Accessibility Percolation

Master's Project
MA40249

Theo Crookes

University of Bath
Project Supervisor: Dr. Matthew Roberts

Contents

Chapter 1

Introduction.

Throughout this paper, we explain the Theorems from Roberts and Zhao [14] and Berestycki, Brunet and Shi [2]. We follow the structure of their proofs, so our arguments will be similar to theirs with some extra explanation.

1.1 Increasing Paths on Regular Trees.

Definition 1.1 (Regular Tree)**.** Fix $\alpha > 0$. A regular n-ary tree $T = (V, E)$ of height h where $n = \lfloor \alpha h \rfloor$, is a connected acyclic graph, upon which a root vertex has n vertices connected to it, called its children, of which it is called their parent. Then, each of those vertices has n children, this continues h times, until the final children have no more children and are called leaves. To each of the edges in the graph, attach a $U[0,1]$ uniform random variable.

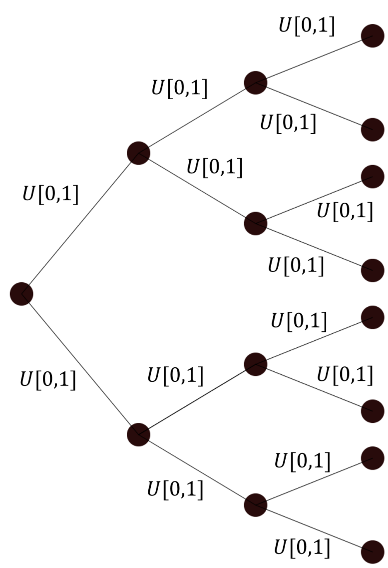

Figure 1.1: The regular 2-ary tree formed when $\alpha = 2/3$ and $h = 3$, with the $U[0,1]$ random variables attached.

Definition 1.2 (Simple Path)**.** We define a *simple path* along the regular tree to be an ordered sequence of distinct vertices $v_0, v_1, ..., v_h$ where v_i is a vertex of the tree for all $i \in \{1, ..., n\}$ such that v_0 is the root of the tree, v_h is a leaf, the edge $\{v_i, v_{i+1}\}$ is an edge in the tree.

Definition 1.3 (Increasing Path)**.** An *increasing path* on the regular tree is a simple path on which the labels attached to the edges produced by the random variables are increasing along the path. An increasing path is also called *accessible* or *open*.

We are concerned with the probability that one of these paths exists. This is called *accessibility percolation* by Nowak and Krug [12]; we will look at this not only on the regular tree but also on the hypercube.

For fixed $\alpha > 0$, we show that in the limit as $h \longrightarrow \infty$, the probability that there exists an increasing path on the corresponding regular n-ary tree depends on the value of α and undergoes a phase transition at $\alpha = 1/e$, as highlighted by the following Theorem, which is the first major Theorem of the project.

Theorem 1.4. Suppose that $n = \lfloor \alpha h \rfloor$. As $h \longrightarrow \infty$,

$$\mathbb{P}[\text{There exists an increasing path}] = \begin{cases} 0 & \text{if } \alpha \leqslant 1/e, \\ 1 & \text{if } \alpha > 1/e. \end{cases} \tag{1.1}$$

We will prove both cases in the Theorem. Nowak and Krug [12] proved the case when $\alpha \leqslant 1/e$, and to prove the case when $\alpha > 1/e$, Roberts and Zhao [14] proved the following slightly stronger result on the number of increasing paths. For any $\alpha > 1/e$, there exists $\delta > 0$ and $\eta > 0$ such that

$$\mathbb{P}[\text{There exist at least } \exp(\delta h) \text{ increasing paths}] \geqslant 1 - \exp(-\eta h). \tag{1.2}$$

1.2 Motivation - Evolution is like a Path on a Tree.

The setup of increasing paths on regular trees is used to model evolution. We have a population of organisms that reproduce asexually, and we are concerned with the fitness of a specific genotype as it mutates. Firstly we should understand how the evolutionary process works; from Andrea, Rice and Townsend [8], we know that evolution occurs through a 'house of cards' model, meaning that from Hathaway [6], mutations 'with large effects effectively reshuffle the genomic deck' and this is a better model than lots of mutations each resulting in a small change. It means that the fitness after each mutation can be defined in an independently and identically distributed way. So we can initially assign a minimal fitness to the gene and then, after each mutation, use a $U[0, 1]$ random variable to assign a new fitness to the gene, similar to the edges on the regular tree.

From Franke [4], we learn that a 'mutational path is considered selectively accessible (or accessible for short) if the fitness values encountered along it are monotonically increasing; thus, along such a path, the population never encounters a decline in fitness.' These selectively accessible paths are similar to the increasing paths on the tree; in both cases, we are not interested in cases where the fitness/path decreases.

The particular type of model we will relate our increasing paths on the tree to is referred to as the strong selection weak mutation (SSWM) model from Gillespie [5]. Under this model, by Kimura [11], it means that disadvantageous mutations are unable to replace the entire population, that is, we are only concerned with increasing/accessible paths like in the tree. Roberts and Zhao [14] further justify the use of this model.

1.3 Further Biological Motivation.

We explore a similar model that also models evolution but is slightly more complicated. The model comes from Berestycki, Brunet and Shi [2]. Consider an organism that has a genome with L_0 sites, each containing one of two possible alleles. Then, the state of the genome at any time can be encoded as a series of L_0 ones and zeros, where a 0 is present at position i if the genome contains the wild starting allele at position i and a 1 at position i if the genome contains the mutated new allele at position i. Therefore before any mutations have occurred the state of the genome is $(0, 0, ..., 0)$, a vector of L_0 zeros. As the organism evolves, the state of the genome moves through a path on the L-dimensional hypercube, and at any point can be written as a point $\sigma \in \{0, 1\}^{L_0}$.

We assume that no gene that has mutated will ever return to the original; that is to say, no 1 will ever return to a 0. The genome evolves and moves along the edges of the L_0-dimensional hypercube from vertex to vertex, and to each vertex σ, we attach a fitness x_σ. It is important to note here we attach the fitness values to each specific node in such a way that two paths leading to the same node σ will result in the same fitness at that node x_σ. This was trivial in the case of our tree before because there is only one possible path up to each vertex and therefore it did not matter whether we attached the fitness values to the nodes or the edges. The fitness label attached to each vertex can be given by a $U[0, 1]$ uniform random variable, as in the case of the tree and is justified by the same reasoning through the House of Cards model [8]. The L_0-dimensional hypercube has $2^{L_0-1}L_0$ edges and since this is finite, there must be a vertex with the greatest fitness, denoted σ_L, where L is the number of 1s in σ_L. We only take an interest in the shortest paths to this point, so we are interested in the sub hypercube of dimension L. Figure 1.2 represents the 4-dimensional hypercube with the $U[0, 1]$ random variables attached to each vertex.

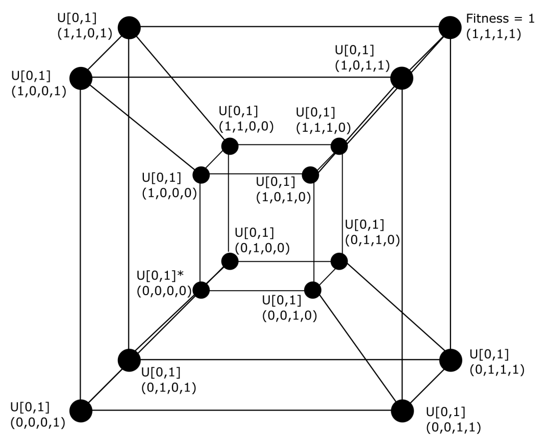

Figure 1.2: The 4-dimensional hypercube with the $U[0, 1]$ random variables attached and the specific value $x_{(1,1,1,1)} = 1$. The * denotes that this value may be fixed, depending on the context.

We assign $x_{\sigma_L} = 1$ and wish to know how many paths from $\sigma_0 = (0, 0, ..., 0)$ to the fittest type $\sigma_L = (1, 1, ..., 1)$ are there on which the attached labels are increasing. We call this number Θ. Again, we call such paths *accessible* or *open*, motivated by [4] calling mutational paths like these *selectively accessible*. We will prove some results both in the case where the starting value x_{σ_0} is randomly picked and where it is not.

1.4 The Irregular Tree.

Working with the L-dimensional hypercube can be complicated as two given paths may not be independent after a certain number of edges. For example, two paths may share the same first k edges, then have different p edges in the middle and then share the same final $L-k-p$ edges. This was not possible on the tree because after two paths split, they could never intersect one another again and share a common edge after the splitting. For this reason, we create a new tree that is similar to the hypercube that can make computations easier.

Definition 1.5. We define the *irregular* $L-tree$, $T_L = (V_L, E_L)$, to be a connected acyclic graph upon which the root vertex has L vertices (children) connected to it. Then, each of those L children have $L-1$ children. This continues until the final nodes have no more children and are called leaves. To the root vertex attach a fixed number x, to the leaves attach the number 1, and to each of the other vertices, we attach a $U[0,1]$ random variable.

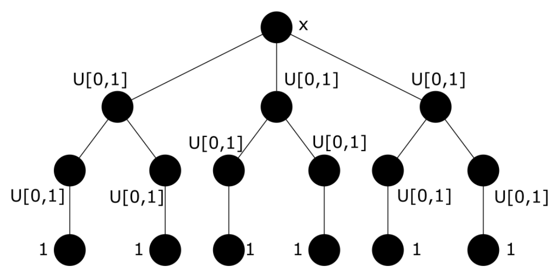

Figure 1.3: The irregular tree for $L = 3$.

Attaching the random variables to the nodes is justified with similar biological motivation before, i.e., motivated by the 'House of Cards' model [8]. This tree is constructed to be very similar to the hypercube. The final vertex of each path has fixed fitness 1 and both the L-dimensional hypercube and this tree have $L!$ possible paths through them. Again, we are particularly interested in directed paths from the root to the leaf on which all of the labels are increasing. Again, motivated by [4] we call these paths accessible, open, or increasing.

1.5 Main Results on the Hypercube and Irregular Tree.

In Chapter 3 we prove the next Theorem from [2] regarding Θ, the number of open paths on the tree.

Theorem 1.6. Fix $X \geqslant 0$; then when the root vertex of the irregular tree is assigned the value X/L, we have that as $L \to \infty$, the random variable Θ/L converges in distribution to a standard exponential random variable multiplied by e^{-X}.

In Chapter 4 we will follow and explain the steps of the proof of the following Theorem on the number of accessible paths Θ on the hypercube from [2]. However, we will ultimately show that the proof of this Theorem contains a mistake that does not appear to be easily rectified.

Theorem 1.7. Fix $X \geqslant 0$; then when the starting node of the hypercube is assigned the value X/L, we have that as $L \to \infty$, the random variable Θ/L converges in distribution to the product of two independent standard exponential random variables multiplied by e^{-X}.

Chapter 2

Increasing Paths on Regular Trees.

2.1 Notation - The Regular Tree.

To prove Theorem 1.4 we follow the proof given in [14] and liberate their notation too. Firstly, we frequently use a double bound on Stirling's approximation, given by the following Lemma.

Lemma 2.1. For $n \geqslant 1$,

$$2 < \frac{n!}{\sqrt{n}(\frac{n}{e})^n} < 3. \tag{2.1}$$

Proof. From Robbins [13] we have for $n \geqslant 1$ that

$$n! = n^n \sqrt{2\pi n} e^{-n} e^{r_n}, \quad \text{where } r_n \text{ satisfies} \quad \frac{1}{12n+1} < r_n < \frac{1}{12n}. \tag{2.2}$$

When solving the equation in (2.2) for e^{r_n} and substituting it into the equality by noticing that $e^{\frac{1}{12n+1}}$ is bounded below by 1 and $e^{\frac{1}{12n}}$ is bounded above by $e^{1/12}$ we obtain the result. □

We define P to be the set of simple paths on the regular tree. Since each vertex has n children and the tree has height h, $|P| = n^h$. For $u \in P$, write $u_0, u_1, ..., u_h$ with u_i an edge of the tree for all $0 \leqslant i \leqslant h$. We write $X(u) = (X(u_1), X(u_2), ..., X(u_h))$ where $X(u_i)$ is the $U[0,1]$ label attached to the edge from u_{i-1} to u_i, for all $1 \leqslant i \leqslant h$. We write $n = \alpha h$ throughout the paper and do not use the floor notation.

Definition 2.2. Let $u, v \in P$. We define $a(u,v)$ to be the number of shared edges between the two paths u and v before they diverge. That is,

$$a(u,v) \coloneqq \max\{k \mid u_k = v_k\}.$$

Note that since to paths u and v they share the same first $a(u,v)$ edges, they share the same labels, so $X(u_k) = X(v_k)$ for $k \leqslant a(u,v)$.

Definition 2.3. Define I to be the set of h-dimensional vectors with increasing components all with a value in $[0,1]$, that is

$$I \coloneqq \{(x_1, x_2, ..., x_h) \in [0,1]^h \mid x_1 < x_2 < ... < x_h\}.$$

Definition 2.4. Fix $\epsilon \in [0,1)$. Define firstly C_ϵ to be the set of vectors in $[0,1]^h$ where each component is greater than ϵ. That is,

$$C_\epsilon \coloneqq \{(x_1, x_2, ..., x_h) \in [0,1]^h \mid x_j \geqslant \epsilon, \ 1 \leqslant j \leqslant h\}.$$

Define D_ϵ to be the set of vectors in $[0,1]^h$ where each component x_j is greater than $\epsilon + (1-\epsilon)\frac{j-1}{h}$ for all $1 \leqslant j \leqslant h$. That is

$$D_\epsilon \coloneqq \left\{(x_1, x_2, ..., x_h) \in [0,1]^h \mid x_j \geqslant \epsilon + (1-\epsilon)\Big(\frac{j-1}{h}\Big), \ 1 \leqslant j \leqslant h\right\}$$

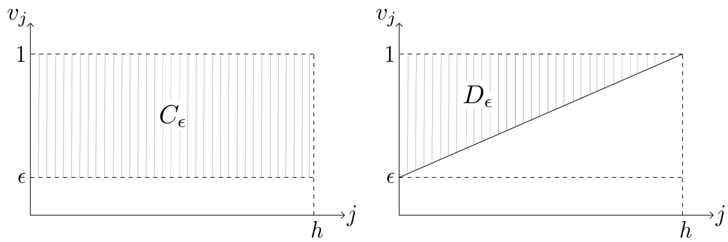

Figure 2.1: The regions with which the points in C_ϵ and D_ϵ can take.

These two regions are visualised in Figure 2.1. Note that for a path to be in C_ϵ and D_ϵ we do not require the edge labels to be increasing, but we are interested in increasing paths so we will often look at the intersection of I with C_ϵ or D_ϵ.

Definition 2.5. Define N to the number of increasing paths on the regular tree, so

$$N \coloneqq \sum_{u \in P} \mathbb{1}_{\{X(u) \in I\}}.$$

Also, define N_ϵ to be the number of increasing paths that also satisfy the condition to be in D_ϵ, so

$$N_\epsilon \coloneqq \sum_{u \in P} \mathbb{1}_{\{X(u) \in I \cap D_\epsilon\}}.$$

2.2 Outline of the Proof of Theorem 1.4.

Firstly, we will prove the first case of the Theorem, when $\alpha \leqslant 1/e$. We will do this by using the first moment method on N. That is, we will prove that $\mathbb{E}[N] \to 0$ as $h \to \infty$. Then we can use Markov's Inequality from Probability Theory [1] with $\alpha = 1$ to see that

$$\mathbb{P}[N \geqslant 1] \leqslant \mathbb{E}[N].$$

From which we see that as $h \to \infty$, $\mathbb{P}[N \geqslant 1] \to 0$. Proving the first part of the Theorem.

To the second case of the Theorem, when $\alpha > 1/e$, we want to use the second moment method to give a lower bound on the probability that there exists an increasing path. However, when we use the second moment method on N, the bound produced is not useful. Instead, we use the second moment method on N_ϵ. The use of N_ϵ is justified in full later, but the key idea is that the second moment method gives a good lower bound for $\mathbb{P}[N \geqslant 1]$.

We will show that

$$\frac{(\alpha(1-\epsilon)e)^h}{3h^{3/2}} \leqslant \mathbb{E}[N_\epsilon] \leqslant \mathbb{E}[N].$$

We will then proceed to show that when $\alpha(1-\epsilon) > 1$ and h is large,

$$\mathbb{E}[N_\epsilon^2] \leqslant \mathbb{E}[N_\epsilon] + \mathbb{E}[N_\epsilon]^2 + c(\alpha(1-\epsilon)e)^{2h}$$

for some constant $c > 0$.

Then, by using the Paley-Zygmund Inequality we obtain a lower bound for $\mathbb{P}[N_\epsilon > \exp(\delta h)]$, this lower bound is 1 minus a constant times h^{-3} when h is sufficiently large. From here, we then use a technical argument which revolves around consider the first 4 levels of the tree separately from the final $h-4$. Through this argument, we will obtain the result.

2.3 Proof of Theorem 1.4 - Case $\alpha \leqslant 1/e$.

Recall that $|P| = n^h$. Then we have that

$$\mathbb{E}[N] = \mathbb{E}[\sum_{p \in P} \mathbb{1}_{\{X(p) \in I\}}] = n^h \mathbb{P}[X(p) \in I] = \frac{n^h}{h!}.$$

In the final equality we used the fact that the probability a path is increasing is $1/h!$, since there are $h!$ orderings edges in each path and only one ordering is increasing.

