= 1.\end{aligned}$$

Now consider

$$\frac{\sqrt{b^2-4ac} - |b|}{2a} \tag{4.2.4}$$

By our assumption that $-c \leq a + |b|$, we have

$$\begin{aligned}\frac{\sqrt{b^2-4ac} - |b|}{2a} &\leq \frac{\sqrt{b^2+4a(a+|b|)} - |b|}{2a} \\ &= \frac{\sqrt{4a^2+4a|b|+b^2} - |b|}{2a} \\ &= \frac{2a + |b| - |b|}{2a} \\ &= 1.\end{aligned}$$

All of the nth roots of (4.2.3) have absolute value at least 1, while all of the nth roots of (4.2.4) have absolute value at most 1. Hence,

$$M(f) = \frac{|b| + \sqrt{b^2 - 4ac}}{2}.$$

Note that

$$\frac{|b| + \sqrt{b^2 - 4ac}}{2} = \frac{|ba - cb| + \sqrt{(ba - cb)^2 - 4(a-c)^2 ca}}{2(a-c)}$$

since a and c have opposite signs. Thus, we obtain our bound.

Note 1.3.6. If we impose the restriction $a_0, a_n = \pm 1$ on this example, then we will have α being even. It is unknown whether the inequality in Theorem 1.3.5 is still sharp if we impose $a_0, a_n = \pm 1$ and α being odd.

Chapter 5

Conclusions and Future Work

We will now give an overview of the thesis and its contributions, as well as discuss possible directions for further research.

In Chapter 2, we examined the first area of study, which was the Random Fibonacci tree. We first examined random paths down the tree and calculated the probability of a random path not hitting any $(1, 1)$ pair except at the root. We came up with a quadratic expression for this probability in terms of p, the probability to selecting a right branch at any intermediate point. We then examined coprime (a, b) pairs in general in the tree. Rittaud had already studied various useful properties of the tree, determining that for any coprime pair (a, b), such a coprime pair occurs as a parent-child relation in the tree infinitely many times. We improved upon Rittaud's result, deriving tight asymptotics for the number of times a specific parent-child pair (a, b) occurred at a specific depth in the tree. These asymptotics were derived from combinatorial identities and recurrence relations that the tree was observed to satisfy.

On counting the number of (a, b) pairs in the Fibonacci tree, there are still many questions that have been left unanswered. Some of these are as follows. Can we get even tighter bounds for $A(n)$? Theorem 1.1.8 was essentially derived from Robbins' bounds for factorials. Since Robbins, however, there have been numerous improvements on bounds for factorials that will probably help us derive even better bounds for $A(n)$. For example, Knopp [10] shows that there exists constants $a_1, a_2, a_3, a_4, \ldots$ such that the sequence

$$r_n := \ln\left(\frac{n!e^n}{\sqrt{2\pi}n^{n+1/2}}\right)$$

is bounded above and below by the partial sums of

$$\frac{a_1}{n} + \frac{a_2}{n^2} + \frac{a_3}{n^3} + \frac{a_4}{n^4} + \cdots$$

where $a_n = 0$ for all even n, $a_n > 0$ if $n \equiv 1 \pmod 4$, and $a_n < 0$ if $n \equiv 3 \pmod 4$. Impens [9] shows how to compute those constants recursively. These constants are also directly related to the Bernoulli numbers. More specifically, for all $k \in \mathbb{N}$, we have

$$a_k = \frac{B_{k+1}}{k(k+1)}$$

where B_{k+1} is the $(k+1)$st Bernoulli number [14]. We may be able to use these results to prove that there exist constants $a_0, a_1, a_2, a_3, \ldots$ such that

$$\frac{A(n) \cdot 4\sqrt{3\pi} n^{3/2}}{243 \cdot (27/4)^n}$$

can be approximated by

$$a_0 + \frac{a_1}{n} + \frac{a_2}{n^2} + \frac{a_3}{n^3} + \cdots$$

We showed that $a_0 = 1$, $a_1 = \frac{-1387}{72}$ and that, if a_2 exists, then $0 \leq a_2 \leq \frac{5548}{9}$. We may be able to use the same procedure as in this Chapter 2 to derive more terms of this series. Analogous questions remain open for $A_{(a,b)}$ for all coprime ordered pairs (a, b). As another direction, what is the probability of a walk in the Fibonacci tree containing exactly k occurrences of $(1, 1)$ where $k \in \mathbb{N}$?

We can also look at variations of the Fibonacci Tree. For example, in taking a left branch from the ordered pair (x, y) do a subtraction $x - y$ instead of taking the mere difference $|x - y|$. More generally, for some $k \in \mathbb{N}$, take k children, all of them being $x + \delta y$, where δ is a different kth root of unity for each one.

In Chapter 3, we examined the second area of study, which was applying the Turán and simple sieves in the area of random graph theory. We made explicit in our work here that the Turán and simple sieves complement each other. That is, for the problems studied, one of them will give a nontrivial or meaningful result if and only if the other gives a trivial or meaningless result. The specific problem we used these sieves on was calculating the probability of a random graph having diameter 2 (or diameter 3 in the case of bipartite graphs) from a closed graph set. We then gave examples such as the set of all graphs with n vertices, sets of k-partite graphs, sets of circulant graphs, as well as analogous sets of directed graphs. Given an edge probability p, we then calculated the respective probabilities of a random graph having diameter 2 (or diameter 3 for bipartite graphs) and gave asymptotic formulas for the probabilities assuming that the edge probability approached 0 as the number of vertices got arbitrarily large. As we saw, it was usually the case that in a given situation exactly one of the sieves would give a non-trivial result, i.e., a bound on the probability that was strictly between 0 and 1.