Now, we apply Stirling's approximation to see that

$$\mathbb{E}[N] = \frac{n^h}{h!} \sim \frac{(\alpha h)^h e^h}{\sqrt{h} h^h} = \frac{(\alpha e)^h}{\sqrt{h}}.$$

If $\alpha \leqslant 1/e$, then $\alpha e \leqslant 1$ and since N is always non-negative, we can bound the expected number of increasing paths above and below in such a way that using the Squeeze Theorem yields $\mathbb{E}[N] \to 0$ as $h \to \infty$. Then using Markov's Inequality [1] we have that $\mathbb{P}[N \geqslant 1] \to 0$, proving the Theorem 1.4 in the case $\alpha \leqslant 1/e$. □

2.4 Justification of the use of N_ϵ.

To understand why we use N_ϵ, we must first understand why using the second moment method for N does not yield a useful lower bound. To use the second moment method and obtain a lower bound on $\mathbb{P}[N \geqslant 1]$ we use a need to calculate $\mathbb{E}[N^2]$,

$$\begin{aligned}
\mathbb{E}[N^2] &= \mathbb{E}[(\sum_{p \in P} \mathbb{1}_{\{X(p) \in I\}})^2] \\
&= \mathbb{E}[\sum_{p \in P} \sum_{q \in P} \mathbb{1}_{\{X(p) \in I\}} \mathbb{1}_{\{X(q) \in I\}}] \\
&= \sum_{p,q \in P} \mathbb{E}[\mathbb{1}_{\{X(p) \in I\}} \mathbb{1}_{\{X(q) \in I\}}] \\
&= \sum_{p,q \in P} \mathbb{E}[\mathbb{1}_{\{X(p) \in I,\ X(q) \in I\}}] \\
&= \sum_{p,q \in P} \mathbb{P}[X(p) \in I,\ X(q) \in I] \qquad (\dagger),
\end{aligned}$$

where in the second equality we use that the square of the number of open paths is the number of pairs of open paths. Now, let $p = p_0, p_1, ..., p_h$ and $q = q_0, q_1, ..., q_h$ be a pair of paths in P satisfying $k = a(p, q)$, this means they share the first k edges, and then diverge after this. We would want most

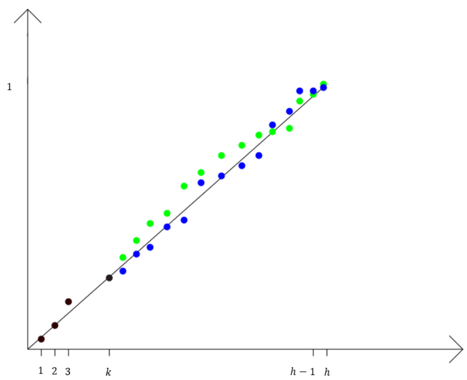

Figure 2.2: An illustration of two increasing paths, increasing linearly like we would expect.

of the pairs of paths to follow the setup in Figure , because most increasing paths increase linearly. However, this is not the case for pairs of paths. The probability that a pair of paths is increasing depends heavily on the value of the label attached to the edge before diverging and $a(p, q)$. Therefore, for the aforementioned p and q, let $X(p_k) = \zeta$ for some $\zeta > 0$. Then since the first k labels must be less than ζ and the final $h - k$ vertices must be greater than ζ on both p and q we see that viewed as a function of ζ,

$$\mathbb{P}[p, q \in I] \propto \zeta^k (1 - \zeta)^{2(h-k)}.$$

Differentiating this function and setting it equal to 0 we find that the turning points of this function are $\zeta = 0$, $\zeta = 1$, $\zeta = \frac{k}{2h-k}$. The second derivative at $\zeta = \frac{k}{2h-k}$ is negative so this is a maximum. Notice that $\zeta = \frac{k}{2h-k}$ is very small when h is large, which we will consider in this Theorem and is, in fact, about twice as small as we would expect the label attached to the k'th edge of a single increasing path to be.

As a result of the above, the pairs of paths that contribute the most in the double sum in (†) are those that keep the labels small while the paths coincide and then have a lot of 'space' to increase for the rest of the path. Figure illustrates two such paths increasing in this way. However, it is unlikely that a single increasing path follows a structure similar to this and the labels are more likely to increase more linearly as in Figure . Using the second moment method we have that $\mathbb{P}[N \geqslant 1] \geqslant \mathbb{E}[N]^2/\mathbb{E}[N^2]$.

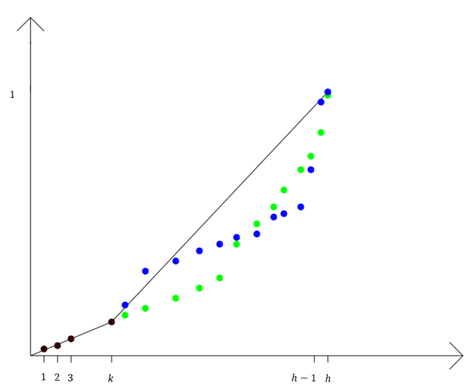

Figure 2.3: An illustration of two increasing paths that are unlikely to occur individually but are likely to be increasing as a pair, pairs of this type contribute the most to $\mathbb{E}[N^2]$.

However, due to the large number of pairs of paths that follow the pattern in Figure 2.3, $\mathbb{E}[N^2]$ is large, and so the RHS of this inequality is close to 0 and is therefore not a useful lower bound.

However, notice that for the reasons described above $\mathbb{E}[N^2]$ counts many pairs of increasing paths that are unlikely to occur individually, but are likely to occur as a pair, by removing these paths we can form a better lower bound. So instead, fix $\epsilon > 0$, and now we are only interested in paths that are increasing *and* the j-th edge on the path to be greater than or equal to $\epsilon + (1-\epsilon)(\frac{j-1}{h})$, $\forall j$, that is to say, we are only interested in paths in N_ϵ.

This means that we only consider paths that are above the red line on Figure 2.4, with two example paths shown. Notice that this solves the problem we had before, because now the labels on the shared

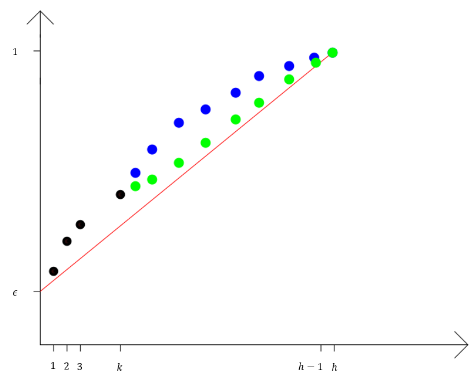

Figure 2.4: An illustration of the types of paths we will look at.

edges of the pair of paths are not allowed to stay small. Then, using a similar method as before and noting that $I \cap D_\epsilon \subseteq I$, we have:

$$\mathbb{P}[N \geqslant 1] \geqslant \mathbb{P}[N_\epsilon \geqslant 1] \geqslant \frac{\mathbb{E}[N_\epsilon]^2}{\mathbb{E}[N_\epsilon^2]}.$$

This lower bound will not be as close to 0 as the previous, and now we have a good lower bound for the probability that there exists an increasing path.

2.5 First Moment Bound on N_ϵ.

Before we prove a lower bound on $\mathbb{E}[N_\epsilon]$ we require the following three Lemmas.

Lemma 2.6. Let $U_1, U_2, ..., U_j$ be i.i.d $U[0,1]$ random variables. Then,

$$\mathbb{P}\Big(U_1 \leqslant ... \leqslant U_j,\ U_1 \geqslant \frac{1}{j+1}, ..., U_j \geqslant \frac{j}{j+1}\Big) = \frac{1}{(j+1)!}$$

Figure 2.5 is a *visualisation* of the event we are interested in calculating the probability of in the Lemma. The dots with the line connecting them are a possible instance in which the event is satisfied, highlighting the fact that the random variables have to be increasing. *Proof of Lemma* 2.6. Firstly, let p denote the probability we are interested in, so

$$p = \mathbb{P}\Big(U_1 \leqslant ... \leqslant U_j,\ U_1 \geqslant \frac{1}{j+1}, ..., U_j \geqslant \frac{j}{j+1}\Big).$$

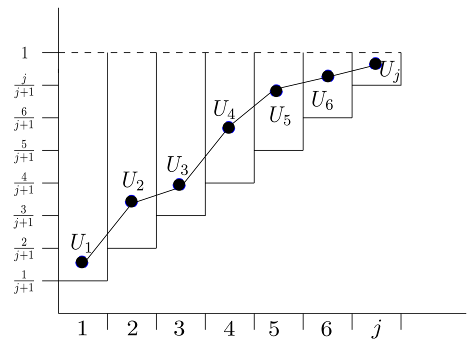

Figure 2.5: An illustration of the types of path we are looking at in Lemma 2.6.

Now define, for $2 \leqslant i \leqslant j$,

$$I_i := \int_{\frac{j}{j+1}}^{1} \int_{\frac{j-1}{j+1}}^{v_j} \dots \int_{\frac{i}{j+1}}^{v_{i+1}} \Big(\frac{v_i^{i-1}}{(i-1)!} - \frac{v_i^{i-2}}{(j+1)(i-2)!} \Big) \, dv_i \dots dv_{j-1} dv_j. \tag{2.3}$$

Notice that

$$p = \int_{\frac{j}{j+1}}^{1} \int_{\frac{j-1}{j+1}}^{v_j} \dots \int_{\frac{1}{j+1}}^{v_2} 1 \, dv_1 \dots dv_{j-1} dv_j$$

because the innermost integral is calculating the probability that U_1 is greater than $1/(j+1)$ and less than some value v_2, then when integrating this over v_2 between $2/(j+1)$ and v_3 this is calculating the probability that $1/(j+1) \leqslant U_1 \leqslant U_2$ and that U_2 is greater than $2/(j+1)$ and less than some value v_3, continuing this argument gives the desired probability. Computing the innermost integral, we see that

$$p = \int_{\frac{j}{j+1}}^{1} \int_{\frac{j-1}{j+1}}^{v_j} \dots \int_{\frac{2}{j+1}}^{v_3} \Big(v_2 - \frac{1}{(j+1)} \Big) \, dv_2 \dots dv_{j-1} dv_j = I_2.$$

Now by computing the innermost integral of (I - 6) for $2 \leqslant i \leqslant j-1$, we can show that $I_i = I_{i+1}$. That is,

$$\begin{aligned}
I_i &= \int_{\frac{j}{j+1}}^{1} \int_{\frac{j-1}{j+1}}^{v_j} \dots \int_{\frac{i}{j+1}}^{v_{i+1}} \Big(\frac{v_i^{i-1}}{(i-1)!} - \frac{v_i^{i-2}}{(j+1)(i-2)!} \Big) \, dv_i \dots dv_{j-1} dv_j \\
&= \int_{\frac{j}{j+1}}^{1} \int_{\frac{j-1}{j+1}}^{v_j} \dots \int_{\frac{i+1}{j+1}}^{v_{i+2}} \Big[\frac{v_i^{i}}{i!} - \frac{v_i^{i-1}}{(j+1)(i-1)!} \Big]_{v_i = \frac{i}{j+1}}^{v_{i+1}} \, dv_{i+1} \dots dv_{j-1} dv_j \\
&= \int_{\frac{j}{j+1}}^{1} \int_{\frac{j-1}{j+1}}^{v_j} \dots \int_{\frac{i}{j+1}}^{v_{i+1}} \Big(\frac{v_{i+1}^{i}}{i!} - \frac{v_{i+1}^{i-1}}{(j+1)(i-1)!} \Big) - \frac{i^i}{(j+1)^i i!} + \frac{i^{i-1}}{(j+1)^i (i-1)!} \, dv_{i+1} \dots dv_{j-1} dv_j \\
&= \int_{\frac{j}{j+1}}^{1} \int_{\frac{j-1}{j+1}}^{v_j} \dots \int_{\frac{i}{j+1}}^{v_{i+1}} \Big(\frac{v_{i+1}^{i}}{i!} - \frac{v_{i+1}^{i-1}}{(j+1)(i-1)!} \Big) \, dv_{i+1} \dots dv_{j-1} dv_j = I_{i+1}.
\end{aligned}$$

So we see that

$$\begin{aligned}
p &= I_2 = I_3 = \cdots = I_j \\
&= \int_{\frac{j}{j+1}}^{1} \Big(\frac{v_j^{j-1}}{(j-1)!} - \frac{v_j^{j-2}}{(j+1)(j-2)!} \Big) \, dv_j = \Big[\frac{v_j^{j}}{j!} - \frac{v_j^{j-1}}{(j+1)(j-1)!} \Big]_{v_j = \frac{j}{j+1}}^{1} \\
&= \frac{1}{j!} - \frac{1}{(j+1)(j-1)!} - \frac{j^j}{j!(j+1)^j} + \frac{j^{j-1}}{(j-1)!(j+1)^j} = \frac{1}{j!} - \frac{1}{(j+1)(j-1)!} = \frac{1}{(j+1)!}.
\end{aligned}$$

□

Lemma 2.7. Let $\epsilon \in [0,1]$. A uniform $U[0,1]$ random variable conditioned to be at least ϵ is a uniform $U[\epsilon,1]$ random variable.

Proof. Fix $\epsilon \in [0,1]$. Let $U \sim U[0,1]$. Note first that in particular for $x \in [\epsilon,1]$ we have the probability that U is less than x given that U is greater than ϵ is $\mathbb{P}[\epsilon < U < x]/\mathbb{P}[U > \epsilon] = (x-\epsilon)(1-\epsilon)$. So, comparing CDFs gives the result. □

Lemma 2.8. Let $\epsilon \in [0,1]$, $j,h \in \mathbb{N}$ with $j < h$. Let $U \sim U[0,1]$ and $U^\epsilon \sim U[\epsilon,1]$, then

$$\mathbb{P}\Big[U^\epsilon \geqslant \epsilon + (1-\epsilon)\Big(\frac{j-1}{h}\Big)\Big] = \mathbb{P}\Big[U \geqslant \frac{j-1}{h}\Big].$$

Proof. Using the CDF of a uniform random variable, we have that

$$\mathbb{P}\Big[U^\epsilon \geqslant \epsilon + (1-\epsilon)\Big(\frac{j-1}{h}\Big)\Big] = 1 - \frac{\Big(\epsilon + (1-\epsilon)\Big(\frac{j-1}{h}\Big)\Big) - \epsilon}{1-\epsilon} = 1 - \frac{j-1}{h} = \mathbb{P}\Big[U \geqslant \frac{j-1}{h}\Big].$$

□

Proposition 2.9.

$$\mathbb{E}[N_\epsilon] \geqslant \frac{(\alpha(1-\epsilon)e)^h}{3h^{3/2}}.$$

Proof. Let $U = (U_1, U_2, ..., U_j)$ where $U_i \sim U[0,1]$ for all $i \in J := \{1,...,j\}$. Then

$$\mathbb{E}[N_\epsilon] = \mathbb{E}\Big[\sum_{p\in P} \mathbb{1}_{\{X(p)\in I\cap D_\epsilon\}}\Big] = \sum_{p\in P} \mathbb{E}[\mathbb{1}_{\{X(p)\in I\cap D_\epsilon\}}] = n^h \mathbb{P}[U \in I \cap D_\epsilon].$$

Now, using Baye's Theorem and noting that $\mathbb{P}[U \in C_\epsilon \mid U \in I \cap D_\epsilon] = 1$ since $\{U \in I \cap D_\epsilon\} \subset \{U \in C_\epsilon\}$, we have that

$$\mathbb{E}[N_\epsilon] = n^h \mathbb{P}[U \in I \cap D_\epsilon \mid U \in C_\epsilon]\mathbb{P}[U \in C_\epsilon].$$

Now, let $U^\epsilon = (U_1^\epsilon, U_2^\epsilon, ..., U_j^\epsilon)$ where $U_i^\epsilon \sim U[\epsilon,1]$ for all $i \in J := \{1,...,j\}$. Then,

$$\begin{aligned}
\mathbb{P}[U \in I \cap D_\epsilon \mid U \in C_\epsilon] &= \mathbb{P}[U_1 \leqslant ... \leqslant U_j,\ U_j \geqslant \epsilon + (1-\epsilon)\Big(\frac{j-1}{h}\Big)\ \forall j \in J \mid U_j \geqslant \epsilon\ \forall j \in J] \\
&= \mathbb{P}[U_1^\epsilon \leqslant ... \leqslant U_j^\epsilon,\ U_j^\epsilon \geqslant \epsilon + (1-\epsilon)\Big(\frac{j-1}{h}\Big)\ \forall j \in J] \\
&= \mathbb{P}[U_1 \leqslant ... \leqslant U_j,\ U_j \geqslant \frac{j-1}{h}\ \forall j \in J] = \mathbb{P}[U \in I \cap D_0],
\end{aligned}$$

where the second inequality we use Lemma 2.7 and in the third inequality we use Lemma 2.8. Using the above, recalling that $n = \alpha h$ and noticing that $\mathbb{P}[U \in C_\epsilon] = (1-\epsilon)^h$ because all h of the $U[0,1]$ random variables must be bigger than ϵ, we obtain that

$$\mathbb{E}[N_\epsilon] = (\alpha h(1-\epsilon))^h \mathbb{P}[U \in I \cap D_0].$$

Now,

$$\begin{aligned}\mathbb{P}[U \in I \cap D_0] &= \mathbb{P}[U_1 \leqslant 1/h]\mathbb{P}[U_1 \leqslant ... \leqslant U_j,\ U_j \geqslant \frac{j-1}{h}\ \ \forall j \in J \mid U_1 \leqslant 1/h] \\ &+ \mathbb{P}[U_1 \geqslant 1/h]\mathbb{P}[U_1 \leqslant ... \leqslant U_j,\ U_j \geqslant \frac{j-1}{h}\ \ \forall j \in J \mid U_1 \geqslant 1/h] \\ &\geqslant \mathbb{P}[U_1 \leqslant 1/h]\mathbb{P}[U_2 \leqslant ... \leqslant U_j,\ U_j \geqslant \frac{j-1}{h}\ \ \forall\, 2 \leqslant j \leqslant h] = \frac{1}{h}\frac{1}{h!} = \frac{1}{h(h!)},\end{aligned}$$

where in the second to last equality we used Lemma 2.6. Finally, using Lemma 2.1, the double bound on Stirling's formula, we have

$$\mathbb{E}[N_\epsilon] = (\alpha h(1-\epsilon))^h \mathbb{P}[U \in I \cap D_0] \geqslant \frac{(\alpha h(1-\epsilon)^h}{h(h!)} \geqslant \frac{(\alpha h(1-\epsilon)e)^h}{3h^{h+3/2}} \geqslant \frac{(\alpha(1-\epsilon)e)^h}{3h^{3/2}}.$$

□

2.6 Second Moment Bound on N_ϵ.

Proposition 2.10. If $\alpha(1-\epsilon)e > 1$, then there exists some constant $c > 0$ such that

$$\mathbb{E}[N_\epsilon^2] \leqslant \mathbb{E}[N_\epsilon] + \mathbb{E}[N_\epsilon]^2 + c(\alpha(1-\epsilon)e)^{2h}.$$

Proof. Firstly, using the fact that the square of the number of increasing paths in D_ϵ is the number of pairs of increasing paths in D_ϵ, we have that

$$N_\epsilon^2 = \Big[\sum_{u \in P} \mathbb{1}_{\{X(u)\in I\}}\Big]^2 = \sum_{u\in P}\sum_{v \in P} \mathbb{1}_{\{X(u)\in I\cap D_\epsilon\}}\mathbb{1}_{\{X(v)\in I\cap D_\epsilon\}} = \sum_{u,v\in P} \mathbb{1}_{\{X(u),X(v)\in I\cap D_\epsilon\}}.$$

Note that one of the defining properties of a pair of increasing paths is the number of edges shared before they diverge. In fact, the sets $A_0 := \{(u,v) \in P \times P \mid a(u,v) = 0\}$, $A_1 := \{(u,v) \in P \times P \mid a(u,v) = 1\}$, ...,$A_h := \{(u,v) \in P \times P \mid a(u,v) = h\}$ form a partition of $A := \{(u,v) \in P \times P\}$. So, define $N_\epsilon^2(k)$ to be the number of increasing paths in D_ϵ such that they share the first k edges, that is

$$N_\epsilon^2(k) := \sum_{(u,v)\in A_k} \mathbb{1}_{\{X(u),X(v)\in I\cap D_\epsilon\}} = \sum_{\substack{u,v\in P,\\ a(u,v)=k}} \mathbb{1}_{\{X(u),X(v)\in I\cap D_\epsilon\}}.$$

Then, since A is a disjoint union of $A_1, A_2, ..., A_h$ we have

$$N_\epsilon^2 = \sum_{(u,v)\in A} \mathbb{1}_{\{X(u),X(v)\in I\cap D_\epsilon\}} = \sum_{k=0}^{h} N_\epsilon^2(k).$$

Note that $N_\epsilon^2(h) = N_\epsilon$, since we are simply counting all the pairs of paths that are identical.

Now, Roberts and Zhao [14] claim that $\mathbb{E}[N_\epsilon^2(0)] = \mathbb{E}[N_\epsilon]^2$. This is a slight inaccuracy, but it does not affect the outcome of the proof because all we require is that $\mathbb{E}[N_\epsilon^2(0)] \leqslant \mathbb{E}[N_\epsilon]^2$, which we now show to be true. Taking the expectation of $N_\epsilon^2(0)$ and noting that if $a(u,v)=0$, then the paths are disjoint,

$$\begin{aligned}
\mathbb{E}[N_\epsilon^2(0)] &= \mathbb{E}[\sum_{\substack{u,v\in P,\\ a(u,v)=0}} \mathbb{1}_{\{X(u),X(v)\in I\cap D_\epsilon\}}] \\
&= \mathbb{E}[\sum_{u\in P} \mathbb{1}_{\{X(u)\in I\cap D_\epsilon\}} \sum_{\substack{v\in P,\\ a(u,v)=0}} \mathbb{1}_{\{X(v)\in I\cap D_\epsilon\}}] \\
&= \sum_{u\in P} \sum_{\substack{v\in P,\\ a(u,v)=0}} \mathbb{P}[X(u),X(v)\in I\cap D_\epsilon] \\
&= \sum_{u\in P} \sum_{\substack{u,v\in P,\\ a(u,v)=0}} \mathbb{P}[X(u)\in I\cap D_\epsilon]\mathbb{P}[X(v)\in I\cap D_\epsilon] \\
&\leqslant \sum_{u\in P}\sum_{v\in P} \mathbb{P}[X(u)\in I\cap D_\epsilon]\mathbb{P}[X(v)\in I\cap D_\epsilon] = \mathbb{E}[N_\epsilon]^2.
\end{aligned}$$

Following the proof in [14], let $U=(U_1,...,U_h)$ and $V=(V_1,...,V_h)$ each be a sequence of i.i.d $U[0,1]$ random variables such that $U_j=V_j$ for all $j\leqslant k$ and U_j and V_j are independent for $j>k$, so U and V are like two paths which share the same edges up to level k.

There are n^k possible paths up to the point where the two paths U and V deviate. Then there are $n(n-1)$ choices for their next step that must be different and then there are $(n^{h-k-1})^2=n^{2h-2k-2}$ possible paths from there to the leaf. So we have for $k=2,...,h-1$,

$$\begin{aligned}
\mathbb{E}[N_\epsilon^2(k)] &= n^k n(n-1)n^{2h-2k-2}\mathbb{P}[U,V\in I\cap D_\epsilon] \\
&= \left(\frac{n-1}{n}\right) n^{2h-k}\mathbb{P}[U,V\in I\cap D_\epsilon|U,V\in C_\epsilon]\mathbb{P}[U,V\in C_\epsilon] \\
&= \left(\frac{n-1}{n}\right) (\alpha h)^{2h-k}(1-\epsilon)^{2h-k}\mathbb{P}[U,V\in I\cap D_0], \qquad (\dagger\dagger)
\end{aligned}$$

where in the final equality we used Lemma 2.8. We now investigate $\mathbb{P}[U,V\in I\cap D_0]$. This is equivalent to the following integral,

$$\begin{aligned}
\mathbb{P}[U,V\in I\cap D_0] &= \int_{\frac{k-1}{h}}^1 \mathbb{P}[U,V\in I\cap D_0, U_k\in dx] \\
&= \int_{\frac{k-1}{h}}^1 \mathbb{P}[U,V\in I\cap D_0|U_k=x]\mathbb{P}[U_k\in dx] \\
&= \int_{\frac{k-1}{h}}^1 \mathbb{P}[U,V\in I\cap D_0|U_k=x]dx.
\end{aligned}$$

Now, within the integral we need the first $U_1,...,U_{k-1}$, that both U and V share, to be in order and less than x. We also require that $U_{k+1},...,U_h$ and $V_{k+1},...,V_h$ are in order and greater than x. We also relax the condition that $U, V \in D_0$, so we integrate over more possible paths and obtain the inequality,

$$\begin{aligned}
\mathbb{P}[U,V \in I \cap D_0] &= \int_{\frac{k-1}{h}}^{1} \mathbb{P}[U,V \in I \cap D_0 | U_k = x] dx \\
&\leqslant \int_{\frac{k-1}{h}}^{1} \mathbb{P}[U_1 < U_2 < ... < U_{k-1} < x]\mathbb{P}[x < U_{k+1} < U_{k+2} < ... < U_h] \\
&\qquad\qquad\qquad\qquad \cdot \mathbb{P}[x < V_{k+1} < V_{k+2} < ... < V_h] dx \\
&= \int_{\frac{k-1}{h}}^{1} \mathbb{P}[U_1 < U_2 < ... < U_{k-1} < x]\mathbb{P}[x < U_{k+1} < U_{k+2} < ... < U_h]^2 dx.
\end{aligned}$$

Since $x \in [0,1]$, probability that the first $k-1$ $U[0,1]$ random variables are less than x is x^{k-1}. Only 1 of the $(k-1)!$ possible orderings of these random variables is in ascending order so,

$$\mathbb{P}(U_1 < U_2 < ... < U_{k-1} < x) = \frac{x^{k-1}}{(k-1)!}. \tag{2.4}$$

Similarly, the probability that the last $h-k$ $U[0,1]$ random variables are greater than x is $(1-x)^{h-k}$. There are $(h-k)!$ possible orderings of these random variables and only 1 is in ascending order, so $\mathbb{P}(x < U_{k+1} < U_{k+2} < ... < U_h)^2 = x^{2h-2k}/(h-k)!$. We can substitute this probability and (2.4) into the integral to obtain,

$$\begin{aligned}
&\int_{\frac{k-1}{h}}^{1} \mathbb{P}(U_1 < U_2 < ... < U_{k-1} < x)\mathbb{P}(x < U_{k+1} < U_{k+2} < ... < U_h)^2 dx \\
&\qquad\qquad\qquad\qquad = \int_{\frac{k-1}{h}}^{1} \frac{x^{k-1}}{(k-1)!} \frac{(1-x)^{2h-2k}}{(h-k)!} dx.
\end{aligned}$$

Now we define the function $f(x) \coloneqq x^{k-1}(1-x)^{2h-2k}$ and investigate its properties. Firstly, see that

$$f'(x) = (k-1)x^{k-2}(1-x)^{2h-2k} - (2h-2k)x^{k+1}(1-x)^{2h-2k-1}.$$

Setting this equal to 0 and solving for x yields the turning points $x = 0,\ x = 1,\ x = (k-1)/(2h-k+1)$. Verifying that the second derivative is negative at $x = (k-1)/(2h-k+1)$ we see that this is indeed a maximum. In particular, for $x \in [(k-1)/(2h-k+1), 1]$, $f(x)$ is a decreasing function. So since $(k-1)/(2h-k+1) < (k-1)/h$ we have that,

$$\begin{aligned}
\int_{\frac{k-1}{h}}^{1} \frac{x^{k-1}}{(k-1)!} \frac{(1-x)^{2h-2k}}{(h-k)!} dx &\leqslant \left(1 - \frac{k-1}{h}\right) \frac{((k-1)/h)^{k-1}}{(k-1)!} \frac{((h-k+1)/h)^{2h-2k}}{(h-k)!^2} \\
&\leqslant \frac{((k-1)/h)^{k-1}}{(k-1)!} \frac{((h-k+1)/h)^{2h-2k}}{(h-k)!^2}.
\end{aligned}$$

Now, combining this with (††) we see that

$$\begin{aligned}
\mathbb{E}[N_\epsilon^2(k)] &= \binom{n-1}{n} (\alpha h(1-\epsilon))^{2h-k} \frac{((k-1)/h)^{k-1}}{(k-1)!} \frac{((h-k+1)/h)^{2h-2k}}{(h-k)!^2} \\
&\leqslant (\alpha h(1-\epsilon))^{2h-k} \frac{((k-1)/h)^{k-1}}{(k-1)!} \frac{((h-k+1)/h)^{2h-2k}}{(h-k)!^2}.
\end{aligned}$$

By bounding the two fractions in the inequality above using Lemma 2.1 Stirling's double bound. We have that for $k = 2, ..., h-1$,

$$\begin{aligned}\mathbb{E}[N_\epsilon^2(k)] &\leqslant (\alpha h(1-\epsilon))^{2h-k}\frac{((k-1)/h)^{k-1}}{(k-1)!}\frac{((h-k+1)/h)^{2h-2k}}{(h-k)!^2} \\ &\leqslant (\alpha(1-\epsilon))^{2h-k}h^{2h-k}\frac{e^{k-1}h^{1-k}}{2(k-1)^{1/2}}\frac{e^{2h-2k+2}h^{-2h+2k}}{4(h-k+1)} = \frac{e}{8}\frac{(\alpha(1-\epsilon)e)^{2h-k}h}{(k-1)^{1/2}(h-k+1)}.\end{aligned}$$

By using Stirling's Approximation again, for the case $k = 1$ we have

$$\mathbb{E}[N_\epsilon^2(1)] \leqslant (\alpha h(1-\epsilon))^{2h-1}\frac{1}{(h-1)!^2} = \alpha^{2h-1}h^{2h-1}\frac{(1-\epsilon)^{2h-1}}{(h-1)!^2} \leqslant \frac{e}{4}(\alpha(1-\epsilon)e)^{2h-1}.$$

Combining all of the above we see that

$$\begin{aligned}\mathbb{E}[N_\epsilon^2] &\leqslant \mathbb{E}[N_\epsilon] + \mathbb{E}[N_\epsilon]^2 + \frac{e}{4}(\alpha(1-\epsilon)e)^{2h-1} + \sum_{k=2}^{h-1}\frac{e}{8}.\frac{(\alpha(1-\epsilon)e)^{2h-k}h}{(k-1)^{(1/2)}(h-k-1)} \\ &= \mathbb{E}[N_\epsilon] + \mathbb{E}[N_\epsilon]^2 + (\alpha(1-\epsilon)e)^{2h}\left(\frac{1}{4\alpha(1-\epsilon)} + \sum_{k=2}^{h-1}\frac{e}{8}.\frac{h}{(k-1)^{(1/2)}(h-k-1)(\alpha(1-\epsilon)e)^k}\right).\end{aligned}$$

We wish to show that what is contained in the large brackets is finite. Clearly, $1/(4\alpha(1-\epsilon))$ is finite. So now we consider the sum and bound it in the following way,

$$\sum_{k=2}^{h-1}\frac{e}{8}.\frac{h}{(k-1)^{(1/2)}(h-k-1)(\alpha(1-\epsilon)e)^k} \leqslant \sum_{k=2}^{h-1}\frac{e}{8}.\frac{h}{(\alpha(1-\epsilon)e)^k}.$$

Now, if $\alpha(1-\epsilon)e > 1$, this is a geometric sum with common ratio $r = 1/(\alpha(1-\epsilon)e)$. By defining $c := \frac{1}{4\alpha(1-\epsilon)} + \frac{he}{8}r^2\left(\frac{1-r^{h-2}}{1-r}\right)$, we see that when $\alpha(1-\epsilon)e > 1$,

$$\mathbb{E}[N_\epsilon^2] \leqslant \mathbb{E}[N_\epsilon] + \mathbb{E}[N_\epsilon]^2 + c(\alpha(1-\epsilon)e)^{2h}.$$

□

2.7 A Bound on $\mathbb{P}[N_\epsilon \geqslant e^{-\delta h}]$ and Completing the Proof.

From Kahane [9] we have the Paley-Zygmund Inequality, that is, let $X > 0$ be a random variable with finite variance and if $\lambda \in [0, 1]$ then,

$$\mathbb{P}[X > \lambda\mathbb{E}[X]] \geqslant (1-\lambda)^2\frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}.$$

We wish to use Paley-Zygmund Inequality on N_ϵ with $\lambda = 1/2$, so first we need to show that N_ϵ has finite variance to satisfy the conditions of the Inequality. Using the standard formula for the variance of a random variable and Proposition 2.10 we readily see that

$$\mathrm{Var}[N_\epsilon] = \mathbb{E}[N_\epsilon^2] - \mathbb{E}[N_\epsilon]^2 \leqslant \mathbb{E}[N_\epsilon] + c(\alpha(1-\epsilon)e)^{2h} \leqslant \mathbb{E}[N] + c(\alpha(1-\epsilon)e)^{2h} \leqslant \frac{n^h}{h!} + c(\alpha(1-\epsilon)e)^{2h},$$

which is finite for finite h as required.

We are now ready to use the Paley-Zygmund Inequality. To satisfy one of the conditions in proof of Proposition 2.10 we choose $\epsilon \in (0,1)$ such that $\alpha(1-\epsilon)e > 1$. Then using the Paley-Zygmund Inequality with $\lambda = 1/2$,

$$\mathbb{P}\left[N_\epsilon \geqslant \frac{\mathbb{E}[N_\epsilon]}{2}\right] \geqslant (1-1/2)^2 \frac{\mathbb{E}[N_\epsilon]^2}{\mathbb{E}[N_\epsilon^2]} \geqslant \frac{\mathbb{E}[N_\epsilon]^2}{4\mathbb{E}[N_\epsilon^2]}. \tag{2.5}$$

Now let $\delta \in (0, \log(\alpha(1-\epsilon)e))$, we aim to show that $\mathbb{E}[N_\epsilon]/2 \geqslant e^{\delta h}$ for sufficiently large h. Now, for some $\eta > 0$ we have that,

$$e^{\delta h} = e^{h\log(\alpha(1-\epsilon)e)-\eta h} = e^{h\log(\alpha(1-\epsilon)e)}e^{-\eta h} = (\alpha(1-\epsilon)e)^h e^{-\eta h}.$$

Now using Proposition 2.9 and using that $(e^{\eta h}/6h^{3/2}) \geqslant 1$ for sufficiently large h,

$$\frac{\mathbb{E}[N_\epsilon]}{2} \geqslant \frac{(\alpha(1-\epsilon)e)^h}{6h^{3/2}} = \frac{e^{\delta h}e^{\eta h}}{6h^{3/2}} = e^{\delta h}\left(\frac{e^{\eta h}}{6h^{3/2}}\right) \geqslant e^{\delta h},$$

Combining this with (2.5) we see that

$$\mathbb{P}\left[N_\epsilon \geqslant e^{\delta h}\right] \geqslant \frac{\mathbb{E}[N_\epsilon]^2}{4\mathbb{E}[N_\epsilon^2]}. \tag{2.6}$$

Using Proposition 2.9 we can see that for large h, $\mathbb{E}[N_\epsilon] \geqslant \frac{(\alpha(1-\epsilon)e)^h}{3h^{3/2}} > 1$. Combining this with Proposition 2.10 we see that,

$$\mathbb{E}[N_\epsilon^2] \leqslant \mathbb{E}[N_\epsilon] + \mathbb{E}[N_\epsilon]^2 + c(3h^{3/2}\mathbb{E}[N_\epsilon])^2 = \mathbb{E}[N_\epsilon] + \mathbb{E}[N_\epsilon]^2 + 9ch^3\mathbb{E}[N_\epsilon]^2 \leqslant (9c+2)h^3\mathbb{E}[N_\epsilon]^2.$$

By defining $c' := 9c + 2$ we have that $\mathbb{E}[N_\epsilon^2] \leqslant c'h^3\mathbb{E}[N_\epsilon]^2$. Therefore,

$$\mathbb{P}\left[N_\epsilon \geqslant e^{\delta h}\right] \geqslant \frac{1}{4c'h^3}. \tag{2.7}$$

2.7.1 Separating the Tree into 2 Segments.

We need a bound on the right hand side of (2.7) that goes to 1 in the limit as $h \to \infty$ to show that the desired probability tends to 1. To do this, we will split the tree into 2 segments, the first 4 levels of the tree and the final $h-4$ levels.

There are $\asymp n^4$ increasing paths from the root to level 4 where each label is less than ϵ. To see why this is true, let $p_1, p_2, ..., p_{n^4}$ be the n^4 paths from the root at level 4 and whether or not a path p_i is increasing is bernoulli distributed, with distribution $V_i \sim \text{Bernoulli}(\epsilon^4/4!)$ for all $i \in \{1, ..., n^4\}$. These random variables are not independent since many of the paths are overlapping. Therefore we cannot use the law of large numbers right away.

However, seeing that for all $i, j \in \{1, ..., n^4\}$, $\mathbb{E}[V_i] = \epsilon^4/4!$, $\text{Cov}(V_i, V_j) \leqslant 1 = 1^{|i-j|}$ and defining $S_{n^4} := \sum_{i=1}^{n^4} V_i$ the number of increasing paths from the root to level 4, we are able to apply Wolpert [18] to S_{n^4} and use the Weak Law of Large Numbers despite the dependence on the paths to obtain

$$\lim_{n\to\infty} \mathbb{P}\left[\left|\frac{S_{n^4}}{n^4} - \mathbb{E}[V_i]\right| > \delta\right] = 0, \tag{2.8}$$

which implies that with high probability, there are $\asymp n^4$ increasing paths from the root to level 4 where each random variable is less than ϵ. For every vertex on level 4 there are n^{h-4} paths to the leaf of length $h-4$ since we are looking at the regular n-ary subtree of length $h-4$. By defining $C := 1/(4c'+1)$ and using (2.7) we see that

$$C\frac{1}{h^3} \leqslant \mathbb{P}\left[N_\epsilon \geqslant e^{\delta h}\right] \leqslant \frac{1}{h^3}, \tag{2.9}$$

that is to say, with probability $\Theta(h^{-3})$ lots of the n^{h-4} paths in the subtree are in $I \cap D_\epsilon$. Thus, the probability there are no paths on the subtree in $I \cap D_\epsilon$ looks like, $(1-h^3)$. So the probability that there are no increasing paths on any of the $\sim n^4$ subtrees is like up to constants, $(1-h^{-3})^{n^4}$.

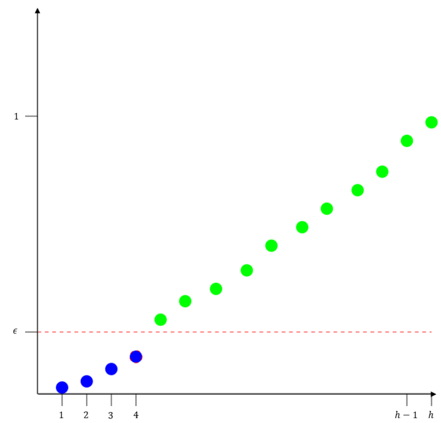

Figure 2.6: An illustration of the paths we are considering, showing how the first 4 levels have labels less than ϵ and the final $h-4$ levels having labels greater than ϵ, and the entire path is increasing.