There are many questions that remain unanswered. For instance, we looked at but a few examples of closed graph sets, namely sets of all graphs, k-partite graphs, and circulant graphs. Are there other important sets of graphs out there that would fit the definition of a closed graph set? If so, what happens when we apply the sieves to these graphs? Also, in the case of circulant graphs, we managed to apply the simple sieve to them, but unfortunately were not able to apply the Turán sieve to them. Are there any counting techniques out there that would help us apply the Turán sieve to circulant graphs? We can also ask if there is any way we can apply the sieves on any or all closed graph sets to give nontrivial upper and lower bounds on a graph have diameter d for some $d \geq 3$. The case of diameter 2 was obtained by noting that a graph is of diameter 2 if and only if it is not the complete graph and every two neighbouring vertices share a common neighbouring vertex. Such a characterisation gets more and more complicated the higher and higher the diameter under consideration is, but perhaps a better diameter can be obtained. Alternatively, we could express the diameter under consideration as a function on the number of vertices. Depending on the function, the probability of a graph having diameter at least or at most that number might be doable to calculate.

Besides diameter, can we use the Turán sieve and/or the simple sieve on other features of a graph, for example, the probability of a graph being k-colourable for some $k \geq 2$? We can also examine if we can use the sieves on similar or analogous problems on hypergraphs.

In Chapter 4, we examined the Mahler measure of a polynomial with integer coefficients. We generalised the notion of reciprocal polynomials to the more general concept of k-nonreciprocal polynomials, which are polynomials where the outermost coefficients mirror each other, but letting this pattern break down for the innermost coefficients. We studied these k-nonreciprocal polynomials, deriving lower bounds for their Mahler measures, using the same proof technique that Borwein, Hare, and Mossinghoff [3] used on their own generalised version of reciprocal polynomials. This proof technique involved ideas from complex analysis, such as complex Taylor expansions and the behaviour of polynomials inside the unit disk. We also managed to show that the achieved bound is tight for certain k-nonreciprocal polynomials for every $k \in \mathbb{N}$.

A few further questions are worth pursuing. For example, can Smyth's bound [20] be further improved for certain other classes of polynomials that are not reciprocal? For instance, there are many different ways of defining a polynomial to be “almost reciprocal”. Can we generalise Smyth’s proof in other ways to obtain even better bounds for polynomials that satisfy two or more definitions of “almost ” reciprocal? Another question is: can we use these ideas on sparse polynomials, which are polynomials of high degree but where almost all the coefficients are 0?

References

[1] A. Bazylewicz. On the product of the conjugates outside the unit circle of an algebraic integer. *Acta Arithmetica*, 30(1):43–61, 1976/77.

[2] P. Blanksby and H. Montgomery. Algebraic integers near the unit circle. *Acta Arithmetica*, 18(1):355–369, 1971.

[3] Peter Borwein, Kevin G. Hare, and Michael J. Mossinghoff. The Mahler measure of polynomials with odd coefficients. *Bulletin of the London Mathematical Society*, 36(3):332–338, 2004.

[4] Edward Dobrowolski. On a question of Lehmer and the number of irreducible factors of a polynomial. *Acta Arithmetica*, 34(4):391–401, 1979.

[5] Harry Furstenberg and Harry Kesten. Products of random matrices. *The Annals of Mathematical Statistics*, 31(2):457–469, 1960.

[6] Edgar N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.

[7] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science. 2nd Edition.* Addison-Wesley Longman Publishing Co., Inc, 1994.

[8] K.G. Hare and J.C. Saunders. On (a, b) pairs in random Fibonacci sequences. *Journal of Number Theory*, 90:352–366, 2018.

[9] Chris Impens. Stirling's series made easy. *The American Mathematical Monthly*, 110(8):730–735, 2003.

[10] Konrad Knopp. *Theory and Application of Infinite Series*. Courier Corporation, 2013.

[11] Derrick H. Lehmer. Factorization of certain cyclotomic functions. *Annals of Mathematics*, 34(2):461–479, 1933.

[12] Yu-Ru Liu and M. Ram Murty. Sieve methods in combinatorics. *Journal of Combinatorial Theory, Series A*, 111(1):1–23, 2005.

[13] Eran Makover and Jeffrey McGowan. An elementary proof that random Fibonacci sequences grow exponentially. *Journal of Number Theory*, 121(1):40–44, 2006.

[14] Victor Namias. A simple derivation of Stirling's asymptotic series. *The American Mathematical Monthly*, 93(1):25–29, 1986.

[15] Benoıt Rittaud. On the average growth of random Fibonacci sequences. *Journal of Integer Sequences*, 10(2):32, 2007.

[16] Herbert Robbins. A remark on Stirling's formula. *The American Mathematical Monthly*, 62(1):26–29, 1955.

[17] J.C. Saunders. Mahler Measure of "almost" reciprocal polynomials. *Bulletin of the Australian Mathematical Society*, 98(1):70–76, 2018.

[18] Andrzej Schinzel. On the product of the conjugates outside the unit circle of an algebraic number. *Acta Arithmetica*, 24(4):385–399, 1973.

[19] Andrzej Schinzel. *Polynomials with special regard to reducibility, volume 77 of Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2000.

[20] Chris J. Smyth. On the product of the conjugates outside the unit circle of an algebraic integer. *Bulletin of the London Mathematical Society*, 3(2):169–175, 1971.

[21] Pál Turán. On a theorem of Hardy and Ramanujan. *Journal of London Mathematical Society*, 9(4):274–276, 1934.

[22] Divakar Viswanath. Random Fibonacci sequences and the number 1.13198824. *Mathematics of Computation*, 69(231):1131–1155, 2000.