2.7.2 Subpaths From the Root to Level 4.

To get an exponentially small bound on probability that there are n^4 increasing paths from the root to level 4 Roberts and Zhao use a different technique to that described previously using Wolpert [18]. Instead, they control the random variables more strongly by forcing the random variables on each edge into a specific range depending on their edge level and using an appropriate Chernoff bound and taking care of the dependence of the paths by considering the paths level by level, they are able to show the stronger condition. To consider the subpaths v from the root to level 4 we consider levels in turn by noting that the number of open paths up to level j is dependent on the number of open paths up to level $j-1$. In fact, we consider a stronger condition than the paths just being in $I \cap D_\epsilon$.

Let P_j denote the set of simple paths from the root to level j. Now, for $j \leqslant 4$ define the set,

$$M_j := \{v \in P_j \mid X(v_i) \in [(i-1)\epsilon/4, i\epsilon/4),\ \forall i \in \{1, ..., j\}\}.$$

One such path is shown in Figure 2.7, where the values of the $X(v_i)$ are in the correct intervals for all $i \in \{1, ..., 4\}$. Now, there are n possible paths from the root to level 1 where the edge has a label that is $< \epsilon/4$, so $\#M_1 = \sum_{i=1}^{n} Z_i$ where Z_i are i.i.d Bernoulli random variables with parameter $\epsilon/4$, for all $i \in \{1, ..., n\}$.

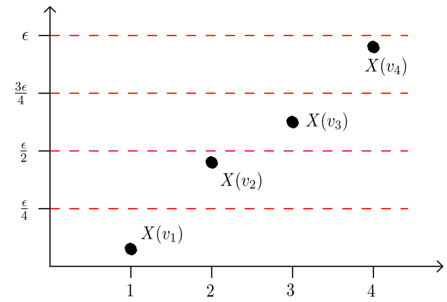

Figure 2.7: A visualisation of the type of paths in M_4.

Again, for M_j with $2 \leqslant j \leqslant 4$, conditioning on that $\#M_{j-1} \geqslant k$ we have $\#M_j \geqslant \sum_{i=1}^{kn} Z_i$ where Z_i are again i.i.d Bernoulli random variables with parameter $\epsilon/4$. From Tsun [17] with $\delta = 1/2$ we obtain the following Lower Tail Chernoff Bound.

Lemma 2.11. Let $Z_1, ..., Z_r$ be independent Bernoulli random variables and define $Z := \sum_{i=1}^{r} Z_i$. Let $\delta \in (0, 1)$. Then,

$$\mathbb{P}(Z \leqslant \frac{\mathbb{E}[Z]}{2}) \leqslant \exp\left(\frac{-\mathbb{E}[Z]}{8}\right). \tag{2.10}$$

The proof is omitted here for space constraints but can be found in [17]. Using this we can prove the following Lemma.

Lemma 2.12.

$$\mathbb{P}(\#M_4 \leqslant (n\epsilon/8)^4) \leqslant 4\exp(-n\epsilon^4/16384).$$

Proof. Due to space constraints we give a sketch proof. Using Lemma 2.11 to show that for each $i \in \{1, ..., 4\}$ we have that $\mathbb{P}(\#M_i \leqslant (n\epsilon/8)^i \mid \#M_{i-1} > (n\epsilon/8^{i-1}) \leqslant \exp\left(\frac{-n(\epsilon/8)^4}{4}\right)$, where the conditioning is not present for the case $i = 1$. Summing each of these bounds gives the result. □

To complete the proof of Theorem 1.4 note that

$$\begin{aligned} \mathbb{P}(N \leqslant \exp(\delta h)) &= \mathbb{P}(N \leqslant \exp(\delta h) \cap \#M_4 \leqslant (n\epsilon/8)^4) + \mathbb{P}(N \leqslant \exp(\delta h) \cap \#M_4 > (n\epsilon/8)^4) \\ &\leqslant \mathbb{P}(\#M_4 \leqslant (n\epsilon/8)^4) + \mathbb{P}(N \leqslant \exp(\delta h),\ \#M_4 > (n\epsilon/8)^4). \end{aligned}$$

Let $u \in M_4$, then the first 4 labels in u are increasing and less than ϵ and consider a subtree of height $h - 4$ that is rooted at u_4. Now, $N_\epsilon \leqslant N$, so for $N \leqslant e^{\delta h}$, there must be no more than $e^{\delta h}$ paths in this subtree that are also in N_ϵ. However, we know from (2.7) that since $n/(h-4) \geqslant n/h = \alpha$, the probability of this event is at most $1 - Kh^{-3}$ for some constant $K > 0$. Combining this with Lemma 2.12, noting that when $\#M_4 > (n\epsilon/8)^4$, for $N \leqslant \exp(\delta h)$ this event must happen at least $(n\epsilon/8)^4$ times and using the inequality $1 + x \leqslant e^x$, we have,

$$\begin{aligned} \mathbb{P}(N \leqslant \exp(\delta h)) &\leqslant 4\exp(-n\epsilon^4/16384) + (1 - c'h^{-3})^{(n\epsilon/8)^4} \\ &\leqslant 4\exp(-n\epsilon^4/16384) + \exp(-c'h^{-3}(n\epsilon/8)^4). \end{aligned}$$

Then for sufficiently small $\eta > 0$ and recalling that $n = \alpha h$ we obtain

$$\mathbb{P}(N \leqslant \exp(\delta h)) \leqslant 4\exp(-(\alpha\epsilon^4/16384)h) + \exp(-c'(\alpha\epsilon/8)^4 h) \leqslant \exp(-\eta h).$$

□

Chapter 3

Accessible Paths on the Irregular Tree.

Throughout this Chapter we look at the irregular tree. We use the notation from [2] and also the structure/arguments in their proofs throughout the rest of the paper.

3.1 Notation on the Irregular Tree and Hypercube.

From now on when we refer to 'the tree' we are referring to the irregular tree, also we let $\mathbb{E}^x[\cdot]$, $\mathbb{P}^x[\cdot]$ and $\mathrm{Var}^x[\cdot]$ denote the expectation and probability of an event and the variance of a random variable where the root value of the irregular tree/hypercube is x. We also denote

$$\mathbb{E}^*[\cdot] = \int_0^1 \mathbb{E}^x[\cdot]dx, \quad \mathbb{P}^*[\cdot] = \int_0^1 \mathbb{P}^x[\cdot]dx, \quad \mathrm{Var}^*[\cdot]$$

to be the expectation and probability of an event and the variance of a random variable where the root value of the irregular tree/hypercube is given by a $U[0,1]$ random variable.

Proposition 3.1. On the hypercube and tree,

$$\mathbb{E}^x[\Theta] = L(1-x)^L, \quad \mathbb{E}^*[\Theta] = 1, \quad \lim_{L\to\infty} \frac{1}{L}\mathrm{Var}^*[\Theta] = 1. \tag{3.1}$$

On the tree,

$$\lim_{L\to\infty} \mathbb{E}^{X/L}[\Theta/L] = e^{-X}, \quad \lim_{L\to\infty} \frac{1}{L}\mathrm{Var}^{X/L}[\Theta/L] = e^{-2X}. \tag{3.2}$$

On the hypercube,

$$\lim_{L\to\infty} \mathbb{E}^{X/L}[\Theta/L] = e^{-X}, \quad \lim_{L\to\infty} \frac{1}{L}\mathrm{Var}^{X/L}[\Theta/L] = 3e^{-2X}. \tag{3.3}$$

Proof. All of the variance results are proven in Chapter 4. Now, note that the first part of (3.1) follows from the fact that there are $L!$ possible paths along the tree/hypercube and the probability that a path is increasing is $(1-x)^{L-1}/(L-1)!$. Integrating this between 0 and 1 gives the second part of (3.1).

Next, the first part of (3.2) follows from the linearity of expectation and using the first part of (3.1). The same is true for the first part of (3.3). □

3.2 A Theorem Given Starting value $x = (\ln L + X)/L$.

Theorem 3.2. On the irregular tree, when the starting value is $x = (\ln L + X)/L$ we have,

$$\lim_{L\to\infty} \mathbb{E}^{\frac{\ln L + X}{L}}(\Theta) = e^{-X}, \qquad \lim_{L\to\infty} Var^{\frac{\ln L + X}{L}}(\Theta) = e^{-2X} + e^{-X}. \tag{3.4}$$

When the starting value x is chosen at random uniformly in $[0,1]$, the probability of having no open path goes to 1 as $L \to \infty$ and,

$$\mathbb{P}^*(\Theta \geqslant 1) \sim \frac{\ln L}{L} \quad \text{as} \quad L \to \infty. \tag{3.5}$$

Proof of Theorem 3.2. To show (3.4) we first use the result from (prop) to see that,

$$\mathbb{E}^{\frac{\ln L + X}{L}}(\Theta) = L\Big(1 - \frac{\ln L + X}{L}\Big)^{L-1}.$$

We will bound $L(1-(\ln L + X)/L)^L$ above and below and then use the Squeeze Theorem from Sohrab [15]. Using the inequality $1 - x \leqslant e^{-x}$ for all $x \in \mathbb{R}$ we see that

$$L\Big(1 - \frac{\ln L + X}{L}\Big)^L \leqslant Le^{-(\ln L + X)} = e^{-X}.$$

We use the inequality $1 - x \geqslant e^{-x-x^2}$, for all $x \in [0, 1/2)$. By noticing that $\frac{\ln L + X}{L} \in [0, 1/2)$ for all $L \geqslant 1$ we have,

$$L\Big(1 - \frac{\ln L + X}{L}\Big)^L \quad \geqslant \quad L(e^{-\frac{\ln L + X}{L} - (\frac{\ln L + X}{L})^2})^L = Le^{-\ln L - X - \frac{(\ln L + X)^2}{L}} = e^{-X}e^{-\frac{(\ln L + X)^2}{L}}.$$

Since $e^{-\frac{(\ln L + X)^2}{L}} \to 1$ as $L \to \infty$, (3.4) follows from (3.1) and the Squeeze Theorem. The result for the bound on the variance is obtained in Chapter 4 of variance results.

Now we prove equation (3.5). Due to space constraints, we explain every step and give details for some of them. First using Markov's Inequality we know $\mathbb{P}^x(\Theta \geqslant 1) \leqslant \mathbb{E}^x(\Theta)$ and hence

$$\mathbb{P}^x(\Theta \geqslant 1) \leqslant \min\{1, L(1-x)^{L-1}\}.$$

Note that, $\mathbb{P}^*(\Theta \geqslant 1) = \int_0^1 \mathbb{P}^x(\Theta \geqslant 1)\, dx$. To compute this, we split the integral at the point x_0 where $\min\{1, L(1-x)^{L-1}\}$ transitions from being 1 to $L(1-x)^{L-1}$. Then, $x_0 = 1 - \exp(\frac{-\ln(L)}{L-1})$ and

$$\mathbb{P}^*(\Theta \geqslant 1) \int_0^{x_0} 1\, dx + \int_{x_0}^1 L(1-x)^{L-1}\, dx = 1 - \exp\Big(\frac{-\ln(L)}{L-1}\Big) + \exp\Big(\frac{-\ln(L)}{L-1}\Big) = \frac{\ln L}{L} + \mathcal{O}(1/L)$$

with the final equality coming from the asymptotic expansion of e^{-x}. We have found an upper bound; next we find a lower bound. Let $f : \mathbb{R} \to \mathbb{R}$ go to ∞ as L does, but such that $f(L) \lll \ln(L)$ and $f(L) < \ln(L)$. It suffices to show that,

$$\lim_{L \to \infty} \mathbb{P}^{\frac{\ln(L) - f(L)}{L}}(\Theta \geqslant 1) = 1 \tag{3.6}$$

because

$$\mathbb{P}^*(\Theta \geqslant 1) \geqslant \int_0^{\frac{\ln(L)-f(L)}{L}} \mathbb{P}^x(\Theta \geqslant 1)\, dx \geqslant \frac{\ln(L) - f(L)}{L}\mathbb{P}^{\frac{\ln(L)-f(L)}{L}}(\Theta \geqslant 1),$$

where in the last inequality we used that $x \mapsto \mathbb{P}^*(\Theta \geqslant 1)$ is a non decreasing function. If the above is true, then combining it with (3) we obtain the result.

Consider the tree starting at $x = \frac{\ln(L) - f(L)}{L}$ and let M be the number of nodes at the first level with a value between x and $\frac{\ln(L)}{L}$. Since there are L nodes and each has a probability of $\frac{f(L)}{L}$ of being between x and $\frac{\ln(L)}{L}$, M follows a binomial distribution, $M \sim \text{Bin}(L, \frac{f(L)}{L})$.

Conditioning on the value of M we see that,

$$\mathbb{P}(\Theta = 0 \mid M = m) \leqslant \mathbb{P}(\text{No open paths through the } m \text{ specific nodes}).$$

Summing over all m we obtain,

$$\mathbb{P}^{\frac{\ln(L)-f(L)}{L}}(\Theta = 0) \quad \leqslant \quad \sum_{m=0}^{L} \binom{L}{m} \Big(\frac{f(L)}{L}\Big)^m \Big(1 - \frac{f(L)}{L}\Big)^{L-m} \Big[\mathbb{P}^{\frac{\ln(L)}{L};L-1}(\Theta = 0)\Big]^m, \tag{3.7}$$

where we used that, for $x \leqslant \ln(L)/L$, $\mathbb{P}^x(\Theta = 0) \leqslant \mathbb{P}^{\frac{\ln(L)}{L}}(\Theta = 0)$, since if the starting value is smaller, there is a greater probability of there existing an open path. The notation $\mathbb{P}^{\frac{\ln(L)}{L};L-1}(\Theta = 0)$ means we are considering a tree of size $L-1$ not L. To sum the equation (4), we use the binomial formula from Spivak [16], by taking $a = \frac{f(L)}{L}\mathbb{P}^{\frac{\ln(L)}{L};L-1}(\Theta = 0)$, $b = 1 - \frac{f(L)}{L}$ and $n = L$ we obtain,

$$\begin{aligned}
\sum_{m=0}^{L} \binom{L}{m} \Big(\frac{f(L)}{L}\Big)^m \Big(1 - \frac{f(L)}{L}\Big)^{L-m} \Big[\mathbb{P}^{\frac{\ln(L)}{L};L-1}(\Theta = 0)\Big]^m &= \Big(\frac{f(L)}{L}\mathbb{P}^{\frac{\ln(L)}{L};L-1}(\Theta = 0) + 1 - \frac{f(L)}{L}\Big)^L \\
&= \Big(1 - (1 - \mathbb{P}^{\frac{\ln(L)}{L};L-1}(\Theta = 0))\frac{f(L)}{L}\Big)^L.
\end{aligned}$$

So,

$$\mathbb{P}^{\frac{\ln(L)-f(L)}{L}}(\Theta = 0) \quad \leqslant \quad \Big(1 - (1 - \mathbb{P}^{\frac{\ln(L)}{L};L-1}(\Theta = 0))\frac{f(L)}{L}\Big)^L.$$

Applying the Cauchy-Schwartz Inequality [1] to Θ and $\mathbb{1}(\Theta \geqslant 1)$ to see that

$$\mathbb{E}[\Theta \mathbb{1}(\Theta \geqslant 1)]^2 \leqslant \mathbb{E}[\Theta^2]\mathbb{E}[\mathbb{1}(\Theta \geqslant 1)^2],$$

so $\mathbb{E}[\Theta]^2 \leqslant \mathbb{E}[\Theta^2]\mathbb{P}(\Theta \geqslant 1)$. Therefore, using this and taking $X = 0$ in (3.4),

$$\mathbb{P}^{\frac{\ln L}{L}}(\Theta \geqslant 1) \geqslant \frac{\mathbb{E}^{\frac{\ln L}{L}}[\Theta^2]}{\mathbb{E}^{\frac{\ln L}{L}}[\Theta]^2} \xrightarrow{L\to\infty} \frac{1}{3}.$$

This means that $\mathbb{P}^{\frac{\ln L}{L};L-1}(\Theta \geqslant 1) \geqslant 0.33$ for sufficiently large L and again for sufficiently large L we have

$$\mathbb{P}^{\frac{\ln L}{L};L-1}(\Theta \geqslant 1) \leqslant \Big(1 - \mathbb{P}^{\frac{\ln(L)}{L};L-1}(\Theta \geqslant 1)\Big)\frac{f(L)}{L}\Big)^L \leqslant \Big(1 - 0.33\frac{f(L)}{L}\Big)^L,$$

which, as $L \to \infty$, goes to 0 since $f(L) \to \infty$, as required. □

3.3 Proof of Theorem 1.6.

To prove this Theorem, we start with a definition.

Definition 3.3. Consider the irregular tree with a starting value $x > 0$. We define the generating function G of Θ to be:

$$G(\lambda, x, L) := \mathbb{E}^x(e^{-\lambda\Theta}). \tag{3.8}$$

Note that although the right-hand-side of (3.8) does not contain L, $G(\lambda, x, L)$ does depend on L implicitly because that determines the height of the irregular tree, which will affect the possible values of Θ. From this definition, we have the following proposition.

Proposition 3.4. Fix $X \geq 0$ and consider a tree with a starting value $x = X/L$. We have a recursive formula for $G(\lambda, x, L)$, given by

$$G(\lambda, x, L) = \Big[x + \int_x^1 G(\lambda, y, L-1)dy\Big]^L, \text{ initialised by } G(\lambda, x, 1) = e^{-\lambda}. \tag{3.9}$$

Proof. We first prove the initialisation of (3.9). We readily see that

$$G(\lambda, x, 1) = \mathbb{E}^X(e^{-\lambda\Theta}) = \sum_{i=0}^{1} ie^{-\lambda i}\mathbb{P}^X[\Theta = i] = e^{-\lambda},$$

where in the last equality we used that the sum is equal to 1 because we are summing over all the probabilities of all the possible values of Θ.

Now we look at the general case for $L \geqslant 1$ of (3.9). We do this by decomposing over level 1. Let σ be one of the L nodes at level 1 and let $\Theta(\sigma)$ be the number of open paths passing through the node σ, since the nodes are independent, we have

$$G(\lambda, x, L) = [\mathbb{E}^x(e^{-\lambda\Theta(\sigma)})]^L. \tag{3.10}$$

Let x_σ be the value attached to the node σ. To compute the above there are 2 cases to consider. Firstly, $x_\sigma < x$ with probability x, and on this event, $\Theta(\sigma) = 0$ because no open path can pass through this, so it contributes $xe^{-\lambda\Theta(\sigma)} = xe^{-\lambda 0} = x$ to the expectation.

Secondly, let $y \in [x, 1]$, then $x_\sigma \in (y, y+dy)$ with probability dy. Since this means $x_\sigma > x$, it is possible for an open path to pass through σ. After vertex σ there is a subtree with root σ, height $L-1$ and on this subtree the number of open paths has generating function $G(\lambda, y, L-1)$. Hence, this contributes $\int_x^1 dyG(\lambda, y, L-1)$ to the expectation. Combining both of these cases we see

$$G(\lambda, x, L) = \Big[x\mathbb{E}^x[e^{-\lambda 0}] + \int_x^1 G(\lambda, y, L-1)dy\Big]^L = \Big[x + \int_x^1 G(\lambda, y, L-1)dy\Big]^L. \tag{3.11}$$

Due to the structure of the tree, the number of levels and the size of each level increase together, so taking limits to find the limiting distribution from this is difficult.

Instead, we will assume that the information of the first k levels of the tree is known and for the later levels we use the law of large numbers. To this end, we begin with the following definition.

Definition 3.5. Let $\mathcal{F}_k$ be the available information up to level k and define Θ_k to be the expectation of Θ given $\mathcal{F}_k$, that is

$$\Theta_k \coloneqq \mathbb{E}(\Theta \mid \mathcal{F}_k). \tag{3.12}$$

So Θ_k is the expected value of Θ given the knowledge of the first k levels. For $k < L$ we have,

$$\Theta_k = \sum_{|\sigma|=k} \mathbb{1}_{\{\sigma \text{ } open\}}(L-k)(1-x_\sigma)^{L-k-1}.$$

To explain this expression, note that since we have all the information up to level k, the event that a path up to a node σ at level k denoted $\{\sigma \text{ } open\}$ is an $\mathcal{F}_{|\sigma|}$-measurable event. Taking the indicator of this event determines whether a possible path through this node along the entire tree could be open. Then we multiply this indicator by $(L-k)(1-x_\sigma)^{L-k-1}$, the expected number of paths along the subtree of height $L-k$ with starting value x_σ at σ, derived in the same way as (3.11). Then we sum over all nodes at level k to obtain the expected number of open paths through the tree, given knowledge up to level k.

If k is small, the number of possible paths up to level k is small so the number of open paths up to level k is small so there is little information given by $\mathcal{F}_k$, therefore the variance of Θ given $\mathcal{F}_k$ is large as Θ will not necessarily be close to Θ_k.

If k is large, there is a lot of information given by $\mathcal{F}_k$, that is, there are many open paths up to level k and many of those paths will be open throughout the entire tree since k is large, that is to say, by the Law of Large Numbers, the variance of Θ given $\mathcal{F}_k$ is small, and therefore Θ_k is a good approximation for Θ.

Note that the number of nodes at level k is $L!/(L-k)!$ and if σ' is a node at level k then $\mathbb{P}(\sigma' \text{ is open}) = 1/k!$. Hence, the expected number of open paths is $\binom{L}{k}$. So we see that since L is much larger than k, as we increase k, the expected number of open paths increases and so by the previously mentioned statement, Θ_k becomes a better approximation for Θ as k increases. Looking at Θ_k/L with starting value X/L, we can take the limit as $L \to \infty$ for fixed k and then take the limit as $k \to \infty$ to ensure that k is large but still smaller than L such that the expected number of open paths at level k is high. The proof consists of 2 parts and involves the idea of taking 2 limits, as explained previously.

Part 1) If $\frac{X}{L}$ is the starting value of the tree then we show that $\frac{\Theta}{L}$ for L large has the same distribution as $\frac{\Theta_k}{L}$ for L large and then for k large. That is:

$$\lim_{L\to\infty} \mathbb{P}^{\frac{X}{L}}\Big(\frac{\Theta}{L} \leqslant z\Big) = \lim_{k\to\infty}\lim_{L\to\infty} \mathbb{P}^{\frac{X}{L}}\Big(\frac{\Theta_k}{L} \leqslant z\Big). \tag{3.13}$$

Part 2) Construct a new generating function similar to the one previously defined to find the distribution of $\frac{\Theta_k}{L}$ as $L \to \infty$ and then $k \to \infty$. By comparing this to previously known generating functions, we see this follows an exponential law.

3.3.1 Proof of Part 1.

Due to space constraints, we only give a sketch proof here. We fix $\delta > 0$ and by considering the cases $\frac{\Theta_k}{L} \leqslant z \pm \delta$ and $\frac{\Theta_k}{L} > z \pm \delta$ we obtain the following inequality,

$$\begin{aligned}\mathbb{1}\Big(\frac{\Theta_k}{L} \leqslant z-\delta \mid \mathcal{F}_k\Big) + \mathbb{P}\Big(\frac{\Theta-\Theta_k}{L} \geqslant \delta \mid \mathcal{F}_k\Big) &\leqslant \mathbb{P}\Big(\frac{\Theta}{L} \leqslant z \mid \mathcal{F}_k\Big)\\ &\leqslant \mathbb{1}\Big(\frac{\Theta_k}{L} \leqslant z+\delta \mid \mathcal{F}_k\Big) + \mathbb{P}\Big(\frac{\Theta-\Theta_k}{L} \geqslant \delta \mid \mathcal{F}_k\Big).\end{aligned} \tag{3.14}$$

Now using Chebyshev's Inequality from Probability Theory [1] on the random variable Θ given $\mathcal{F}_k$ we see that

$$\mathbb{P}\Big(\frac{|\Theta-\Theta_k|}{L} \geqslant \delta \mid \mathcal{F}_k\Big) \leqslant \frac{\mathrm{Var}(\Theta|\mathcal{F}_k)}{L^2\delta^2}. \tag{3.15}$$

Substituting this into (3.14) and taking the expectation over $\mathcal{F}_k$ yields

$$\mathbb{P}^{\frac{X}{L}}\Big(\frac{\Theta_k}{L} \leqslant z-\delta\Big) - \frac{\mathbb{E}^{\frac{X}{L}}[\mathrm{Var}(\Theta|\mathcal{F}_k)]}{L^2\delta^2} \leqslant \mathbb{P}\Big(\frac{\Theta}{L} \leqslant z\Big) \leqslant \mathbb{P}^{\frac{X}{L}}\Big(\frac{\Theta_k}{L} \leqslant z+\delta\Big) + \frac{\mathbb{E}^{\frac{X}{L}}[\mathrm{Var}(\Theta|\mathcal{F}_k)]}{L^2\delta^2}. \tag{3.16}$$

Therefore to show (3.13), it is sufficient to show

$$\lim_{L\to\infty} \frac{1}{L^2}\mathbb{E}^{\frac{X}{L}}[\mathrm{Var}(\Theta|\mathcal{F}_k)] = \frac{e^{-2X}}{2^k}, \tag{3.17}$$

because as $k \to \infty$ this goes to 0 and forces the desired inequality by (3.16). In Chapter 4 we give a proof of (3.17).

3.3.2 Proof of Part 2.

In a similar way to the start of the proof of this Theorem, we define the generating function for Θ_k.

Definition 3.6. Fix $x \geqslant 0$ and consider a tree with height L a starting value x. We define the generating function G_k of Θ_k, for $k \leqslant L$ as

$$G_k(\lambda, x, L) := \mathbb{E}^x(e^{-\lambda\Theta_k}). \tag{3.18}$$

When $k = 0$, we have have no information about the tree so by (num), $\Theta_0 = \mathbb{E}^x(\Theta) = L(1-x)^{L-1}$. Thus,

$$G_0(\lambda, x, L) = \exp[-\lambda L(1-x)^{L-1}], \tag{3.19}$$

and using a very similar technique to before, we have that

$$G_k(\lambda, x, L) = \Big[x + \int_x^1 G_{k-1}(\lambda, y, L-1)dy\Big]^L = \Big[1 - \int_x^1 (1 - G_{k-1}(\lambda, y, L-1))dy\Big]^L. \tag{3.20}$$

First we take the limit as $L \to \infty$ in (3.19) and (3.20) for k fixed and then take the limit as $k \to \infty$. There are 2 steps to showing this, firstly, we show that following limit exists and then the second step is showing that this generating function is the same as that of an exponential random variable. We only consider the case for $\lambda \geqslant 0$ as it is sufficient to characterise the distribution.

Step 1. Let $a, b, \mu \in \mathbb{R}$ and $\mu \geqslant 0$. Then we show that the following limit exists,

$$G_k\Big(\frac{\mu}{L+a}, \frac{X}{L+b}, L\Big) \xrightarrow{L\to\infty} \tilde{G}_k(\mu, X), \tag{3.21}$$

where

$$\tilde{G}_k(\mu, X) = \exp\Big[-\int_X^\infty 1 - \tilde{G}_{k-1}(\mu, Y)dY\Big], \quad \tilde{G}_0(\mu, X) = \exp\Big[-\mu e^{-X}\Big]. \tag{3.22}$$

We show (3.21) and (3.22) are true by induction. For the base case ($k = 0$), from (3.19), the algebra of limits and the definition of e as a limit, it is clear that (3.21) is true with limit given in (3.22).

For the inductive hypothesis, let $k > 0$ and assume that (3.21) and (3.22) hold for G_{k-1}. Now we consider G_k. Using (3.20) we have,

$$G_k\Big(\frac{\mu}{L+a}, \frac{X}{L+b}, L\Big) = \Big[1 - \frac{1}{L+b}\int_X^{L+b} (1 - G_{k-1}\Big(\frac{\mu}{L+a}, \frac{Y}{L+b}, L\Big)dY\Big]^L \tag{3.23}$$

Now using the inductive hypothesis, we know that the G_{k-1} term on the right hand-side has a limit and from (3.18) and the fact that $e^{-z} \geqslant 1 - z$ for $z \in \mathbb{R}$,

$$1 \geqslant G_k(\lambda, x, L) \geqslant 1 - \lambda\mathbb{E}^x(\Theta_k) = 1 - \lambda\mathbb{E}^x(\Theta) = 1 - \lambda L(1-x)^{L-1} \tag{3.24}$$

Then, using the above, that $\mu > 0$ and that for L sufficiently large, $L/(L+a) < \sqrt{2}$ and $1-Y/(L+b) < 1$ we have that

$$1 \geqslant G_{k-1}\Big(\frac{\mu}{L+a}, \frac{Y}{L+b}, L-1\Big) \geqslant 1 - \frac{\mu L}{L+a}\Big(1-\frac{Y}{L+b}\Big)^{L-2} \geqslant 1 - \sqrt{2}\mu\Big(1-\frac{Y}{L+b}\Big)^{L}. \tag{3.25}$$

Now, $(1 - L/(L + b))^L \to e^{-Y}$ as $L \to \infty$, so for sufficiently large L, $(1 - L/(L + b))^L < \sqrt{2}e^{-Y}$. Therefore,

$$1 \geqslant G_{k-1}\Big(\frac{\mu}{L+a}, \frac{Y}{L+b}, L-1\Big) \geqslant 1 - 2\mu e^{-Y} \geqslant 1 - 2\mu e^{-\frac{Y}{2}}, \tag{3.26}$$

where to obtain the final inequality, we use that e^{-Y} is a decreasing function in Y. Now, note that since $G_{k-1}(\frac{\mu}{L+a}, \frac{Y}{L+b}, L-1) \leqslant 1$, for all $L \in \mathbb{R}$, they are bounded above and since they converge we can use the Dominated Convergence Theorem from Bogachev [3] to see that

$$\int_X^{L+b} \Big(1 - G_{k-1}\Big(\frac{\mu}{L+a}, \frac{Y}{L+b}, L\Big)\Big)dY \xrightarrow{L\to\infty} \int_X^{\infty} (1 - \tilde{G}_{k-1}(\mu, Y))dY. \tag{3.27}$$

So we have,

$$\begin{aligned} G_k(\frac{\mu}{L+a}, \frac{X}{L+b}, L) &= \Big[1 - \frac{1}{L+b}\int_X^{L+b}(1 - G_{k-1}(\frac{\mu}{L+a}, \frac{Y}{L+b}, L)dY\Big]^L \\ &\xrightarrow{L\to\infty} \exp\Big[-\int_X^{\infty} 1 - \tilde{G}_{k-1}(\mu, Y)dY\Big] \end{aligned}$$

as required.

Step 2. From step 1, we know that by setting $b = 0$ (i.e. when the tree has root value X/L), there exists a limit as $L \to \infty$ of $\frac{\Theta_k}{L}$ and this limit has generating function $\tilde{G}_k$. Now we take the limit as $k \to \infty$ of this generating function and by comparison to the known generating function of an exponential random variable and using the uniqueness of generating functions for random variables, we see that $\lim\limits_{L\to\infty} \frac{\Theta_k}{L}$ converges when $k \to \infty$ to an exponential random variable.

To simplify computations we write $\tilde{G}_k$ as a function of one variable, $\tilde{G}_k(\mu, X) = F_k(\mu e^{-X})$ where,

$$F_k(z) = \exp\Big[-\int_0^z \frac{1 - F_{k-1}(z')}{z'}dz'\Big], \quad F_0(z) = e^{-z}. \tag{3.28}$$

If $Z \sim \text{Exp}(\beta)$, then for $t < \beta$ the moment generating function of Z is given by:

$$M_Z(t) = \frac{\beta}{\beta - t}. \tag{3.29}$$

We will show that the solution to (3.28) satisfies, for $z > -1$,

$$F_k(z) \xrightarrow{k\to\infty} \frac{1}{1+z}. \tag{3.30}$$

By comparison with (3.29), this implies that $\lim\limits_{L\to\infty} \Theta_k/L$ converges weakly to an exponential distribution with expectation e^{-X}. Note this solution is true for $z > -1$, however, we will only show this is true for $z \geqslant 0$ since we only proved (**??**) in that case. Define a function $\delta_k : (-1, \infty)\backslash\{0\} \to \mathbb{R}$ such that

$$F_k(z) = \frac{1}{1+z} - \frac{z^2}{(1+z)^3}\frac{\delta_k(z)}{2^k}. \tag{3.31}$$

Notably, we will show that this function is bounded above by a constant and below by 0, that is to say, there exists $B \in \mathbb{R}$ such that for all k,

$$0 \leqslant \delta_k(z) \leqslant B. \tag{3.32}$$

Using this we will then bound $F_k(z)$ above and below and upon taking the limit as $k \to \infty$, we will obtain the desired result. Firstly, we show (3.32) by induction. For the base case, $k = 0$, using (3.31) and (3.28) we see that

$$F_0(z) = e^{-z} = \frac{1}{1+z} - \frac{z^2}{(1+z)^3}\frac{\delta_0(z)}{2^k}, \tag{3.33}$$

rearranging gives

$$\delta_0(z) = \frac{(1+z)^3}{z^2}\Big(\frac{1}{1+z} - e^{-z}\Big). \tag{3.34}$$

Then, by expressing e^z as an infinite series it is clear that it is greater than the first 2 terms, i.e., $e^z \geqslant 1 + z$ for $z \geqslant 0$. Therefore, $\delta_0(z) \geqslant 0$ as required. In addition to this, we have that,

$$\lim_{z\to\infty} \delta_0(z) = \lim_{z\to\infty} \frac{(1+z)^3}{z^2}\Big(\frac{1}{1+z} - e^{-z}\Big) = \lim_{z\to\infty} \frac{(1+z)^3}{z^2(1+z)} + \lim_{z\to\infty} \frac{(1+z)^3}{z^2 e^z} = 1 \tag{3.35}$$

This limit is finite. Also, $\delta_0(-1) = 0$ and by continuity we can define $\delta_0(0)$ to be a finite value. Therefore, because the function is continuous, there exists a constant $B \in \mathbb{R}$ that $\delta_k(z)$ is bounded above by for all $k \in \mathbb{N}$ and $z > -1$.

Now for the inductive hypothesis, assume that (3.32) holds for $k - 1$, that is to say

$$F_k(z) = \frac{1}{1+z}\exp\Big[-\int_0^z \frac{z'}{(1+z')^3}\frac{\delta_k(z')}{2^k}dz'\Big]. \tag{3.36}$$

So, again using the fact that for all $x \in \mathbb{R}$, $e^x \geqslant 1 + x$ and the bounds (3.28) we have

$$\frac{1}{1+z} \geqslant F_k(z) \geqslant \frac{1}{1+z}\Big[1 - \frac{B}{2^{k-1}}\int_0^z \frac{z'}{(1+z')^3}dz'\Big] = \frac{1}{1+z}\Big[1 - \frac{B}{2^k}\frac{z^2}{(1+z)^2}\Big]. \tag{3.37}$$

So using (3.31) and multiplying through by 1/(1+z) which is possible since this is always positive, we see that

$$1 \geqslant 1 - \frac{z}{(1+z)^2}\frac{\delta_k(z)}{2^k} \geqslant 1 - \frac{B}{2^k}\frac{z^2}{(1+z)^2}, \tag{3.38}$$

which yields

$$0 \leqslant \delta_k(z) \leqslant B, \tag{3.39}$$

completing the inductive step, and proving the bounds on $\delta_k(z)$ for all $k \in \mathbb{N}$ and $z > -1$. Taking the limit in (num) by using the Squeeze Theorem [15], we see that $F_k(z) \xrightarrow{k\to\infty} \frac{1}{1+z}$, completing the proof of Theorem 1. □

Chapter 4

Second Moment Results on the Irregular Tree.

4.1 Results on the Second Moment of the Tree

In this section we will prove that (3.1), (3.2), (3.4), (3.17) that were used in the proofs of Theorem 1.3 and Theorem 3.2. These calculations will involve the variance so before we prove these results, we will first investigate $\mathbb{E}[\Theta^2]$.

4.2 Exact Expression of the Second Moment

We wish to find an exact expression for $\mathbb{E}[\Theta^2]$. To this end, we begin with the following definition.

Definition 4.1. Define P_t to be the set of all paths along the tree and Q_t to be the set of all pairs of paths along the tree, that is $Q_t := P_t \times P_t$.

Then we have that,

$$\mathbb{E}[\Theta^2] = \sum_{p \in P_t} \sum_{p' \in P_t} \mathbb{P}[p \text{ and } p' \text{ are open}] = \sum_{\{p,p'\} \in Q_t} \mathbb{P}[p \text{ and } p' \text{ are open}]. \tag{4.1}$$

The number of possible pairs of paths is $L!^2$, since there are $L!$ choices for the first path and $L!$ choices for the second. So we see that $|Q_t| = L!^2$.

The probability that a given pair of paths are accessible depends on the number of edges the pair of paths share, call this number q. By the structure of the irregular tree, it is impossible for 2 paths to share only $L-1$ edges, because then they must also share the final edge too, since for a node at level $L-1$ there is only one edge going to level L. So we see that have that $q \in \{0, 1, ..., L-2, L\}$.

The number of pairs of paths that are identical (i.e. $q = L$) is $L!$, since this is simply the same as the number of paths along the tree, which is $L!$.

The probability that each of these paths in the pair is open is the same as the probability one is open (because they completely overlap), and when the starting value is x, this is $(1-x)^{L-1}/(L-1)!$ since the first $L-1$ bonds must have value between x and 1 and these $L-1$ bonds have to be increasing (i.e there is only one combination that is increasing out of the $(L-1)!$ possible combinations).

The number of pairs of paths that share $q \in \{0, 1, ..., L-2\}$ nodes is

$$L!(L-q-1)(L-q-1)!. \tag{4.2}$$

This is because there are $L!$ choices for the first path. Then the second path is the same for the first q edges, the next edge must be different and there are $L-q-1$ choices for this, since one of the $L-q$

edges attached to this edge is already in the first path and we know it is not in the second. Then there are $(L-q-1)!$ possible choices for the final $L-q-1$ edges in the second path.

We now compute the probability that when the tree has starting value x, both of the paths are open. Now, both of the paths combined contain $L+L-q=2L-q$ distinct edges. We know the final edges always have value one and we need each of the other $2L-q-2$ edges to have a value in between x and 1. Now, we need these $2L-q-2$ edges to be in the correct order. There are $(2L-q-2)!$ possible combinations of these values. We need the q smallest terms to be in increasing order in the shared segment of length q of both of the paths, there is only one possible way for this to happen. Next, we need both of the separate sections of length $L-q-1$ to be in increasing order also. There are $\binom{2L-2q-2}{L-q-1}$ ways to separate these final $2L-2q-2$ terms into 2 equally sized blocks, one for each path, combining all this, we see that for two paths that share the first q edges, the probability that both of these paths are increasing is,

$$\frac{(1-x)^{2L-q-2}}{(2L-q-2)!}\binom{2L-2q-2}{L-q-1}, \tag{4.3}$$

so we have that,

$$\mathbb{E}^x[\Theta^2] = L(1-x)^{L-1} + \sum_{q=0}^{L-2} a(L,q)(1-x)^{2L-q-2}. \tag{4.4}$$

The first term of (4.4) is for when the paths in the pair of paths are the same, so is the same as $\mathbb{E}^x[\Theta]$. In the sum in the second half of (4.4), we have,

$$a(L,q) = \frac{L!(L-q-1)(L-q-1)!}{(2L-q-2)!}\binom{2L-2q-2}{L-q-1} = \frac{L!(2L-2q-2)!}{(L-q-2)!(2L-q-2)!}.$$

4.3 Estimates and Bounds on $a(L,q)$.

The following proposition gives some useful results regarding $a(L,q)$.

Proposition 4.2. Firstly,

$$a(L,q) = \frac{L^2}{2^q}\frac{(1-\frac{q+1}{L})(1-\frac{q}{L})\cdots(1-\frac{1}{L})}{(1-\frac{2q+1}{2L})(1-\frac{2q}{2L})\cdots(1-\frac{q+3}{2L})(1-\frac{q+2}{2L})}. \tag{4.5}$$

Then, also if if $q \ll \sqrt{L}$,

$$a(L,q) = \frac{L^2}{2^q}\Big[1+\mathcal{O}\Big(\frac{q^2}{L}\Big)\Big]. \tag{4.6}$$

Finally, for L sufficiently large and $q_0(L) := \lceil 1+\ln(L^2)/\ln(2)\rceil$,

$$a(L,q) = \begin{cases} L^2 1.99^{-q} & \text{if } 0 \leqslant q \leqslant q_0(L), \\ 2 & \text{if } q_0(L) \leqslant q \leqslant L-2. \end{cases} \tag{4.7}$$

Proof. To compute (4.5) we expand the factorials, divide each term on the numerator by L and each term on the denominator by $2L$ and then simplify.

Now we show (4.6). If $q \ll \sqrt{L}$ then $q^2 \ll L$ which means $q^2/L \to 0$ as $L \to \infty$. From above we can write $a(L,q)$ in the following way, where when we expand the brackets we only show the leading term,

$$\begin{aligned} a(L,q) &= \frac{2^q}{L} \frac{((L-q-1)(L-q)\cdots(L-1)L)}{(((2L-2q-1)(2L-2q)\cdots(2L-q-3)(2L-q-2))} \frac{L^2}{2^q} \\ &= \frac{2^q}{L} \frac{(L^{q+1}+\cdots)}{((2L)^q+\cdots)} \frac{L^2}{2^q} \\ &= \frac{2^q L^{q+1}}{L(2L)^q} \frac{1+\frac{1}{L^{q+1}}(-L^{\frac{(q+2)(q+1)}{2}}+\cdots)}{1+\frac{1}{(2L)^q}(-(2L)^{q-1}\frac{3q}{2}(q+1)+\cdots)} \frac{L^2}{2^q} = [1+\mathcal{O}(q^2/L)]\frac{L^2}{2^q}. \end{aligned}$$

The final equality above comes from noticing the numerator and denominator are $1+\mathcal{O}(q^2/L)$ and then using the geometric series expansion, for $y \in (0,1)$ of $1/(1+y) = 1-y+y^2-y^3+...$ (on the denominator).

The above is when $q \ll \sqrt{L}$. If q is close L for example, $a(L,L-2)=2$, $a(L,L-3)=24/(L+1)$ and $a(L,L-4)=360/((L+1)(L+2))$. Note that $a(L,L-1)$ is not needed given as an example because recall it is impossible for $q=L-1$ because of the structure of the tree.

The full proof of (4.7) is omitted here due to space constraints, but we explain how to prove the result. Firstly, show that $\ln(h(q))$ is a convex function by showing that when $L \geqslant q+3$, the discrete second derivative is non-negative, in turn, this implies that $a(L,q)$ is convex. We then find an upper bound for $a(L,q)$ using (4.5) by noticing that the numerator is less than 1 and the denominator is bigger than $(1-\frac{2q+1}{2L})^q$. Defining $q_0(L)$ as in the statement of the proposition, using the definition of convexity and looking as L gets sufficiently large yields the result.

4.4 Outline of Proofs of the Limits (3.1), (3.2), (3.4) and (3.17).

The general method to prove (3.1), (3.2), (3.4) and (3.17) of $\mathrm{Var}(\Theta)$ is based around each of these expressions containing a sum similar (4.4) and we separate this into 2 separate sums. The first sum from $q=1$ to $q_o(L)$ we will bound appropriately using (4.7) and use a discrete form of the dominated convergence theorem to determine its limit as $L \to \infty$. Regarding the second sum from $q = q_0(L)+1$ to $L-2$, we bound the running term above by $2/L^2$ which shows the whole sum is bounded above by $1/L$ which goes to 0 as $L \to \infty$.

To prove (3.1) we use the standard definition of variance and integrate over all the possible values of the starting node. Then use the technique described above. To prove (3.2) we calculate $\mathbb{E}^{\frac{X}{L}}[(\Theta/L)^2]$ and use the same technique above.

To prove (3.4) bounding the second sum from $q = q_0(L)+1$ to $L-2$ is slightly more difficult. We use (4.7) and by extending it to the sum from $q=0$ to ∞ we can use the geometric series formula to obtain an upper bound that goes to 0 in the large L. Then, we use the standard formula for variance to obtain the result.

The calculation to prove (3.17) is more involved. We write $\mathrm{Var}(\Theta|\mathcal{F}_k)$ as a sum over the vertices at level k of the variances of the subtrees of height $L-k$ multiplied by the indicator function of whether or not the root vertex of that subtree is open, that is,

$$\mathrm{Var}(\Theta|\mathcal{F}_k) = \sum_{|\sigma|=k} \mathbb{1}_{\{\sigma \text{ open}\}} v(x_\sigma, L-k),$$

where $v(x_\sigma, L-k)$ is the variance of Θ on the subtree of height $L-k$, with root σ with value x_σ. Using (4.4) and taking the $q=0$ term out of the sum we obtain,

$$v(x_\sigma, L) := \mathbb{E}^{x_\sigma}(\Theta^2) - \mathbb{E}^{x_\sigma}(\Theta)^2 = L(1-x_\sigma)^{L-1} + \sum_{q=1}^{L-2} a(L,q)(1-x_\sigma)^{2L-q-2} - L(1-x_\sigma)^{2L-2}.$$

Now, by the structure of the tree there are $L!/(L-k)!$ nodes at level k and the probability that a given node σ at level k is open is $(x_\sigma - x)^{k-1}/(k-1)!$ since the $(k-1)$ labels need to be increasing and need to all be between x and x_σ. Hence,

$$\mathbb{E}^x[\mathrm{Var}(\Theta|\mathcal{F}_k)] = \frac{L!}{(L-k)!}\int_0^1 \frac{(x_\sigma - x)^{k-1}}{(k-1)!} v(x_\sigma, L-k)dx_\sigma.$$

To compute this integral we use the fact that $\int_x^1 (y-x)^m(1-x)^n dy = m!n!(1-x)^{m+n+1}/(m+n+1)!$. This fact is obtained through a change of variables to $z = (y-x)/(1-x)$ and then comparing this to the PDF of a Beta Distribution gives the fact. After lots of manipulation and simplifying, we obtain,

$$\mathbb{E}^x[\mathrm{Var}(\Theta|\mathcal{F}_k)] = \frac{a(L,k)}{L-k-1}(1-x)^{2L-k-2} + \sum_{q=k+1}^{L-2} a(L,q)(1-x)^{2L-q-2} + L(1-x)^{L-1}. \tag{4.8}$$

Then we complete the proof by dividing through by L^2 and using the general technique used previously.

□

Chapter 5

Accessible Paths on the Hypercube.

We follow the steps of the proof given in [2] and explain some sections in more detail. Unfortunately, there is a mistake in the proof that does not appear easy to rectify.

5.1 Outline of the Proof of Theorem 1.7.

We will attempt to use a similar technique as used before on the tree to obtain the distribution of Θ/L as $L \to \infty$ on the hypercube, that is to say, ultimately finding a generating function for this and comparing it to known generating functions to give us the distribution of Θ/L in this limit.

Recall before we looked at the number of open paths given knowledge of the first k levels, we use a similar technique but now define

$$\Theta_k := \mathbb{E}[\Theta|\mathcal{F}_k], \tag{5.1}$$

where $\mathcal{F}_k$ is knowledge of the nodes with distance less than or equal to k from the root (the starting point) of the hypercube and also the nodes with distance less than or equal to k from the node $(1, 1, ..., 1)$.

There are $\binom{L}{k}$ nodes that are k steps away from the root, this is because each node at level k is a vector of length L containing k ones. When describing a node at this level, we will denote it σ, and the value of the node will be denoted x_σ. By symmetry of the hypercube, there are $\binom{L}{k}$ nodes that are k steps away from the final vertex of each path, the node $(1, 1, ..., 1)$. Similarly, when describing a node at this level, we will denote it τ, and the value of the node will be denoted $1 - y_\tau$.

Let n_σ be the number of open paths from the root to σ. Note that there can be no open paths and also that there are most $k!$ open paths since there are only $k!$ possible paths from the root to a node at distance k from the root. Similarly, m_τ be the number of open paths from the vertex τ the final node $(1, 1, ..., 1)$. Note again that $0 \leqslant m_\tau \leqslant k!$.

We let $\mathbb{1}(\sigma \leftrightarrow \tau)$ be the indicator function on the event that there exists a directed path from σ to τ. It will make sense why we have defined this when we begin to forbid paths going through certain nodes.

Now we look at the expected number of increasing paths between σ and τ, given x_σ and y_τ. This is given by $\mathbb{1}(\sigma \leftrightarrow \tau)\mathbb{1}(x_\sigma + y_\tau \leqslant 1)(L-2k)(1 - y_\tau - x_\sigma)^{L-2k-1}$. Firstly note that $\mathbb{1}(\sigma \leftrightarrow \tau)$ is necessary because if there is no path possible we would expect there to be no increasing paths. Next, $\mathbb{1}(x_\sigma + y_\tau \leqslant 1)$ is necessary because were $x_\sigma > 1 - y_\tau$, no increasing path would be possible because the start node value would be greater than the final node value. This is all multiplied by $(L - 2k)$ because that is the length of the path between the two nodes. Therefore, knowing this expectation we have that the expected number of paths on the hypercube is given by,

$$\Theta_k = \sum_{|\sigma|=k} \sum_{|\tau|=L-k} n_\sigma m_\tau \mathbb{1}(\sigma \leftrightarrow \tau)\mathbb{1}(x_\sigma + y_\tau < 1)(L-2k)(1 - y_\tau - x_\sigma)^{L-2k-1}. \tag{5.2}$$

From here we can outline the proof into 3 steps.

Step 1) Show that Θ/L as $L \to \infty$ has the same distribution as that of Θ_k/L as $L \to \infty$ and then as $k \to \infty$. This is similar to what we did on the tree.

Step 2) We then show that as $L \to \infty$, Θ_k/L has the same distribution as a new random variable $\tilde{\Theta}_k/L$, which can be rewritten as a product of two independent sums, one over σ and one over τ.

Step 3) We then show that these two sums each have the same limit as $L \to \infty$ and that these follow the same distribution as to what we found on the tree. However, we will show that the proof at this step fails to be correct.

5.2 Proof of Theorem 1.7 - Step 1

Fix $X \geqslant 0$, and let the root node have value $x = X/L$. Then we wish to show that weakly as $k \to \infty$ we have that

$$\lim_{L\to\infty} \frac{\Theta_k}{L} \longrightarrow \lim_{L\to\infty} \frac{\Theta}{L}. \tag{5.3}$$

As on the tree by the same reasoning, all we need to show is that

$$\lim_{k\to\infty} \limsup_{L\to\infty} \frac{1}{L^2} \mathbb{E}^x[\mathrm{Var}(\Theta|\mathcal{F}_k)] = 0. \tag{5.4}$$

To this end, note that because $\mathbb{E}^x[\Theta|\mathcal{F}_k] = \Theta_k$ we have,

$$\mathbb{E}^x[\mathrm{Var}(\Theta|\mathcal{F}_k)] \;=\; \mathbb{E}^x[\mathbb{E}^x[\Theta^2|\mathcal{F}_k] - \mathbb{E}^x[\Theta|\mathcal{F}_k]^2] = \mathbb{E}^x[\mathbb{E}^x[\Theta^2|\mathcal{F}_k] - \Theta_k^2] = \mathbb{E}^x[\Theta^2] - \mathbb{E}^x[\Theta_k^2]. \tag{5.5}$$

To compute this, we compute both the $\mathbb{E}^x[\Theta^2]$ and $\mathbb{E}^x[\Theta_k^2]$. Define P_h to be the set of paths on the hypercube and let $\alpha \in P_h$. We denote x_i^α to be the value of the i'th node on α for $i \in \{0, ..., L\}$, where $x_0^\alpha = x$ as this is the value of the root of the hypercube and $x_L^\alpha = 1$ as this is the value of the vertex $(1, 1, ..., 1)$. Next, for $0 \leqslant i \leqslant j \leqslant L$, define $\xi_{i,j}^\alpha$ be the indicator function on the event hat α is open between node i and node j. That is,

$$\xi_{i,j}^\alpha := \mathbb{1}(x_i^\alpha \leqslant x_{i+1}^\alpha \leqslant \cdots \leqslant x_j^\alpha) = \begin{cases} 1 & \text{if } x_i^\alpha < x_{i+1}^\alpha \leqslant \cdots \leqslant x_j^\alpha, \\ 0 & \text{otherwise.} \end{cases} \tag{5.6}$$

Then since Θ is the number of open paths on the hypercube and $\xi_{0,L}^\alpha$ is the indicator on the event that the entire path α is increasing we have that $\Theta = \sum_{\alpha\in P_h} \xi_{0,L}^\alpha$, leading to

$$\mathbb{E}^x[\Theta^2] = \sum_{\alpha,\beta\in P_h} \mathbb{E}^x[\xi_{0,L}^\alpha \xi_{0,L}^\beta] = L! \sum_{\alpha\in P_h} \mathbb{E}^x[\xi_{0,L}^\alpha \xi_{0,L}^0], \tag{5.7}$$

where in the final equality we arbitrarily pick a path fix it, denoting it κ, and so denoting the corresponding indicator function $\xi_{0,L}^\kappa$. In the final equality, we also use the fact that there are $L!$ possible paths on the tree.

In a similar way, Θ_k is the number of open paths on the hypercube given $\mathcal{F}_k$ so $\Theta_k = \sum_{\alpha\in P_h} \mathbb{E}[\xi_{0,L}^\alpha|\mathcal{F}_k]$, this leads to

$$\mathbb{E}^x[\Theta_k^2] = \sum_{\alpha,\beta\in P_h} \mathbb{E}^x[\mathbb{E}[\xi_{0,L}^\alpha|\mathcal{F}_k]\mathbb{E}[\xi_{0,L}^\beta|\mathcal{F}_k]] = L! \sum_{\alpha\in P_h} \mathbb{E}^x[\mathbb{E}[\xi_{0,L}^\alpha|\mathcal{F}_k]\mathbb{E}[\xi_{0,L}^\kappa|\mathcal{F}_k]]. \tag{5.8}$$

Now, note that a path is increasing from 0 to L if and only if it is increasing from 0 to k, increasing from k to $L-k$ and increasing from $L-k$ to L. Because of this, we can write $\xi_{0,L}^\alpha = \xi_{0,k}^\alpha \xi_{k,L-k}^\alpha \xi_{L-k,L}^\alpha$.

Also, note that $\xi^{\alpha}_{0,k}$ and $\xi^{\alpha}_{L-k,L}$ are $\mathcal{F}_k$-measurable, so we can use the property of taking out what is known to obtain

$$\mathbb{E}^x[\Theta_k^2] = L! \sum_{\alpha \in P_h} \mathbb{E}^x[\xi^{\alpha}_{0,k}\xi^{\kappa}_{0,k}\mathbb{E}[\xi^{\alpha}_{k,L-k}|\mathcal{F}_k]\mathbb{E}[\xi^{\kappa}_{k,L-k}|\mathcal{F}_k]\xi^{\alpha}_{L-k,L}\xi^{\kappa}_{L-k,L}]. \tag{5.9}$$

By the tower rule [19], $\mathbb{E}^x[\cdot] = \mathbb{E}^x[\mathbb{E}[\cdot|\mathcal{F}_k]$, we can similarly modify (5.9), again by taking out what is known to obtain

$$\mathbb{E}^x[\Theta_k^2] = L! \sum_{\alpha \in P_h} \mathbb{E}^x[\xi^{\alpha}_{0,k}\xi^{\kappa}_{0,k}\mathbb{E}[\xi^{\alpha}_{k,L-k}\xi^{\kappa}_{k,L-k}|\mathcal{F}_k]\xi^{\alpha}_{L-k,L}\xi^{\kappa}_{L-k,L}]. \tag{5.10}$$

Substituting (5.7) and (5.9) into (5.5) and using the linearity of expectation, we obtain

$$\begin{aligned}\mathbb{E}^x[\mathrm{Var}(\Theta|\mathcal{F}_k)] &= \mathbb{E}^x[\Theta^2] - \mathbb{E}^x[\Theta_k^2] \\ &= L! \sum_{\alpha \in P_h} \mathbb{E}^x[\xi^{\alpha}_{0,k}\xi^{\kappa}_{0,k}(\mathbb{E}[\xi^{\alpha}_{k,L-k}\xi^{\kappa}_{k,L-k}|\mathcal{F}_k] - \mathbb{E}[\xi^{\alpha}_{k,L-k}|\mathcal{F}_k]\mathbb{E}[\xi^{\kappa}_{k,L-k}|\mathcal{F}_k])\xi^{\alpha}_{L-k,L}\xi^{\kappa}_{L-k,L}].\end{aligned}$$

The term, $\mathbb{E}[\xi^{\alpha}_{k,L-k}\xi^{\kappa}_{k,L-k}|\mathcal{F}_k] - \mathbb{E}[\xi^{\alpha}_{k,L-k}|\mathcal{F}_k]\mathbb{E}[\xi^{\kappa}_{k,L-k}|\mathcal{F}_k]$ is a covariance. If two paths are non-intersecting between step k and step $L-k$ on the paths then $\xi^{\alpha}_{k,L-k}$ and $\xi^{\kappa}_{k,L-k}$ are independent random variables, so the covariance term is 0.

Let P_h^{κ} be the set of all paths that intersect (at least once) the path we denoted κ in between step k and step $L-k$. Then, on these paths the covariance term is not 0. When summing over these paths, by removing all the negative terms, recomposing $\xi^{\kappa}_{k,L-k}$ and $\xi^{\alpha}_{k,L-k}$ and using the tower law we have that

$$\mathbb{E}^x[\mathrm{Var}(\Theta|\mathcal{F}_k)] \leqslant L! \sum_{\alpha \in P_h^{\kappa}} \mathbb{E}^x[\xi^{\alpha}_{0,k}\xi^{\kappa}_{0,k}\mathbb{E}[\xi^{\alpha}_{k,L-k}\xi^{\kappa}_{k,L-k}|\mathcal{F}_k]\xi^{\alpha}_{L-k,L}\xi^{\kappa}_{L-k,L}] = L! \sum_{\alpha \in P_h^{\kappa}} \mathbb{E}^x[\xi^{\alpha}_{0,L}\xi^{\kappa}_{0,L}].$$

We now find another upper bound for this. First let $0 \leqslant p, q \leqslant L$. Then, define $I_{p,q}$ to be the set of all paths such that they coincide completely with κ for the first $p+1$ nodes (including the root of the hypercube), then they are completely different to κ for the next $L-p-q-1$ nodes and then they coincide completely with κ for the final $q+1$ nodes (including the node $(1,1,...,1)$.

Using this we have that

$$\mathbb{E}^x[\mathrm{Var}(\Theta|\mathcal{F}_k)] \leqslant L! \sum_{\alpha \in P_h^{\kappa}} \mathbb{E}^x[\xi^{\alpha}_{0,L}\xi^{\kappa}_{0,L}] \leqslant L! \sum_{\alpha \in P_h} \mathbb{E}^x[\xi^{\alpha}_{0,L}\xi^{\kappa}_{0,L}] - L! \sum_{p=0}^{k-1}\sum_{q=0}^{k-1}\sum_{\alpha \in I_{p,q}} \mathbb{E}^x[\xi^{\alpha}_{0,L}\xi^{\kappa}_{0,L}], \tag{5.11}$$

where the second inequality is true because in the sum at first we are removing all possible paths that are different to κ between step k and $L-k$, whereas in the sum on the right hand side we are just removing some of the possible paths that are different to κ. For example, fix $s \in \{1, ..., k-2\}$, in the sum on the right-hand-side we are not removing paths that are the different to κ for the first s nodes, the nodes $\{k, ..., L-k\}$ and the final s nodes, but coincide everywhere else.

Recall that $L! \sum_{\alpha \in P_h} \mathbb{E}^x[\xi^{\alpha}_{0,L}\xi^{\kappa}_{0,L}] = \mathbb{E}^x[\Theta^2]$, also recall that we are interested in calculating $\limsup_{L\to\infty} \frac{1}{L^2}\mathbb{E}^x[\mathrm{Var}(\Theta|\mathcal{F}_k)]$, so dividing by L^2 using [2] we have that

$$\lim_{L\to\infty} \frac{L!}{L^2} \sum_{\alpha \in P_h} \mathbb{E}^x[\xi^{\alpha}_{0,L}\xi^{\kappa}_{0,L}] = \lim_{L\to\infty} \frac{1}{L^2}\mathbb{E}^x[\Theta^2] = \lim_{L\to\infty} \mathbb{E}^x[\frac{\Theta^2}{L^2}] \to 3e^{-2X} + e^{-2X} = 4e^{-2X}. \tag{5.12}$$

Now, we turn our attention to the second term on the right hand side of (5.11). We will first show that

$$\mathbb{E}^x[\xi^\alpha_{0,L}\xi^\kappa_{0,L}] = \frac{(1-x)^{2L-p-q-2}}{(2L-p-q-2)!}\binom{2L-2p-2q-2}{L-p-q-1}. \tag{5.13}$$

Note that $\mathbb{E}^x[\xi^\alpha_{0,L}\xi^\kappa_{0,L}]$ is simply the probability that both of these paths are open. Excluding the root and the end node, the paths α and κ each contain $L-1$ nodes and they overlap on $p+q$ nodes, so the number of different nodes in total in both the paths combined is $2L-p-q-2$. For this path to be increasing, the label attached to each vertex must be greater than x, which happens with probability $(1-x)^{2L-p-q-2}$ since the labels are independent and follow a $U[0,1]$ uniform distribution.

Next, we need the first p nodes to be the p smallest nodes in increasing order of which there is only one possible way for this to happen, we need the final q nodes to be the q largest nodes and in increasing order and there is only one possible way for this to happen. We need the 2 sections of the paths where they do not coincide, each of length $L-p-q-1$ to also be in increasing order. This can happen in $\binom{2L-2p-2q-2}{L-p-q-1}$ ways because that is the number of ways to split the middle $2L-2p-2q-2$ labels into 2 evenly sized blocks the $L-p-q-1$ and for each way to split the labels there is only one way for both of sets of labels to be ordered in an increasing way.

From Hegarty and Martinsson [7], we define the following useful number.

Definition 5.1. For $n,k \in \mathbb{N}$ with $k \leqslant n$, we define $T(n,k)$ to be the number of ways of permuting the set $\{1,...,n\}$ with k components, where for a permutation $\pi_1,\pi_2,...,\pi_n$, the number of components is the largest $s \in \mathbb{N}$ such that $\pi_1,\pi_2,...,\pi_s$ is a permutation of $\{1,...,s\}$.

Remark 5.2. From [7], on the n-dimensional hypercube, let i the identity path denote the path in which we add a 1 in position i. Let j be any other path represented by $\nu_1,\nu_2,...,\nu_n$, where for step s we add a 1 in position ν_s. Then if $\nu_1,\nu_2,...,\nu_n$ has k components, i and j intersect in $k-1$ interior nodes. Under a suitable isomorphism, we can see that $T(n,k)$ is the number of possible paths that share $k-1$ vertices. Therefore, by viewing the $L-p-q$ steps after the first p shared nodes of two paths α and κ as 2 smaller sub paths, to calculate the number of paths in $I_{p,q}$, it is equivalent to calculate $B((L-p-q) := T_\kappa(L-p-q,1) = T(L-p-q,1)$.

We use the following 2 propositions from [7] to show that $B(n) \sim n!$.

Proposition 5.3. $B(n)$ is uniquely defined by,

$$n! = \sum_{i=1}^{n} B(i)(n-i)!. \tag{5.14}$$

Proof. There are $n!$ possible paths along the n-dimensional hypercube. Let γ be a path on the n-dimensional hypercube. Then you can view this sum as a sum over the steps of the path, where we are summing the number of paths that intersect γ for the first time at that step. Since then the number of paths ν on the n-dimensional hypercube that for the first time, intersect γ at step i is $B(i)(n-i)! = T(i,1)(n-i)!$, this gives the result. □

Proposition 5.4.

$$n!(1-\mathcal{O}(1/n)) \leqslant B(n) \leqslant n!. \tag{5.15}$$

Proof. Since $n!$ is the number of possible paths on the n-dimensional hypercube, it follows immediately that $B(n) \leqslant n!$. So it remains to prove the lower bound. Rearranging Proposition 5.3 by pulling the

$B(n)$ term out of the sum we see that

$$B(n) = n! - \sum_{i=1}^{n-1} B(i)(n-i)! \geqslant n! - \sum_{i=1}^{n-1} i!(n-i)! = n!(1 - \sum_{i=1}^{n-1} \frac{i!(n-i)!}{n!}).$$

For $i = 1$, we have that $\frac{i!(n-i)!}{n!} = 1/n$. Since $\frac{i!(n-i)!}{n!}$ is symmetric in i, without loss of generality assume that $i \leqslant n - i$. For $i > 1$ notice that

$$\frac{i!(n-i)!}{n!} = \left(\frac{1}{1} \cdot \frac{2}{2} \cdots \frac{i}{i}\right)\left(\frac{1}{n} \cdot \frac{2}{n-1} \cdot \frac{3}{i+1} \cdot \frac{4}{i+2} \cdots \frac{n-i}{n-2}\right) \leqslant \frac{1}{n} \cdot \frac{2}{n-1} \leqslant \frac{4}{n^2} = \mathcal{O}(1/n^2),$$

we see that every term in $\sum_{i=2}^{n-2} \frac{i!(n-i)!}{n!}$ is $\mathcal{O}(1/n^2)$. Since the sum contains $n-3$ terms, the sum itself is $\mathcal{O}(1/n)$ and hence,

$$B(n) \geqslant n!(1 - 2\frac{1!(n-1)!}{n!} - \sum_{i=2}^{n-2} \frac{i!(n-i)!}{n!}) = n!(1 - \mathcal{O}(1/n)).$$

□

Now, we know that $|I_{p,q}| = B(L-p-q)$ and by expanding the binomial coefficient in the following we obtain,

$$\begin{aligned} L! \sum_{\alpha \in I_{p,q}} \mathbb{E}^x[\xi^\alpha_{0,L} \xi^\kappa_{0,L}] &= L! \frac{(1-x)^{2L-p-q-2}}{(2L-p-q-2)!} \binom{2L-2p-2q-2}{L-p-q-1} \\ &= \frac{L!}{(L-p-q-1)!} \cdot \frac{(2L-2p-2q-2)!}{(2L-p-q-2)!} \cdot \frac{B(L-p-q)}{(L-p-q-1)!} \cdot (1-x)^{2L-p-q-2}. \end{aligned}$$

This has been written in a very specific way, so we can look at what each of the 4 terms look like individually when L is large. The first is like L^{p+q+1}, the second like $(2L)^{-p-q}$, the third like L by Proposition 5.4 and the fourth is like e^{-2X}. Using this, we have that in the following limit, all the L terms cancel,

$$\lim_{L\to\infty} \frac{L!}{L^2} \sum_{\alpha \in I_{p,q}} \mathbb{E}^x[\xi^\alpha_{0,L} \xi^\kappa_{0,L}] = \frac{e^{-2X}}{2^{p+q}}.$$

By the algebra of limits and the geometric series formula,

$$\begin{aligned} \lim_{L\to\infty} \frac{L!}{L^2} \sum_{p=0}^{k-1} \sum_{q=0}^{k-1} \sum_{\alpha \in I_{p,q}} \mathbb{E}^x[\xi^\alpha_{0,L} \xi^\kappa_{0,L}] &= e^{-2X} \sum_{p=0}^{k-1} \sum_{q=0}^{k-1} \frac{1}{2^{p+q}} = e^{-2X} \left(\frac{1 - \left(\frac{1}{2}\right)^k}{1 - \frac{1}{2}}\right)^2 \\ &= 4e^{-2X}(1 - (1/2)^{k+1} + (1/4)^k) \end{aligned}$$

Combining this with (5.12) through (5.11) yields,

$$\limsup_{L\to\infty} \frac{1}{L^2} \mathbb{E}^x[\mathrm{Var}(\Theta|\mathcal{F}_k)] \leqslant 4e^{-2X} - 4e^{-2X} - \frac{4e^{-2X}}{2^{k+1}} + \frac{4e^{-2X}}{4^k} = 8e^{-2X}(\frac{1}{2^k} - \frac{1}{2 \cdot 4^k}) \leqslant \frac{8e^{-2X}}{2^k},$$

as required.

5.3 Proof of Theorem 1.7 - Step 2.

Recall we have that

$$\Theta_k = \sum_{|\sigma|=k} \sum_{|\tau|=L-k} n_\sigma m_\tau \mathbb{1}(\sigma \leftrightarrow \tau)\mathbb{1}(x_\sigma + y_\tau < 1)(L-2k)(1-y_\tau - x_\sigma)^{L-2k-1}. \tag{5.16}$$

We now define the following quantity and show that it can be written as the product of a sum over nodes of distance k away from the root and a sum over the nodes of distance k away from the end node. To this end, define

$$\tilde{\Theta}_k := \sum_{|\sigma|=k} \sum_{|\tau|=L-k} n_\sigma m_\tau L(1-y_\tau - x_\sigma + x_\sigma y_\tau)^{L-2k-1}. \tag{5.17}$$

Note that this expression, in comparison to (5.16) has neither of the indicator functions in the summand, it also has a factor L which is greater than the factor of $L-2k$ in the other, it also has an extra $x_\sigma y_\tau$ in the term in brackets. Therefore, the summand of (5.17) is greater than the summand of (5.16). So, $\tilde{\Theta}_k \geqslant \Theta_k$.
Also, $\mathbb{E}^{\frac{X}{L}}[\Theta_k/L] = \mathbb{E}^{\frac{X}{L}}[\Theta] = (1-\frac{X}{L})^{L-1}$, so

$$\lim_{L\to\infty} \mathbb{E}^{\frac{X}{L}}\left[\frac{\Theta_k}{L}\right] = e^{-X}. \tag{5.18}$$

Now we compute $\lim_{L\to\infty} \mathbb{E}^{\frac{X}{L}}\left[\frac{\tilde{\Theta}_k}{L}\right]$ and show that this the same as for Θ_k. Firstly, note that

$$\mathbb{E}^x[n_\sigma | x_\sigma] = k(x_\sigma - x)^{k-1}\mathbb{1}(x_\sigma \geqslant x),$$

this is because there are k possible paths to σ, for any path to be increasing we need $x_\sigma \geqslant x$ hence the indicator function and finally the probability a path is increasing is $(x_\sigma - x)^{k-1}$. Similarly we see that

$$\mathbb{E}^x[m_\tau | y_\tau] = k(y_\tau)^{k-1}.$$

From this we see that

$$\begin{aligned}
\mathbb{E}^{\frac{X}{L}}\left[\frac{\tilde{\Theta}_k}{L}\right] &= \binom{L}{k}^2 \int_x^1 dx_\sigma \int_0^1 k^2 (x_\sigma - x)^{k-1}(y_\tau)^{k-1}(1-y_\tau - x_\sigma + x_\sigma y_\tau)^{L-2k-1} dy_\tau \\
&= \left[\frac{L!(L-2k-1)!}{(L-k)!(L-k-1)!}\right]^2 (1-x)^{L-k-1}.
\end{aligned}$$

The integral is computed through the use of the formula in Section 4.4. Taking the limit as $L \to \infty$ we obtain,

$$\lim_{L\to\infty} \mathbb{E}^{\frac{X}{L}}\left[\frac{\tilde{\Theta}_k}{L}\right] = e^{-X}. \tag{5.19}$$

Now, note that $\tilde{\Theta}_k/L - \Theta_k/L$ is a non-negative random variable where, by the linearity of expectation, $\lim_{L\to\infty}[\tilde{\Theta}_k/L - \Theta_k/L] = e^{-X} - e^{-X} = 0$. Now, we can use Slutsky's Theorem from Karr [10] because $X_L = \tilde{\Theta}_k/L - \Theta_k/L$ converges in probability to 0 and $Y_l = \Theta_k/L$ converges in distribution to e^{-X} we have that the sum of these 2 random variables, $\tilde{\Theta}_k/L$, converges to e^{-X} in distribution.

We see that

$$\begin{aligned}
\frac{\tilde{\Theta}_k}{L} &= \sum_{|\sigma|=k} \sum_{|\tau|=L-k} n_\sigma m_\tau (1 - y_\tau - x_\sigma + x_\sigma y_\tau)^{L-2k-1} \\
&= \sum_{|\sigma|=k} \sum_{|\tau|=L-k} n_\sigma m_\tau ((1-y_\tau)(1-x_\sigma))^{L-2k-1} \\
&= \left(\sum_{|\sigma|=k} n_\sigma (1-x_\sigma)^{L-2k-1} \right) \left(\sum_{|\tau|=L-k} m_\tau (1-y_\tau)^{L-2k-1} \right). \qquad (5.20)
\end{aligned}$$

Notice that this is the product of 2 independent sums, the first of which is over the first k steps of the hypercube and the second of which is over the final k steps. Note that because of the presence of the n_σ term, the first sum is dependent on the starting value of the root vertex. The second sum has the same distribution as the first sum with root vertex value $x = 0$.

5.4 Proof of Theorem 1.7 - Step 3.

We begin by defining ϕ_k to be the first sum in (5.20).

Definition 5.5.

$$\phi_k := \sum_{|\sigma|=k} n_\sigma (1-x_\sigma)^{L-2k-1}. \qquad (5.21)$$

The plan is to show that this converges weakly to an exponential distribution multiplied by e^{-X} by comparing the generating function of this to the one found on the irregular tree for Θ_k/L. Constructing a recursive definition for the generating function of ϕ_k is not possible because two paths on the hypercube may share an edge after they split for the first time, whereas this is not possible on the tree. To get around this problem, we define a new random variable.

Definition 5.6.

$$\tilde{\phi}_k(b) := \sum_{|\sigma|=k} \tilde{n}_\sigma(b)(1-x_\sigma)^L, \qquad (5.22)$$

where b is a set of '*forbidden bits*', a subset of the set of all L dimensional binary vectors, and $\tilde{n}_\sigma(b)$ is either 1 or 0 depending on whether or not there is an open '*interesting path*' (defined below) through node σ. We denote B to be the size of b.

We define an '*interesting path*' to σ in the following way. At the origin there are $L - B$ nodes that are not forbidden, the nodes with label less than $\ln(L)/L$ we are interested in. Defining b' (with size B') to be the set of all these nodes, we now forbid these bits and at the next level for each of the interesting nodes, we look at the $L - B - B'$-dimensional sub-hypercube and repeat this construction until we reach level k. All the possible paths found are called interesting.

By this construction, note that if $B \geqslant L - k$, there can be no interesting paths because there are too many forbidden bits. Also, there are no interesting paths if $x_\sigma > \ln(L)/L$. Since there are L terms in a path from the origin to $(1, 1, ..., 1)$ and we expect an increasing path to increase linearly, we expect x_σ to be like k/L on an increasing path, so we are not concerned that there are no interesting paths when $x_\sigma > \ln(L)/L$, because we would not expect there to be many increasing paths. Also by this construction, the interesting paths are all independent.

For each interesting path, we expect there to be $k\mathcal{O}(\ln(L))$ forbidden bits, but as $L \to \infty$ this is insignificant because $L \gg \ln(L)$.

Note also that

$$\begin{aligned}
\tilde{\phi}_k(b) &= \sum_{|\sigma|=k} \tilde{n}_\sigma(b)(1-x_\sigma)^L &&= \sum_{\rho\in b'} \sum_{\substack{|\sigma|=k,\\ \rho\leqslant\sigma}} \tilde{n}_\sigma(b)(1-x_\sigma)^L \\
&= \sum_{\rho\in b'} \mathbb{1}_{\{x\leqslant x_\rho\}} \sum_{\substack{|\sigma|=k,\\ \rho\leqslant\sigma}} \tilde{n}_\sigma(b)(1-x_\sigma)^L &&= \sum_{\rho\in b'} \mathbb{1}_{\{x\leqslant x_\rho\}} \tilde{\phi}_{k-1}(B+B', x_\rho),
\end{aligned} \tag{5.23}$$

where $\tilde{\phi}_{k-1}(B+B', x_\rho)$ means the root value is now x_ρ. Using the above we will construct the recursion on $\tilde{\phi}_k(b)$ initialised by $\tilde{\phi}_0(b) = (1-x)^L$. Note also that $\tilde{n}_\sigma \leqslant n_\sigma$. We will show that in the large L limit with starting value X/L, the expected value of $\tilde{\phi}_k(b)$ is equal to the expected value of ϕ_k. Through the use of Slutsky's Theorem again tell us that when the large L limit of $\tilde{\phi}_k(b)$ exists, the large L limit of ϕ_k also exists and they follow the same distribution.

To this end, note that since there are $L-B$ potential interesting nodes at level 1 and they are interesting with probability $\ln(L)/L$, $B' \sim \text{Binomial}(L-B, \ln(L)/L)$. Also note that which nodes are forbidden is not necessarily important, only the number of nodes that are, so we can equally write $\tilde{\phi}_k(B)$ instead of $\tilde{\phi}_k(b)$. Now from the PMF of a binomial distribution we see that

$$\mathbb{P}[B'=i] = \binom{L-B}{i}\Big(\frac{\ln(L)}{L}\Big)^i\Big(1-\frac{\ln(L)}{L}\Big)^{L-B-i}. \tag{5.24}$$

Then, noting that if $k > L-B$ then there are no interesting paths of length k. Taking the expectation over (5.23) with starting value x we obtain

$$\mathbb{E}^x[\tilde{\phi}_k(b)] = \sum_{i=0}^{L-B-k+1} \mathbb{P}[B'=i]\cdot i \int_0^{\frac{\ln(L)}{L}} \frac{L}{\ln(L)} \mathbb{1}_{\{x\leqslant y\}} \tilde{\phi}_{k-1}(B+i, y)dy. \tag{5.25}$$

The proof for this is a simplified version of the proof of Proposition 5.8 which we will prove later. Note that were $k > L-B$ then the left hand side of the above would be 0. Now, we show that we can write this as

$$\mathbb{E}^x[\tilde{\phi}_k(b)] = \frac{(L-B)!}{(L-B-k)!L^k}\psi_k(x, L) \tag{5.26}$$

for some function $\psi_k(x, L)$. We prove this by induction, for $k=0$ this holds because $\tilde{\phi}_0(b) = (1-x)^L$. For the inductive hypothesis, assume 5.26 holds for $k-1$, so that

$$\mathbb{E}^x[\tilde{\phi}_k(b)] = \frac{1}{L^{k-1}} \sum_{i=0}^{L-B-k+1} \mathbb{P}[B'=i]\cdot \frac{(L-B)!}{(L-B-k)!L^k} i \int_0^{\frac{\ln(L)}{L}} \frac{L}{\ln(L)} \psi_k(y, L)dy. \tag{5.27}$$

Now look closer at the sum to see that

$$\begin{aligned}
\sum_{i=0}^{L-B-k+1} \mathbb{P}[B'=i]\cdot \frac{(L-B)!}{(L-B-k)!L^k} i &= \frac{\ln(L)}{L}\Big(1-\frac{\ln(L)}{L}\Big)^{k-1} \frac{(L-B)!}{(L-B-k+1)!} \\
&\quad\cdot \sum_{i=0}^{L-B-k+1} \frac{(L-B-k+1)!}{(L-B-k+1-i)!i!}\Big(\frac{\ln(L)}{L}\Big)^i\Big(1-\frac{\ln(L)}{L}\Big)^{L-B-k+1-i} \\
&= \frac{\ln(L)}{L}\Big(1-\frac{\ln(L)}{L}\Big)^{k-1} \frac{(L-B)!}{(L-B-k+1)!},
\end{aligned}$$

where the final equality comes from the use of the binomial formula. The rest of this section of the proof is omitted due to space constraints but we explain the steps to complete it. From above we have the following recurrence with initialisation,

$$\psi_k\Big(\frac{X}{L}, L\Big) = \Big(1 - \frac{\ln(L)}{L}\Big)^{k-1} \int_X^{\ln(L)} \psi_{k-1}\Big(\frac{Y}{L}, L\Big)dy, \quad \psi_0\Big(\frac{X}{L}, L\Big) = \Big(1 - \frac{X}{L}\Big)^L. \tag{5.28}$$

This recurrence yields an upper bound of e^{-X} for $\psi_k(X/L, L)$. Then using the dominated convergence theorem and another recurrence that for a function $b(L) = o(L)$, the limit of $\mathbb{E}^{X/L}[\tilde{\phi}_k(b(L))]$ is e^{-X} in the large L limit. Showing that in the large L limit, $\tilde{\phi}_k(b)$ and $\phi_k(b)$ have the same distribution.

All that remains now is to compute the generating function of $\tilde{\phi}_k(b)$. With a similar aim as in the proof of Theorem 1.6, we begin with the following definition,

Definition 5.7. Consider the L-dimensional hypercube with origin label x. Then, for $\mu \geqslant 0$ define the generating function of $\tilde{\phi}_k(B)$ to be

$$G_k(\mu, x, L, b) := \mathbb{E}^x[e^{-\mu\tilde{\phi}_k(B)}]. \tag{5.29}$$

Proposition 5.8. Fix $X \geqslant 0$ and consider the L-dimensional hypercube with origin label $x = X/L$. Then,

$$G_k\Big(\mu, \frac{X}{L}, L, b\Big) = \sum_{i=0}^{L-B-k+1} \mathbb{P}[B' = i]\Bigg[1 - \frac{1}{\ln(L)}\int_X^{\ln(L)} [1 - G_{k-1}\Big(\mu, \frac{Y}{L}, L, B+i\Big)]dy\Bigg]^i. \tag{5.30}$$

Proof. Using 5.23 and taking the expectation we see that

$$\begin{aligned}
\mathbb{E}^x\Big[\exp\{-\mu\tilde{\phi}_k(B)\}\Big] &= \mathbb{E}^x\Big[\exp\{-\mu\sum_{\rho\in b'} \mathbb{1}_{\{x\leqslant x_\rho\}}\tilde{\phi}_{k-1}(B+B', x_\rho)\}\Big] \\
&= \sum_{i=0}^{L-B-k+1} \mathbb{P}[B' = i]\mathbb{E}^x\Big[\exp\{-\mu\sum_{j=0}^{i} \mathbb{1}_{\{x\leqslant x_j\}}\tilde{\phi}_{k-1}(B+j, x_j)\} \\
&\qquad\qquad \mid x_1, x_2, ..., x_j \leqslant \ln(L)/L\Big] \\
&= \sum_{i=0}^{L-B-k+1} \mathbb{P}[B' = i]\mathbb{E}^x\Big[\prod_{j=0}^{i}\exp\{-\mu\mathbb{1}_{\{x\leqslant x_j\}}\tilde{\phi}_{k-1}(B+j, x_j)\} \\
&\qquad\qquad \mid x_1, x_2, ..., x_j \leqslant \ln(L)/L\Big] \\
&= \sum_{i=0}^{L-B-k+1} \mathbb{P}[B' = i]\prod_{j=0}^{i}\int_0^{\frac{\ln(L)}{L}} \frac{L}{\ln(L)}\Big[\exp\{-\mu\mathbb{1}_{\{x\leqslant y\}}\tilde{\phi}_{k-1}(B+j, y)\}\Big]dy \\
&= \sum_{i=0}^{L-B-k+1} \mathbb{P}[B' = i]\Big[\frac{xL}{\ln(L)} + \frac{L}{\ln(L)}\int_x^{\frac{\ln(L)}{L}} \exp\{-\mu\tilde{\phi}_{k-1}(B+i, y)\}dy\Big]^i \\
&= \sum_{i=0}^{L-B-k+1} \mathbb{P}[B' = i]\Big[1 - \frac{L}{\ln(L)}\int_x^{\frac{\ln(L)}{L}} [1 - G_{k-1}(\mu, y, L, B+i)]dy\Big]^i.
\end{aligned}$$

Using a sufficient substitution gives the result of the proposition. □

Now, [2] claim the following upper and lower bounds for $G_k(\mu, \frac{X}{L}, L, b)$. Then the strategy from there would be to use a similar technique as in the proof of Theorem 1.6 to show that $G_k(\mu, X/L, L, b)$ has a limit as $L \to \infty$, and the equation for this generating function is the same as that in the proof of Theorem 1.6. We show the bounds given by [2] and explain why they do not hold. Firstly, [2] give the following lower bound,

$$G_k(\mu, X/L, L, b) \geqslant [1 - \frac{1}{L}\int_X^{\ln(L)} [1 - G_{k-1}(\mu, \frac{Y}{L}, L, B)]dy]^{L-B}. \tag{5.31}$$

Now, viewing $G_k(\mu, X/L, L, b)$ as a function of b, we see that it is increasing. Therefore, the smallest term in the large square brackets in (5.30) is when $i = 0$. Since the large square brackets is between 0 and 1, raising this to the largest power in the sum makes it smallest. This is shown by the following equalities,

$$\begin{aligned}
G_k(\mu, X/L, L, b) &\geqslant \sum_{i=0}^{L-B-k+1} \mathbb{P}[B' = i]\left[1 - \frac{1}{\ln(L)}\int_X^{\ln(L)} [1 - G_{k-1}\Big(\mu, \frac{Y}{L}, L, B+i\Big)]dy\right]^i &(5.32)\\
&\geqslant \sum_{i=0}^{L-B-k+1} \mathbb{P}[B' = i]\left[1 - \frac{1}{\ln(L)}\int_X^{\ln(L)} [1 - G_{k-1}\Big(\mu, \frac{Y}{L}, L, B\Big)]dy\right]^i &(5.33)\\
&\geqslant \sum_{i=0}^{L-B-k+1} \mathbb{P}[B' = i]\left[1 - \frac{1}{\ln(L)}\int_X^{\ln(L)} [1 - G_{k-1}\Big(\mu, \frac{Y}{L}, L, B\Big)]dy\right]^{L-B-k+1} &(5.34)\\
&\geqslant (1 - o(1))\left[1 - \frac{1}{\ln(L)}\int_X^{\ln(L)} [1 - G_{k-1}\Big(\mu, \frac{Y}{L}, L, B\Big)]dy\right]^{L-B-k+1}. &(5.35)
\end{aligned}$$

Now, note the differences between this expression and (5.31). Firstly, the integral is multiplied by $1/\ln(L)$ in our equation and is multiplied by $1/L$ in (5.31). Now, the $1/L$ in (5.31) makes the inside of the large square brackets larger (because we are subtracting something smaller from 1) instead of smaller which is a problem. It is not immediately clear where the $1/L$ came from in (5.31), but it is necessary for their argument as when taking the limit as $L \to \infty$ we would obtain e to the power of the negative of the integral, which is what they need to complete their argument. In an attempt to remedy the problem of the presence of the $1/L$ making inside brackets too large, if we kept it as $1/\ln(L)$ we would now obtain the incorrect limit for their consequent arguments.

Furthermore, in (5.31) there is no sum present. This is because the limits in the sum are slightly incorrect, running from $i = 0$ to $L - b$ and then using that the sum does indeed sum to 1. However, sum runs from $i = 0$ to $L - B - k + 1$, so in fact is $1 - o(1)$. This does converge to 1 in the large L limit though, so would not affect further calculations.

Next, [2] give an upper bound. This is obtained by splitting the sum into two sections, one sum from $i = 0$ to $\lfloor(\ln(L))^2\rfloor$ and the second running from $i = \lfloor(\ln(L))^2\rfloor + 1$ to $L - B - k + 1$. They then bound the first sum above to what they desire in the limit as $L \to \infty$ for their subsequent arguments and they show the second sum is goes to 0 in the large L limit. So, [2] claim that

$$G_k(\mu, X/L, L, b) \leqslant [1 - \frac{1}{L}\int_X^{\ln(L)} [1 - G_{k-1}\Big(\mu, \frac{Y}{L}, L, B\Big)]dy]^{L-B} + \sum_{i=\lfloor(\ln(L))^2\rfloor+1}^{L-B} \mathbb{P}[B' = i]. \tag{5.36}$$

Now, through our own calculations we have that

$$G_k(\mu, X/L, L, b) \leqslant \sum_{i=0}^{\lfloor(\ln(L))^2\rfloor} \mathbb{P}[B'=i]\left[1-\frac{1}{L}\int_X^{\ln(L)}[1-G_{k-1}\Big(\mu,\frac{Y}{L},L,B+\lfloor(\ln(L))^2\rfloor\Big)]dy\right]^i + \sum_{i=\lfloor(\ln(L)\rfloor)^2+1}^{L-B-k+1}\mathbb{P}[B'=i]. \quad (5.37)$$

Similar to [2], the second sum does indeed become vanishingly small. However, it is the first expression that is different. Note that unlike before, here it is okay to multiply the integral by $1/L$ because we want the largest possible term in the square brackets. Also, because $G_k(\mu, \frac{X}{L}, L, b)$ is an increasing function we can substitute the largest value of the sum into it to obtain the largest value in the square brackets. This largest value will be between 0 and 1, so unlike (5.36) we do not raise the large brackets to a high power, as this would make the term smaller. Were we to remove the power this would not work as it would no longer give the correct limit for the subsequent arguments in [2].

We could attempt to bound this again in a different way, by raising the square brackets to the power $\ln(L)$ and keeping the integral also multiplied by $1/\ln(L)$ because this would give the desired limit. Note that also $\mathbb{E}[B'] = \ln(L)$ which is another reason why this would be a good approach. However, defining z to be what is inside the square brackets and taking L to be sufficiently large we see that

$$\begin{aligned} G_k(\mu, X/L, L, b) &\leqslant \sum_{i=0}^{\lfloor(\ln(L))^2\rfloor} \mathbb{P}[B'=i]\frac{z^{\ln(L)}}{z^{\ln(L)-i}} + \sum_{i=\lfloor(\ln(L)\rfloor)^2+1}^{L-B-k+1}\mathbb{P}[B'=i] \\ &= \sum_{i=0}^{\lfloor(\ln(L))^2\rfloor}\binom{L-B}{i}\Big(\frac{\ln(L)}{L}\Big)^i\Big(1-\frac{\ln(L)}{L}\Big)^{L-B-i}\frac{z^i}{z^{\ln(L)}}z^{\ln(L)} + \sum_{i=\lfloor(\ln(L)\rfloor)^2+1}^{L-B-k+1}\mathbb{P}[B'=i], \end{aligned}$$

looking at the $i = \ln(L)$ term specifically we see that at least,

$$\binom{L-B}{\ln(L)}\Big(\frac{\ln(L)}{L}\Big)^{\ln(L)}\Big(1-\frac{\ln(L)}{L}\Big)^{L-B-\ln(L)}\frac{z^{\ln(L)}}{z^{\ln(L)}} = \mathcal{O}\Big(L^{\ln(L)}\frac{1}{(\ln(L))^{\ln(L)}}\Big) \quad (5.38)$$

which is clearly too large and will not give the correct limit. This fix will not work.

Thus we have shown that the proof given is incorrect and examples of simple fixes that fail to work. This completes the section on the proof, and thus the project.

Bibliography

[1] Marcelo Sampaio de. Alencar. *Probability theory*. eng. Communications and signal processing collection. New York, [New York] (222 East 46th Street, New York, NY 10017): Momentum Press, 2016. ISBN: 1-60650-748-6.

[2] Julien Berestycki, Éric Brunet, and Zhan Shi. "The number of accessible paths in the hypercube". In: *Bernoulli* 22.2 (2016), pp. 653–680.

[3] Vladimir Igorevich Bogachev and Maria Aparecida Soares Ruas. *Measure theory*. Vol. 1. Springer, 2007.

[4] Jasper Franke et al. "Evolutionary accessibility of mutational pathways". In: *PLoS computational biology* 7.8 (2011), e1002134.

[5] John H Gillespie. "Some properties of finite populations experiencing strong selection and weak mutation". In: *The American Naturalist* 121.5 (1983), pp. 691–708.

[6] Bill Hathaway. *In evolution, 'house of cards' model wins*. 2015. URL: `https://news.yale.edu/2015/05/13/evolution-house-cards-model-wins` (visited on 04/19/2022).

[7] Peter Hegarty and Anders Martinsson. "On the existence of accessible paths in various models of fitness landscapes". In: *The Annals of Applied Probability* 24.4 (2014), pp. 1375–1395.

[8] Andrea Hodgins-Davis, Daniel P Rice, and Jeffrey P Townsend. "Gene expression evolves under a house-of-cards model of stabilizing selection". In: *Molecular Biology and Evolution* 32.8 (2015), pp. 2130–2140.

[9] Jean-Pierre Kahane. *Some random series of functions*. Vol. 5. Cambridge University Press, 1993.

[10] Alan F Karr. *Probability*. eng. Springer texts in statistics. New York, NY: Springer New York : Imprint: Springer, 1993. ISBN: 9781461208914.

[11] Motoo Kimura. "On the probability of fixation of mutant genes in a population". In: *Genetics* 47.6 (1962), p. 713.

[12] Stefan Nowak and Joachim Krug. "Accessibility percolation on n-trees". In: *EPL (Europhysics Letters)* 101.6 (2013), p. 66004.

[13] Herbert Robbins. "A remark on Stirling's formula". In: *The American mathematical monthly* 62.1 (1955), pp. 26–29.

[14] Matthew Roberts and Lee Zhao. "Increasing paths in regular trees". In: *Electronic Communications in Probability* 18 (2013), pp. 1–10.

[15] Houshang H Sohrab. *Basic real analysis*. Vol. 231. Springer, 2003.

[16] Michael Spivak. *Calculus*. eng. 3rd ed. Cambridge: Cambridge University Press, 1994.

[17] Alex Tsun. *Chapter 6. Concentration Inequalities*. URL: `https://courses.cs.washington.edu/courses/cse312/20su/files/student_drive/6.2.pdf` (visited on 04/19/2022).

[18] Robert L Wolpert. *8 The Laws of Large Numbers*. URL: `https://www2.stat.duke.edu/courses/Fall17/sta711/lec/wk-08.pdf` (visited on 04/19/2022).

[19] Robert L Wolpert. *Conditional expectation*. URL: `https://www2.stat.duke.edu/courses/Fall15/sta711/lec/wk-10.pdf` (visited on 04/19/2022).